Ethnic Wage Inequality in Vietnam: Empirical Evidence from 2002

T. Hung Pham and Barry Reilly

Department of Economics, University of Sussex

November 2007

Ethnic Wage Inequality in Vietnam:

Empirical Evidence from 2002

HUNG T. PHAM and BARRY REILLY

Department of Economics

University of Sussex, Brighton, United Kingdom

Abstract

Purpose – This paper complements earlier studies on ethnic minority underdevelopment in Vietnam by empirically examining the ethnic wage gap in the Vietnamese labour market, using data from a large-scale household survey conducted in 2002.

Methodology – The paper uses the 'index number' decomposition method suggested by Oaxaca (1973) to decompose the ethnic wage gap into treatment and endowment effects at both the mean and at selected quantiles of the conditional wage distribution.

Findings – The results confirm the existence of an ethnic wage gap in the labour market, through this gap is found to be substantially narrower than the ethnic gap observed using household living standard measures for Vietnam. Decomposition results reveal that the ethnic wage gap is largely attributable to differentials in the returns to endowments, a finding invariant to whether the mean or selected quantiles of the conditional wage distribution is examined.

Limitations – In the absence of superior alternatives, the paper uses an *ad hoc* procedure to correct for selectivity into wage employment for the quantile regression models. In addition, due to data constraints in regard to earnings, the paper does not examine the ethnic wage gap for the self-employed.

Originality – This paper is the first to analyse the ethnic wage gap in the Vietnam labour market and one of the few to examine ethnic pay differentials at selected points of the conditional wage distribution using quantile regression analysis.

Keywords Wage inequality, ethnic minority, quantile regression, Vietnam

Paper type Research paper

November, 2007

Address for correspondence: Department of Economics, University of Sussex, Brighton, BN1 9SN, UK. Tel: +44 (0)1273 606755 (Ext 2209); Fax: +44 (0)1273 673563. Email: t.pham@sussex.ac.uk or hungpt@neu.edu.vn

1. Introduction

Vietnam is a multi-ethnic country comprising about 54 ethnic groups. The largest group is the Kinh, which accounts for nearly 84 percent of the total population. The other 53 ethnic minorities range from some populous groups of around one million (such as the Tay, Thai, Muong, Khmer) to some smaller groups with populations measured in their hundreds (see Dang *et al.* 2000). The country as a whole has registered impressive progress in terms of overall poverty reduction over the past two decades. However, poverty is still widespread within the ethnic minority groups. Between 1993 and 2004, the national poverty headcount fell by nearly two-thirds (from 58 percent to 19.5 percent), while ethnic minority poverty decreased by less than one-third (from 86.4 percent to 60.7 percent). Furthermore, the gap in living standards between, on the one hand, the Kinh majority and Hoa (the Chinese), and on the other, the remaining ethnic minorities has grown over time (Baulch *et al.*, 2004). More worryingly, recent evidence demonstrates that hunger among ethnic minorities remains widespread, even when located in those parts of the country enjoying rapid economic growth (Swinkels and Turk, 2006).

Previous studies on ethnic under-development in Vietnam have attempted to investigate this gap through examining differences in endowment (i.e., characteristics) and treatment (i.e., returns to characteristics) effects between the majority (Kinh and Chinese) and the other ethnic minority groups using household welfare measures. The differences in both components are found to favour the Kinh and Chinese (Van de Walle and Gunewardana, 2001; Baulch *et al.*, 2004). Though the poor endowment of the ethnic minority groups can be linked to the fact that most ethnic minorities reside in remote and mountainous areas, this explains only part of the gap. Baulch *et al.* (2004) report that "[…] even if ethnic minority households had the same endowments as the Kinh and Hoa (Kinh majority and the Chinese), this would close no more than a third of the gap in their living standards" (p. 274). This suggests that the ethnic minorities secure considerably lower returns to their characteristics than the majority.

Instead of investigating the gap in living standards between the majority and minority at the household-level, the current paper adds to a modest literature on ethnic disadvantage in Vietnam by empirically examining the nature and extent of ethnic labour market wage differentials at the level of the individual. We believe this emphasis is apposite given that those in wage employment in the Vietnamese labour market accounted for about 28 percent of total employment in 2002. In addition, the waged labour market is likely to become a more important institution as Vietnamese economic development proceeds.

Our primary research theme will be investigated using both mean and quantile regression analysis. In contrast to the mean regression approach, which sheds important light on the magnitude of the ethnic wage gap at the average, the use of quantile regression techniques provides insights on how the ethnic wage gap varies across the conditional wage distribution.

