Spatial patterns of poverty in Vietnam and their implications for policy

Nicholas Minot[*a], Bob Baulch[b]

[a] *International Food Policy Research Institute, 2033 K St. NW, Washington, D.C., U.S.A. Tel: +1 202 862-8199*. Fax: +1 202-467-4439. *E-mail address*: n.minot@cgiar.org

[b] *Institute of Development Studies, University of Sussex, Brighton, BN1 9RE, U.K. Tel: 44-1273-606261. Fax: +44(0)1273 621202. E-mail address*: b.baulch@ids.ac.uk

Abstract

This study examines the geographic distribution of poverty in Vietnam by applying small area estimation methods to household budget data and population census data. The resulting district-level poverty estimates suggest that the incidence of poverty is highest in the remote northern and central highlands and lowest in the south-east and in large urban centres. However, mapping the density

[*] Corresponding author.

of poverty reveals that most poor people do not live in the poorest districts but in the two lowland deltas, where poverty incidence is intermediate. The policy implications of these findings present an important trade-off between targeting poor areas and poor people that can only be resolved with better information on the relative costs of delivering different programmes and their expected impact. Existing government estimates of poverty at the district level are not closely correlated with our poverty estimates, perhaps because of regional variation in their methods of collecting poverty data.

Keywords: Spatial distribution; Poverty incidence; Programme targeting; Vietnam

Introduction

Information on the spatial distribution of poverty is of interest to policy makers and researchers for a number of reasons. First, it can be used to quantify suspected regional disparities in living standards and identify the areas falling behind in the process of economic development. Second, the information facilitates targeting programmes such as education, health, credit and food aid whose purpose, at least partly, is to alleviate poverty. Third, the information may shed light on geographic factors, such as topography or market access, which are associated with poverty.

Typically, household income and expenditure surveys are the main sources of information on spatial patterns of poverty. These surveys generally have sample sizes of 2000 to 8000 households, which typically only allow estimates of poverty for 3 to 12 regions within a country. Research has shown that geographic targeting is most effective when the geographic units are small, such as a village or district (Baker and Grosh, 1994, Bigman and Fofack, 2000). Census data are the only household information usually available at this level of disaggregation but census questionnaires are generally limited to household characteristics and rarely include questions on income or expenditure.

In Vietnam, there are two main sources of information on the spatial distribution of poverty. First are the estimates of poverty based on household income-expenditure surveys conducted by the General Statistical Office (GSO). The GSO carried out two Vietnam Living Standards Surveys (VLSS), one in 1992-93 with samples of 4800 households and the other in 1997-98 with 6000 households. The VLSS generate poverty estimates for each of the seven to eight regions of Vietnam.[1] More recently, the GSO set up its own household survey programme, known as Vietnam Household Living Standards Survey (VHLSS). The VHLSS has samples of between 25,000 and 75,000 households but a more limited core questionnaire than the VLSS and is intended to generate representative estimates for each of the 61 provinces in the country.[2]

A second source of information on the spatial distribution of poverty is the Ministry of Labour, Invalids and Social Affairs (MOLISA), which prepares a yearly list of poor households in each commune based on information gathered by local officials using MOLISA-established criteria. This information is used to identify the poorest communes, making them eligible for special programmes and subsidies to reduce poverty. Although the system is relatively inexpensive and provides annual estimates, different provinces use different poverty lines and different data collection guidelines in implementing this analysis. Furthermore, the use of thousands of unpaid local officials to collect household-level data makes it difficult to ensure consistent application of those guidelines in the field (see Conway, 2001).

Our study builds on two earlier poverty mapping studies of Vietnam. Minot (1998, 2000) used district-level means of the 1994 agricultural census data and the 1993 VLSS to rank districts by the incidence of rural poverty. The use of aggregated census data leads to biased poverty estimates, although Minot and Baulch (2005) show that the errors are modest, particularly if used for ranking. Minot and Baulch (2000) used household-level data from the 1999 Population and Housing Census (PHC) and the 1997-98 VLSS to generate provincial poverty estimates but could not generate reliable district-level estimates because the analysis was based on a 3% sample of the census data.

Data and methods

Data

The poverty mapping portion of our study makes use of two household data sets: the 1998 VLSS and the 1999 PHC. The GSO of Vietnam implemented the VLSS with funding from the Swedish International Development Agency and the United Nations Development Programme (UNDP), and with technical assistance from the World Bank. A stratified, clustered random sample of 6000 households in Vietnam was collected, which includes 4270 households in rural areas and 1730 households in urban areas.

The GSO, with financial and technical support from the United Nations Family Planning Agency and the UNDP, also carried out the 1999 PHC, which refers to the situation on 1 April 1999. The GSO does not make available full unit record level census results but we were able to obtain a 33% sample. The GSO selected this sample using linear systematic sampling of every third household in each census enumeration area. The sample includes 5,553,811 households.

The poverty line used in our study is the "overall poverty line" used for the analysis of the 1997-98 VLSS (see Poverty Working Group, 1999 and GSO, 2000). The poverty line corresponds to the expenditure (including the value of home production) required to purchase 2100 Kcal per person per day, plus a modest allowance for non-food expenditures. The food component of this poverty line is based on the food consumption pattern of households in the third quintile of the expenditure distribution, while the non-food component is equal to what a typical household at the poverty line spends on non-food items. This poverty line was set at 1,789,871 VND/person/year, with consumption expenditures from the VLSS adjusted using monthly and regional price indices to compensate for differences in the cost of living across regions and over the year spanned by the survey.[3]

Methods to estimate the incidence of poverty

Poverty mapping is one application of the small area estimation (SAE) method. Following Hentschel et al. (2000), the method is typically divided into three stages.[4] The first step compares the household survey and the census to identify common variables that can be used to predict poverty. In our study, we first compared the questionnaires and data values of the 1998 VLSS and the 1999 PHC. We selected 16 household characteristics for inclusion in the poverty mapping analysis, represented by 39 variables (see Table 1).

