



Errors of Indices in Household Surveys of Punjab Urban through Principal Components

Muhammad Imran¹, Fatima Seher Zaidi²

¹ Nanjing University of Science and Technology, China, Email: imrankharal@njust.edu.cn

² University of Nantes, France, Email: fatima.zaidi@etu.univ-nantes.fr

ARTICLE INFO

Article History:

Received: 08 December 2020

Revised: 28 June 2021

Accepted: 30 June 2021

Available Online: 30 June 2021

Keywords:

Wealth Indices

Standard Error of Indices

Household Survey

Principal Component Analysis

Factor Analysis

JEL Classification Codes:

C38, C43, D10, D19, H31, R20,

ABSTRACT

In this research article, household wealth indices are calculated to estimate the sampling errors, which gave us complete information on the quality and reliability of published data upon the household surveys. Estimates are calculated based on simple random sampling, which contains sampling errors. Here principal components analysis (PCA) estimate standard errors of wealth indices as orthogonal transformation to develop solid measures of individual economic status. These measures evaluate the significance and explain the living status and economic dissimilarity of Punjab urban. Instrumental variables are used here to the enlightened social status of the Punjab urban area of Pakistan by using PCA of household surveys. Twenty-five variables are included in this study. Total variance analysis explains the variation of total components. A comparison study (PCA) approaches to estimating the standard error of indices of the household survey are presented in this paper. We conclude that errors of indices in household surveys through PCA, when compared to direct measures of estimating household wealth indices, are an efficient and reliable method.

© 2021 The Authors, Published by iRASD. This is an Open Access Article under the Creative Common Attribution Non-Commercial 4.0

Corresponding Author's Email: imrankharal@njust.edu.cn

1. Introduction

When index series are estimated from sample surveys, the indexes are subject to sampling error, and an important statistical question is how to best estimate their variances. Wealth indices are used to build the socio-economic status. It is important for the report that the data should be accurate and the error should be minimal; our focus is to have minimum variation and error caused by Sampling design and Sampling error to reduce error (Gerhart, Wright et al. 2000, Biemer 2010, Vinaixa, Samino et al. 2012). I used cluster sampling and sampling design. Wealth indices are the socio-economic status. So I constructed the wealth indices by PCA and used cluster sampling to reduce the error caused by sampling design and sampling error (Wamani, Tylleskär et al. 2004, Abeyasekera 2005, Rarani, Rashidian et al. 2017). (Crown 1998, Webley and Nyhus 2006) statistical measures that show alteration in a variable or associated variables concerning time or other characters such as income, expenditure, etc.

Valliant (1992) studied that index series are characterized by seasonal and irregular fluctuations in addition to underlying trends. The literature is replete with methods for decomposing and smoothing such time series. The general idea is to use models to approximate variances. Given a set of survey variables whose variances all follow the same model, parameters of the model are estimated by least squares studied by (Lillacci and Khammash 2010, Lohmöller 2013). Household data of Uganda has been calculated from sources about expenditure patterns in environmental data by Food and Agriculture Organization (FAO)(Smith, Alderman et al. 2006). different estimation techniques have been applied to find out the oblique solution within socioeconomics data to calculate indices of the household survey(Ng 2015). (Filmer, Pritchett et al. 1999) describe different methods and reasons to distinguish between richness and poverty level; for said purpose, household data is taken from Demographic and health surveys. They gave different reasons between rich and

poor peoples, people of an urban area rich then the people of rural area because of discriminations of facilities and vital reassures.(Rutstein and Rojas 2003) suggest that considered Demographic and Health Survey (DHS) surveys to gather household data. They collected household data of different category, like main sources of drinking water, kind of facilities in toilet usage,type of material used in wall, roof etc., main fuel used for cooking , main source of lighting, Public transport, Health clinic or hospital, Family planning Centre etc. Exploratory analysis of the variables you have collected: Rules of thumb: variables Recode the household amenities variable into improved / not improved. The recoding between improved/ not improved is just one possibility (Howard and Bogh 2002, Lillacci and Khammash 2010). The analyst can include only a certain variable/category (only households with a 'household water connection' y/n) to extract/highlight households with a very good water source. (Montana and Spencer 2004) designed DHS survey to show global positioning system (GPS) coordinates at the cluster level, where a cluster is usually a census enumeration area, sometimes a village. The authors claim that "the DIHOPIT model DRAFT used to estimate permanent income has the potential to be modified so that estimates of permanent income can be directly compared across countries" and present three possible technique.