The structure of this paper is now outlined. The next section provides some background to the ethnic wage gap and describes the data used in this study. It is followed by a description of the empirical methodology employed to decompose the ethnic wage gap at the mean and selected quantiles of the conditional wage distribution. The empirical results from this decomposition are discussed in section four. Conclusions and some policy implications are provided in the final section of the paper.

2. Background and Data

2.1 Data

This paper draws on data from the Vietnam Household Living Standards Survey 2002 (hereafter VHLSS 2002), the third in the series of living standard measurement surveys undertaken for Vietnam to date. This survey series is conducted by the General Statistic Office (GSO) of Vietnam, under the technical guidance of the World Bank, with funding from the United Nations Development Program (UNDP) and the Swedish International Development Agency (Sida). The overall approach adopted in these surveys is compliant with the framework

used in the World Bank's Living Standard Measurement Surveys. These surveys are widely recognized as of high quality and nationally representative.

The first two surveys, which are commonly referred to as the VLSS 1992/93 and VLSS 1997/98, were undertaken using samples of 4,800 and 6,000 households respectively (see World Bank, 2000 and 2001 for more details). Although these surveys have been widely used to explore the impact of the Vietnamese reform process of *Doi moi*, they possess the drawback of having relatively small sample sizes. This was a primary motive for the introduction of the second phase of household surveys in 2002 designed to cover the period 2002 – 2010. The first survey of this second phase, the VHLSS 2002, collected information from a sample of 30,000 households. However, this relatively large sample size combined with less technical support from donors than was the case for the two earlier surveys represented a technical challenge for the GSO resulting in problems with the administration and commune coverage of the survey [1]. Given the large administrative costs incurred and the potential for non-sampling errors in the implementation of VHLSS 2002, the sample size used in the subsequent survey in 2004 was restricted to about 9,000 households (see Phung and Nguyen, 2006 for more details).

In spite of some potential limitations in terms of its coverage, the VHLSS 2002 is selected for this study primarily because it contains a relatively large sample size. In the other surveys conducted, the ethnic minority wage workers comprise relatively small sub-samples (i.e., 239 in the VLSS 1992/93, 289 in the VLSS 1997/98, and 349 in the VHLSS 2004). However, a sub-sample of 968 ethnic minority wage earners out of a total sample of 16,170 is available for the VHLSS 2002. This provides a feasible sub-sample of wage employees for a meaningful analysis of the ethnic wage gap [2]. The sample size is also important given our focus on investigating, *inter alia*, wage gaps at selected quantiles of the conditional wage distribution. Aside from information on various aspects of household well-being and a detailed array of household-level characteristics, the survey also provides detailed individual-level labour market information including data on earnings, hours worked, educational attainment, age, employment sector, and location.

2.2 Overview of the Ethnic Wage Gap in Vietnam

The nature and scale of ethnic minority wage employment in Vietnam remains partly linked to the government's erstwhile manpower allocation policy, which was a characteristic feature of the central planning regime. Under this system jobs were allocated by the government and wages were fixed by the authorities. Graduates with secondary and tertiary education were assigned jobs in the public sector (Moock *et al.*, 1998). This policy was widely used as part of a concerted effort to promote the socio-economic development of ethnic minority groups. As a result, graduates from the majority ethnic group were allocated to the remote and mountainous areas, while ethnic minority graduates were placed in different organizations in urban areas, where the majority predominantly resides. In addition, the government also formulated a separate policy for educating ethnic minority graduates at public colleges and universities. Through the tertiary education system, which is mainly dependent on the State budget, the Ministry of Education and Training (MoET) allocates a target number for ethnic minority students. The ethnic minority students under this arrangement are fully funded and subject to a separate university entrance examination, which is regarded as easier than the national university entrance examination (MoET, 2006). Though the number of these graduates is relatively small, comprising about 8,000 ethnic minority students from a total of more than 1.3 million Vietnamese university students in the 2005/6 academic year, they were obligated to return to their home provinces after graduation to take jobs in wage employment in certain remote and mountainous regions.

The government, in recognition of the importance of improving the living standards of the ethnic minorities and reducing the household income gap between different ethnic groups, also introduced a number of supportive policies and programs. The Comprehensive Poverty Reduction and Growth Strategy (CPRGS), which was approved by the Prime Minister of Vietnam in 2002, outlined a number of objectives designed to promote the provision of basic infrastructure and social services for ethnic minorities (e.g., in education and healthcare). In conjunction with this general strategy, there have also been national programs that specifically

target ethnic minorities and the poorest households, which are overwhelmingly rural based and focused on ethnic minority groups. Most notably, Program 135 was launched in 1998 (see MOLISA and UNDP, 2004) and facilitated the development of infrastructure in villages and communes (e.g., in terms of the construction and maintenance of roads, small irrigation dams, clean water systems, schools, health centers and other infrastructure projects) in the mountainous and remote areas of Vietnam. The Program covered 1715 communes with a population of around 1.1 million households comprising more than six million people.

