Revised Draft

ETHNICITY AND HOUSEHOLD WELFARE IN VIETNAM: EMPIRICAL EVIDENCE FROM 1993 TO 2004

Bob Baulch*, Hung T. Pham*,** and Barry Reilly**

November 2007

Abstract

The primary purpose of this paper is to determine the relative welfare position of different ethnic groupings in Vietnam using data spanning a twelve-year period corresponding to radical economic transformation in Vietnam. The analysis reported in this study confirms that the Kinh majority has been the primary beneficiaries of the Doi moi reform process. The living standards of Kinh-headed households have risen relative to the rural average over the period 1993 to 2004, whether we examine the poorest, richest or the average Kinh-headed household. The relative position of the Khmer and Cham in recent times has also improved in rural areas while that of the Chinese (Hoa) has declined, so that by 2004 these groups were found to be statistically indistinguishable from the national average. However, sizeable and persistent gaps in household welfare are found to remain for the Northern and Central Highlands Minorities. Our findings also suggest that the disadvantaged position of Vietnam's ethnic minority groups cannot be attributed exclusively to the role of geography and the concentration of ethnic minorities in the more remote parts of the country.

* Institute of Development Studies, University of Sussex, Brighton, BN1 9RE
** Department of Economics, University of Sussex, Brighton, BN1 9SN

Acknowledgement: This research was funded by the UK Economic and Social Research Council and Department for International Development under their Joint Research Scheme (Award Number RES-167-25-0157). We thank Nguyen Thang and Nguyen Thi Thu Phuong and seminar participants at the Committee for Ethnic Minorities, NGO Resource Centre, Vietnam Academy of Social Sciences and the World Bank for their comments on an earlier version of this paper.

Introduction

The rapid economic growth experienced in Vietnam during the 1990s and early 2000s resulted in unprecedented reductions in household poverty incidence. The 54 officially recognized ethnic groups within Vietnam's diverse society have not, however, shared equally from the benefits of this growth. Poverty, life expectancy, nutritional status, and other living standard measures remain stubbornly low among Vietnam's ethnic minorities despite numerous policies introduced to assist these groups.

Previous studies for Vietnam (Van de Walle and Gunewardena, 2001; Baulch *et al*. 2004, Hoang *et al*. 2007), using a per capita household expenditure measure, have investigated the gap in living standards between majority and broadly defined minority groups at specific points in time using mean regression analysis in conjunction with standard Oaxaca (1973) decompositions. Although the aggregation of distinct groups is necessary and inevitable in such an exercise, the simple majority-minority dichotomy used in these studies may distort important differences that may exist between individual ethnic groups. Therefore, this paper exploits a different empirical approach to the analysis of inter-ethnic differences in living standards and uses information on seven different ethnic groupings, six of which could be interpreted as relatively homogenous in composition.

The primary purpose of our research is to determine the relative welfare position of seven ethnic groups in Vietnam using household survey data covering a twelve-year period that corresponds to an era of radical economic transformation in Vietnam. In order to satisfy this objective, we exploit an approach originally suggested by Krueger and Summers (1988) to assess the nature and persistence of the inter-industry wage structure in the United States, an approach subsequently modified and enhanced by Haisken-De New and Schmidt (1997). The methodology has heretofore not been applied, to the authors' knowledge, to interrogate the magnitude or persistence of ethnic differentials in living standards.

An exclusive focus on the mean will, however, provide an incomplete account of the nature of inter-ethnic living standard differentials in Vietnam. So we also estimate a set

of conditional quantile functions allows for a more detailed portrait of the relationship between the welfare measure and selected covariates (including the ethnic controls) than that provided by the mean regression function. Thus, another theme of the current paper is to examine inter-ethnic differences in household welfare at different points of the conditional welfare measure's distribution using quantile regression functions. The use of these functions requires modification to the mean-based deviation approach popularized by Krueger and Summers (1988) and Haisken-De New and Schmidt (1997), and we suggest a potentially useful refinement to their approach. We also take the analysis further by using inter-quantile regression models to investigate the presence of intra-ethnic differences in household welfare within Vietnam.

The structure of the paper is now outlined. The next section provides a contextualization for our empirical analysis by way of a review of recent economic events in Vietnam. It also details the nature of Vietnam's ethnic diversity. This is then followed by sections outlining the data sources and the econometric methodology used. A subsequent section discusses the empirical results and is followed by a final section that outlines the policy implications and offers some concluding remarks.

Background

The *Doi moi* (economic renovation) reforms of the late 1980s stimulated rapid economic growth in Vietnam over the last two decades and this has impacted strongly on poverty and welfare at the household level. Between 1993 and 2004, Vietnam's national poverty headcount fell from 58.1 to 19.5 percent, while educational enrolments, life expectancy and other measures of human development increased dramatically (VASS, 2007). Though the different groups within Vietnam's ethnically diverse society have reaped rewards from such growth, benefits have generally not been shared equally. For instance, despite numerous policies and programmes designed to assist minority groups, the poverty headcount rate among Vietnam's broadly defined ethnic minorities fell from 86.4 to 60.7 percent between 1993 and 2004 (VASS, 2006). School enrollments, nutritional indicators and life expectancy also remain low among the minorities.

Vietnam has 54 officially recognized ethnic groups, of which the Kinh (the Việt or mainstream Vietnamese) accounted for 86.7 percent in 1999 (Dang *et al*., 2000). With the exception of the Hoa (ethnic Chinese) and the Khmer and Chăm, the remaining 50 ethnic groups mostly reside in remote, mountainous rural areas and are economically and socially disadvantaged across a range of dimensions (Poverty Task Force, 2002; Hoang *et al*., 2007). The members of ethnic minority groups are estimated to be four-and-a-half times more likely to be poor than the Kinh-Hoa, and are also more likely to be malnourished, illiterate, and suffering from poor health. Despite comprising just over one-tenth of the national population, the minorities accounted for about 40 percent of the poor in 2004 (VASS, 2006). Some government agencies forecast that by 2010, the ethnic minorities will constitute more that half of Vietnam's poor (MOLISA, 2005).

Previous studies investigating ethnic minority issues in Vietnam (Van de Walle and Gunewardana, 2001; Baulch *et al*., 2004; Hoang *et al*., 2007) have used the household surveys conducted in 1993, 1998 and 2004, and have relied on a simple dichotomy between the Kinh-Hoa and all other ethnic minority groups. Although, this is a useful and convenient device for demonstrating disparities in living standards between majority and minority ethnic groups, such simple comparisons potentially conceal important distinctions across individual ethnic groups.

Traditionally the Kinh have inhabited lowland and coastal areas in and around Vietnam's two densely populated deltas (the Red River Delta and the Mekong River Delta) while the Hoa tend to live in urban areas (especially Ho Chi Minh City). With the exception of the Khmer (who are concentrated in the Mekong Delta) and the Chăm (who are located along the southern coast), most other ethnic groups are scattered across Vietnam's upland and highland areas. These stretch from the northern border with China through the Annamite Cordillera in North-Central Vietnam to the Central Highlands, which borders the South Central Coast (McElwee, 2004). Within the upland and highland areas, some ethnic groups (in particular, the Tày, Thái, Mường, Nùng – each of which have populations of close to one million) specialize in wet-rice cultivation and usually live in

the flat, lower areas along the valley bottoms (the 'midlands'). Other less populous groups (such as the Hmong, Dao and Kho-mu in the Northern Uplands and the Ede, Bana and Hre in the Central Highlands) tend to live in higher, more mountainous areas where rice often cannot be grown. There are also 17 ethnic groups with populations of less than 10,000, some of which are likely to disappear in the absence of special measures to protect them (CEM, 2006). All ethnic groups have their individual identities which embody diverse and unique cultures.

The widening disparities in living standards observed between the different ethnic groups have been driven by a combination of factors. The most frequently cited relates to the role of geography and the associated specialization of economic activities described above. Until recently, this has been exacerbated by Vietnam's *ho khau* (or household registration) system, which aims to restrict internal migration, particularly between rural and urban areas. Nonetheless, Vietnam's population in recent years has been urbanizing quite rapidly growing from 20 percent of the population in 1993 to over one-quarter in 2004 (GSO, 2006). In addition, there has been both government sponsored and 'spontaneous' migration from the deltas into what are generally perceived as thinly populated regions, such as the Northern Uplands and the Central Highlands.[1] Such migration has served to both increase the ethnic diversity of many rural areas and accentuate inter-regional inequality, though the populations of Vietnam's urban areas remain predominantly Kinh and Hoa. The highly skewed pattern of industrialization and foreign direct-investment, which has been concentrated around the growth poles of Ho Chi Minh City in the South and Hanoi in the North, have served to further exacerbate regional differences (VASS, 2006). As a consequence, many ethnic minority groups are thus on the margins of the rapid economic growth experienced by the urban and coastal areas of Vietnam, and consequently constitute a growing share of the country's extreme poor.

