Micro Data Analysis of the Living Standards and Inequalities in Dar Es Salaam

Satoru KANOH[1)]
&
Nobuyuki YASUI[2)]

This article investigates the change in the living standards and inequalities in Dar Es Salaam. Two publicly available micro data sets are used to estimate the household expenditures in 1991 and 1996. Based on the estimated expenditures, living standards and inequalities are quantitatively assessed for each ward, as well as for the city as a whole. These analyses show that both the living standard and the inequality in the city as a whole stayed more or less the same over those five years. However, the living standard of individual ward changed greatly, and as a result, the inequality among wards increased considerably. In 1992, Japan made a grant aid to a wealthy district, which further raised the district's living standard, thereby resulting in the increase of inequality among the wards. This fact suggests that Japanese policy regarding aid to developing countries should be carefully considered.

JEL Classification: C49; C81; I32;
Keywords: Micro Data; Living Standard; Inequality; Bootstrap;

1) Institute of Economic Research, Hitotsubashi University
2) Embassy of Japan in Tanzania

The authors are grateful to Professors H. Akabayashi, Y. Ikemoto, T. Kurosaki and seminar participants at Hitotsubashi University for their constructive comments and suggestions.

1. Introduction

Many large cities in developing countries, particularly those in the poorest countries, are facing the problems of urbanization and abrupt increase of population size with annual rate as much as 7%. Those developing countries are also under the pressure of the structural adjustment policy proposed by the IMF and the World Bank. As a result of the policy, various aspects of the urban economies in those countries have drastically changed. Under such circumstances, the importance of the infrastructure management has become more prominent than in the past and not a few industrialized countries including Japan have carried out development aid for this purpose. In spite of such cooperation on the part of other countries, the living standards of the large cities in the developing countries have not improved and the inequalities in those cities seem to have become even greater.

The World Bank repeatedly discussed the relationship between the living standard of the people in urban area and the infrastructure aids of developed countries. The World Development Report (1990), whose main theme was 'poverty', pointed out that past projects in urban areas related to poverty improvement were of more value to the middle class than to the lower class. Further, the World Bank mentioned that one of the best ways to mitigate the unemployment problem in urban areas, particularly in the informal sector, is 'provision of appropriate infrastructure to the urban area.' The World Development Report (1991) stated that, since the public investments have strong redistribution effects, the government investment in infrastructure must improve the efficiency of resource distribution, stimulate the growth of the economy, and reduce the inequality in the country.

There have been few empirical research works on inequalities in developing countries, mainly because of the lack of availability and reliability of micro data. Among them, Fields (1979), based on the household survey data of four cities in Colombia including its capital Bogotá, tried to factorize the cause of inequalities using the analysis of variance technique. He then pointed out that the age of the head of the household did not explain income inequality; instead, the educational level of the head greatly influenced inequality. Moreover, he argued that inequality among urban habitants in Colombia was due to inequality within cities rather than between cities. Knight and Song (1991) analyzed the data of the urban household survey in 1986 and estimated the importance of the various determinants of income inequalities such as region, gender, age, educational background, occupation and ownership of a company. Though both Fields (1979) and Knight and Song (1991) have analyzed the micro data

quantitatively, neither focused on the change of the living standard nor the change of inequality over time. As Fields (1979) pointed out, the conventional research works on development aid have not paid much attention to the change of income distribution among people. That is, a large scale of development aid has been made without proper assessments of their distributional effects.

The objective of this paper is to present an empirical approach to estimate the change of the living standards and inequalities in developing countries, making most of the available data sets and using simulation techniques. The analysis is applied to estimate the change of the living standards and inequalities in Dar Es Salaam, the de facto capital of Tanzania that is one of the poorest countries in East Africa. The empirical approach adopted here is not specific to the city, but applicable to other large cities in developing countries. Based on the results, this paper evaluates the role of the infrastructure assistance made by the Japanese government in 1992 in regard to the change of the living standard and inequality in Dar Es Salaam.