The National Program on Hunger Eradication and Poverty Reduction (HEPR) also provided another tool for reducing the gap between the disadvantaged and more advantaged regions (SRV, 2001). It has recently provided credits, support services for poor households, and capacity building for local authorities not covered under Program 135. Although the effectiveness of these programs in achieving their targets is subject to some criticism (Baulch *et al.* 2004), the recent evaluation by MOLISA and UNDP (2004) suggests success in terms of improving the healthcare and educational services for the poor in general and the ethnic minorities in particular.

Despite the above array of government policies to assist ethnic minorities, the proportion of ethnic groups represented in total wage employment remains modest in Vietnam. Though accounting for nearly 14 percent of the total population in 2002, the ethnic minority groups comprised only six percent of total wage employment and 12 percent of the total labour force (authors' calculations based on the VHLSS 2002). The incidence of wage employment within the ethnic minorities is substantially lower than the Kinh and Chinese, henceforth known as the majority [3]. On average, only 15 percent of the ethnic minority labour force is wage-employed while the corresponding figure for the majority is nearly twice that rate. At the household level, 14 percent of ethnic minority households have wage earner(s) among their household members, while 41 percent of majority households receive wage incomes from the employment of their household members.

Table 1, using data from the VHLSS 2002, reveals that the household welfare gap between the majority and the minority is much smaller if ethnic minority household members are employed in the wage employment sector. More interestingly, the educational levels of the ethnic minority wage workers are considerably closer to their majority counterparts (except in the proportion of workers who have graduated from universities/colleges). Although the overall majority/minority gap in living standards remains high as suggested in earlier studies (see above), this gap narrows substantially for those who are in the wage employment sector. This can be taken to suggest tentatively that the ethnic minority wage earners have perhaps become more assimilated (or '*Vietnamized*' in the language of some ethnologists (see Dang *et al.*, 2000; Baulch *et al.*, 2004)) within the majority group compared to non-wage earners from the same ethnic groups.

[Take in Table 1]

Despite apparent evidence of similarities between the Kinh/Chinese and the ethnic minority wage earners, table 2 reveals an ethnic wage gap across a number of different dimensions. The raw data on the hourly wage rate reveals a statistically significant minority/majority wage gap of about 11 percent. Though the raw earnings differentials vary across sectors of employment, education groups, and location (i.e., urban or rural settlement areas), the data consistently reveal an earnings gap between the majority and minority groups. This is re-affirmed by the kernel density plots of the wages depicted in Figure 1.

[Take in Table 2 and Figure 1]

3. Econometric Methodology

3.1 Decomposition Methodologies

Following the seminal work of Mincer (1974), it is conventional to specify log wages as a function of a set of wage determining characteristics with the most austere form including only controls for human capital. The specification is then augmented to capture other variables

7

viewed as important in the wage determination process. In the context of the ethnic wage gap, the ethnic-specific labour market earnings equations for an ith individual are specified as follows:

$$w_m = x_m'\beta_m + \mu_m \tag{1}$$

$$w_e = x_e'\beta_e + \mu_e \tag{2}$$

where x_j is a ($k \times n$) matrix of human capital and other characteristics (e.g., education, labour force experience, employment sector *etc.*) and j is the ethnic group subscript; b is a ($k \times 1$) vector of unknown parameters capturing the effect of various covariates on the natural log wage (w); m is a ($n \times 1$) vector of random error terms; and m and e denote the majority and minority groups respectively.

Applying the conventional Oaxaca (1973) decomposition, the estimated mean ethnic difference in log wages is generally expressed as:

$$\overline{w}_m - \overline{w}_e = (\overline{x}_m - \overline{x}_e)'\hat{\beta}_m + \overline{x}_e'(\hat{\beta}_m - \hat{\beta}_e) \tag{3}$$

where the 'bar' denotes mean values and the 'hat' denotes coefficient estimates. This allows the overall average differential in wages between the two ethnic groups to be decomposed into a part attributable to differences in characteristics (known as the 'explained' or 'endowment' effect) and a part attributable to differences in the estimated returns to characteristics between majority and minority workers (known as the 'unexplained', 'treatment' or 'residual' effect). The final part of expression (3) is sometimes taken to capture the effect of unequal treatment within the labour market.