[1] However, once the 'carrying-capacity' of the generally more fragile uplands is taken into account, this perception is probably erroneous (see Jamieson, Cuc and Rambo (1998))

Data

The empirical analysis in this paper uses data drawn from household-level surveys conducted in three separate years. These surveys were implemented by Vietnam's General Statistical Office (GSO) under funding and technical support from UNDP, the World Bank and other donors. The Vietnam Living Standards Surveys (VLSS) of 1992/93 and 1997/98 are multi-topic surveys patterned after the World Bank's Living Standard Measurement Surveys with nationally representative samples of 4,800 and 6,000 households respectively (see World Bank, 2000; 2001). These surveys were superseded in 2002 and 2004 by a new biennial household survey programme known as the Vietnam Household Living Standards Surveys (VHLSS), which uses a rotating core-and-module designed survey with an expanded sample size intended to provide statistics that are representative at the provincial level (Phung and Nguyen, 2006).[2] Since various concerns have been expressed regarding the level of non-sampling errors in the first of the VHLSS surveys, we restrict attention to the 1993, 1998 and 2004 surveys in this paper. We also confine our attention to rural areas, as this is where the vast majority of ethnic minority people live in Vietnam and it is well-known that the VLSS and VHLSS surveys have under-sampled migrant households (many of whom will come from the ethnic minorities) in urban areas (Pincus and Sender, 2006).

There is a panel of 4,300 households between the two earlier VLS surveys, and a separate panel of around 4,000 households between the two more recent VHLS surveys. However, there is no panel linking the VLSS and VHLSS. Both the VLSS and VHLS surveys have clustered, stratified sampling designs. Though the content of the questionnaire has evolved over time, the core information contained within the four surveys facilitates the construction of a set of variables that are compatibly defined across the relevant years.

Following Van de Walle and Gunewardena (2001) and Baulch *et al*. (2004), we use per capita household expenditure as the metric to examine inter-ethnic differences in welfare

[2] The number of households surveyed in the expenditure modules of the VHLSS 2002 and VHLSS 2004 was 29,530 and 9,189, respectively.

in Vietnam. Although per capita expenditures are an incomplete measure of welfare, there is considerable evidence to suggest that many of the more commonly used non-monetary measures of well-being are highly correlated with expenditures in Vietnam.[3] The measure is defined as real household expenditure per capita, with the expenditure computed on the basis of household food and non-food consumption over the past 12 months. The living standard variable is expressed in real January 2004 prices using the GSO's monthly CPI price deflator.

The surveys also include information on the household head and their spouse, education and age of all household members, the household's demographic structure, physical assets (particularly access to different types of land), geographical location, the date of interview, and general infrastructure and socio-economic conditions in the communes where the sampled households are situated. The survey data thus provide a rich set of variables over time that can be used to model household welfare. Table A1 provides both a description of the variables used and selected summary statistics.

The ethnic status of a particular household is defined in terms of the ethnicity of the household head but no other household member. The list of 54 ethnic groups used in the survey questionnaires from 1998 onwards is similar to those proposed by the Vietnamese Institute of Ethnology.[4] In this paper, it has been necessary to combine some of the smaller ethnic minority groups into composite categories, in order to ensure that sufficiently large cell sizes are available for the subsequent analysis of inter-ethnic welfare effects. The groupings we have used are intended to be functional and are based on discussions with Vietnamese anthropologists and local NGOs. The ethnic categories comprise the (1) Kinh (Việt) majority; (2) Chinese (Hoa); (3) Chăm and Khmer; (4) Tày, Thái, Mường, Nùng; (5) Other Northern minorities; (6) Central Highland minorities; (7) an 'others' or miscellaneous category comprising the remaining smaller ethnic groups, which are mostly located in the North and South Central Coasts. Table A2 of the appendix provides a detailed description of the allocation of the 54 official recognized

[3] See for example the chapters on education, health and nutrition in Glewee, Agrawal and Dollar (2004).
[4] Note that the VLSS conducted in 1993 only recorded the ethnicity of household heads using a narrower list of ten ethnic codes, which partly dictates the nature of the aggregation undertaken.

ethnic groups across these seven categories. A sufficient number of observations is available for each of the ethnic minority groups defined above for each year available to us, though small cell sizes in particular cases warrants interpretational caution. This caveat is particularly relevant to the Chinese and 'All others' categories and to the Central Highlands Minorities in 1993 (see table A3 of the appendix).

Table A3 reveals an improvement in the real per capita expenditure levels across all the ethnic groups between 1993 and 2004. The largest increases were observed for the Kinh majority and the Chinese, while other ethnic groups benefited significantly less. In this context, while the average poverty headcount has substantially reduced between 1993 and 2004, it is evident poverty is still widespread among the ethnic minority groups.

Figure 1 about here

We can glean some further insights on the changes in poverty and welfare over time by plotting the kernel densities for per capita household expenditure for the three years using the conventional dichotomy between the Kinh and the Chinese, and all other ethnic groups. The plots are reported in Figure 1. The poverty line, using the GSO and World Bank criterion, is also super-imposed on these densities.[5] The unbroken plot represents the Kinh and Chinese and the broken line that of the other ethnic minority groups. In general, the densities for the majority group are strongly right-skewed compared to the minority group. The inter-ethnic differences in headcount poverty rates are also evident from an inspection of these plots. The contraction in headcount poverty is fairly apparent in conducting a direct comparison between 1993 and 2004, though it is also the case that the magnitude of the difference in regard to this improved welfare outcome remains sizeable between the two broadly defined groups.

[5] There are two poverty lines in common use within Vietnam: the GSO-World Bank poverty line (which is based on a standard cost-of-basic-needs methodology and estimated from the V(H)LSS) and the MOLISA poverty lines (which are used for targeting and monitoring the number of poor households at the commune level). The 'international' $1 and $2 a day poverty lines are rarely used for poverty analysis within Vietnam due to PPP conversion factor issues.

Econometric Methodology

The logarithm of per capita household expenditure is defined by y_i and the following equation is then specified:

$$y_i = \mathbf{x}_i'\boldsymbol{\beta} + \mathbf{d}_i'\boldsymbol{\gamma} + u_i \qquad [1]$$

where $\mathbf{x}_i$ is a k×1 vector of household-specific and other characteristics with the first element a one, $\mathbf{d}_i$ is a 6×1 vector containing binary variables that capture the ethnic affiliation of the head of household[6], $\boldsymbol{\beta}$ is a k×1 vector of unknown parameters, $\boldsymbol{\gamma}$ is a 6×1 vector of unknown parameters corresponding to six of the ethnic groups, and u_i is an error term assumed to satisfy standard properties.

The above model can be estimated using OLS to obtain the set of mean regression coefficients. The ethnic effects can be interrogated more thoroughly by normalizing the estimated ethnic effects as a deviation from a hypothetical overall sample weighted average. The transformation has appeal in that the estimated differences are then expressed relative to an overall sample-based average rather than an arbitrary base group and are thus more easily interpretable. If we define the effect for the k^{th} ethnic group as γ_k, the deviation for the k^{th} group (D_k) is expressed as:

$$D_k = \gamma_k - \sum_{j=1}^{7} \pi_j\gamma_j \qquad [2]$$

where π_j is the sample average proportion for the j^{th} ethnic group.[7] The sampling variances are computed using the approach suggested by Zanchi (1998) as:

$$V(\mathbf{D}) = (I - e\pi')V(\boldsymbol{\gamma})(I - e\pi')'$$

[6] One ethnic category is thus omitted in estimation, though it should be stressed that the results are invariant to including all ethnic groups and setting the constant term to zero in [1].