2. Data and Analyses

2.1 Summary of the Data

Two data sets are available in the form of micro data. They are the Living Standard Measurement Survey (LSMS) published by the World Bank and the Demographic and Health Survey (DHS) published by Macro International Inc. with the support of USAID. The LSMS is a sampling household survey commenced in 1979 by the World Bank and has been conducted in 17 countries throughout the world. Since the objective of the survey is to obtain detailed information on the living standards, questions are made about various economic aspects of households. In Tanzania, however, the survey has so far been conducted only once in the past. The DHS survey is also a sampling household survey on demography and health commenced in 1984 with the help of USAID. To the present, the survey has been conducted in 33 countries in Africa, 15 in Asia and 13 in Latin America. The characteristics of the DHS are that common questionnaires are made for all countries so that international comparisons are possible and that the surveys have been conducted several times in some of the countries and therefore survey results can be compared over years in those countries. However, economic information such as income or expenditure of households is not provided in the data set.

In this research, by combining these two data sets, the monthly per capita expenditures for consumption in 1991 and 1996 are estimated. The monthly per capita expenditure is regarded as a proxy of the living standard of a household. By using the estimated expenditure data, analyses on the change of the living standard and inequality during these five years are carried out.

2.2 Estimation of the living standard of households

As stated above, the DHS data do not contain the information of the per capita expenditure for consumption. The per capita expenditures for consumption are estimated for the households sampled in the DHS data set by using the LSMS data as is shown below. The analyses are confined to 19 wards where several households are obtained as the sample in both 1991 and 1996. This excludes three wards where less than five samples are obtained. Table 1 lists the selected wards and the sample size there in the three surveys (LSMS, DHS91 and DHS96).

As the first step, the monthly per capita expenditure for consumption Pce_h is created for the sampled households in the LSMS data set by dividing the total household expenditure by the number of adults in the family. The total expenditure is calculated by aggregating expenditures for individual goods and services after adjusting them to monthly basis. Here the expenditure for rent was not included in the consumption expenditure in this analysis, because the imputed house rent prices for the house-owners can not be estimated easily and objectively. Following Deaton (1997) and Yamazaki (1998), a child under the age of five is counted as 0.4 of an adult.

The regression model for $\log(Pce_h)$ is then constructed using explanatory variables that are selected from the candidates that appear both in the LSMS data and in the DHS data and that are regarded useful in explaining the household's income level. Dummy variables to express wards are introduced for all wards even if they are insignificant in order to set the average value of the residuals in each ward zero, which is necessary in the following simulation analysis. Instead, the constant term is dropped.

There are several points to be taken care of in carrying out the regression analysis. First, the LSMS is a sampling survey and the sampling weights differ depending on the population size of the wards. This has to be counted in the estimation process. Second, the conditional variance of $\log(Pce_h)$ cannot be assumed homogenous among wards since inequality in a ward is closely related to its variance and this is to be investigated rather

than assumed. From these, the Weighted Least Squared method (WLS) instead of the OLS is used to obtain the regression plane. Third, the data set is expected to contain outliers, as is often the case with data on developing countries. There may also exist outliers in explanatory variables. The standardized residuals of the WLS have to be examined carefully in order to avoid distortion possibly stemming from such influential points.

Taking these factors into consideration, the regression estimation steps the following process.

1. There are 841 households sampled from the 19 wards in the LSMS data. Using the given sample weights s_k (k= 1,2,…,19), the first WLS is carried out. In this regression, the differences of variances among the wards, whose values are unknown, are ignored.
2. From the residuals obtained in 1, the variance σ_k^2 of the k-th ward(k=1,2,…,19) are estimated.
3. From the obtained σ_k^2 and the sampling weights s_k, new weights for the wards are defined as: $w_k = s_k / \sigma_k^2$.
4. The second WLS is carried out with these new weights w_k and residuals r_i are obtained.

Let S and Ω be 841x841 diagonal matrices whose elements consist of σ_k^2 and the weights w_k attached to each individual, respectively. Then the variance matrix of the residuals from the second WLS regression is expressed as:

$$M = [\, I - X(X'\Omega^{-1}X)^{-1} X'\Omega^{-1}] \, S[\, I - X(X'\Omega^{-1}X)^{-1} X'\Omega^{-1}]' \quad .$$

Note the fact that the variance of y is not Ω but S slightly complicates the expression.

5. The residuals r_i's are transformed to the standardized residuals according to the formula: $e_i = r_i / \sqrt{m_i}$, where m_i is the i-th diagonal element of the matrix M.
6. When e_i is larger than 2.0, such households are not used to estimate the regression plane in the next step. 39 households are omitted according to this rule. This does not mean, however, those households are completely excluded from the following analysis.