The second step involves using the household survey data and regression analysis to estimate household welfare as a function of household characteristics. In our study, we estimate real per capita consumption expenditure from the 1997-98 VLSS as a function of the 39 variables, adopting the conventional log-linear functional form to reduce heteroskedasticity:

$$\ln(y_i) = \mathbf{X}_i^{'}\boldsymbol{\beta} + \varepsilon_i \qquad (1)$$

where y_i is the real per capita consumption expenditure of household i, $\mathbf{X}_i'$ is a $1 \times k$ vector of household characteristics of household i, $\boldsymbol{\beta}$ is a $k \times 1$ vector of estimated coefficients and ε_i is a random disturbance term distributed as $N(0,\sigma)$. Because our main interest is predicting the value of $\ln(y)$ rather than assessing the impact of each explanatory variable, we are not concerned about the possible endogeneity of some of the explanatory variables. Elbers et al. (2003) show that the probability that household i with characteristics $\mathbf{X}_i$ is poor can be expressed as:

$$\mathrm{E}[p_i \mid \mathbf{X}_i, \boldsymbol{\beta}, \sigma^2] = \Phi\left[\frac{\ln z - \mathbf{X}_i'\boldsymbol{\beta}}{\sigma}\right] \tag{2}$$

where p_i is a variable taking a value of 1 if the household is poor and 0 otherwise, z is the "overall poverty line" (see GSO, 2000, p. 260) and Φ is the cumulative standard normal function.

In the third step, the estimated regression coefficients from the first step are combined with census data on the same household characteristics to predict the probability that each household in the census is poor. This is accomplished by inserting the household characteristics for household i from the census, $\mathbf{X}_i^c$, into Eq. (2). The expected probability that household i is poor can be calculated as:

$$\mathrm{E}[p_i \mid \mathbf{X}_i^c, \boldsymbol{\beta}, \sigma^2] = \Phi\left[\frac{\ln z - \mathbf{X}_i^c\boldsymbol{\beta}}{\sigma}\right] \tag{3}$$

This estimate is not very accurate for an individual household but becomes more so when aggregated over many households. For a given area (such as a district or province), Elbers et al. (2003) show that the proportion of the population living in households that are below the poverty line is estimated as the mean of the probabilities that individual households are poor.

Provided that (a) the error term is homoskedastic, (b) there is no spatial auto-correlation and (c) the full census data are used, the variance of the estimated poverty headcount can be calculated as:

$$\operatorname{var}(p^*) = \left(\frac{\partial p^*}{\partial \hat{\boldsymbol{\beta}}}\right)' \operatorname{var}(\hat{\boldsymbol{\beta}}) \frac{\partial p^*}{\partial \hat{\boldsymbol{\beta}}} + \left(\frac{\partial p^*}{\partial \hat{\sigma}^2}\right)^2 \frac{2\hat{\sigma}^4}{n-k-1} + \sum_{i=1}^{N} \frac{m_i^2 p_i^* (1-p_i^*)}{\mathrm{M}^2} \qquad (4)$$

where n is the sample size in the regression model. Thus, n, k and σ^2 are from the regression analysis, while m_i, M and N are obtained from the census data. The partial derivatives of p^* with respect to the estimated parameters can be calculated using equations provided by Hentschel et al. (2000).

As noted above, Eq. (4) is valid only if the full census data are available for the second stage of the mapping procedure. Since a 33% sample of the census data was available for use in this study, Eq. (4) must be expanded to include the sampling error associated with working with only part of the census data:

$$\operatorname{var}(p^*) = \left(\frac{\partial p^*}{\partial \hat{\boldsymbol{\beta}}}\right)' \operatorname{var}(\hat{\boldsymbol{\beta}}) \frac{\partial p^*}{\partial \hat{\boldsymbol{\beta}}} + \left(\frac{\partial p^*}{\partial \hat{\sigma}^2}\right)^2 \frac{2\hat{\sigma}^4}{n-k-1} + \sum_{i=1}^{N} \frac{m_i^2 p_i^* (1-p_i^*)}{\mathrm{M}^2} + \mathrm{V}_s \qquad (5)$$

where V_s represents the variance associated with the sampling error in the census and takes into account the design of the sample.

Qualifications

Two qualifications need to be made regarding the implementation of the poverty mapping method in the case of Vietnam. First, the regression analysis in Stage 1 does not explicitly take into account heteroskedasticity (differences in the variance of the dependent variable across the sample). On the other hand, by expressing the dependent variable (per capita expenditure) as a logarithm, we reduce the degree of heteroskedasticity. In addition, we use an estimator for the variance-covariance matrix (the Huber-White sandwich estimator), which is robust to heteroskedasticity and takes into account stratification and clustering in the sample. The estimated coefficients in Table 1 therefore are not biased, although they are "inefficient" in that they do not use all possible information (see StataCorp, 2001, Volume 4 "svyreg").

Second, the Stage 1 regression coefficients do not take into account spatial autocorrelation, which exists when either the dependent variable (or the error term) of the regression in one cluster of households in the VLSS is correlated with the dependent variable (or error term) in nearby clusters. If the error terms are correlated, the coefficients are again unbiased but inefficient. This would be the case if some other factors (such as distance to a major city) were excluded from the regression model and spatially correlated. Exploratory data analysis indicates the presence of some spatial autocorrelation of this type that we were unable to eliminate by including community-level variables in the regression analysis. On the other hand, if the dependent variable (per capita expenditure) in one cluster is affected by that in nearby clusters, then the estimated regression coefficients will be biased. Given the distance between neighbouring clusters in the VLSS, this seems less likely.

Spatial patterns in poverty

Household characteristics correlated with per capita expenditure

As described above, the first step in constructing a poverty map is to estimate econometrically per capita consumption expenditure as a function of variables that are common to the 1999 PHC and the 1997-98 VLSS. These household characteristics include household size and composition, ethnicity, education of the head of household and his or her spouse, occupation of the head of household, housing size and type, access to basic services and ownership of selected consumer durables.