The use of an 'index number' approach is subject to the conventional 'index number' problem. Expression (3) could therefore be re-expressed in terms of average majority characteristics and yield numerically different values for the component parts compared to (3) above [4]. The re-formulation yields:

$$\overline{w}_m - \overline{w}_e = (\overline{x}_m - \overline{x}_e)'\hat{\beta}_e + \overline{x}_m'(\hat{\beta}_m - \hat{\beta}_e) \tag{4}$$

The foregoing decompositions associated with (3) and (4) are traditionally cast within the mean regression framework. An exclusive focus on the mean, however, may provide an incomplete account of the ethnic pay gap. The estimation of a set of conditional quantile functions allows for a more detailed portrait of the relationship between the conditional wage distribution and selected covariates than that provided by mean regression analysis. In contrast to the OLS approach, the quantile regression 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. Using this methodology, the log wage equation may be estimated conditional on a given specification for various percentiles of the residuals (e.g., 10th, 25th, 50th 75th or 90th (see Chamberlain, 1994)).

In the current case, the quantile regression for the majority and minority sub-samples can be defined as:

$$w_m = x_m ' \beta_{qm} + \mu_{qm} \tag{5}$$

$$w_e = x_e ' \beta_{qe} + \mu_{qe} \tag{6}$$

If $Q_q(\cdot)$ is taken to denote the conditional q^{th} quantile operator, then $Q_q(w_j | x_j) = x_j ' \beta_{qj}$, where β_{qj} is the unknown parameter vector for the qth quantile with q representing the selected quantile of interest (i.e., 0.1, 0.25, 0.5, 0.75 and 0.9 in the current application) ; m_{qj} denotes the error term, the distribution of which is left unspecified but for which $Q_q(m_{qj} | x_j) = 0$ is assumed; and j is the subscript for the ethnic group.

From equations (5) and (6) the conditional q^{th} quantile of the distribution of wages for the two groups are then expressed as:

$$Q_q(w_m) = E(x_m | w_m = Q_q(w_m))' \hat{\beta}_{qm} + E(m_{qm} | w_m = Q_q(w_m)) \tag{7}$$

$$Q_q(w_e) = E(x_e|w_e = Q_q(w_e))'\hat{\boldsymbol{\beta}}_{qe} + E(m_{qe}|w_e = Q_q(w_e)) \tag{8}$$

where the 'hats' now denote quantile regression estimates and $E(\cdot)$ denotes the expectations operator. In the expressions (7) and (8), the characteristics are evaluated conditionally at the unconditional quantile log wage value and not unconditionally as in the case of the mean regression. The terms $E(m_{qj}|w_j = Q_q(w_j))$ are thus non-zero. From (7) and (8), the ethnic pay gap at the qth quantile is defined as D_q and this can be decomposed into three parts:

$$\Delta_q = \Delta\boldsymbol{\Omega}_q'\hat{\boldsymbol{\beta}}_{qm} + \boldsymbol{\Omega}_{qe}'\Delta\hat{\boldsymbol{\beta}}_q + \Delta\boldsymbol{R}_q \tag{9}$$

where $\Delta\hat{\boldsymbol{\beta}}_q = (\hat{\boldsymbol{\beta}}_{qm} - \hat{\boldsymbol{\beta}}_{qe})$ and $\Delta\boldsymbol{\Omega}_q = \boldsymbol{\Omega}_{qm} - \boldsymbol{\Omega}_{qe}$

with $\boldsymbol{\Omega}_{qm} = E(x_m|w_m = Q_q(w_m))$ and $\boldsymbol{\Omega}_{qe} = E(x_e|w_e = Q_q(w_e))$

and $\Delta\boldsymbol{R}_q = [E(\boldsymbol{\mu}_{qm}|w_m = Q_q(w_m)) - E(\boldsymbol{\mu}_{qe}|w_e = Q_q(w_e))]$

The term $\Delta\boldsymbol{R}_q$ is best interpreted as a difference in unobservables between the two ethnic groups. Given the 'index number' approach noted earlier, the ethnic pay gap can also be decomposed as:

$$\Delta_q = \Delta\boldsymbol{\Omega}_q'\hat{\boldsymbol{\beta}}_{qe} + \boldsymbol{\Omega}_{qm}'\Delta\hat{\boldsymbol{\beta}}_q + \Delta\boldsymbol{R}_q \tag{10}$$

Using mean characteristics in the computation of (9) and (10) may provide unrepresentative realizations for the characteristics at points other than the conditional mean wage to which they relate. Therefore, it is necessary to use realizations of the characteristics that more accurately reflect the relevant points on the conditional wage distribution. In order to address this issue, we use an approach originally suggested by Machado and Mata (2005) to derive the realizations for the relevant characteristics at different quantiles of the conditional wage distribution. The procedure involves drawing 100 observations at random and with replacement from each of the majority and minority sub-samples. Each observation once ranked comprises a percentile point on the wage distribution. The full set of characteristics for the observation at the qth wage quantile is then retrieved. This process is then replicated 200

times to obtain 200 observations at the selected q^{th} quantile. The mean characteristics of these observations at each quantile are then used to construct the realizations for $\boldsymbol{\Omega}_{qm}$ and $\boldsymbol{\Omega}_{qe}$ used in equations (9) and (10) [5].