Rogers, Wint et al. (2011) and Townsend (2014) studied that the same poverty index to environmental data begins to break out of this circularity and looks for causes of poverty rather than its consequences. The underlying assumption in this approach is that people in rural settings are poor because their environments fail to provide the goods and services available to richer people. Soil fertility, good health, access to fuel and water all have environmental correlates for which satellite data may act as proxies: people are often poor because of an inadequate supply of these vital resources (McManus, Meltzer et al. 2009, Townsend 2014). By incorporating the driving factors connected with the exceptional deficiency levels, the modelling approach allows by (McManus, Meltzer et al. 2009) (Chambers and urbanization 1995, Sachs 2008).

The paper focuses on indices that are computed using data from a sampling survey. By definition, all estimates from such a survey contain sampling error, and survey statisticians are tasked with estimating these errors. In most cases, indices are calculated using data from panel surveys, in which the same sample is surveyed at successive time intervals. The study concentrated on some of the index's properties as a statistical estimate derived from sampling survey data. The study's main goal was to investigate the index estimator's sampling distribution and compare three different methods for estimating the estimate's sampling variance and standard error. In the first section of the paper, we will introduce some fundamental concepts from the index, sampling theory and PCA. The main findings of the study will then be presented, followed by some conclusions.

Dunteman (1989), Jong and Kotz (1999) studied two basic approaches like component and factor analysis. (PCA) is used to create an oblique solution and orthogonal transformation. We know that correlated variables created by component analysis are due to oblique solution and uncorrelated variables in PCA are due to orthogonal transformation. That first component shows most of the variation and eigenvalue, and the first component consists of most of the correlated variables in the result of the oblique solution are purposed by (Filmer and Pritchett 1998).(Matsunaga 2010, Yong and Pearce 2013) studied the difference between principal and factor analysis; there is little difference between component and factor analysis (FA) and PCA total variance while factor analysis shares variance in the data set. (Jolliffe, Cadima et al. 2016) suggested that different components explained most of the variance and standard error through eigenvalues used to identify the pattern of data.(Wiktorowicz 2016, Soosalu, Henwood et al. 2019) studied that FA are conceptualized as "real world" entities such as depression, anxiety, and disturbed thought.

In PCA, all of the observed variance is analyzed while in factor analysis, only the shared variances are analyzed. The paper focuses on indices that are computed using data from a sampling survey. By definition, all estimates from such a survey contain sampling error, and survey statisticians are tasked with estimating these errors. In most cases, indices are estimated using data from panel surveys, in which the same sample is surveyed at successive time intervals. The study concentrated on some of the index's properties as a statistical estimate derived from sampling survey data. The study's main goal was to investigate the

index estimator's sampling distribution and compare three different methods for estimating the estimate's sampling variance and standard error. In the first section of the paper, we will introduce some fundamental concepts from the index, sampling theory and PCA. The main findings of the study will then be presented, followed by some conclusions.

The overall goal of this research is to create errors in household wealth indices for Punjab. The specific objectives of this study are to extract communities for errors of indices of households through PCA. Wealth is a household feature that significantly influences people's health, economic, and economical position in a particular economy. Accurate wealth assessment is essential in many fields of study. Despite its importance, wealth is an ill-defined term that is difficult to quantify. As a result, several studies have relied on tangible measures of household income and consumption to capture wealth. Still, income and consumption data are typically plagued by measurement error and systematic biases related to memory and sensitivity to questions. More significantly, many surveys do not gather reliable income and spending data, particularly in developing nations with domestic production and informal transactions, as well as in health and education surveys. The measurement problems, as well as the cost of installing long spending modules, have driven a quest for an accurate and conveniently collected wealth proxy.