It is reasonable to expect that the coefficients to “predict” expenditure in rural areas may be different from those predicting expenditure in urban areas. Indeed, statistical tests indicate that the coefficients in the urban model are significantly different from those in the rural model[5]. This implies that separate analyses should be carried out on rural and urban samples.[6]

Table 1 shows results of the rural and urban regression analysis. Both models explain more than half of the variation in per capita expenditure. This is a relatively good result for cross-sectional data but it is useful to keep in mind that other factors, not included in the model, explain almost half of the variation. For example, climate may influence housing characteristics, availability of television signals will affect television ownership and local infrastructure investment partly determines electrification, thus making these characteristics imperfect indicators of per capita expenditure.

According to the results in Table 1, large households are strongly associated with lower per capita expenditure in both urban and rural areas. The negative sign of the coefficient on household size implies that, other factors being equal, each additional household member is associated with a 7% to 8% reduction in per capita expenditure. In rural areas, households with a large proportion of elderly members, children and females are likely to be poorer. In urban areas, the proportion of females and elderly members per household is not significant, although a large share of children is still associated with poverty. After controlling for other household characteristics, ethnicity[7] is a surprisingly weak predictor of per capita expenditure. In rural areas, the coefficient on ethnicity was significant only at the 10% level, while in urban areas where few ethnic minority people live, it was not statistically significant from zero.

In both urban and rural areas, the level of education completed by the head of household is a good predictor of a household's per capita expenditure. The five variables that represent the education of the head of household are jointly significant at the 1% level in both rural and urban areas. In rural areas, each level of education completed by the head is associated with significantly higher levels of per capita expenditures relative to the omitted category of not completing primary school. In urban areas, households whose head has completed primary or lower secondary school are no better off than those whose head has not completed primary school, although higher levels of education are associated with significantly higher expenditures.

In general, the educational level of the spouse is less significant than that of the household head as a predictor of per capita expenditure. In rural areas, only the highest two levels of education of the spouse show any significant effect relative to the spouse not completing primary school. The education of the spouse is a better predictor in urban areas than in rural areas (see Table 1). It is worth noting that the education of the spouse may yield benefits that are not captured by our welfare measure (per capita expenditure), such as better child nutrition or health care.

Table 1 also shows that the occupation of the head of household is a statistically significant predictor of per capita expenditure in both rural and urban areas. In rural areas, the first three occupational categories are significantly better off than households in which the head of household is not working (the omitted category). In urban areas, households whose head is a leader/manager are significantly better off than those with non-working heads, while those whose head is an unskilled worker are significantly worse off.

Various housing characteristics are good predictors of expenditures. Living in a dwelling made of permanent rather than temporary materials is associated with 23% higher per capita expenditure in rural areas and 27% higher in urban areas. Similarly, having a house of semi-permanent rather than temporary materials implies a 14% to 15% higher level of per capita expenditure. The living area of houses is also a useful predictor of household well-being: each 10% increase in area is associated with a 13% to 34% increase in per capita expenditure, depending on the area of residence (urban or rural) and the type of house (permanent or semi-permanent).

Access to basic services is also a useful predictor of household per capita expenditures. Electrification[8] is a statistically significant predictor of household welfare in rural areas, where 71% of households have access to electricity. By contrast, in urban areas, where 98% of households are already electrified, electricity is not a significant predictor of expenditures. Similarly, households with access to well water in rural areas have a higher level of per capita expenditure than households using river or lake water (the omitted category). In urban areas, more than half the sample households (58%) have access to tap water compared to just 2% in rural areas, so this variable is a good predictor of urban per capita expenditures.

Sanitation facilities can also be used to separate poor from non-poor households. In rural areas, flush toilets and latrines are statistically significant indicators of higher per capita expenditure at the 5% level. In urban areas, having a flush toilet is a significant predictor of expenditures at the 5% level but having a latrine is not.

Television ownership is one of the strongest predictors of per capita expenditures, being a statistically significant predictor at the 1% level in both urban and rural areas. Radio ownership is almost as good a predictor, although, as would be expected, the coefficients for radio ownership are lower than those for television ownership.

Finally, regional dummy variables were included in the urban and rural regression models, with the Northern Uplands as the base region. Even after controlling for other household characteristics, rural households in the four southern regions are shown to be better off than those in the Northern Uplands. The coefficient in the South-east is the largest, implying that households in this region have expenditure levels 72% higher than similar households in the Northern Uplands. A similar pattern holds for urban households. The regional dummy variables are jointly significant at the 1% level in both urban and rural areas.

Incidence of poverty

The incidence of poverty (also known as the poverty rate or poverty headcount) is defined here as the proportion of the population living in households whose per capita expenditure is below the "overall poverty line," as defined above. This is the Foster-Greer-Thorbecke measure of poverty when $\alpha = 0$, also known as p_0. The small sample estimation technique allows poverty to be estimated at the national, provincial, district and commune level but we focus on the district-level estimates of poverty incidence.

At this point, it is appropriate to note that the national incidence of poverty, as estimated in this application of the SAE method using a 33% sample of the 1999 PHC data, is 35.9%. By comparison, the 1997-98 VLSS produces an estimate of 37.4% (see Poverty Working Group, 1999), thus confirming that the household characteristics in the PHC have similar values to those in the 1998 VLSS.

