Economics Bulletin

Volume 29, Issue 3

# Updating Poverty Maps for Ho Chi Minh City of Vietnam using a Small Area Estimation Method

Cuong Viet Nguyen
*National Economics University, Hanoi, Vietnam*

## Abstract

Ho Chi Minh City (HCMC) is the biggest and richest city in Vietnam with the population of over 7 million people. This study combines the HCMC Mid-Census 2004 and panel data from the Vietnam Household Living Standard Surveys 2004 and 2006 to produce poverty maps for HCMC's districts in 2006 using a method of small area estimation. It is found that the poverty incidence is very high in the rural districts. However, the poverty density is higher in the urban districts, since these districts have much higher population density.

I would like to thank Roy Van Der Weide from World Bank for his helpful comments on this study. I am also grateful for helpful comments provided from an anonymous reviewer of Economics Bulletin.

**Citation:** Cuong Viet Nguyen, (2009) "Updating Poverty Maps for Ho Chi Minh City of Vietnam using a Small Area Estimation Method ", *Economics Bulletin*, Vol. 29 no.3 pp. 1971-1980.

**Submitted:** Apr 01 2009. **Published:** August 17, 2009.

## 1. Introduction

Ho Chi Minh City (HCMC) is the biggest city in Vietnam with the population of over 7 million people. It is also the richest city with an annual economic growth rate of around 11 percent during the past 10 years. Basically, there are no poor households who are classified by the national income poverty line in HCMC.[1] However, HCMC still implements socioeconomic programs and policies to support low-income households. To identify poor households, HCMC applies its own income poverty line. More specifically, the income poverty line was set up equal to 6000 million VND/person/year during the 2004-2008 period. Since the early 2009, the income poverty line is increased to 12000 million VND.

An important tool for poverty monitoring is the poverty map, which is geographical visualization of disaggregated poverty measures. However, generation of poverty maps is not simple. Household surveys which contain income and expenditure data for poverty estimation are not representative for small areas such as districts and communes. Censuses cover all households but do not have data on expenditure and income and thus cannot be used directly for poverty estimation. Fortunately, Elbers et al. (2002, 2003) propose a small area estimation method which combines a household survey and a population census to estimate poverty rate at small areas. This method estimates an equation of income (or expenditure) from a household survey, and uses this estimated equation to predict income (or expenditure) for households in a census given households' characteristics. Then, the predicted income of households in the census will be used to estimate welfare indicators at small areas.

Up to now, there are two maps of district poverty of HCMC which are constructed using the small area estimation method of Elbers et al. (2002, 2003). The first poverty map of HCMC is constructed using Vietnam Household Living Standard Survey (VHLSS) in 2002 and a 10 percent sample of the HCMC Mid-Census for 2004 (Nguyen et al., 2005). The second map is also estimated using the 10 percent HCMC Mid-Census, but the household survey is the 2004 VHLSS instead of the 2002 VHLSS (Nguyen et al., 2007). Both the maps refer to the poverty estimates for the year 2004.

The availability of the 2006 VHLSS allows for updating the poverty map of HCMC to the year 2006. Thus this study combines the HCMC Mid-Census 2004 and panel data of the VHLSS 2004-2006 to produce the poverty map for HCMC's districts in 2006 using the method of Elbers *et al.* (2003). The paper consists of 6 sections. Section 2 introduces the data used for the analysis. Section 3 describes the method of small area estimation. The income model regressions and the poverty estimates are presented in sections 4 and 5, respectively. Finally, section 6 concludes.

## 2. Data Sources

The research relies on two data sources to estimate poverty and inequality for the districts of HCM city. The first is the Vietnam Household Living Standard Surveys (VHLSS) conducted by the General Statistics Office of Vietnam (GSO) in 2004 and 2006. The surveys contain information on household characteristics including basic demography, employment and labor force participation, education, health, income, expenditure,

[1] During the period 2005-2008, the national income poverty line is 2400 and 3120 thousand VND/person/year for the rural and urban areas, respectively.

housing, fixed assets and durable goods, and the participation of households in the most important poverty alleviation programs.

VHLSSs 2004 and 2006 cover around 9188 and 9189 households, respectively. These samples are representative at the regional level, but not at the provincial level. It is interesting that the surveys set up panel data of 4216 households.

The second data source is a 10 percent sample of the Population and Housing Mid-Census of HCM city in 2004. The census collects information on basic demography, education of people, unemployment status, and several characteristics on housing and assets. The census sample is designed to be representative at the district level.