[7] The ethnic base group in estimation attracts a zero coefficient in this exercise if the constant term is included in equation [1].

where V(**D**) represents the variance-covariance matrix for the deviations from the average expressed in [2], I is a (7×7) identify matrix; e is a (7×1) vector of ones; π is a (7×1) vector with elements comprising the sample proportions for the seven ethnic groups; and $V(\boldsymbol{\gamma})$ is the (7×7) estimated variance-covariance matrix for γ_j obtained from [1] suitably modified to cater for the zero coefficient corresponding to the ethnic category that provides the base group in estimation.[8]

A measure of overall variability is also calculated and, following Haisken-De New and Schmidt (1997), is computed as the square root of the weighted adjusted standard deviations in [2], where the weights are again provided by the ethnic sample proportions and the adjustment is introduced to correct for the sampling variance associated with using sample estimated coefficients.

An exclusive focus on the mean, however, may provide an incomplete account of the nature of inter-ethnic differentials in log per capita household expenditures. The estimation of a set of conditional quantile functions allows for a more detailed portrait of the relationship between the conditional household expenditure distribution and selected covariates. In contrast to the OLS approach, the quantile procedure is less sensitive to outliers and provides a more robust estimator in the face of departures from normality (Koenker, 2005; Koenker and Basset, 1978). In addition, Deaton (1997, pp.80-85) notes that quantile regression models may also possess better properties than OLS in the presence of heteroscedasticity.

In the quantile regression framework, the median regression coefficients can be obtained by choosing the coefficient values that minimize L:

$$L = \sum_{i=1}^{n} \left| y_i - \mathbf{x}_i'\boldsymbol{\beta} - \mathbf{d}_i'\boldsymbol{\gamma} \right| = \sum_{i=1}^{n} \left(y_i - \mathbf{x}_i'\boldsymbol{\beta} - \mathbf{d}_i'\boldsymbol{\gamma} \right) \mathrm{sgn}\left(y_i - \mathbf{x}_i'\boldsymbol{\beta} - \mathbf{d}_i'\boldsymbol{\gamma} \right) \qquad [3]$$

where sgn(a) is the sign of a, 1 if a is positive and –1 if a is negative or zero, and n equals the sample size. The estimator is known as the Least Absolute Deviations (LAD)

[8] See Haisken-De New and Schmidt (1997) for further details.

estimator and the median regression estimates can be obtained by minimizing the absolute sums of the errors. This ensures the estimates are less sensitive to extreme observations than in the case of the mean regression. However, it is also desirable to explore quantile regressions at points other than the median. Thus, the log household expenditure equation may be estimated conditional on a given specification and then calculated at various percentiles of the residuals (e.g., the 10^{th}, the 25^{th}, the 75^{th} or the 90^{th}) by minimizing the sum of absolute deviations of the residuals from the conditional specification.

In the current application, the quantile regression model in its general form can be expressed as:

$$y_i = \mathbf{x}_i'\boldsymbol{\beta}_\theta + \mathbf{d}_i'\boldsymbol{\gamma}_\theta + u_{\theta i} \qquad [4]$$

where $Q_\theta(y_i \mid \mathbf{x}_i, \mathbf{d}_i) = \mathbf{x}_i'\boldsymbol{\beta}_\theta + \mathbf{d}_i'\boldsymbol{\gamma}_\theta$ and $Q_\theta(u_{\theta i} \mid \mathbf{x}_i, \mathbf{d}_i) = 0$; and $\boldsymbol{\beta}_\theta$ and $\boldsymbol{\gamma}_\theta$ comprise the unknown parameter vectors corresponding to the $\mathbf{x}_i$ and $\mathbf{d}_i$ vectors respectively for the θ^{th} quantile. Thus, quantile regression analysis provides estimates for the elements of the $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ parameter vectors at, for example, $\theta = 0.1$, $\theta = 0.25$, $\theta = 0.75$, or $\theta = 0.90$ (i.e., the 10^{th}, 25^{th}, 75^{th} and 90^{th} percentiles respectively).

It has become conventional to obtain the sampling variances corresponding to the quantile regression model estimates using bootstrapping procedures. This is the approach adopted in this study and the sampling variances reported for the point estimates corresponding to the ethnic (and other) coefficients are based on a bootstrapping procedure with 200 replications.

Given that the investigation of inter-ethnic differences in the log of per capita household expenditures is a primary focus of our analysis, the normalization approach undertaken in [2] for the mean regression requires refinement. In expression [2] the weights used to compute the deviation measures were based on sample average proportions. However, using the sample mean proportions in conjunction with quantile regression estimates may

provide unrepresentative realizations for the weights at points other than the conditional mean to which they relate. Therefore, it is necessary to use weights that more accurately reflect the relevant points on the conditional household expenditure distribution.

In order to address this issue, we use a variation of an approach originally suggested in a different context by Machado and Mata (2005) to derive the necessary weights. The procedure uses bootstrap sampling and draws at random with replacement 100 observations from the original sample of households. Each observation, once ranked, is taken to comprise a percentile point on the log household expenditure distribution. The ethnic characteristic of the head of household in terms of being in one of the seven mutually exclusive ethnic groupings is then recorded in binary form for each percentile. This process is then replicated 500 times to obtain 500 observations at each of the θ^{th} quantiles. The average of the observations corresponding to the ethnic status at each selected quantile is then used to construct the realizations for the ethnic proportions, which are then used as the sample weights in the computation of the deviations. The quantile regression formula for the ethnic deviations can then be expressed as:

$$D_{\theta k} = \gamma_{\theta k} - \sum_{j=1}^{7} \pi_{\theta j}\gamma_{\theta j} \qquad [5]$$

where $\gamma_{\theta k}$ is the estimate for the k^{th} ethnic group using the θ^{th} quantile function and $\pi_{\theta j}$ is the sample proportion in the j^{th} ethnic group conditional on being at the θ^{th} quantile of the log of per capita household expenditure, obtained using the bootstrap sampling method outlined above. The sampling variance for the point estimates in [5] and measures of overall variability can be constructed in a manner analogous to the mean regression case.

Empirical Results

The log per capita household expenditure equations are estimated by OLS for the 1993, 1998 and 2004 surveys using a sample restricted to the rural households. The specifications used are broadly based on Baulch *et al*. 2004, though the educational variables are specified on the basis of Van de Walle and Gunewardena, 2001 (see table A1). Given our research interest in effecting a meaningful comparison between mean

and quantile regression model estimates, district dummies are not included in these initial specifications. Instead, a more aggregate set of controls for commune type is included to capture potential spatial and/or geographical variation in household expenditure effects. In addition, commune level dummy variables reflecting access to transportation, markets, electrification and non-farm employment were included the regressions. The sensitivity of the OLS estimates to the inclusion of the district effects is examined separately below.

As a prelude to the empirical analysis, we first report in table 1 raw deviations computed using expression [2]. These deviations are calculated without introducing any controlling variables or district dummies. The table confirms the advantage enjoyed by the Kinh majority and reveals that their raw advantage relative to the average has more than doubled time. The relative position of both the Chinese, Khmer and Chăm has narrowed relative to the mean over time, with the substantial advantage experienced by the Chinese in 1993 becoming statistically indistinguishable from the mean in 1998 and 2004. In contrast, the position of the Tày, Thái, Mường and Nùng, Other Northern Minorities and Central Highlands minorities have worsened. The table also reveals evidence of an increased variability in inter-ethnic differences over time.

Table 1 about here

Attention now turns to the regression-based estimates that include a variety of controls. The estimated coefficients for the mean regression equations are not the subject of independent discussion here but are reported for completeness in Appendix A4. It is worth noting that the estimates are generally plausible in terms of both sign and magnitude, and the equation fits are satisfactory. Table 2 provides *ceteris paribus* estimates based on computing deviations from the sample average for the seven Vietnamese ethnic groups again using expression [2]. Kinh-headed households are still found to secure an advantage relative to the sample average with strong evidence again that their advantage has more than doubled over time.

Table 2 about here

The sizeable advantage enjoyed by Chinese-headed households has exhibited a steady decline between 1993 and 2004, and by the final year of our study, is found to be statistically indistinguishable from the rural average. Some interpretational caution is required here given the limited number of Chinese-headed households reported surveyed in all three years (see table A3 of the appendix).[9] Table 2 also reveals a diminution over time in the relative disadvantage observed for the Khmer and Chăm. This reflects the Khmer and Cham's residence in the faster growing lowland areas in southern Vietnam, which they tend to live side-by-side with the Kinh.