Figure 1 shows the residual plots against the predicted values of $\log(Pce_h)$ in the 19

wards. There are 841 households among which 39 households are beyond ±2.0.

7. The third WLS is applied to the remaining 802 households with the weight matrix Ω.

The final estimation results are summarized in Table 2. The variables to express wards are retained in the equation even if they are insignificant. As stated above, this is necessary to make the average of the residuals in each ward zero. The variable expressing the number of children is retained in the equation though insignificant since its level is nearly 10%. All the other variables are significant.

Then, the regression equation obtained in this way is applied to the DHS data. The predicted values of $\log(Pce_h)$($\log(PPce_h)$) are calculated for the sampled households included in the 1991 and 1996 DHS data sets. Note that $PPce_h$ are measured in terms of the 1993 price level and, therefore, those predicted values are also expressed by the 1993 price level. Price adjustment using a deflator is unnecessary, since no nominal variables are used as the explanatory variables in the regression model.

Then, one residual from the third WLS regression and one $\log(PPce_h)$ are separately, randomly selected in the same ward and they are summed up. The resulting values are exponentially transformed and regarded as household expenditure per capita, called Artificial Pce_h ($APce_h$). That is, it holds

$$\log(APce_h) = \log(PPce_h) + r_h ,$$

where r_h is the randomly selected residual from the third regression. Such $\log(APce_h)$ are created in a ward as many as $n_k' = n_k \times w_k / w_0$, where w_0 is the minimum weight value among wards, i.e. $w_0 = \min\{ w_k; k=1,2,\ldots,19\}$. This is to adjust the sample size proportional to the population size by maintaining the minimum sample size among the 19 wards.

Figure 2-a shows the histogram of the actual values of $\log(Pce_h)$ in the LSMS data. Figure 2-b shows the histogram of $\log(PPce_h)$ created above for 1996 and Figure 2-c shows the histogram of the corresponding $\log(APce_h)$. Note that the dispersion of $\log(Pce_h)$ in Figure 1 is smaller than that of $\log(PPce_h)$ in Figure 2-a and closer to that of $\log(APce_h)$ in Figure 2-b. This is confirmed by the coefficients of variations for the three histograms and this illustrates why the above-mentioned procedure is necessary. Apparently, any indicators calculated from $PPce_h$ will underestimate the inequality.
In the next section, in order to take account of the statistical properties of the various

indicators calculated from $APce_h$, the above procedures are repeated 500 times.

3. Comparison of the living standards and inequalities between 1991 and 1996

In this section, based on the produced $APce_h$ for 1991and 1996, comparisons are made to grasp the changes of the living standard and inequality in Dar Es Salaam during the five years. First, in order to see the change of people's living standards, the transitions of the quartile points are examined. These quartiles are calculated from each set of the produced $APce_h$ and, after 500 repetitions, the median values of the 500 sets of estimated quartiles are regarded as the final estimates of the corresponding population quartiles. Table 3 summarizes the results. For illustrative purposes, the histograms of 500 estimates of the median(the second quartile) for 1991 and 1996 are shown in Figure 3-a and 3-b, respectively. In order to evaluate the differences between the two years, the probability of the statistic for 1996 being larger than that for 1991 is estimated by randomly sampling 1000 pairs from the obtained 500 estimates of each parameter. The results are also shown in Table 3 as $P(X_{96} > X_{91})$. If the conventional hypothesis testing approach were taken, the standard errors of those statistics would be used to test if the two populations had the same values of parameters under the assumption that the statistics were distributed independently and normally. The approach in this paper assumes neither normality nor independence of the statistics and tries to evaluate the differences non-parametrically. From the table, it is seen that the living standards of Dar Es Salaam had slightly risen at all quartiles during the five years, though the differences were not large enough to be statistically endorsed.