## 3. Methodology

The method of "small area estimation", which is developed by Elbers et al. (2002, 2003), combines a household survey and a population census to estimate poverty rate at small areas. The main idea is to estimate equation of income or expenditure from a household survey, and use this equation to predict income or expenditure for households in a census given households' characteristics. Once predicted income or expenditures are available, poverty indexes can be estimated at small areas. The method is described in details in Elbers et al. (2002, 2003). Practically, this method can be described by three steps:

Step 1: Select common explanatory variables in a household survey and a population census, which will be used in regressions of income.

Step 2: Run regression of logarithm of per capita income on selected explanatory variables using data of the household survey:

$$\ln(y_i) = X_i\beta + \varepsilon_i \quad (1)$$

where:

- $y_i$ is per capita income of household *i*.
- $X_i$ are explanatory variables of household *i*.
- $\varepsilon_i$ is a random disturbance term distributed as $N(0,\sigma)$.[2]

Step 3: Apply this equation into the population census to predict the expected probability that household *i* is poor:

$$E[P_i \mid X_i, \beta, \sigma^2] = \Phi\left[\frac{\ln z - X_i^C\beta}{\sigma}\right] \quad (2)$$

where:

- $P_i$ is a variable taking a value of 1 if the household is poor and 0 otherwise.
- $z$ is the poverty line.
- $\Phi$ is the cumulative standard normal function.
- $X_i^C$ are explanatory variables for household i from the census.

Then the poverty rate for an area can be estimated:

$$E[P \mid X^C, \beta, \sigma^2] = \frac{m_i}{M}\sum_{i=1}^{N}\Phi\left[\frac{\ln z - X_i^C\beta}{\sigma}\right] \quad (3)$$

where:

- $m_i$ is the size of household *i.*

[2] In Elber et al. (2000, 2003), the error terms are decomposed into a household idiosyncratic component and a cluster component to capture correlation of the error terms within clusters.

- $M$ is the total population of the area in question.
- $N$ is the number of households in the area.

The standard errors of the poverty estimates are estimated using Monte Carlo simulations.

In this study, the HCM mid-census refers to 2004. To estimate the poverty map for the year 2006, we will use the feature of panel data from VHLSSs 2004-2006 to estimate per capita income in 2006 as a function of the explanatory variables in 2004. More specifically, our model is as follows:

$$\ln(y_i^{2006}) = X_i^{2004}\beta + \varepsilon_i \quad (4)$$

Then, the estimated model is applied to the HCMC mid-census to predict income per capita and poverty indexes for the year 2006.

In addition to the poverty rate, other two popular measures of poverty including the poverty gap index and the poverty severity index (see Foster et al., 1984) and Gini coefficient are also estimated in this paper.

## 4. Income Models

A main problem in the estimation of income models is selection of explanatory variables. The explanatory variables used in the income models should meet the following criteria:

- Available in both the survey and the census.
- Comparable between the survey and census, i.e., they are constructed in similar definitions and have similar distribution.
- Correlated with household income.

Common variables that are selected include household composition, education of household heads and members, the ratio of working members, housing characteristics, household assets such as telephone, television, radio, computer etc.

Once the common variables are selected, the model of income can be estimated using panel data from VHLSSs 2004 - 2006. The dependent variable is logarithm of per capita income in 2006, while the independent variables are household characteristics in 2004. Since the sample size of HCMC in the panel data of VHLSSs is very small and not representative (only 113 households), we use the sample of region "South East" for the regression of the income models. The South East region includes HCMC and other provinces which have similar economic and natural conditions. The number of households in the "South East" sample is 512.

Table 1 presents results from the GLS regressions of income per capita. The value of adjusted R-squared is 0.52.[3] All the explanatory variables have expected signs. Being ethnic minority households reduces the per capita income. Households who have a large number of members are more likely to have lower income. Households with no toilets or temporary households tend to have low income. Having telephone is positively associated with income per capita. Education variables also have reasonable relation with

3 To examine the sensitivity of the poverty estimates to income model specifications, we compared 3 different models, which mostly vary in the number of explanatory variables they included. The models are called large, medium, and small models. The large model is presented in Table 1. In general, there is not much difference in poverty estimates between the three models. The large model produces lowest standard errors. Thus we incline to use the estimates from the large model to interpret the poverty for HCMC.

income. Finally, the spatial correlation is accounted for in the regression (the location error as a part of the total model error is presented in the bottom row in Table 1).[4]