The relative disadvantage experienced by the other four ethnic categories have remained persistent across the twelve year period. The relative position of the Tày, Thái, Mường, Nùng and other Northern Minorities has deteriorated over time by a statistically significant factor, and by 2004 households headed by members of these ethnic groups had per capita expenditures that were both about 25 percent below the rural averages. In contrast, the position of the Central Highlands minorities has improved substantially since 1993. Nevertheless, per capita expenditure in households headed by those attached to this ethnic category are lower than all other categories except for the miscellaneous all others category in 2004. The Central Highlands Minorities had expenditures 30 percent lower than the sample average in the most recent survey year. The all others category miscellaneous ethnic category also experienced a steady widening in their welfare gap.

The analysis reported for table 2 indicates that the rank order in the relative household welfare position of the seven ethnic groups has exhibited strong persistence over time, though the primacy of the Chinese ethnic group within rural areas has been subject to substantial erosion in recent years. The Khmer and Chăm group's third placed ranking is consistent both across years, and there is little to distinguish between the rankings of the Tày, Thái, Mường, Nùng and the other Northern Uplands minorities. The Central

[9] It may also be speculated that part of the Kinh advantage noted above may reflect the effects of 'self-identification' of the Chinese with the Kinh. This issue was investigated further by using the panel dimensions of the VLSS (from 1993 to 1998) to identify the incidence of change in the self-reported ethnic status of the head of households over these two separate periods. Approximately one-tenth of the panel heads, who identified themselves as Chinese in 1993, re-classified themselves as Kinh in 1998.

Highlands minorities are always ranked last, with the exception of 2004 when their gap is similar to the all others category. There is also evidence from the mean regressions of increased variability over time in inter-ethnic household welfare differences. However, the magnitude of this *ceteris paribus* variability is considerably less than that computed using the raw differentials of table 1.

Table 3 about here

Table 3 repeats the above exercise using mean regression models that replace the controls for the geographic type of commune with a more extensive set of district dummies. This allows for a richer specification of the unobservable factors influencing household welfare that may have hitherto been absorbed within the ethnicity measures reported in table 2. In particular, these controls might capture more effectively the role of geography, which is usually considered to be an important factor in the determination of ethnic disadvantage in Vietnam. The introduction of these controls attenuates the estimated ethnic effects to some extent, though the foregoing narrative respect to the improving advantage of the Kinh, the disappearing relative advantage of the Chinese and the worsening position of the Tay-Thai-Muong-Nung remains broadly intact. However, with district dummies, the Khmer and Chăm are found not be statistically different from the average in any of the three survey years. With the inclusion of district effects, the Central Highlands Minorities relative disadvantage has decline to approximately one-third of its level in 1993, while the gap for the others category is only significantly different from the overall mean in 1998.

Not surprisingly, given the introduction of the district fixed effects, the overall variability in inter-ethnic household expenditure is reduced for two of the three years, declining by two-and-a-half percentage points between 1993 and 2004. However, any temporal comparisons with respect to this table carries the important caveat that the number of district fixed effects is considerably larger in 2004 than the two earlier years, so the

conditioning set cannot be considered constant across the time period reviewed here.[10] Nevertheless, geography does appear to play a modest role in moulding some of the inter-ethnic differences in welfare observed for Vietnam. ** Can we say something about the role of the commune level access/infrastructure variables here **

We now turn to the estimates for the ethnic deviations based on the quantile regression models. These models are estimated again using both the pooled samples and the rural samples. The specification of the log per capita household expenditure model is identical to that reported for the mean regression case in table 2 and their explanatory variables contains the dummies for the geographic type of commune and its socio-economic characteristics instead of the extensive district dummies used in table 3. Tables 4 to 6 report the deviations based on the quantile regression estimates for the four years using expression [5].[11] The overall picture that emerges both in terms of magnitude and rankings is broadly similar to that suggested by the mean regression estimates.

Tables 4 to 6 about here

Although, some of the sample sizes on which these quantiles are estimates are rather small (especially for the Chinese), several features stand out from these tables. First, the relative advantage of the Kinh, while growing over time, remains more or less constant across the selected quantiles of the conditional expenditure distribution. Thus, it is reasonable to infer that the differential favouring the majority ethnic group has widened over this time period for both rich and poor Kinh-headed households alike relative to the corresponding conditional average. Second, the relative disadvantage experienced by the Other Northern and Central Highland Minorities narrows as one moves up the conditional expenditure distribution. Indeed in 2004, the 90th percentile for these ethnic categories is statistically indistinguishable from the Tay-Thai-Muong-Nung, while those at the 10th percentile are considerably poorer. This suggest that those members of the Other

[10] The number of district fixed effects in rural areas is 120, 150 and 574 for 1993, 1998, and 2004 surveys respectively.
[11] In order to conserve space, the full set of estimates for the quantile regression models are not reported here. They are available from the authors on request.

Northern and Central Highland Minorities who are able to get ahead fare as well as their Tay-Thai--Muong-Nung counterparts, but the poorest members of these groups are seriously disadvantaged. Third, there is weak evidence of Khmer and Cham households converging toward the mean for all rural households, though this is occurring more slowly for household above the median than those below it. Finally, Tables 4 to 6 also confirm the erosion of the sizeable advantage held by Chinese-headed households in 1993. By 2004 any advantage in respect of this group versus the overall mean has disappeared at all percentiles of the conditional welfare distribution.

In regard to the overall variability measure, the pattern over time is somewhat mixed. There is evidence for the first and particularly the last year of a greater degree of inter-ethnic variability among the poorest households at the bottom end of the welfare distribution than among the richest households near the top. In contrast, the evidence from the quantile regression models for the two intermediate years suggests a more stable pattern, although variability is higher at the 25^{th} and 90^{th} percentiles.

Table 7 about here

Finally, we investigate the stability of the inter-ethnic differences across the distribution of the conditional welfare measure using inter-quantile regression models based on differences between the 90^{th} and 10^{th} percentiles. The ethnic group inter-quantile regression model estimates are then used to compute deviations from a sample weighted baseline using a variant of expression [5] above. The weights used are based on the ethnic proportions at the median of the conditional household expenditure distribution obtained using the bootstrap sampling method outlined in the methodology section. Table 7 reports the estimates for this exercise. For most ethnic groups and in most years, there is a substantial degree of stability in the estimated inter-quantile ethnic effects across the conditional household expenditure distribution. However, the pattern for the Tày, Thái, Mường, Nùng and the other Northern Uplands minorities is more uneven. For the former, the estimated difference between the rich (at the 90^{th} percentile) and poorest (at the 10^{th} percentile) households within this ethnic grouping is below the

median-weighted baseline in 1998 but indistinguishable from this baseline in the other two years. For the Northern Uplands minorities, the first year reveals inter-quantile differences that are above the median-weighted baseline, while in the second year this difference is below the median, and in the third year statistically indistinguishable from This suggests that the Other Northern Minorities are characterised by a greater degree of intra-ethnic welfare differences than most other ethnic groups in Vietnam. There is also some evidence of variability in inter-quantile difference between years for the Chinese and All Others categories. However, not too much should be read into these estimates as the inter-quantile regressions are based on samples of 50 observations or less.

Conclusions

In 1946, Ho Chi Minh famously asserted that:

> 'As people born from the same womb, whether Kinh or Tho, Muong or Man, Gia Rai or Ede, Xedang or Bana, or any other ethnic minority, all of us are the children of Vietnam, all of us are brothers and sisters. We live and die together, share happiness and sorrow together, [and] whether hungry or full, we help each other.'

It is now over thirty years since the re-unification of Vietnam and almost twenty years since the *Doi moi* economic reform process was first initiated. The process of rapid economic growth has certainly been of central importance in removing many households from poverty and improving the well-being of the Vietnamese people across a broad range of dimensions. However, on the basis of the empirical analysis reported in this paper and elsewhere, it is clear that not all ethnic groups have benefited from this process to the same extent and there are sizeable and temporally persistent inter-ethnic differentials in household welfare.