Incidence of poverty (p_0) at the district level

The three-step poverty mapping method described above was used here to generate poverty estimates at the district level. The spatial patterns in the incidence of poverty can be seen in Fig. 1. The incidence of poverty is highest in the North-east and North-west regions, especially among the districts along the border with China and the Lao People's Democratic Republic, the interior areas of the central coast and the northern parts of the Central Highlands. The incidence of poverty is intermediate in the two main deltas of Vietnam, the Red River and Mekong River Deltas, and lowest in the urban areas, particularly Hanoi and Ho Chi Minh City and the South-east. Pockets of poverty exist elsewhere, however, including in the North Central and South Central Coast plus the Mekong River Delta.[9]

Fig. 2 shows the 95% confidence intervals for the incidence of poverty at the district level. The solid line represents our estimates of poverty incidence at the district level, with districts ordered from the least to most poor. The small horizontal dashes above and below the solid line show the upper and lower 95% confidence limits for each district-level estimate. These confidence intervals range from ±1.3 to ±22 percentage points, with an average value of ±5.8 percentage points. Half the districts have confidence intervals between ±4.4 and ±6.9 percentage points (the inter-quartile range). The confidence intervals are smaller (and the poverty estimates more accurate) when the incidence of poverty is close to 0% or to 100%. When the incidence of poverty is in the middle range (40% to 50%), the confidence intervals tend to be ± 5 to ±10 percentage points.

The least reliable district estimate is for Bach Long Vi District in Hai Phong Province: the incidence of poverty is estimated as 19% ± 22%. This poverty estimate is imprecise because it is based on a sample of just 18 households in this island district. This district is an exception, however. The second highest confidence interval is ±12 percentage points. Furthermore, only six of the 614 districts have fewer than 1000 households in our sample of the PHC data, while 90% of them have more than 2500 households.

Poverty density

Fig. 1 shows the incidence of poverty, defined as the percentage of the population living below the poverty line. Another way to look at the spatial distribution of poverty is to examine the poverty density, defined as the number of poor people living in a given area. This is produced by multiplying the incidence of poverty by the population in each area.[10] Fig. 3 shows the poverty density in Vietnam, where each dot represents 1500 poor people. At first glance, it is rather surprising to find that the number of poor people per square km is greatest in the more prosperous parts of Vietnam—the Red River Delta, Mekong River Delta and along the coastal plains. In contrast, poverty density is lowest in the areas where the incidence of poverty is the highest. This is because the areas with the highest incidences of poverty tend to be remote and sparsely populated areas, so their lower population densities more than offset the higher percentages of the population that are poor.

An important implication of Fig. 3 is that if all poverty alleviation efforts are concentrated in the areas where the incidence of poverty is the highest, including the North-east, North-west, Central Highlands and the interior of the central coast, most of the poor will be excluded from the benefits of these programmes. We discuss the implications of this map further in the concluding section.

Relationship with MOLISA poverty estimates

As mentioned above, MOLISA prepares a yearly list of poor households in each commune, based on information gathered by local officials. The welfare indicator is per capita income, expressed in terms of the number of bags of rice it will buy at local prices. The poverty line is defined in terms of the number of kilograms of rice per person per month: 15 kg in mountainous areas, 20 kg in the deltas and midlands and 25 kg in the mountainous areas. In addition, some provinces have adopted different poverty lines.

The MOLISA district-level poverty estimates are generally lower than those generated by the SAE method used (Fig. 4). The median value of the MOLISA estimates of the incidence of poverty is 15%, compared to 41% for our poverty estimates. This is not surprising, given that the poverty line calculated for the 1998 VLSS used in our analysis is 1.79 million VND/person/yr, whereas the monetary equivalent of the MOLISA poverty line ranges from 0.66 to 1.08 million VND/person/yr.

Fig. 4 also shows little correlation between the district-level poverty estimates produced by MOLISA and those generated by our study: the R^2 of a linear trendline is just 0.17. The difference between income and expenditure probably cannot explain this difference. The weakness of our method is that we are estimating per capita expenditure based on household characteristics rather than measuring it directly. The main weakness of the MOLISA approach is that it relies on the efforts of thousands of local officials who may not use the same criteria and may have strategic reasons for citing certain poverty rates. For example, local officials may overstate poverty in order to qualify for anti-poverty programmes or understate it to meet poverty reduction targets.

To illustrate the disagreement in the estimates, we consider two districts in which the contrast between the two methods is the greatest. In the upper left corner of Fig. 4 is a dot representing Bat Xat District in Lao Cai. Our poverty estimate for the district is almost 82%, while that of MOLISA is less than 6%. Lao Cai is one of the poorest provinces in Vietnam by any measure and Bat Xat is a remote mountainous district on the Chinese border, less well off than the rest of Lao Cai in terms of various indicators: urbanization, occupation, electrification, education, water source and ownership of radio and television. These characteristics suggest that a high poverty rate in Bat Xat is quite plausible.

Towards the lower right corner of Fig. 4 is the district of Nha Trang in Kanh Hoa Province. The MOLISA poverty estimate for Nha Trang is 68%, while the estimate produced in our report is just 15%. Nha Trang is a provincial capital on the South Central Coast whose economy benefits from substantial local and international tourism as Vietnam's best-known beach resort. These facts and the relatively good household indicators imply that a low poverty rate is more plausible.

Clearly, the choice of poverty estimates can make a large difference in terms of the targeting of poverty alleviation programmes. Further research is needed to resolve the discrepancies between these two poverty estimates. One approach would be to select districts where the two estimates vary widely (such as the two cited above) and collect primary or secondary data to determine which estimate conforms more closely to reality. Casual inspection of the two districts with the greatest discrepancies supports our estimates.

Summary and conclusions

It is well known that Vietnam, like most countries, has a spatially distinct pattern of poverty. The incidence of poverty in Vietnam is highest in the Northern Uplands and lowest in the cities and the South-east. However, our poverty estimates show that poverty also varies widely across districts. In some districts, particularly remote ones in the upland areas, over 90% of the population lives below the poverty line. In others, particularly in or near the large urban centres, less than 5% of the population is poor.