2. Data Collection and Analysis

To the construction of wealth indices of Pakistan as well as Punjab # urban coding scheme is used for assigning processing codes for social & living standards measurement

2.1 District Level Survey

Coding scheme in urban sub-universe, big city and other urban areas of an admin district of Punjab has been considered independent strata. All rural area and admin districts have been grouped to constitute separate strata within an admin district of the four provinces.

3. Results and Discussion

To run an exploratory analysis on the variables that have been collected: Rules of thumb: variables Recode the household amenities variable into improved / not improved the recoding between improved/ not improved is just one possibility. Everitt and Dunn; (2001) calculated that analysts could include only a certain variable/category (only households with a 'household water connection' y/n) to extract/highlight households with a very good water source. (Lawley and Maxwell 1962) proposed pattern about factor analysis to investigate relationships among variables by testing of hypothesis. The analyst can include only a particular variable/category (only households) of ' household survey that oblique solution.

Because of the oblique solution, correlated variables are extracted which value in PCA is more significant than 0.4 and valuables that value is less than 0.4 are treated uncorrelated variables extracted due to orthogonal transformation in principal component analysis correspondence study of (Amatete 2016). Communalities of Punjab urban are extracted in table 1, which consist of variance and standard error of Punjab urban area.

Table 1: Communalities Punjab Urban

		Variance	Standard error
1.	What is your present occupancy status?	.057	0.238747
2.	How many separate rooms are there in your dwelling?	.464	0.681175
3.	Type of roof material	.288	0.536656
4.	Type of wall material	.272	0.521536
5.	Drinking water sources	.203	0.450555
6.	Toilet Facility used in household?	.118	0.343511
7.	What is the main fuel used for cooking?	.029	0.170294
8.	What is the main source of lighting?	.213	0.461519
9.	Working telephone connection?	.473	0.68775
10.	Supply of drinking water	.560	0.748331
11.	Supply of drinking water (usual mean of transport)?	.649	0.805605
12.	Groceries shops (in minutes)?	.123	0.350714
13.	Groceries shops (usual mean of transport)?	.023	0.151658
14.	Public transport (in minutes)?	.361	0.600833

15.	Public transport (usual mean of transport)?	.083	0.288097
16.	Primary school (in minutes)?	.460	0.678233
17.	Primary school (usual mean of transport)?	.221	0.470106
18.	Middle school (in minutes)?	.757	0.870057
19.	Middle school (usual mean of transport)?	.685	0.827647
20.	High school (in minutes)?	.742	0.861394
21.	High school (usual mean of transport)?	.728	0.853229
22.	Health clinic or hospital (in minutes)?	.597	0.772658
23.	Health clinic or hospital (usual mean of transport)?	.611	0.781665
24.	Family planning Centre (in minutes)?	.566	0.75233
25.	Family planning Centre (usual mean of transport)?	.585	0.764853

Variance and standard errors of household Indices point out the part of each variable's variance in principal components analysis. Variables which variance is more significant than 0.5 are well-represented variance and standard errors of household Indices of the household survey. In contrast, variables with low values are not a well-represented variance and standard error of wealth Indices of the household survey. Variables that contain low values are present occupancy status, separate rooms in dwelling, material of roof, the material of wall, source of drinking, toilet facility, fuel used for cooking, source of lighting, telephone connection, nearest supply of drinking water(in minutes) and mean of transport.