Our analysis confirms that the Kinh majority has been the primary beneficiaries of the *Doi moi* reform process and the living standards of those households headed by members of this group have widened sharply relative to a hypothetical national average over the period 1993 to 2004. This particular finding is invariant to whether we examine the poorest, richest or the average Kinh-headed household. The average advantage in household welfare traditionally enjoyed by the traditionally more prosperous Hoa (ethnic Chinese) appears to have disappeared over time, while the Khmer and Chăm have

recorded modest improvements in their relative position. By 2004, the welfare position of these two ethnic categories is found to be statistically indistinguishable from the rural average.

Sizeable and persistent inter-ethnic gaps in household welfare are found to remain for the other four groupings. For the Tay-Thai-Muong-Nung and Other Northern Minorities categories, the size of the inter-ethnic gap has been growing over time with a deficit of about one-quarter of the rural average in 2004. For the Central Highlands Minorities, the gap has declined substantially but expenditure levels for both this group and the All Others categories were still approximately a third lower than the rural average in 2004.

The findings of this paper also suggest that the disadvantaged position of Vietnam's ethnic minority groups cannot be attributed exclusively to the role of geography and the concentration of ethnic minorities in the more inhospitable locations within the country. The addition of an elaborate set of district-specific fixed effects to capture, *inter alia*, the role of geographical location fails to alter the underlying narrative that emphasized persistent gaps in living standards between the Kinh and most other ethnic minority groups. Thus, although geography is an important factor which helps explain the disadvantaged position of the ethnic minorities in Vietnam, it cannot be interpreted as the key driver for the gaps observed.

The Government of Vietnam, over the last twenty years, has introduced a plethora of programmes and policies designed to assist the country's ethnic minority groups (see Nguyen and Baulch, forthcoming). It is always difficult to conceptualize the counter-factual but it may well be the case that, in the absence of such interventionss, poverty among minority-headed households would have been greater than is currently the case. Nevertheless, these policies and programs appear to have done little to narrow the gap between the majority Kinh and most other ethnic groups. If anything, the passage of time has seen a widening in the advantaged position enjoyed by the majority group.

Ethnic minority disadvantage is, of course, not peculiar to Vietnam. Studies in other countries show the underlying causes for these wide and persistent gaps in living standards are multi-faceted and vary from one ethnic group to another.[12] Within Vietnam, a less than exhaustive set of factors would include poor endowments of human capital among the minorities, the unavailability of land of an adequate quality in upland areas, more limited access to credit, product and labour markets, and poor competence in the dominant Viet language among the minorities. All of these factors will influence the nature of the economic activities undertaken by the ethnic minorities. In addition, some ethnic groups probably receive lower returns to endowments, both at the mean and the lower end of the distribution, than their Kinh counterparts—a topic we shall be investigating further in a follow-up paper.

The time may therefore be apposite for the Government of Vietnam, donors and NGOs to re-appraise their policies and programmes they have designed to assist the ethnic minorities and, in particular, to recognize that interventions designed to reduce poverty and inequality may need to have an ethnic group specific focus. The diversity in the socio-economic development experiences of the different ethnic groups demands greater diversity in the anti-poverty and other policies designed to assist them. For such an approach to be effective, this will require greater inputs from ethnic minority households, and more decentralization in the anti-poverty interventions than has occurred hitherto.

[12] See, for example, Psacharopoulos and Patrinos (1994) for a review of this issue from a Latin American perspective.

References

Baulch, B., Truong, C., Haughton, D. and Haughton, J., 2004, 'Ethnic minority development in Vietnam: a socio-economic perspective', Chapter 8 in Glewwe, P., Agrawal, N., and Dollar, D.

Committee for Ethnic Minorities (CEM), 2006, *Annual Report*, Hanoi: Committee for Ethnic Minorities

Dang, N.V., Son, Chu.,T.S., and Lu,H.,, 2000, *Ethnic Minorities in Vietnam*, Hanoi: The Gioi Publishers.

Deaton, A., 1997, *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*, Baltimore: John Hopkins University Press

Glewwe, P., Agrawal, N., and Dollar, D. (eds), *Economic Growth, Poverty and Household Welfare in Vietnam*, World Bank Regional and Sectoral Studies, Washington DC: World Bank

Haisken-DeNew, J. P. and Schmidt, C. M. (1997) Inter-industry and inter-region differentials: mechanics and interpretation. *Review of Economics and Statistics*, Vol. 79, pp. 516-521.

Hoang, H., Pham, G., Tran, M., and Hansen, H., 2007. 'Ethnicity and poverty reduction' in Nguyen, T. and Hansen, H. (eds) *Markets, Policy and Poverty Reduction in Vietnam,* Hanoi: Vietnam Academy of Social Sciences/National Political Publishing House

Jamieson, H., Cuc, L.T., and Rambo, T., 1998, *The Development Crisis in Viet Nam's Mountains*, Honolulu: East West Centre

Koenker, R., 2005, *Quantile Regression*, Econometric Society Monographs, Cambridge University Press

Koenker, R. and G. Bassett, 1978, 'Regression quantiles', *Econometrica* , Vol. 46: 33-50.

Krueger, A. and Summers, L, 1988, 'Efficiency wages and the inter-industry wage structure', *Econometrica* 56(2): 259-293

Machado, J. and J. Mata, 2005, 'Counterfactual decomposition of changes in wage distributions using quantile regression models', *Journal of Applied Econometrics, 20: 445-465* .

McElwee, P, 2004, 'Becoming socialist or becoming Kinh? Government policies for ethnic minorities in the Socialist Republic of Vietnam", in Duncan. C. (ed.) *Civilizing the Margins: Southeast Asian Government Policies for the Development of Minorities*, Ithaca: Cornell University Press

Ministry of Labour Invalids and Social Assistance (MOLISA), 'A new poverty line for Vietnam', Mimeo

Nguyen, P. and Baulch, B., forthcoming, 'A review of ethnic minority programmes and policies in Vietnam', *IDS Working Paper* in preparation

Oaxaca, R., 1973, ' Male-female wage differentials in urban labour market', I*nternational Economic Review*, Vol. 9(3), pp. 693-709

Pincus, J. and Sender, J,, 2006, 'Quantifying poverty in Viet Nam: Who Counts?' Mimeo, United Nations Development Program, Hanoi

Poverty Task Force, 2002, *Promoting Ethnic Minority Development*, Hanoi: United Nations Development Program

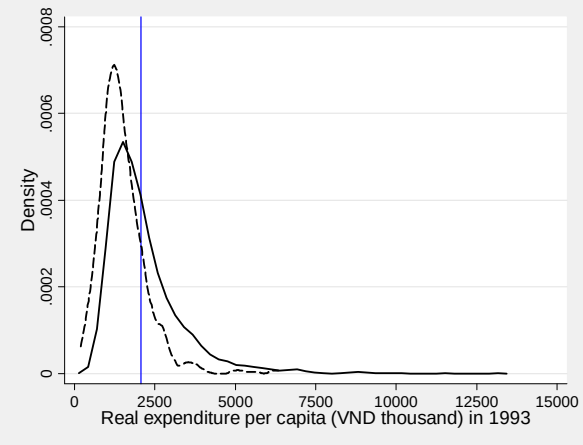
Psacharopoulos, G. and Patrinos, H., 1994, Indigenous People and Poverty in Latin America: An Empirical Analysis, The World Bank Regional & Sectoral Studies, Washington DC.

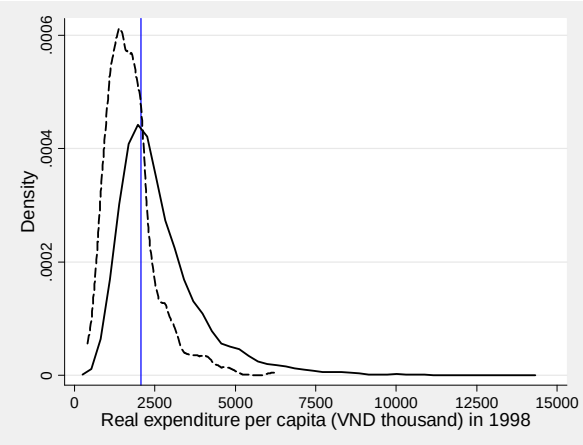
Van de Walle, D. and Gunewardena, D., 2001, 'Sources of ethnic inequality in Viet Nam', *Journal of Development Economics*, 65: 177-207

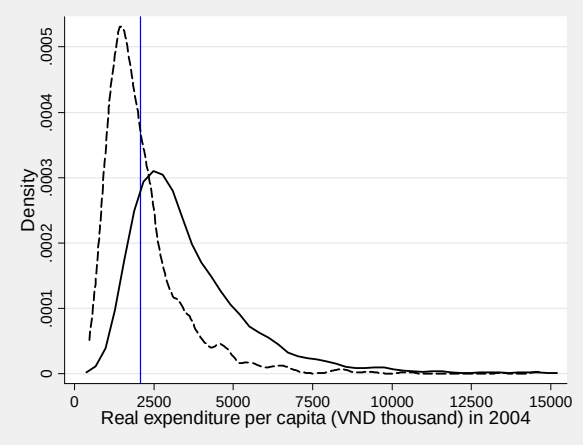
Vietnam Academy of Social Sciences, 2006, *Vietnam Poverty Update Report: Poverty and Poverty Reduction in Vietnam*, 1993-2004, Hanoi: National Political Publishing House

Zanchi, L., 1998, 'Inter-industry wage differentials in dummy variable models', *Economics Letters*, Vol. 60, pp. 297-301.