Table 1. GLS regression on log of income per capita for South East

| **Explanatory variables** | Coef. | Std. Err. |
|---|---|---|
| Intercept | 9.4076 | 0.0921 |
| Ethnic minority (yes = 1) | -0.3173 | 0.1595 |
| Household size | -0.0472 | 0.0133 |
| Temporary house | -0.2764 | 0.0755 |
| Have no toilet | -0.2212 | 0.1028 |
| Ratio of primary school members | -0.4822 | 0.0947 |
| Ratio of post-secondary school members | 1.2793 | 0.4754 |
| Using telephone | 0.4488 | 0.0562 |
| HCM city * Ratio of primary school members | | |
| HCM city * Ethnic minority | -0.9952 | 0.3902 |
| Urban areas * HCM city | 0.2826 | 0.0896 |
| Urban areas * Head primary school | -0.2146 | 0.0667 |
| Urban areas * Temporary house | 0.3684 | 0.1221 |
| Urban areas * Using mobile phone | 0.2030 | 0.0810 |
| Urban areas * Ratio of lower secondary-school members | -0.4343 | 0.1324 |
| Urban areas * Ratio of post secondary-school members | -1.0487 | 0.5130 |
| Number of observations | | 512 |
| Adjusted R squared | | 0.516 |
| Number of clusters in survey | | 68 |
| Number of clusters in census | | 24 |
| $\hat{\sigma}_\eta^2 / \hat{\sigma}_u^2$ | | 0.128 |

*Source: Author's estimation*

## 5. Welfare Estimates

Once the income equation is estimated, it can be applied in the Mid-Census sample to estimate the poverty indexes of districts of HCM city for the year 2006. The poverty line used in this study is equal to 6000 thousand VND. This poverty line comes from HCM City People's Com. - Decision No. 145/2004/QĐ-UB on 25/5/2004 on poverty reduction strategy of HCMC.

According to the poverty estimates, the poorest district is Can Gio, followed by Nha Be. Many other districts have poverty rates lower than 10 percent. The urban households have much lower poverty than the rural ones. One important aspect of poverty is the poverty density and the number of poor households per district. Table 2 presents these estimates for HCMC in 2006. It shows that the poverty density, which is expressed as the number of poor per kilometer squared, is highest in urban districts and lowest in rural districts.

[4] Districts are specified as clusters.

Table 2. The number of poor households and the poverty density of districts of HCMC