Table 2: Total Variance of Punjab Urban

Component	Initial Eigenvalues			Total Extraction		
	Total Variance	Cumulative	Variance	Total Variance		Cumulative Variance
1	4.939	19.755	19.755	4.939	19.755	19.755
2	1.853	7.412	27.166	1.853	7.412	27.166
3	1.586	6.343	33.509	1.586	6.343	33.509
4	1.489	5.957	39.466	1.489	5.957	39.466
5	1.364	5.456	44.922			
6	1.317	5.267	50.190			
7	1.141	4.565	54.755			
8	1.130	4.519	59.274			
9	1.021	4.085	63.359			
10	.957	3.828	67.187			
11	.912	3.649	70.836			
12	.877	3.507	74.343			
13	.813	3.252	77.595			
14	.789	3.156	80.750			
15	.708	2.831	83.582			
16	.657	2.627	86.209			
17	.628	2.511	88.720			
18	.601	2.406	91.125			
19	.548	2.190	93.315			
20	.480	1.920	95.236			
21	.462	1.850	97.085			
22	.295	1.182	98.267			
23	.189	.758	99.025			
24	.143	.573	99.598			
25	.100	.402	100.000			

Table 2 explains the variation of components; the first four principal components explained most variance in variables sets. First component always explanation significant variation, which is 4.939, and the next component will explanation for as much of the leftover variation as it can, and so on. Hence each successive component will account for less and less variance; this finding is consistent with (Måge, Smilde et al. 2019).

Table 3: Component Matrix of Punjab Urban

		Component			
		PC1	PC2	PC3	PC4
1.	What is your present occupancy status?	-.072	-.219	.061	-.020
2.	How many separate rooms are there in your dwelling?	-.026	.558	.136	-.364
3.	Type of roof material	.189	-.356	-.026	.353
4.	Type of wall material	.022	-.380	.180	.308
5.	Drinking water sources	.091	.163	.378	.161
6.	Toilet Facility used in household?	.121	-.162	-.220	.169
7.	What is the main fuel used for cooking?	.078	-.061	-.139	.012
8.	What is the main source of lighting?	-.005	-.331	.167	.274
9.	Household Telephone connection	-.139	.567	.066	-.358
10.	supply of drinking water	.213	-.071	.706	.103
11.	Supply of drinking water (usual mean of transport)?	.194	.184	.752	.109
12.	Groceries shops (in minutes)?	.184	-.211	.027	-.210
13.	Groceries shops (usual mean of transport)?	.096	-.013	.111	-.033
14.	Public transport (in minutes)?	.396	-.253	.192	-.321
15.	Public transport (usual mean of transport)?	.189	.104	.178	.066
16.	Primary school (in minutes)?	.470	-.216	.066	-.434
17.	Primary school (usual mean of transport)?	.340	.303	-.009	.118
18.	Middle school (in minutes)?	.781	-.206	-.107	-.304
19.	Middle school (usual mean of transport)?	.675	.316	-.246	.264
20.	High school (in minutes)?	.788	-.229	-.069	-.253
21.	High school (usual mean of transport)?	.697	.297	-.240	.311
22.	Health clinic or hospital (in minutes)?	.750	-.121	.045	-.137
23.	Health clinic or hospital (usual mean of transport)?	.681	.266	-.002	.276
24.	Family planning Centre (in minutes)?	.730	-.102	.032	-.145
25.	Family planning Centre (usual mean of transport)?	.679	.243	-.018	.252

Extraction Method: Principal Component Analysis

Table 3 describes wealth indices of household surveys. Variables which score are greater than 0.5 in PC1 are 16, 18,19,20,21,22,23,24, and 25 and these variables are well representing wealth indices. As we can see that only PCA1 explain a large amount of variation that why PCA1 is enough for an orthogonal and oblique solution. The PCA1 correlated with items 16, 18,19,20,21,22,23,24, and PCA1 suggest that if any items are focused, other items will also automatically focus, such as if the primary school education system is developed, other correlated items will also be developed. These above items are increasing and decrease together because the above-said items are positively correlated with each other If one increases, then the remaining will also increase. This component can be viewed as a measure of the quality of primary school (in minutes), middle school(in minutes), middle school(usual mean of transport), high school (in minutes), high school (usual mean of transport), health clinic or hospital (in minutes), health clinic or hospital (usual mean of transport), family planning Centre(in minutes) and family planning Centre (usual mean of transport)

Furthermore, we see that these results are consistent with the findings of (Salchow 2001, Tayalı and Tolun 2018).