Figure 1: Kernel Density Plots of the Majority/Minority Expenditure Per Capita, Rural samples, 1992-2004







Source: drawn from the VLSS 1992/93 and 1997/98, VHLSS 2004

Notes: Expenditures per capita are given in Jan 2004 prices; the solid line represents the kernel density of the per capita household expenditures for the Kinh and Chinese; the dash line represents that of the other ethnic minority groups.

Table 1: Weighted Mean Real Expenditures Per Capita of the Ethnic Groups 1993-2004, Raw Deviations

	1993	1998	2004
Kinh majority	0.0379***	0.0518***	0.0789***
	(0.004)	(0.006)	(0.003)
Chinese	0.3588**	0.0470	-0.0476
	(0.169)	(0.148)	(0.098)
Khmer and Cham	-0.2108***	-0.1577***	-0.0889*
	(0.051)	(0.059)	(0.053)
Tay, Thai, Muong, Nung	-0.2326***	-0.2711***	-0.4037***
	(0.021)	(0.039)	(0.022)
Other Northern Uplands minorities	-0.5052***	-0.6083***	-0.6153***
	(0.066)	(0.101)	(0.042)
Central Highlands minorities	-0.5927***	-0.7977***	-0.6615***
	(0.077)	(0.102)	(0.054)
All others	-0.6532***	-0.5288***	-0.5866***
	(0.11)	(0.098)	(0.145)
Overall variability	0.1381	0.1571	0.1970

Sources: Estimates based on use of the VLSS 1992/93, VLSS 1997/98, and VHLSS 2004.

Note: These deviations are calculated without controlling for any characteristics or dummies

Table 2: Inter-ethnic Household Welfare Differentials for Vietnam 1993-2004 using Mean Regressions Models without District Effects

	1993	1998	2004
Kinh majority	0.0221***	0.0407***	0.0544***
	(0.004)	(0.007)	(0.004)
Chinese	0.4455***	0.0654	-0.0042
	(0.143)	(0.106)	(0.070)
Khmer and Cham	-0.1727***	-0.0858	-0.0536
	(0.053)	(0.078)	(0.045)
Tay, Thai, Muong, Nung	-0.0992***	-0.1807***	-0.2505***
	(0.026)	(0.049)	(0.022)
Other Northern Uplands minorities	-0.0673	-0.0776	-0.2448***
	(0.058)	(0.053)	(0.039)
Central Highlands minorities	-0.4695***	-0.4101***	-0.3301***
	(0.074)	(0.090)	(0.053)
All others	-0.1453	-0.2937***	-0.3680***
	(0.097)	(0.096)	(0.135)
Overall variability	0.0737	0.1007	0.1179

Sources: Estimates based on use of the VLSS 1992/93, VLSS 1997/98, and VHLSS 2004.

Notes:

(a) The deviations are computed using expression [2] in the text; ***, **, and * denotes statistically significant at the 0.01, 0.05 and 0.1 levels respectively.

(b) The welfare measure used is the logarithm of real per capita household expenditure and full set of OLS regression estimates are reported in table A3 of the appendix.

(c) The standard errors are computed using Zanchi's (1998) formula and are reported in parentheses; the overall variability measure is calculated using expression (9) reported in Haisken-DeNew and Schmidt (1997).

(d) The explanatory variables include dummies for commune type but not district dummies (see table A1 of the appendix). For the VLSS 1992/93, no information on the geographical types of communes was reported so information from the 1993-1998 panel was used to identified commune types.

Table 3: Inter-ethnic Household Welfare Differentials for Vietnam 1993-2004 using Mean Regression Models with District Effects

	1993	1998	2004
Kinh majority	0.0237***	0.0459***	0.0322***
	(0.005)	(0.007)	(0.005)
Chinese	0.3212***	0.0802	-0.0501
	(0.118)	(0.064)	(0.069)
Khmer and Cham	-0.0879	-0.0729	-0.0313
	(0.068)	(0.051)	(0.041)
Tay, Thai, Muong, Nung	-0.0863***	-0.1165***	-0.136***
	(0.034)	(0.039)	(0.029)
Other Northern Uplands minorities	-0.0345	0.0065	-0.1464***
	(0.089)	(0.070)	(0.046)
Central Highlands minorities	-0.7648***	-0.6232***	-0.2593***
	(0.097)	(0.117)	(0.052)
All others	-0.1575	-0.6697***	0.0365
	(0.130)	(0.127)	(0.081)
Overall variability	0.0956	0.1414	0.0708

Sources: Estimates based on use of the VLSS 1992/93, VLSS 1997/98, and VHLSS 2004.

Notes: See notes (a) to (c) for Table 2

Table 4: Inter-ethnic Household Welfare Differentials for Vietnam using Quantile Regressions Models – 1993

	10th	25th	Median	75th	90th
Kinh majority	0.0331***	0.0322***	0.0209***	0.0289***	0.0144***
	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)
Chinese	0.3638**	0.3589*	0.4713***	0.6668***	0.4275***
	(0.165)	(0.211)	(0.172)	(0.195)	(0.138)
Khmer and Cham	-0.1009	-0.199***	-0.2298**	-0.1407**	-0.1942***
	(0.068)	(0.063)	(0.092)	(0.070)	(0.067)
Tay, Thai, Muong, Nung	-0.0511	-0.0643**	-0.0878***	-0.1455***	-0.144***
	(0.042)	(0.03)	(0.028)	(0.030)	(0.051)
Other Northern Uplands minorities	-0.3494***	-0.3073***	-0.2182**	-0.0071	0.0955
	(0.126)	(0.079)	(0.101)	(0.093)	(0.094)
Central Highlands minorities	-0.5264***	-0.6199***	-0.4259***	-0.4172***	-0.4144***
	(0.091)	(0.119)	(0.118)	(0.085)	(0.099)
All others	-0.7318***	-0.4993***	-0.1640*	-0.1302	-0.0727
	(0.187)	(0.178)	(0.101)	(0.101)	(0.142)
Overall variability	0.1188	0.1005	0.0761	0.0795	0.0609

Sources: Estimates based on use of the VLSS 1992/3.

Notes:

(a) The deviations are computed using expression [5] in the text.

(b) The sampling variances for the quantile regression model estimates are based on bootstrapping with 200 replications used.

(c) ***, **, and * denotes statistically significant at the 0.01, 0.05 and 0.1 levels respectively.

(d) The realization of the proportion of each ethnic group at the selected quantiles is obtained by bootstrapping the sample at relevant percentiles with 500 replications.

(e) The explanatory variables include dummies for commune type but omit district dummies (see table A1 of the appendix).

(f) The full set of quantile regression estimates is available from the authors on request.