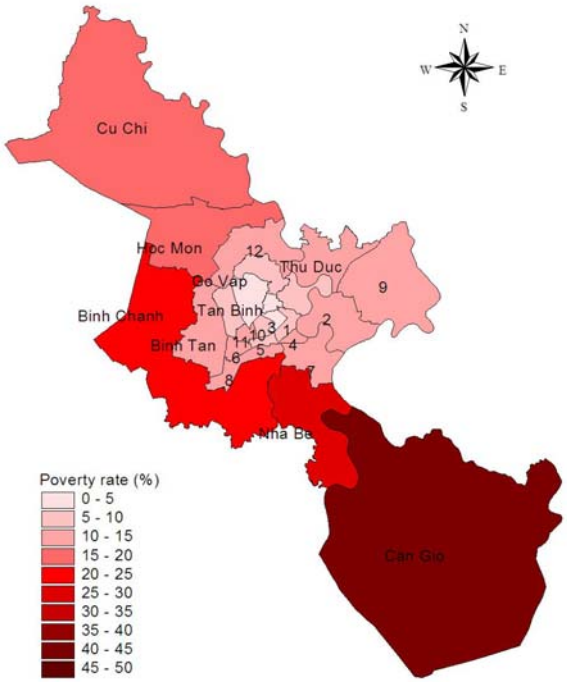
| Name | Urban Rural | No. of sampled households | Total households | Area (km2) | Percentage of poor households | No. of poor households | Poverty density (no. of poor households per km2) |
|---|---|---|---|---|---|---|---|
| 1 | Urban | 4490 | 44231 | 7.74 | 5.5 | 2428 | 314 |
| 2 | Urban | 3106 | 30616 | 48.89 | 10.8 | 3300 | 68 |
| 3 | Urban | 3576 | 42180 | 4.75 | 5 | 2101 | 443 |
| 4 | Urban | 2635 | 34100 | 4.06 | 12.2 | 4164 | 1026 |
| 5 | Urban | 3455 | 36020 | 4.08 | 7.2 | 2593 | 635 |
| 6 | Urban | 3734 | 47792 | 6.97 | 10.1 | 4837 | 693 |
| 7 | Urban | 4211 | 37348 | 34.60 | 10.1 | 3765 | 109 |
| 8 | Urban | 3902 | 71673 | 18.79 | 14 | 10056 | 535 |
| 9 | Urban | 4496 | 50958 | 111.96 | 10.1 | 5142 | 46 |
| 10 | Urban | 3630 | 47478 | 5.62 | 5.6 | 2635 | 469 |
| 11 | Urban | 3703 | 44555 | 4.98 | 10.4 | 4616 | 927 |
| 12 | Urban | 4332 | 70468 | 51.64 | 10.2 | 7202 | 139 |
| Go Vap | Urban | 4007 | 106079 | 19.59 | 6.1 | 6428 | 328 |
| Tan Binh | Urban | 3820 | 87110 | 22.12 | 4.5 | 3929 | 178 |
| Tan Phu | Urban | 4396 | 83365 | 15.47 | 7.7 | 6436 | 416 |
| Binh Thanh | Urban | 3844 | 88541 | 20.44 | 6.2 | 5498 | 269 |
| Phu Nhuan | Urban | 4126 | 37477 | 4.71 | 4.8 | 1803 | 383 |
| Thu Duc | Urban | 4221 | 86657 | 47.29 | 10.4 | 9038 | 191 |
| Binh Tan | Urban | 3750 | 105137 | 50.98 | 11.4 | 11944 | 234 |
| Cu Chi | Rural | 4254 | 70515 | 424.98 | 18.6 | 13116 | 31 |
| Hoc Mon | Rural | 4039 | 52943 | 106.94 | 16.7 | 8841 | 83 |
| Binh Chanh | Rural | 4318 | 77192 | 247.55 | 22.9 | 17692 | 71 |
| Nha Be | Rural | 3054 | 16277 | 98.32 | 30 | 4890 | 50 |
| Can Gio | Rural | 3268 | 14730 | 732.55 | 42.9 | 6321 | 9 |
| All the city | | 92367 | 1383442 | 2095.01 | 10.8 | 148775 | 71 |

*Source: Author's estimation*

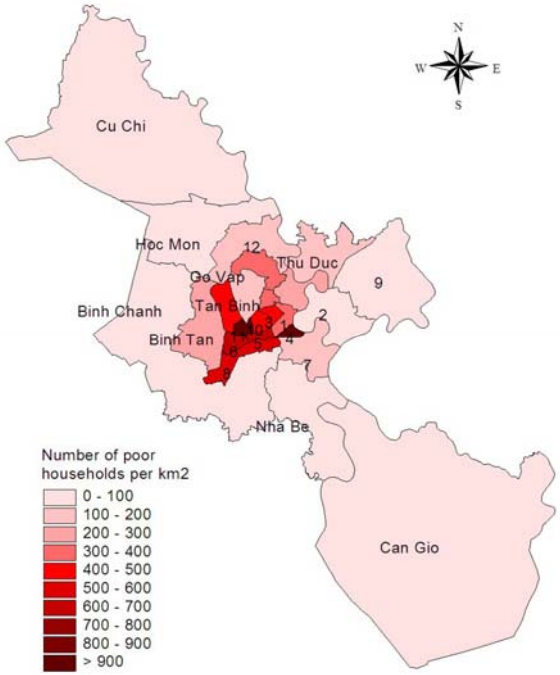
Figure 1 graphs the poverty incidence map and the poverty density map estimated from the large model, respectively. The pictures of poverty incidence and poverty density are opposites, since the population density in the rich districts is much higher than in the poor districts. The number of poor households of HCMC's districts is also presented in Figure 1.