Subsequently, the average values of $APce_h$ in wards that are regarded as the living standards of the district are calculated for the 19 wards. The results are summarized in Table 4. The table shows the following characteristics. The total living standard rose slightly. The average growth rate of per capita consumption expenditure during the five years is estimated as approximately 1.4%, that is 0.3% annually. According to the World Bank Developing Indicators, the average value of the per capita private consumption growth rate between 1991 and 1996 was –0.2%. The 0.3% average growth rate in its capital seems an intuitively reasonable value. The living standard of each ward changed greatly during the five years. The percentage changes range from –48% to 56%. Among them, five positive changes are significantly large and the probabilities are greater than 0.9. In addition, three negative changes are also significant. Figure 4-a and 4-b illustrate the change of the living standards by

separating the wards into two groups where the living conditions have improved and deteriorated. The wards where the living standard rose remarkably are Makurumla(1), Ndugumbi(2), Mabibo(8), Vingunguti(11) and Mtoni(17). Note that Msasani(5), with the highest living standard in 1991, recorded a sizable increase, though not significant. On the other hand, the living standard sharply dropped in Tendale(3), Buguruni(13) and Kurasini(19). Then, the changes of inequalities in the city are analyzed in a similar way. For this purpose, the Gini Index and the Thile Index are used as measures of inequality. Concretely, the Gini Index and the Thile Index are defined as follows.

Let X_{ik} be the individual value of $APce_h$ for the j-th household in the k-th ward, n_k be the number of sampled households in the k-th ward, s_k be the sampling weight attached to the k-th ward, μ_k be the average value of $APce_h$ in the k-th ward, and μ_0 be overall weighted average of $APce_h$ in the whole city. Then the Gini Indices for the k-th ward and for the whole city are respectively defined as:

$$G_k = \sum_{i=1}^{n_k}\sum_{j=1}^{n_k} \frac{|X_{ie} - X_{jk}|}{2m_k n_k^2} \quad ,$$

where $m_k = \sum_{i=1}^{n_k} X_{ik} / n_k$ and

$$G_0 = \frac{\sum_{k=1}^{19}\sum_{l=1}^{19}\left(s_k s_l \sum_{i=1}^{n_h}\sum_{j=1}^{n_l}|X_{ik} - X_{jl}|\right)}{2m_0(\sum_{k=1}^{19} n_k s_h)^2} \quad ,$$

where $m_0 = \frac{\sum_{k=1}^{19}\left(s_k \sum_{i=1}^{n_k} X_{ik}\right)}{\sum_{k=1}^{19} n_k s_k}$.

The Thile Index for the k-th ward and that for the whole city are respectively defined as:

$$T_k = \left(\frac{1}{n_k}\right)\sum_{i=1}^{n_k}\left(\frac{X_{ik}}{m_k}\right)\log\left(\frac{X_{ik}}{m_k}\right)$$

and

$$T_0 = \left(\frac{1}{\sum_{k=1}^{19} n_k s_k} \right) \sum_{k=1}^{19} \sum_{i=1}^{n_k} \left(\frac{X_{ik}}{m_0} \right) \log \left(\frac{X_{ik}}{m_o} \right) \quad .$$

Figure 5 and Figure 6 illustrate the histograms of 500 Gini Indices for 1991 and 1996 and those of Thile Indices for 1991 and 1996, respectively.

The advantage of the Gini Index is in its general acceptability as a measure of inequality. On the other hand, the advantage of the Thile Index is in its additivity, which is useful in the following analysis when the source of inequality is decomposed into the factors of the within-wards and the between-wards. That is, it holds that

$$T_o = \sum_{k=1}^{19} \left(\frac{p_k m_k}{m_0} \right) \log \left(\frac{m_k}{m_0} \right) + \sum_{k=1}^{19} \left(\frac{p_k m_k}{m_0} \right) T_k \quad ,$$

with $p_h = \dfrac{n_k s_k}{\sum_{k=1}^{19} n_k s_h}$ and,

$$T_0 = T_1 + T_2$$

where T_1 denotes the inequality between wards and T_2 denotes the sum of the inequalities within wards. Table 5 summarizes the two inequality indices in the two years.

The Gini indices and the Thile indices are also calculated 500 times from randomly selected 1991 and 1996 $APce_h$. Figure 7 and Figure 8 illustrate the histogram of T_1 and T_2 respectively. The differences between the two years are again evaluated by the probability $P(X_{91} < X_{96})$. From Table 5, it is remarked that both the Gini Index and the Thile Index for the total showed a negligible increase, implying that the inequality in the city stayed more or less the same. Table 5 also presents that the inequalities within wards were four to five times as large as those between wards. From this fact, inequality in Dar Es Salaam is mainly due to inequality within wards rather than between wards. After five years, however, inequality within wards had slightly decreased, whereas inequality between wards had increased more than 30%.