Table 5: Inter-ethnic Household Welfare Differentials for Vietnam using Quantile Regressions Models – 1998

	10th	25th	Median	75th	90th
Kinh majority	0.032***	0.0386***	0.0313***	0.0364***	0.0469***
	(0.004)	(0.004)	(0.003)	(0.005)	(0.007)
Chinese	-0.0409	0.1740	0.1313	0.1174	0.6412***
	(0.172)	(0.132)	(0.137)	(0.272)	(0.226)
Khmer and Cham	-0.113*	-0.1004*	-0.0969*	-0.065	-0.0951*
	(0.067)	(0.057)	(0.054)	(0.061)	(0.051)
Tay, Thai, Muong, Nung	-0.1416***	-0.1513***	-0.1897***	-0.1843***	-0.2519***
	(0.038)	(0.026)	(0.034)	(0.029)	(0.04)
Other Northern Uplands minorities	-0.0346	-0.0514	-0.0560	-0.1287***	-0.2525***
	(0.048)	(0.059)	(0.046)	(0.047)	(0.087)
Central Highlands minorities	-0.5142***	-0.5286***	-0.4922***	-0.4646***	-0.3714***
	(0.043)	(0.045)	(0.052)	(0.075)	(0.082)
All others	-0.2165***	-0.267***	-0.2653***	-0.3057***	-0.2143
	(0.101)	(0.083)	(0.060)	(0.091)	(0.162)
Overall variability	0.0989	0.1165	0.0977	0.0981	0.1170

Sources: Estimates based on use of the VLSS 1997/8.

Notes: see notes to table 4.

Table 6: Inter-ethnic Household Welfare Differentials for Vietnam using Quantile Regressions Models– 2004

	10th	25th	Median	75th	90th
Kinh majority	0.0683***	0.046***	0.0516***	0.0513***	0.0519***
	(0.005)	(0.004)	(0.004)	(0.004)	(0.006)
Chinese	0.0062	0.0630	-0.0233	-0.0251	-0.1186
	(0.107)	(0.120)	(0.076)	(0.097)	(0.122)
Khmer and Cham	-0.076	-0.0317	-0.0428	-0.0772*	-0.2038***
	(0.079)	(0.054)	(0.050)	(0.045)	(0.065)
Tay, Thai, Muong, Nung	-0.2047***	-0.2384***	-0.2325***	-0.2359***	-0.2314***
	(0.025)	(0.022)	(0.021)	(0.027)	(0.036)
Other Northern Uplands minorities	-0.3276***	-0.3472***	-0.3501***	-0.2467***	-0.2829***
	(0.047)	(0.043)	(0.031)	(0.049)	(0.055)
Central Highlands minorities	-0.3872***	-0.3725***	-0.3483***	-0.3257***	-0.3019***
	(0.051)	(0.048)	(0.048)	(0.039)	(0.070)
All others	-0.5721***	-0.3380	-0.2077	-0.3424***	-0.3956**
	(0.170)	(0.218)	(0.177)	(0.062)	(0.174)
Overall variability	0.1398	0.1136	0.1198	0.1134	0.1101

Sources: Estimates based on use of the VHLSS 2004.

Notes: see notes to Table 4.

Table 7: Inter-ethnic Household Welfare Differentials for Vietnam 1993-2004 using 90th – 10th Inter-Quantile Regressions Models

	1993	1998	2004
Kinh majority	0.0033	-0.0035	0.0041*
	(0.003)	(0.003)	(0.003)
Chinese	0.0856	0.6636***	-0.1044
	(0.254)	(0.279)	(0.143)
Khmer and Cham	-0.0714	-0.0006	-0.1073
	(0.083)	(0.091)	(0.095)
Tay, Thai, Muong, Nung	-0.071	-0.1288**	-0.0062
	(0.081)	(0.063)	(0.043)
Other Northern Uplands minorities	0.4668***	-0.2363**	0.0651
	(0.144)	(0.101)	(0.079)
Central Highlands minorities	0.134	0.1243	0.1058
	(0.139)	(0.098)	(0.08)
All others	0.6811***	-0.0163	0.1969
	(0.223)	(0.184)	(0.229)
Overall variability	0.0028	0.0023	0.0053

Notes:

(a) The deviations are computed using a variant of expression [5] in the text where the coefficients are the 90th – 10th inter-quantile coefficients and the weights used are based on ethnic proportions at the median of the conditional distribution (see (d) below).

(b) The sampling variances for the inter-quantile regression model estimates are based on bootstrapping with 200 replications used.

(c) ***, **, and * denotes statistically significant at the 0.01, 0.05 and 0.1 levels respectively.

(d) The realization of the proportion of each ethnic group at the selected quantiles is obtained by bootstrapping the sample at relevant percentiles with 500 replications. The regression models include controls for the geographical type of commune instead of district effects (see table A1 of the appendix).

(e) The full set of inter-quantile regression estimates is available from the authors on request.

Appendix

Table A1: Summary Statistics

Variable	1993	1998	2004
Per capita household expenditure (VND1000; Jan 04 prices)	2,080	2,514	3,400
Per capita household expenditure (natural logarithm)	7.5079	7.7836	8.0124
Kinh majority	0.8369	0.8157	0.8379
Chinese	0.0055	0.0048	0.0041
Khmer and Cham	0.0226	0.0246	0.0145
Tay, Thai, Muong, Nung	0.0810	0.0931	0.0871
Other Northern Uplands minorities	0.0207	0.0263	0.0237
Central Highlands minorities	0.0170	0.0275	0.0289
All other ethnic groups	0.0164	0.0081	0.0039
Household size	5.9219	5.5615	5.0553
Proportion of children aged 6 years and less	0.1796	0.1263	0.0916
Proportion of children aged from 7 to 16 years	0.2616	0.2761	0.2434
Proportion of male adults	0.2599	0.2811	0.3228
Proportion of female adults	0.2988	0.3165	0.3421
Household type 1: head or head and spouse	0.0192	0.0266	0.0327
Household type 2: parents and one child	0.0659	0.0498	0.0822
Household type 3: parents and two children	0.1314	0.1584	0.2405
Household type 4: parents with > three children	0.4551	0.4289	0.3277
Household type 5: three-generation household	0.0626	0.0718	0.1630
Household type 6: other household structures	0.2658	0.2644	0.1539
Age of household head	44.803	46.243	48.344
Age of head squared (divided by 100)	21.968	23.027	25.198
Household head is female	0.1806	0.1685	0.1662
Most educated member: no schooling	0.0248	0.0153	0.0685
Most educated member: primary education	0.1920	0.1582	0.2430
Most educated member: lower secondary	0.4728	0.4642	0.3415
Most educated member: upper secondary	0.1861	0.2457	0.1789
Most educated member: vocational/technical	0.1017	0.0753	0.1233
Most educated member: college/university	0.0226	0.0414	0.0448
Irrigated annual crop land (1000 m^2)	2.1127	3.1903	3.0288
Non-irrigated annual crop land (1000 m^2)	2.7121	1.3494	1.2711
Perennial land (1000 m^2)	0.8591	1.4173	1.2331
Forest plot (1000 m^2)	0.3555	1.2335	1.2946
Water surface (1000 m^2)	0.1375	1.2558	0.3727
Other cultivated lands (1000 m^2)	0.2864	0.7308	0.5956
Commune having access to road that car can travel	0.8418	0.8417	0.9630
Commune having access to public transport	0.5222	0.5623	0.4920
Commune having access to post office	0.3435	0.2335	0.3070
Commune having access to daily market	0.5362	0.4936	0.3053
Commune having access to electricity	0.4305	0.8951	0.9671
Commune having factories located within 10km	0.4472	0.5594	0.6703
Geographical types of commune: rural coastal	0.0845	0.0691	0.0704
Geographical types of commune: rural inland delta	0.5541	0.5145	0.5610
Geographical types of commune: rural midlands	0.0493	0.0620	0.0725
Geographical types of commune: rural low mountain	0.1738	0.2120	0.1497
Geographical types of commune: rural high mountain	0.1382	0.1423	0.1465
Number of observations	3,839	4,270	6,937

Sources: Estimates based on use of the VLSS 1992/93, VLSS 1997/98, and VHLSS 2004.