Figure 1. Poverty estimates of districts of HCM city in 2006

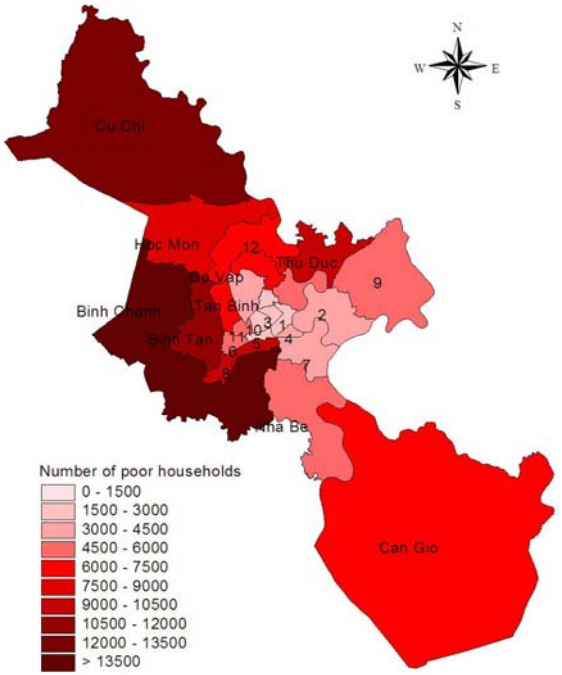
The poverty incidence




Poverty density

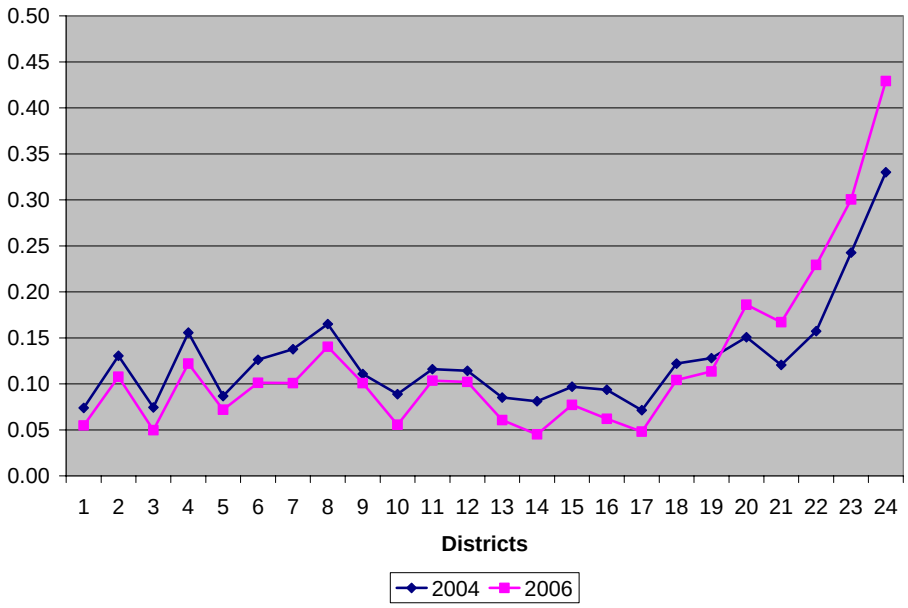



The number of poor households



It is worth noting that the estimates of the poverty incidences in 2006 are just slightly lower than the poverty incidence estimates in 2004 (Figure 2). The point estimates of the poverty rate of HCMC are 12 and 11 percent for 2004 and 2006, respectively. In addition, the poverty estimates of the rural district are even higher in 2006. However, due to associated standard errors, we cannot conclude whether the poverty reduced or increased during the period 2004-2006 based on the poverty estimates.

Figure 2. The poverty headcount index during 2004-2006




Finally, Table 3 presents the estimates of poverty gap and severity indexes and the Gini coefficient for the districts of HCMC. It shows that poverty gap and severity are much higher in rural districts than in urban districts. District Can Gio has highest poverty gap and severity indexes. Inequality within a district is not high. The Gini coefficient of districts ranges from around 0.28 to 0.31.

Table 3. Poverty estimates of districts of HCM city in 2006

| District name | Poverty gap index | | Poverty severity index | | Gini | |
|---|---|---|---|---|---|---|
| | Estimate | Std. error. | Estimate | Std. error. | Estimate | Std. error. |
| 1 | 0.0113 | 0.0090 | 0.0037 | 0.0033 | 0.2970 | 0.0093 |
| 2 | 0.0215 | 0.0164 | 0.0068 | 0.0058 | 0.3100 | 0.0093 |
| 3 | 0.0108 | 0.0077 | 0.0038 | 0.0028 | 0.2953 | 0.0094 |
| 4 | 0.0267 | 0.0182 | 0.0090 | 0.0070 | 0.3172 | 0.0089 |
| 5 | 0.0154 | 0.0114 | 0.0052 | 0.0044 | 0.3049 | 0.0095 |
| 6 | 0.0227 | 0.0162 | 0.0082 | 0.0062 | 0.3007 | 0.0088 |
| 7 | 0.0205 | 0.0151 | 0.0065 | 0.0054 | 0.3046 | 0.0090 |
| 8 | 0.0314 | 0.0184 | 0.0110 | 0.0070 | 0.3113 | 0.0088 |
| 9 | 0.0199 | 0.0149 | 0.0062 | 0.0053 | 0.2998 | 0.0089 |
| 10 | 0.0115 | 0.0083 | 0.0037 | 0.0030 | 0.2942 | 0.0096 |
| 11 | 0.0220 | 0.0166 | 0.0073 | 0.0062 | 0.3098 | 0.0091 |