3.2 Selection Bias Issues

There is a potential selection issue governing the analysis of the ethnic pay gap in this study. Either through a process of self-selection by individuals or sample selection by employers, the engagement of individuals in wage employment activity may not be interpretable as the outcome of a random process. Heckman (1979) and Lee (1983) provide well-known solutions to this problem.

The mean wage decomposition in the presence of selectivity correction favoured by many authors is expressed as [6]:

$$\overline{w}_m - \overline{w}_e = (\overline{x}_m - \overline{x}_e)'\hat{\beta}_m + \overline{x}_e'(\hat{\beta}_m - \hat{\beta}_e) + (\hat{t}_m \overline{I}_m - \hat{t}_e \overline{I}_e) \tag{11}$$

where the definitions outlined for (3) above apply but with $\hat{t}_j$ representing the OLS estimate of the j selection parameter for each ethnic group, and $\overline{I}_j$ is the corresponding sample averaged selection variable for the jth ethnic group obtained from either a probit model's estimates (if there are two work choices or outcomes as *per* Heckman, 1979) or a multinomial logit model's estimates (if there are more than two work choices or outcomes as *per* Lee, 1983). Given the 'index number' problem, the ethnic pay gap can also be re-expressed as follows:

$$\overline{w}_m - \overline{w}_e = (\overline{x}_m - \overline{x}_e)'\hat{\beta}_e + \overline{x}_m'(\hat{\beta}_m - \hat{\beta}_e) + (\hat{t}_m \overline{I}_m - \hat{t}_e \overline{I}_e) \tag{12}$$

There is currently little consensus regarding the most appropriate treatment for selectivity bias in quantile regression models. Buchinsky (1999) uses the work of Newey (1999) to approximate the selection term through use of a higher order series expansion. The power series is based on the inverse Mills ratio or transformations of it. This approach has

implications, however, for the identification of the wage regression intercept term given its conflation with the constant term associated with the higher order series providing the proxy for the selection effect. However, if the choice outcomes are multiple rather than binary, as in the current application (see below), the appropriate correction is less straightforward. Given a lack of consensus and the potential intercept identification problem associated with the introduction of higher order selection terms into quantile regression models, we simply insert a conventionally computed selection correction term into the quantile regression models. This could be taken to provide a reasonably good approximation for the selection effect in the quantile regression model and was, for example, used by Arcand and D'Hombres (2004) in their analysis of ethnic differences in labour market earnings in Brazil. We take the view that it is better to introduce an albeit imperfect proxy for the selection effect in this application rather than ignoring the problem completely. However, as a robustness check, we do explore the implication for our results if the selection terms are excluded from the specifications.

The ethnic pay gap at the qth quantile is thus now decomposed into four parts:

$$\Delta_q = \Delta\boldsymbol{\Omega}_q'\hat{\boldsymbol{\beta}}_{qm} + \boldsymbol{\Omega}_{qe}'\Delta\hat{\boldsymbol{\beta}}_q + \Delta\boldsymbol{R}_q + \Delta t_q \tag{13}$$

with definitions as in (9) above, and where the fourth component represents the difference in the selection effect defined as $\Delta t_q = [\hat{t}_{qm}E(l_m|w_m = Q_q(w_m)) - \hat{t}_{qe}E(l_e|w_e = Q_q(w_e))]$. The ethnic pay gap can also be decomposed as:

$$\Delta_q = \Delta\boldsymbol{\Omega}_q'\hat{\boldsymbol{\beta}}_{qe} + \boldsymbol{\Omega}_{qm}'\Delta\hat{\boldsymbol{\beta}}_q + \Delta\boldsymbol{R}_q + \Delta t_q \tag{14}$$

In order to address the problem of selectivity bias in this paper we exploit the approach developed by Lee (1983). The procedure is two-step and, in our application, exploits estimates from a three-category multinomial logit model (MNL) for the majority and minority sub-samples to construct the set of relevant selection correction terms. The three employment outcomes used in our case are (i) farming, (ii) wage employment, and (iii) self-employment in non-farm activities.