Mapping of poverty density (the number of poor people per unit of area) provides a different perspective of the spatial distribution of poverty in Vietnam. This mapping reveals that the density of poverty is greatest where the incidence of poverty is lowest. In other words, only a relatively small percentage of Vietnam's poor live in the poorest areas. The absolute number of poor people that live in areas with a high incidence of poverty is relatively low because the population density in these areas is also low. By contrast, most of the rural poor live in the Mekong Delta and the Red River Delta. Although these areas have relatively low poverty incidence compared to other rural areas, the population density ensures that most of the poor live in the two deltas.

If most poor people live in less poor areas, what are the implications for targeting poverty alleviation programmes? In particular, should poverty alleviation programmes concentrate their efforts on areas with the greatest poverty density? The answer to these questions depends on the type of poverty alleviation programme. Some programmes are relatively untargeted and benefit all households in an area. Examples of such programmes might be better roads, better health care and financial support to local government. If such programmes have a fixed cost per inhabitant, they will have a greater effect on poverty if they are concentrated on areas with high poverty incidence. In these areas, a higher percentage of the population is poor, so a higher percentage of the beneficiaries of untargeted programmes will be poor.[11] Other programmes are specifically targeted to poor households (e.g. income transfers, food for work or social service fee exemptions). If the goal is to provide the same level of assistance to each poor person, such programmes should spend more overall in areas with many poor people. Indeed, if the aim is to help poor households escape poverty, it may even be argued that more should be spent on targeted programmes in areas with favourable endowments and moderate numbers of poor people (as these will be the most likely to be able to escape poverty).

Of course, these guidelines assume the cost of providing the programme is constant in per capita terms, implying that population density does not affect the cost. Some programmes, such as electrification and extension, will cost more in per capita terms in low-density areas. Other programmes, particularly land-intensive ones such as roads and parks, may be more expensive in a high-density area. It is therefore essential to consider both the spatial distribution of poverty and the relative costs of delivering untargeted and targeted programmes.

Acknowledgements

We would like to thank the New Zealand Agency for International Development Principal for funding for this study and the World Bank and the Swiss Agency for Development and Cooperation for providing additional support. We would also like to thank Michael Epprecht, formerly with the International Food Policy Research Institute, for his work in preparing the poverty maps.

References

Baker, J., Grosh, M., 1994. Poverty reduction through geographic targeting: how well does it work? World Development 22 (7), 983-995.

Bigman, D., Fofack, H., 2000. Geographic targeting for poverty alleviation: methodology and applications. World Bank Regional and Sectoral Studies, Washington D.C.

Conway, T., 2001. Using government data to target activities to poor communes and monitor poverty reduction: a review of options for the Cao Bang-Bac Kan Rural Development Project. Commission of the European Communities, Hanoi.

Elbers, C., Lanjouw, J., Lanjouw, P., 2003. Micro-level estimation of poverty and inequality. Econometrica 71 (1), 355-364.

GSO (General Statistical Office). 2000. Vietnam living standards survey 1997-1998. Statistical Publishing House, Hanoi.

Henninger, N., Snel, M., 2002. Where are the poor? Experiences with the development and use of poverty maps. World Resources Institute, Washington DC and United Nations Environment Programme/Global Resources Information Database (UNEP-GRID) Arendal, Norway.

Hentschel, J., Lanjouw, J., Lanjouw, P., Poggi, J., 2000. Combining census and survey data to trace the spatial dimensions of poverty: a case study of Ecuador. World Bank Economic Review 14 (1), 147-165.

Minot, N., 1998. Generating disaggregated poverty maps: an application to Viet Nam. Markets and Structural Studies Division, Discussion Paper no. 25. International Food Policy Research Institute, Washington DC.

Minot, N., 2000. Generating disaggregated poverty maps: an application to Vietnam. World Development 28 (2), 319-331.

Minot, N., Baulch, B., 2002. The spatial distribution of poverty in Vietnam and the potential for targeting. Markets and Structural Studies Division, Discussion Paper no. 42. International Food Policy Research Institute, Washington DC.

Minot, N., Baulch, B., 2005. Poverty mapping with aggregate census data: what is the loss in precision? Review of Development Economics 9 (1), 5-25.

Poverty Working Group, 1999. Attacking poverty: Viet Nam development report 2000. Joint report of the Government of Vietnam-Donor-NGO Poverty Working Group. World Bank, Hanoi.

StataCorp, 2001. Stata reference manual, release *7*. Stata Press, College Station, Texas.

---

[1] When this study was carried out, Vietnam was administratively divided into 61 provinces, 614 districts and 10,474 communes. Until 1998, the provinces were grouped into seven regions for statistical purposes: the Northern Uplands (also called the “North Mountains and Midlands”), the Red River Delta, the North Central Coast, the South Central Coast, the Central Highlands, the South-east (also called the “North-east South”) and the Mekong River Delta. In 1998, the regional system was revised by splitting the Northern Uplands into the North-east and North-west regions, moving three provinces out of the Central Highlands.

[2] After this study was completed, two of these 61 provinces were split, making currently 63 provinces in Vietnam.

[3] The average exchange rate for 1997-98 was 12,476 VND/US$, so the poverty line is equivalent to US$143 per capita. The food represents 72% of the value of the basket of goods used to define the poverty line.

[4] This technique for poverty mapping has been applied in more than 12 developing countries including Cambodia, Ecuador, Indonesia, Panama, Thailand, South Africa and Uganda (see Henninger and Snel, 2002).

[5] More specifically, the Chow test strongly rejects the hypothesis that the coefficients for the urban subsample are the same as those for the rural subsample ($F = 6.16$, $p < 0.001$).

[6] In an earlier analysis, we tried estimating models for two urban and seven rural strata (see Minot and Baulch, 2002). The regression results were not satisfactory, with lower values of R^2, more coefficients that were statistically insignificant and some coefficients that were the "wrong" sign.
[7] Following the classification commonly used in Vietnam, ethnic minorities are defined as all ethnic groups except for *Kinh* (ethnic Vietnamese) and *Hoa* (ethnic Chinese).
[8] More specifically, this variable refers to the main type of lighting used by the households.
[9] Note that the sharp line between the low incidence of poverty in the South-east (as defined in 1998) and the higher rates in the Central Highlands and South Central Coast are partly an artificial result of the use of regional dummy variables in Stage 1.
[10] The dots are distributed randomly within each commune.
[11] This is certainly true if the goal is to reduce the depth of poverty (p_1) and it is probably true if the goal is to reduce the incidence of poverty (p_0).