In PC2, only two items are correlated, which are separates rooms and phone connections. This component can be viewed as a measure of how the numbers of separates and phone connections increase in terms of available government policy. No one variable is correlated in the third principal component (PC3) are consistent with the study of (Yu, Quinn et al. 1998, Kolenikov, Angeles et al. 2009).

4. Conclusions

In the paper, we presented the preliminary findings of errors of indices of households conducted based on a large dataset of Punjab urban data; estimates of household indices are calculated based on simple random sampling to check the complete information on the quality and reliability of published data. Here principal components are used to estimate standard errors of household indices as orthogonal transformation and oblique solution to develop concrete measures of individual economic status. These measures are used to evaluate the significance and explain Punjab urban's living status and economic dissimilarity. The study goal was to find the errors of indices for estimating the sampling variance of the index estimator. The study's main findings could be summarized based on household indices education and health system can improve the living standards of urban peoples. In Punjab, urban variables 16, 18,19,20,21,22,23,24, and 25 describe household surveys' wealth indices, which score greater than 0.5 due to the oblique solution on PCA1 consistent with the study of (Adler and Golany 2001, Sanders, Zuidgeest et al. 2015). That's why we can suggest that people of Punjab urban government should prioritize education and health system to improve living standards because in table 4 high correlated variables related to education and health system.

It is impossible to overstate the importance of conducting studies to understand and quantify errors of indices in household surveys. This is especially important if the survey concepts being measured are novel and complex. Users' analyses rely on having both high-quality data and an understanding of the data's nature and limitations. Because errors of indices in household surveys studies are costly and time-consuming, they require an explicit commitment from the survey program. However, the commitment does not end with the implementation and execution of the studies.

Finally, the findings of errors of indices in household surveys studies are critical for improving the subsequent survey fielding. Significant measurement improvements are heavily reliant on prior survey knowledge and results. Future enhancements in survey data quality will necessitate the dedication of survey research professionals. There are following recommendations;

- This study built the wealth index using principal components analysis. However additional research is needed to compare (PCA) with other approaches such as (FA)
- Wealth is connected to many other socioeconomic characteristics such as the family head's gender, degree of education, and age. A correlation analysis was not performed in this study to establish the link between the wealth index and such parameters. More investigation is required to discover the link between the wealth index and socioeconomic variables.

References

- Abeyasekera, S. J. R. U. o. R. h. (2005). "Chapter XVIII Multivariate methods for index construction." 377-378.
- Adler, N. and B. J. E. J. o. O. R. Golany (2001). "Evaluation of deregulated airline networks using data envelopment analysis combined with principal component analysis with an application to Western Europe." 132(2): 260-273.
- Amatete, B. W. (2016). Evaluation of critical factors affecting pricing of real estate among low income people in Nairobi, Kenya, Strathmore University.
- Biemer, P. P. J. P. o. q. (2010). "Total survey error: Design, implementation, and evaluation." 74(5): 817-848.
- Chambers, R. J. E. and urbanization (1995). "Poverty and livelihoods: whose reality counts?" 7(1): 173-204.