4. Empirical Analysis

The sample data used comprise all those in wage employment from both gender groups aged between 18 and 60 years and covers employees across a range of enterprise types (e.g., public, SOE, private *etc.*). The wage rate used is an hourly measure relating to the individual's primary or main job and includes basic rates and other payments in terms of bonuses, allowances, subsidies both in cash and in kind. This definition of the wage has been widely used in other labour market studies for Vietnam such as Liu (2004), Gallup (2004), and Pham and Reilly (2007). In contrast to the standard Mincerian specification, the current study uses a set of educational dummies to capture human capital effects. In order to allow for variation in the returns to these human capital measures across urban and rural settlement types, a set of variables interacting urban residence with the educational controls were also included in the regression models estimated [7]. In addition, as there is insufficient information in the VHLSS 2002 to compute actual labour force experience, the age of an individual rather than a potential labour force measure is used to proxy for labour market experience [8]. The age variable is parameterised in the wage specification through a set of age dummies. The other characteristics included in the wage regression models comprise controls for gender, marital status, sector of employment defined according to ownership type, a health status measure based on whether the individual recently had medical treatment [9], a control for the quarter in which the interview occurred to capture any seasonal effects that may impact on the wage determination process, and a set of regional controls. A brief description of these variables (and their corresponding summary statistics) is provided in table A1 of the Appendix.

A set of household-level variables were used to identify the selection effects in the wage equations. These include the number of children and the dependency ratio within the household, the educational attainment and occupation of the head of household, and whether the household had access to non-labour income sources. The MNL estimates for the three-category employment model are neither reported nor discussed here in order to conserve space [10].

Table 3 reports the estimates for the pooled regression model where ethnicity enters as an intercept shift and provides an estimate of the *ceteris paribus* 'mark-up' for the majority wage workers over their ethnic minority counterparts. The table reveals that an ethnic majority worker earns over 11 percent more in hourly wages relative to an ethnic minority worker, considerably more modest than the ethnic differentials computed for Vietnam using household welfare measures [11]. The estimates from the quantile regression model at the median are comparable to the mean regression results perhaps suggesting that outliers do not represent a significant problem in this application. The results from the quantile regression models reveal a narrowing *ceteris paribus* ethnic pay gap with movement up the conditional wage distribution. Around the bottom decile of the conditional wage distribution, the 'mark-up' is nearly 21 percent compared to only four percent at the top decile. These findings at the extremes of the conditional wage distribution could tentatively be argued as being consistent with the notion of 'sticky floors' at the bottom end of the conditional wage distribution, where ethnic minority workers are crowded into low-paying jobs, and indicating an absence of a 'glass-ceiling' for ethnic minority workers at the top end of the distribution [12]. However, these twin inferences, in the absence of additional supporting evidence, are best interpreted as suggestive rather than compelling.

[Take in Table 3]

Attention now turns to the decomposition of the ethnic wage gap into treatment and endowment effects. The coefficient estimates for the separate ethnic wage equations are reported in tables A2 and A3 of the Appendix but are not the subject of separate discussion here. The null hypothesis of common parameters across the two ethnic groups is decisively rejected on the basis of appropriately computed Wald tests in all relevant cases (see table 3 and the corresponding table notes). The separation of the data points across the ethnic groups used here is thus empirically justified.

It is worth noting that the estimated selection effects in the mean regressions for both the majority and minority groups are both well determined statistically (see tables A2 and A3). Given the construction of the Lee (1983) selection term, the estimated negative effect suggests intuitively plausible positive selection into wage employment for both groups (see Gyorko and Tracy, 1988; Reilly, 1991). The average selection effects for the majority and minority workers are computed at 0.22 and 0.23 respectively (i.e., the estimated selection coefficients (from tables A2 and A3) multiplied by the sample average selection value (from table A1)). This suggests that an average individual from either one of the two ethnic groups that selects into wage employment earns $[e^{0.22} - 1] \times 100 \approx 25\%$ more in hourly wages than someone drawn at random from the relevant ethnic population group with comparable observable characteristics. This could be interpretable as a labour market premium to unobservables, which appears broadly comparable in magnitude across the two groups.

The decomposition results for the mean regressions are reported in table 4. As selection effects are present in all reported regression models, we decompose the ethnic wage gap using expressions (11) to (14). As the estimates appear insensitive to the set of characteristics used to compute the pay gaps, interpretation will focus on the decomposition estimates based on equations (11) and (13) only.

The point estimates for all the treatment and endowment effects are found to be precisely determined at a conventional level of statistical significance. In raw terms the average gap in log hourly wages between the majority and minority ethnic groups is dimensionally comparable to *ceteris paribus* estimates based on the pooled OLS regression model (see table 3). On average, majority workers earn nearly 11 percent more than their minority counterparts. However, the estimates based on the mean regression models reveal that about one-third of this differential is accounted for by differentials in the average characteristics between the two groups (i.e., the endowment effect). The treatment effect accounts for just under two-thirds of the earnings differential, with the remaining negligible amount attributable to selection effects. This suggests that most of the ethnic wage gap is due to differentials in returns to a given

'basket' of characteristics. This could be interpreted as representing unequal treatment in favour of the wage employed majority group with the average 'mark-up' of the order of about seven percent.