Figure Legends

Fig. 1. Map of the incidence of poverty (p_0) of each district in Vietnam

Fig. 2. Incidence of poverty and confidence intervals for each district in Vietnam

Fig. 3. Map of the density of poverty in Vietnam

Fig. 4. Comparison of the incidence of poverty (p_0) for Vietnam estimated by the Ministry of Labour, Invalids, and Social Affairs (MOLISA) and by the small area estimation method

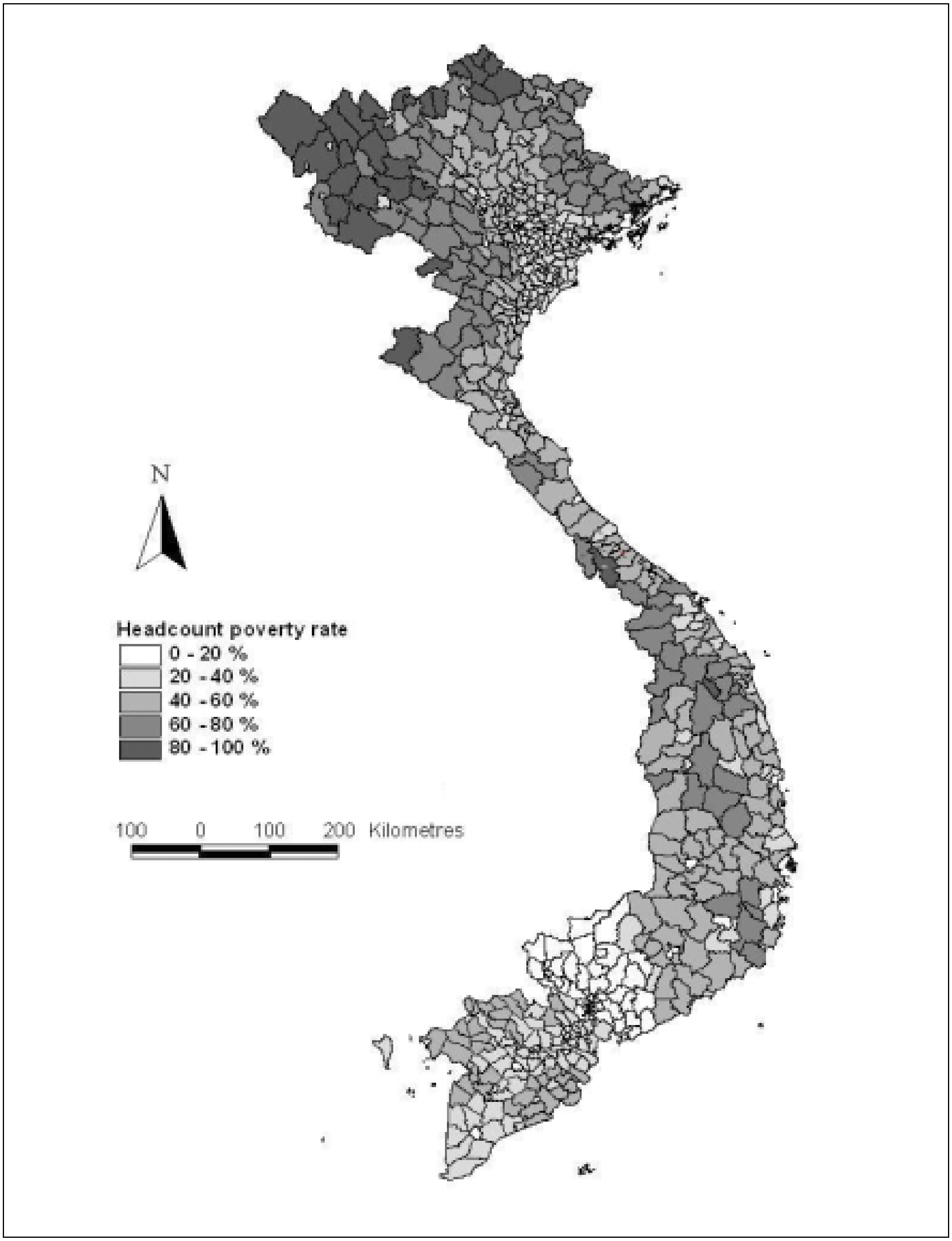

N
Headcount poverty rate
0 - 20 %
20 - 40 %
40 - 60 %
60 - 80 %
80 - 100 %
100 0 100 200 Kilometres