| District name | Poverty gap index | | Poverty severity index | | Gini | |
|---|---|---|---|---|---|---|
| | Estimate | Std. error. | Estimate | Std. error. | Estimate | Std. error. |
| 12 | 0.0204 | 0.0153 | 0.0065 | 0.0053 | 0.2904 | 0.0095 |
| Go Vap | 0.0120 | 0.0098 | 0.0038 | 0.0035 | 0.2874 | 0.0098 |
| Tan Binh | 0.0090 | 0.0066 | 0.0030 | 0.0022 | 0.2835 | 0.0098 |
| Tan Phu | 0.0160 | 0.0124 | 0.0052 | 0.0046 | 0.2928 | 0.0097 |
| Binh Thanh | 0.0138 | 0.0102 | 0.0050 | 0.0039 | 0.3015 | 0.0097 |
| Phu Nhuan | 0.0101 | 0.0076 | 0.0033 | 0.0028 | 0.2918 | 0.0096 |
| Thu Duc | 0.0221 | 0.0167 | 0.0076 | 0.0062 | 0.3023 | 0.0093 |
| Binh Tan | 0.0229 | 0.0183 | 0.0072 | 0.0066 | 0.2849 | 0.0094 |
| Cu Chi | 0.0359 | 0.0297 | 0.0110 | 0.0105 | 0.2741 | 0.0117 |
| Hoc Mon | 0.0344 | 0.0272 | 0.0110 | 0.0105 | 0.2872 | 0.0100 |
| Binh Chanh | 0.0526 | 0.0326 | 0.0193 | 0.0128 | 0.2809 | 0.0102 |
| Nha Be | 0.0773 | 0.0387 | 0.0291 | 0.0170 | 0.3114 | 0.0100 |
| Can Gio | 0.1109 | 0.0589 | 0.0415 | 0.0270 | 0.2903 | 0.0100 |
| All the city | **0.0230** | **0.0055** | **0.0078** | **0.0021** | **0.3250** | **0.0120** |

## 6. Conclusions

This study estimates the poverty and inequality indexes for the districts in HCM city using the method of small area estimation. There are two data sources used for this estimation. The first is panel data from VHLSSs 2004 and 2006, which are used to run income regressions income for HCM city. The second is the 10 percent mid-census sample of HCM city. The estimates refer to the year 2006.

It is found that poverty estimates are much higher in rural districts than in urban districts. However, the poverty density is smaller in the poorest districts and higher in the richest districts, since the population density is much lower in the poorest districts than in the richest districts. The standard errors of the poverty estimates are relatively high, making the comparison of poverty between districts difficult, especially for districts with poverty rates less than 10 percent.

Compared with the poverty estimates in 2004, the poverty estimates in 2006 are slightly smaller. However, due to the high standard errors, it is difficult to compare the poverty indexes between 2004 and 2006.

## References

Cao, M. N. (2006) “Analysis of Economic Growth and Population Changes in Ho Chi Minh City during the Period 1995 – 2005”, Ho Chi Minh Institute for Development Studies (in Vietnamese).

Elbers, C., Lanjouw, J. and Lanjouw, P. (2002) “Micro-Level Estimation of Welfare” Policy Research Working Paper No. WPS 2911. The World Bank.

Elbers, C., Lanjouw, J. and Lanjouw, P. (2003) “Micro-Level Estimation of Poverty and Inequality” *Econometrica* 71(1), 355-364.

Foster, J., Greer, J., Thorbecke, E. (1984) “A Class of Decomposable Poverty Measures” *Econometrica* 52(3), 761-766.

Nguyen, V. C., Tran N. T., and Nguyen H. (2007) “Construction of poverty map for HCM city using the 2004 VHLSS and the 2004 HCM Mid-Census” Research Report, Ministry of Labor, Invalid and Social Affairs, Vietnam.

Nguyen, V. C., Nguyen V. P., and Ngo Q. T. (2005) “Construction of poverty map for HCM city using the 2002 VHLSS and the 2004 HCM Mid-Census” Research Report, Institute of Economic Research in HCM City, Vietnam.