In order to investigate the robustness of the findings for the mean regression models, the wage specifications for both majority and minority groups are augmented by a set of district effects [13]. The motivation for this is driven by the fact that ethnic minorities in Vietnam tend to be geographically differentiated. A failure to control for a finely disaggregated degree of regional heterogeneity may bias, among other things, the estimated treatment effects. The inclusion of these controls is thus designed to attenuate any potential bias in regard to this source of heterogeneity. Although the estimated effects for the treatment components reported in table 4 are subject to a very mild contraction with the introduction of the district effects, the magnitude of the change is modest.

[Take in Table 4]

The decomposition of the pay gap between the two ethnic groups based on the quantile regression models is now reported in table 5 [14]. The raw earnings differentials are widest at the bottom decile of the conditional wage distribution, decline with movement up this distribution, and are narrowest around the top decile. The computed treatment effects are more modest in magnitude than those suggested by the pooled regression models reported in table 3. However, the suggestive narrative offered above in regard to a 'sticky-floor' at the bottom end of the distribution (where the 'mark-up' is about 13%) and the apparent absence of a 'glass-ceiling' at the top end (where the 'mark-up' is about 1.6%) remains intact.

The portion of the ethnic wage gap that is accounted for by differences in endowments monotonically increases with movement up the conditional wage distribution. At the 10th percentile, the endowment effect comprises just nine percent of the raw earnings differential, while the corresponding figure at the 90th percentile is close to one-half. Despite this increasing endowment effect, the ethnic wage differentials are still largely attributable to differentials in

the prices the Vietnam labour market uses to value these characteristics for the different ethnic groups. The minority workers tend to have lower estimated returns to their endowments across the selected points of the conditional wage distribution examined here [15], [16].

[Take in Table 5]

The amalgamation of the ethnic minority groups into one homogeneous category for the purposes of analysis disguises potential variation in the magnitude of ethnic pay disparity between groups within this conflated category. The treatment effects, and their corresponding sampling variances, can be computed for each individual in the ethnic minority sub-sample. Thus, as a final illustrative exercise, we investigate for the presence of variation in the treatment effects across a number of broadly defined ethnic minority sub-groups. Given the constraints associated with using adequately sized cells, the groups we use comprise five in total. The groupings we use are intended to be functional and are based on discussions with Vietnamese anthropologists and local NGOs. The ethnic categories comprise the (i) Khmer and Cham; (ii) Tay, Thai, Muong, Nung; (iii) Other Northern minorities; (iv) Central Highland minorities; (v) an 'other' or miscellaneous category comprising the remaining smaller ethnic groups, which are mostly located in the North and South Central Coasts. For each ethnic minority individual we can compute their treatment effect based on what they would earn if they belonged to the majority group and what they earn given their minority group affiliation broadly defined. A standard error can also be computed for each individual's estimated treatment effect and this can be used to undertake a simple weighted least squares (WLS) regression of the treatment effects on a set of four ethnic minority dummies using only the data for the minority sub-sample [17]. The Khmer and Cham group provide the base category in the estimation of this regression model. Table 6 reveals a degree of variation in the treatment effects across ethnic groups but suggests that only about one-quarter of the total variation in the treatment effects are actually explained by these controls. In particular, the treatment effects for the Tay, Thai, Muong and Nung are

statistically smaller than the base group at the mean and at all selected points of the conditional wage distribution. Thus, this group could be taken to experience the smallest ethnic wage disadvantage relative to the Kinh-Hoa. The largest treatment effects, and thus the largest degree of labour market ethnic disadvantage, are reserved for those wage employed individuals that belong to the Central Highland minorities.

[Take in Table 6]

5. Conclusions

The portrait that emerges from the above analysis suggests evidence of wage inequality between the majority and minority groups in Vietnam [18]. The degree of ethnic minority wage disadvantage is considerably more modest than the ethnic gap in livings standards reported for Vietnam in 2002 (see footnote 11). At first glance, this finding appears to conflict with the widening gap in the living standards between the Kinh (and Chinese) and other ethnic minority groups reported in the literature (Winrock International, 1996; Poverty Task Force (PTF), 2002; Baulch *et al.* 2004). However, the labour market earnings differential is based on a sample-selected sub-group of the ethnic minority that appears to perform reasonably well within a Vietnamese labour market historically dominated by the Kinh majority. For the ethnic minorities in this selective sub-population, the data demonstrate that their living standards, and other observable characteristics associated with well-being (e.g., their education levels), are almost indistinguishable from those of the majority.