Fig. 1

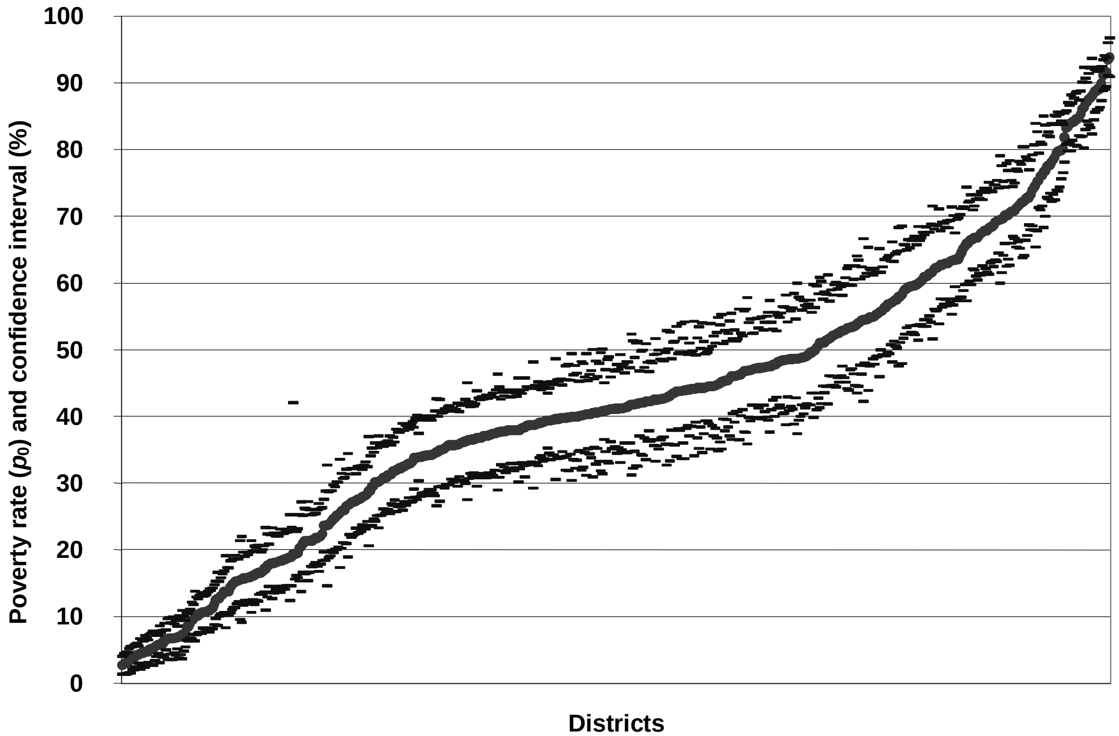

100
90
80
70
60
50
40
30
20
10
0
Poverty rate (p_0) and confidence interval (%)
Districts

Fig. 2

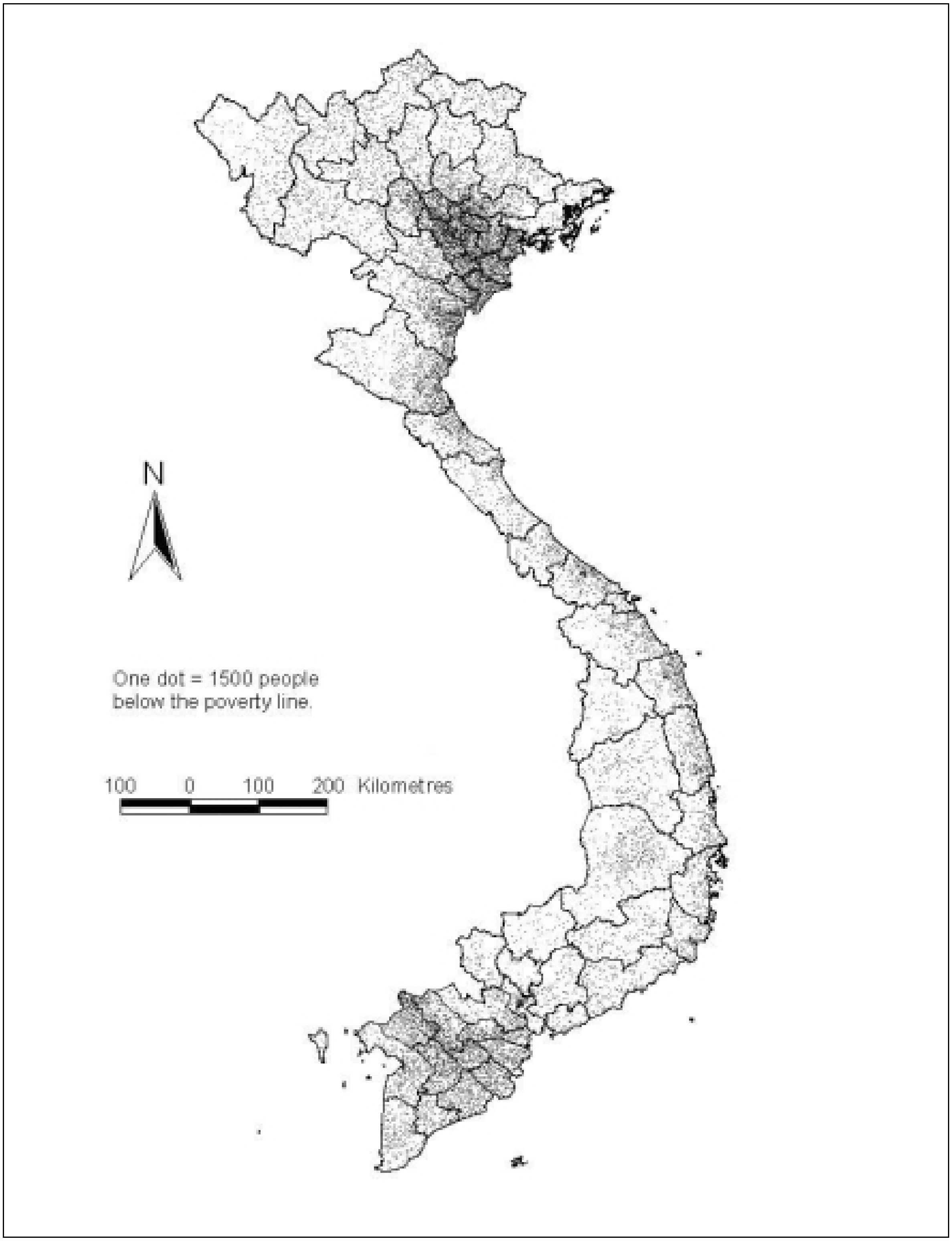
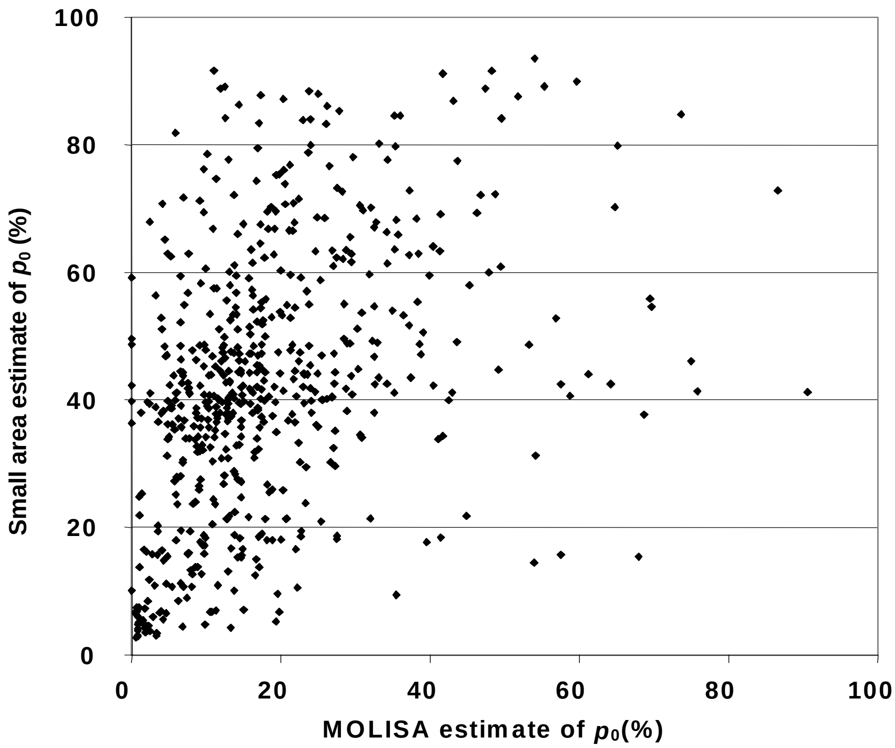