Table 6 shows the changes of inequalities in the 19 wards. It is seen that the

inequalities in the wards were less variable than the living standards. In Kurasini (19), a significant reduction of inequality was seen. The reductions in inequalities in Kipawa(12) and Msasani(5) were also large but insignificant. Several wards showed an increase in inequality, but all of them were insignificant. As a total, the percentage changes of inequality indices were less than 1%.

To summarize, the improvement of the living standard was observed with a minor margin during these five years. On the other hand, the inequality in the city seemed unchanged as a whole. Kulaba(1989) pointed out, based on the survey on economic and social conditions conducted in 1986-7 for 660 households in Dar Es Salaam, that a multiplication of income source had widened the income gap among people. The facts found here tell a slightly different story in the first half of the '90's. However, by breaking down inequality of the city into within wards and between wards, it was seen that the former did not change much, whereas the latter seemed to have increased considerably. Especially, the improvement of the living standard in a few specific wards contributed to the increase in inequality between wards.

4. The Role of the Japanese Foreign Aid

In 1992, the Japanese government carried out a project of approximately 1billion yen aiming at reinforcing the network of electric power distribution in Msasani ward in Dar Es Salaam as a grant aid. Note that the year 1992 lies between the two years analyzed above. Using the information on electricity given in the three data sets (DHS 91, LSMS 93 and DHS 96) and the estimated $APce_h$ in 1991 and 1996, this section tries to evaluate the effect of this infrastructure assistance of the Japanese government.

Table 7 shows the estimated saturation rates of electric power in Dar Es Salaam. They are calculated as the sample proportion of households equipped with electricity in the three data sets. As side information, the 1988 census reported that the number of households equipped with electricity increased to 106,196 but the saturation rate remained as low as 37.8% because of an abrupt population increase. The reason for the DHS 91 figure being smaller than the 1988 census value can be due to sampling error. Summarizing these, the table shows that the saturation rate had steadily increased in the city by more than 16 points from 1991 to 1996. Note, however, that the saturation rate of approximately 50% in the city as a whole is relatively low in comparison with other cities in the poorest countries. For example, in Addis Ababa that is regarded as being a city at a similar poverty level, more than 90% of households are

equipped with electricity. The reason for such low electricity in Dar Es Salaam is pointed out by Lugalla(1995) as follows. The Tanzania Electric Supplying Company raised the price of electricity by 50% following the recommendation of the World Bank. Due to this, it cost 35,000 Tsh for an individual household to be equipped with electricity. Since the electric charge is set at such an unaffordable level, people tended to steal electricity rather than make a legal contract.

Then the saturation rates of electricity in each ward are calculated and the same comparisons are made between 1991 and 1996. Table 8 shows the saturation rates for electricity in the 19 wards. A large increase of the saturation rates is seen in Kurasini(19) and Msasani(5) wards, whereas a slight decrease is observed in Ilala(10) and Buguruni(13) wards.

Here we review the changes of the living standard and inequality in Msasani ward. The living standard of Msasani ward, that was the highest in 1991, greatly improved. The estimated increase was more than 10,000(Tsh), the second largest among the 19 wards. On the other hand, the inequality of Msasani ward, that was also largest when measured either by the Gini index or the Thile index, also decreased 10% to 20%. Unfortunately, the significances of those figures were not statistically confirmed. At the same time, the saturation rate of electricity increased by 70% during the five years.

Figure 9 depicts the change of the saturation rates in the 19 wards in relation to their living standards. The living standard is taken as the X-axis and the saturation rate of electricity is taken as the Y-axis. In the year 1991, the living standard of Msasani ward was already high whereas the saturation rate of electricity stayed low. Along with the improvement of the supply of electricity in this ward, the living standard drastically rose. As a result, Msasani ward became an even wealthier ward in 1996. These facts do not directly imply, however, that the living standards of the people in Msasani ward improved to the same degree, since the migrations of the people are not investigated in this paper. It is possible that the people who used to live in the ward were crowded out by an inflow of richer people after the living conditions improved.