- Crown, W. H. (1998). *Statistical models for the social and behavioral sciences: Multiple regression and limited-dependent variable models*, Greenwood Publishing Group.
- Dunteman, G. H. (1989). *Principal components analysis*, Sage.
- Filmer, D., L. J. P. Pritchett and d. review (1999). "The effect of household wealth on educational attainment: evidence from 35 countries." 25(1): 85-120.
- Filmer, D. and L. J. W. B. Pritchett (1998). "The effect of household wealth on educational attainment around the world: Demographic and health survey evidence."
- Gerhart, B., P. M. Wright, G. C. MC MAHAN and S. A. J. P. p. Snell (2000). "Measurement error in research on human resources and firm performance: how much error is there and how does it influence effect size estimates?" 53(4): 803-834.
- Howard, G. and C. Bogh (2002). *Healthy villages: a guide for communities and community health workers*, World Health Organization.
- Jolliffe, I. T., J. J. P. T. o. t. R. S. A. M. Cadima, *Physical and E. Sciences* (2016). "Principal component analysis: a review and recent developments." 374(2065): 20150202.
- Jong, J.-C. and S. J. T. A. S. Kotz (1999). "On a relation between principal components and regression analysis." 53(4): 349-351.
- Kolenikov, S., G. J. R. o. I. Angeles and Wealth (2009). "Socio-economic status measurement with discrete proxy variables: Is principal component analysis a reliable answer?" 55(1): 128-165.
- Lawley, D. N. and A. E. J. J. o. t. R. S. S. S. D. Maxwell (1962). "Factor analysis as a statistical method." 12(3): 209-229.
- Lillacci, G. and M. J. P. c. b. Khamash (2010). "Parameter estimation and model selection in computational biology." 6(3): e1000696.
- Lohmöller, J.-B. (2013). *Latent variable path modeling with partial least squares*, Springer Science & Business Media.
- Måge, I., A. K. Smilde and F. M. J. J. o. C. van der Kloet (2019). "Performance of methods that separate common and distinct variation in multiple data blocks." 33(1): e3085.
- Matsunaga, M. J. I. j. o. p. r. (2010). "How to Factor-Analyze Your Data Right: Do's, Don'ts, and How-To's." 3(1): 97-110.
- McManus, S., H. Meltzer, T. Brugha, P. E. Bebbington and R. Jenkins (2009). *Adult psychiatric morbidity in England: Results of a household survey*, Health and Social Care Information Centre.
- Montana, L. and J. J. C. H. Spencer, NC: The University of North Carolina at Chapel Hill, Carolina Population Center (2004). "Incorporating geographic information into MEASURE surveys: a field guide to GPS data collection."
- Ng, S. J. E. R. (2015). "Constructing common factors from continuous and categorical data." 34(6-10): 1141-1171.
- Rarani, M. A., A. Rashidian, A. Khosravi, M. Arab, E. Abbasian, E. K. J. I. j. o. h. p. Morasae and management (2017). "Changes in socio-economic inequality in neonatal mortality in Iran between 1995-2000 and 2005-2010: an Oaxaca decomposition analysis." 6(4): 219.
- Rogers, D., G. Wint, N. Alexander, F. Pozzi and T. P. Robinson (2011). "Wealth index mapping in the Horn of Africa."
- Rutstein, S. and G. J. C. Rojas, Maryland (2003). "Guide to DHS Statistics (p. 151)."
- Sachs, J. (2008). *Common wealth: Economics for a crowded planet*, Penguin.
- Salchow, M. E. (2001). *An erosion-productivity study of three fields in west central Ohio with simple digital terrain analysis*, The Ohio State University.
- Sanders, P., M. Zuidgeest and K. J. H. I. Geurs (2015). "Liveable streets in Hanoi: A principal component analysis." 49: 547-558.
- Smith, L. C., H. Alderman and D. Aduayom (2006). *Food insecurity in sub-Saharan Africa: new estimates from household expenditure surveys*, Intl Food Policy Res Inst.
- Soosalu, G., S. Henwood and A. J. S. O. Deo (2019). "Head, heart, and gut in decision making: development of a multiple brain preference questionnaire." 9(1): 2158244019837439.
- Tayali, H. A. and S. J. E. S. w. A. Tolun (2018). "Dimension reduction in mean-variance portfolio optimization." 92: 161-169.
- Townsend, P. (2014). *International analysis poverty*, Routledge.
- Valliant, R. J. J. o. O. S. (1992). "Smoothing variance estimates for price indexes over time." 8(4): 433.