Table A2: Allocation of the 54 Ethnic Groups into Broad Categories

Ethnic Groups	Main Groups and Percent of Population	Population	Main locations
Kinh	Kinh (86.7%)	66,188.844	All over the country
Chinese	Chinese (1.1%)	861,977	Northern Uplands, Southeast, Mekong River Delta
Khmer and Cham	Khmer (1.2%), Chăm (0.1%)	968,566	Mekong River Delta
Tay, Thai, Muong, Nung	Tày (2.1%), Thái (1.7%), Mường (1.6%), Nùng (1.1%)	4,924,100	Northern Uplands, North Central Coast, Central Highlands
Other Northern Uplands	Hmông (0.9), Dao (0.8%), Sán Chay (0.2%), Sán Dìu (0.2%), Giáy (+), Khơ Mú (0.2%), Khang(+), Hà Nhì (+), Lào (+), La Chí (+), Xinh Mun (+), Bo Y(*), Co Lao (*), Cống (*) La Ha (*), La Hu (*), Lao(*) Lo Lo (*), Lự (*), Mang (*), Ngai (*), Pathen (*), Pu Peo (*), Si La(*)	1,847,331	Northern Uplands
Central Highlands minorities	Jarai (0.4%), Ê-Đê (0.2%), Ba-Na (0.3%), Xơ Đăng (0.2%), Cờ Ho (0.1%), Mnông (+), Gie Tiêng (0.1%), Mạ (+), Chu Ru (+), Hre (0. 2%), Ra-Glai (0.1%), Xtieng (0.1%), Cờ Tu (+), Co (+), Cho Ro (+), Brau (*), Ro-Mam (*)	1,384,819	Central Highlands, South Central Coast, Southeast
All others	Bru-Vân Kiều (0.1%), Thổ (0.1%), Tà Ôi (0.1%), Ngái (+), Lô lô (+), Chứt (*), O-Du (*) unspecified (+)	152,285	North Central Coast

Source: Population and main location based on the 3% sample of the Population and Housing Census 1999, using the classification of this paper.

Notes:

a. (*) ethnic groups with populations of less then 10,000 persons;
b. (+) ethnic groups with the population of more than 10,000 persons but accounts for less than 0.1% of the total population.

Table A3: Poverty Headcount Rates by Ethnic Categories

	Household expenditure pc. (VND 000s)	**Headcount** (%)	**Poverty gap** (%)	**Sample size**
1992/93				
Kinh	2,174	59.96	8.73	3,279
Chinese	2,703	46.67	3.01	15
Khmer and Cham	1,729	71.43	23.85	77
Tay, Thai, Muong, Nung	1,583	83.50	24.92	297
Other Northern Uplands minorities	1,248	85.51	41.05	69
Central Highlands minorities	1,176	96.15	43.37	52
All other	1,298	86.00	42.06	50
Average (Total rural)	2,080	63.25	11.75	3,839
1997/98				
Kinh	2,682	38.81	7.86	3,571
Chinese	2,709	37.45	8.02	18
Cham and Khmer	2,135	60.65	14.73	86
Tay, Thai, Muong, Nung	1,851	72.91	18.38	326
Other Northern Uplands minorities	1,665	82.49	25.09	76
Central Highlands minorities	1,194	92.00	46.84	153
All other	1,467	88.04	32.37	39
Average (Total rural)	2,514	45.52	10.74	4,269
2004				
Kinh	3,637	18.05	3.56	5,665
Chinese	3,199	25.15	5.14	30
Cham and Khmer	3,036	30.57	6.87	98
Tay, Thai, Muong, Nung	2,204	57.94	16.53	706
Other Northern Uplands minorities	1,786	73.44	25.71	216
Central Highlands minorities	1,911	73.34	28.44	201
All other	1,838	70.32	24.17	22
Average (Total rural)	3,400	24.89	6.09	6,938

Source: Calculations from the VLSS 1992/93 and 1997/98, and 2004

Notes: Expenditures per capita are given in January 2004 prices; Poverty headcounts are computed using the poverty lines set by World Bank and GSO (often referred as GSO poverty lines)

Table A4. OLS Estimates for Log per Capita Household Expenditure Regression Models

	1993	1998	2004
Chinese	0.4234***	0.0247	-0.0586
	(0.144)	(0.107)	(0.071)
Khmer and Cham	-0.1948***	-0.1264*	-0.108**
	(0.054)	(0.079)	(0.046)
Tay, Thai, Muong, Nung	-0.1213***	-0.2214***	-0.3049***
	(0.028)	(0.055)	(0.026)
Other Northern Uplands minorities	-0.0893	-0.1183**	-0.2991***
	(0.06)	(0.056)	(0.041)
Central Highlands minorities	-0.4916***	-0.4508***	-0.3844***
	(0.075)	(0.096)	(0.056)
All other ethnic groups	-0.1674*	-0.3343***	-0.4224***
	(0.099)	(0.101)	(0.136)
Household size	-0.033***	-0.0579***	-0.0491***
	(0.008)	(0.008)	(0.006)
Proportion of children aged from 7 to 16 years	0.4043***	0.4081***	0.3356***
	(0.054)	(0.06)	(0.049)
Proportion of male adults	0.5762***	0.5684***	0.813***
	(0.069)	(0.075)	(0.06)
Proportion of female adults	0.576***	0.4662***	0.7069***
	(0.073)	(0.072)	(0.064)
Household type 2: parents and one child	-0.0525	-0.0451	-0.039
	(0.04)	(0.037)	(0.03)
Household type 3: parents and two children	-0.0823**	-0.1044***	-0.0383
	(0.042)	(0.038)	(0.031)
Household type 4: parents + > three children	-0.1581***	-0.1624***	-0.1082***
	(0.046)	(0.043)	(0.035)
Household type 5: three-generation household	-0.1978***	-0.1379***	-0.0811**
	(0.056)	(0.049)	(0.035)
Household type 6: other household structures	-0.1452***	-0.1441***	-0.0688*
	(0.05)	(0.045)	(0.037)
Age of household head	0.0034	0.0066*	-0.0008
	(0.004)	(0.004)	(0.003)
Age of head squared (divided by 100)	-0.0041	-0.0062*	-0.0009
	(0.004)	(0.004)	(0.003)
Household head is female	0.0084	-0.0158	0.026
	(0.022)	(0.021)	(0.017)
Most educated member: primary education	-0.3833***	-0.1528***	-0.1888***
	(0.061)	(0.035)	(0.023)
Most educated member: lower secondary	0.1275***	0.1164***	0.0847***
	(0.02)	(0.019)	(0.014)
Most educated member: upper secondary	0.2398***	0.2621***	0.2336***
	(0.025)	(0.022)	(0.018)
Most educated member: vocational/technical	0.2805***	0.2972***	0.3515***
	(0.028)	(0.028)	(0.02)
Most educated member: college/university	0.4779***	0.5582***	0.6094***
	(0.06)	(0.035)	(0.03)

Irrigated annual crop land (1000 m^2)	0.0165***	0.0076***	0.0086***
	(0.002)	(0.002)	(0.001)
Non-irrigated annual crop land (1000 m^2)	0.0116***	0.0026	0.0051***
	(0.002)	(0.002)	(0.001)
Perennial land (1000 m^2)	0.027***	0.0144***	0.0062**
	(0.005)	(0.002)	(0.003)
Forest plot (1000 m^2)	0.0037	0.0057***	0.0004
	(0.003)	(0.001)	(0)
Water surface (1000 m^2)	0.0165*	0	0.0111***
	(0.009)	(0)	(0.002)
Other cultivated lands (1000 m^2)	0.0045*	0.0068***	0.0174***
	(0.002)	(0.002)	(0.004)
Commune having access to road that car can travel	-0.2419***	0.0295	0.0235
	(0.026)	(0.047)	(0.039)
Commune having access to public transport	0.1931***	0.0401	0.057***
	(0.018)	(0.04)	(0.014)
Commune having access to post office	-0.1312***	0.0495	0.0287*
	(0.019)	(0.042)	(0.015)
Commune having access to daily market	0.1325***	0.0735**	0.0986***
	(0.017)	(0.036)	(0.016)
Commune having access to electricity	-0.0177	0.0843	0.0791**
	(0.019)	(0.062)	(0.033)
Commune having factories located within 10km	0.1076***	0.0524	0.0648***
	(0.016)	(0.037)	(0.014)
Geographical types: rural coastal	0.0059	-0.0273	-0.0064
	(0.032)	(0.06)	(0.03)
Geographical types: rural midlands	-0.1133***	-0.0451	0.0266
	(0.032)	(0.094)	(0.026)
Geographical types: rural low mountain	-0.0802***	-0.1133**	-0.0386**
	(0.022)	(0.048)	(0.019)
Geographical types: rural high mountain	-0.0477*	-0.0409	-0.0416
	(0.029)	(0.063)	(0.03)
Constant	6.5891***	7.0949***	7.4374***
	(0.088)	(0.133)	(0.09)
R^2	0.40297	0.4053	0.4158
Number of observations	3,839	4,270	6,712

Notes: ***, **, and * denotes statistically significant at 0.01, 0.05 and 0.1 levels respectively;