There is a lesson to be learned. Msasani ward was one of the wards where the saturation rate of electricity rose remarkably. In that respect, the nominal objective of the Japanese grant aid was certainly achieved. Note, however, that the living standard of Msasani ward was already high in 1991 in spite of its low saturation rate of electricity. It can be assumed, therefore, that there was a potential demand for electricity in the district. One possibility of the low saturation rate was that there were some geographical or physical hindrances that prevented the spread of electricity in the

district until then. As a consequence of the electric supply, the living standard improved drastically. At the same time, the Japanese aid boosted the already wealthy district to an even wealthier state. In Figure 4, it was observed that the improvement of the living standard in Msasani caused the expansion of inequality among wards.

This particular case suggests the possibility that a certain type of infrastructure planning aid may increase the inequality between districts and may result in the expansion of inequality in a city as a whole (See Figure 5). Such incidence may not be specific to Dar Es Salaam. Many large cities in southern Africa where foreign aid capital has been invested might have more or less the similar phenomena.

5. Conclusion

In this article, the changes of the living standards and inequalities in Dar Es Salaam were investigated. From the analyses, it was shown that both the living standard and inequality in the city as a whole stayed more or less the same between 1991 and 1996. It was also shown that the living standard of individual ward changed greatly and, as a result, the inequality among wards increased considerably during the period.

It is often said that an unequal social structure hinders poverty mitigation and may also be a restraint to foreign capital investment because of its social instability. The World Development Reports stress the importance of government investment in infrastructure to reduce inequality. This paper showed, however, that a certain type of infrastructure investment might expand the regional gap. From the viewpoint of controlling the inequality expansion, we should keep it in mind to incorporate infrastructure services that benefit wider region and every walk of people into the development package. At this stage, it is not known for sure what types of infrastructure aid is preferable in consideration of regional inequality gaps. Infrastructure services for inferior goods such as public water facilities or roads to uncultivated areas may be among the possibilities.

Finally, there are some topics that are overlooked in this article and are left for future research. First, as was already mentioned above, the expenditure for rent was not included in consumption expenditure in this analysis. This is because the imputed house rent prices for the house owners could not be estimated easily and objectively. This exclusion results in an underestimation of inequality of the city. It may be desirable to estimate the rent price function by hedonic price approach and include the

estimated rent prices in the consumption expenditure. Second, in this analysis, the inflow and outflow of the people from one district to another was not considered. This is due to lack of proper data. Needless to say, it is important to grasp the transition of the people to evaluate the change of the living standard of districts. In this respect, the analysis carried out in this paper is merely a first step and thorough investigation on this topic is awaited.

[References]

Deaton, A., 1997, *The Analysis of Household Surveys: Microeconometoric Analysis for Development Policy*, Johns Hopkins Univ. Press

Efron B. and Tibshirani R. J. , 1993 *An Introduction to the Bootstrap*, Chapman & Hall.

Fields, G. S., 1979 "Income Inequality in Urban Colombia: A Decomposition Analysis," *Review of Income and Wealth*, Vol. 25, No.3, pp. 327-341.

Fields, G. S., 1994, "Data for Measuring Poverty and Inequality Changes in the Developing Countries," *Journal of Development Economics*, Vol. 44. No.1, pp. 87-102.

Knight, J. and Song, L., 1991, "The Determinants of Urban Income Inequality in China," *Oxford Bulletin of Economics and Statistics*, Vol. 53, No.2, pp. 123-154.

Kulaba, S., 1989, *Local Government and the Management of Urban Services in Tanzania*, (in Stren R. E. eds.) *African Cities in Crisis*.

Lugalla, J., 1995, *Crisis, Urbanization, and Urban Poverty in Tanzania*, Univ. Press of America.

Todaro, M. P., 1996, *Economic Development*, 6th edition, Addison-Wesley Publishing Co.

World Bank World Development Report 1990, 1991, Oxford University Press

Yamazaki, K., 1998, "Measuring Poverty and Policy to Resolve Poverty," Economic Development and Poverty, Towards Economic Analysis of Poverty, Institute of Developing Economics.

Yotopooulos, P. A. and Nugent, J. B.,1976, *Economics of Development: Empirical Investigations*, New York : Harper and Row.