N
One dot = 1500 people
below the poverty line.
100 0 100 200 Kilometres

Fig. 3



100
80
60
40
20
0
0
20
40
60
80
100
Small area estimate of p_0 (%)
MOLISA estimate of p_0(%)

Fig. 4

Table 1

Rural and urban regression models of log per capita expenditure[a]

Category[b]	Rural model		Urban model	
N	4269		1730	
R^2	0.536		0.550	
	Coefficient	t[c]	Coefficient	t[c]
Household size (members)	-0.0772	-19.5 ***	-0.0785	-8.1 ***
Proportion >60 yr (fraction)	-0.0831	-2.4 **	-0.1026	-1.6
Proportion <15 yr (fraction)	-0.3353	-9.4 ***	-0.2368	-3.6 ***
Proportion female (fraction)	-0.1177	-3.5 ***	0.0386	0.5
Household head is ethnic minority	-0.0765	-1.9 *	0.0142	0.2
Head has completed primary school	0.0585	3.4 ***	0.0616	1.7
lower secondary school	0.0883	4.5 ***	0.0338	1.3
upper secondary school	0.0884	3.3 ***	0.1368	3.2 ***
advanced technical degree	0.1355	4.2 ***	0.1603	3.5 ***
Head has post-secondary education	0.2552	4.9 ***	0.1843	3.7 ***
Head has no spouse	0.0173	1.0	0.0344	0.8
Spouse has completed primary school	0.0049	0.3	0.0642	1.9 *
lower secondary school	0.0132	0.6	0.0987	2.6 **
upper secondary school	0.0107	0.3	0.1912	2.7 **
advanced technical degree	0.0921	2.3 **	0.1285	3.2 ***
Spouse has post-secondary education	0.1571	2.7 ***	0.1752	3.1 ***
Head is a political leader or manager	0.1414	3.5 ***	0.2312	3.0 ***
a professional or technical worker	0.1350	3.3 ***	0.0576	1.2
a clerk or service worker	0.1362	3.4 ***	0.0357	0.9
in agriculture, forestry, or fishing	-0.0163	-0.6	-0.0093	-0.2
a skilled worker	0.0701	1.9 *	0.0071	0.2
an unskilled worker	-0.0586	-1.7 *	-0.1599	-2.9 ***
House made of permanent materials	-0.9228	-4.3 ***	-0.5194	-3.4 ***
semi-permanent materials	-0.3120	-3.6 ***	-0.4001	-3.8 ***
Interaction of log(house area) and permanent house	0.2958	5.7 ***	0.2001	5.4 ***
semi-permanent house	0.1180	5.2 ***	0.1403	4.6 ***
House has electricity	0.0765	2.7 ***	-0.0026	0.0
uses water from a public or private tap	0.0828	1.4	0.2289	5.3 ***
uses well water	0.1157	4.4 ***	0.0340	0.6
has flush toilet	0.2700	5.5 ***	0.1311	2.2 **
has latrine	0.0556	2.6 **	0.0049	0.1
Household has television	0.2124	15.1 ***	0.2167	5.5 ***
radio	0.1009	7.0 ***	0.1599	6.2 ***
Household lives in Red River Delta	0.0314	0.6	0.0693	0.7
North Central Coast	0.0485	0.8	0.0445	0.6
South Central Coast	0.1373	2.2 **	0.1460	1.9 *
Central Highlands	0.1708	2.1 **	omitted	
South-east	0.5424	9.4 ***	0.4151	5.5 ***
Mekong River Delta	0.3011	5.1 ***	0.1895	2.1 **
Constant	7.5327	108.7 ***	7.7538	64.7 ***

[a]*Source*: Regression analysis of 1997-98 Vietnam Living Standard Survey, taking into account clustering and stratification and using robust estimates of standard errors.

[b]Omitted categories are: head has no education; spouse has no education; head is not working; house is made of temporary materials; household has other water source; household has no sanitation facilities; and household lives in the Northern Uplands.

[c]Student's t-test: * significant at $p = 0.05$, ** at $p = 0.01$ and *** at $p = 0.001$.