- Vinaixa, M., S. Samino, I. Saez, J. Duran, J. J. Guinovart and O. J. M. Yanes (2012). "A guideline to univariate statistical analysis for LC/MS-based untargeted metabolomics-derived data." 2(4): 775-795.
- Wamani, H., T. Tylleskär, A. N. Åström, J. K. Tumwine and S. J. I. j. f. e. i. h. Peterson (2004). "Mothers' education but not fathers' education, household assets or land ownership is the best predictor of child health inequalities in rural Uganda." 3(1): 1-8.
- Webley, P. and E. K. J. J. o. E. P. Nyhus (2006). "Parents' influence on children's future orientation and saving." 27(1): 140-164.
- Wiktorowicz, J. J. E. (2016). "Exploratory factor analysis in the measurement of the competencies of older people." (54): 48-60.
- Yong, A. G. and S. J. T. i. q. m. f. p. Pearce (2013). "A beginner's guide to factor analysis: Focusing on exploratory factor analysis." 9(2): 79-94.
- Yu, C.-C., J. T. Quinn, C. M. Dufournaud, J. J. Harrington, P. P. Rogers and B. N. J. J. o. E. M. Lohani (1998). "Effective dimensionality of environmental indicators: a principal component analysis with bootstrap confidence intervals." 53(1): 101-119.

Appendix

Appendix 1: Descriptive Statistics of Punjab Urban

	N	Mean	Std. Deviation	Variance	Skewness	Kurtosis			
	Statistic	Statistic	Std. Error	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
What is your present occupancy status?	2935	2.39	.017	.932	.869	1.589	.045	2.117	.090
How many separate rooms are there in your dwelling?	2935	2.61	.028	1.504	2.261	1.815	.045	6.733	.090
Type of roof material	2935	2.09	.024	1.300	1.691	.629	.045	-1.384	.090
Type of wall material	2935	1.04	.005	.254	.064	10.142	.045	133.410	.090
Drinking water sources	2935	3.22	.042	2.248	5.056	1.221	.045	1.541	.090
Toilet Facility used in household?	2935	2.59	.015	.820	.672	.784	.045	.599	.090
What is the main fuel used for cooking?	2935	2.09	.021	1.114	1.241	3.682	.045	15.592	.090
What is the main source of lighting?	2935	1.03	.005	.244	.060	10.697	.045	144.131	.090
Working telephone connection?	2935	2.96	.013	.708	.501	-1.415	.045	2.831	.090
supply of drinking water	2935	1.03	.003	.184	.034	7.056	.045	54.703	.090
Supply of drinking water (usual mean of transport)?	2935	1.03	.004	.215	.046	8.316	.045	70.138	.090
Groceries shops (in minutes)?	2935	1.00	.001	.045	.002	22.061	.045	484.997	.090
Groceries shops (usual mean of transport)?	2935	1.00	.001	.055	.003	26.041	.045	757.549	.090
Public transport (in minutes)?	2935	1.03	.003	.176	.031	6.359	.045	43.205	.090
Public transport (usual mean of transport)?	2935	1.01	.002	.111	.012	15.967	.045	267.365	.090
Primary school (in minutes)?	2935	1.02	.003	.139	.019	7.913	.045	66.893	.090
Primary school (usual mean of transport)?	2935	1.01	.003	.151	.023	12.034	.045	149.014	.090
Middle school (in minutes)?	2935	1.08	.006	.307	.094	4.379	.045	21.360	.090
Middle school (usual mean of transport)?	2935	1.07	.007	.355	.126	4.984	.045	23.560	.090
High school (in minutes)?	2935	1.11	.007	.360	.130	3.560	.045	13.850	.090
High school (usual mean of transport)?	2935	1.10	.008	.416	.173	4.040	.045	14.967	.090
Health clinic or hospital (in minutes)?	2935	1.12	.007	.389	.151	3.739	.045	17.196	.090
Health clinic or hospital (usual mean of transport)?	2935	1.16	.010	.530	.281	3.062	.045	7.589	.090
Family planning Centre (in minutes)?	2935	1.19	.009	.480	.231	2.755	.045	7.977	.090
Family planning Centre (usual mean of transport)?	2935	1.25	.012	.647	.418	2.242	.045	3.162	.090
Valid N (list wise)	2935								