Despite evidence of a relatively good labour market performance, our analysis reveals that, on average, the ethnic minority group secures lower returns in the labour market for their endowments than the majority group. The raw wage premium for the majority is of the order of 11 percent but only one third of this is ascribed to differences in the average characteristics between the majority and minority groups. The remaining portion is largely attributable to differentials in returns to endowments given negligible differences in selection effects between

18

the two groups. In addition, although the ethnic 'mark-up' was found to be relatively high among those at the bottom end of the conditional wage distribution, it declined monotonically with movement up the distribution and little evidence of a 'glass-ceiling' effect for the ethnic minority was detected at the top end of the wage distribution.

The narrow gap in wages between the majority and minority groups could be taken to suggest that economic progress through the labour market may offer a way out of poverty. However, we need to temper this inference with some caution. Firstly, our econometric estimates suggest that the narrow average pay gap may be attributable to the fact that the ethnic minority workers currently in wage employment, on average, possess unobservables that predispose them to secure favourable rewards in the labour market. And even with these favourable unobservables, the ethnic minority group earns less, on average, than the majority due to lower labour market returns to their characteristics. The inevitable expansion of wage employment in Vietnam in response to economic growth is likely to encourage the participation of those with poorer levels of unobservables and, if these are drawn disproportionately from ethnic minority groups, the ethnic pay gap is likely to widen. Secondly, the existence of a large treatment effect at the bottom end of the conditional wage distribution suggests that ethnic minorities in the lower paid jobs are most at risk of unequal treatment. There is no guarantee with the future anticipated expansion of wage employment that this effect will not encroach further up the conditional wage distribution.

The issue of ethnic wage inequality has not been highlighted in Vietnam's most recently announced economic strategy or national programs. Given our findings, and an increasingly important putative role for the wage employment sector in the future, we believe that an emphasis on labour market ethnic wage inequality should be a core theme highlighted within Vietnam's national program for ethnic development.

Finally, it should be noted that this study limited itself to the ethnic earnings gap of the wage employed, which in 2002 accounted for roughly 28 percent of total employment in Vietnam

(see table 2). The remaining employment category includes the self-employed operating in either agriculture or non-farm activities. The self-employed group is largely beyond the influence of official labour market regulations (ADB, 2004; Brassard, 2004), so the magnitude of the ethnic wage gap for this employment category could be wider than that reported here. However, investigating the ethnic earnings differential for the self-employed is constrained by the data currently available but is an issue that warrants consideration as part of an agenda for future research.

Acknowledgment

Notes

1. It is not clear how, if at all, the non-sampling errors associated with the VHLSS 2002 impact on the data for the wage employment sector used here. However, it is worth noting that the ethnic representation within this survey compares favourably with the earlier and the later surveys that were not subject to non-sampling error problems.

2. The ethnic status of the individual is determined by the ethnic status of the head of household within which the individual is located.

3. In common with other studies on ethnic underdevelopment in Vietnam, the Chinese ethnic group, which ranks 6th in terms of population size in Vietnam, is merged with the Kinh majority group for the purposes of our analysis. The Chinese households are generally well-off in Vietnam (see Van de Walle and Gunewardana, 2001). As the objective of this study is to examine the plight of disadvantaged ethnic groups in the waged labour market, the Chinese group is conflated with the Kinh to define the majority, while all other ethnic groups are taken to represent the minority.

4. This is commonly referred to as the 'index number' approach given its similarity to the calculation of index numbers in the field of statistics. It is well known from undergraduate statistics that an index number computed using the Laspeyres formula does not yield a numerical value that is identical to an index number computed using the Paasche formula. This conflict is the essence of the 'index number' problem reported here.

5. The emphasis on decomposing ethnic wage differences using quantile regression models is in the spirit of work undertaken for Brazil on racial pay gaps by Arias, Yamada and Tejerina (2004) and Arcand and D'Hombres (2004).

6. See Neuman and Oaxaca (2004) for a more extensive discussion of issues related to 'index number' decompositions and selectivity correction.

7. We also explored using both mean and quantile regression models whether the estimated returns for other variables varied across the urban/rural divide for both the majority and minority groups. Wald tests revealed no evidence of differential returns using either sub-samples for variables other than the human capital measures. In addition, we also investigated whether estimated returns varied between the north and the south of the country but again found no evidence for this. The test results for both these exercises are available from the corresponding author on request.

8. Other studies on Vietnam have computed potential labour force experience by using the individual's age minus estimated years in schooling minus the school enrolment age. In addition to the potential measurement error embodied in any 'years-in-schooling' variable, this approach might also introduce an additional error if it fails to take into account discrepancies in the enrolment ages and years of school repetition, which is quite common in the more remote areas of Vietnam.

9. It could be argued that the health status measure and wages are endogenously determined rendering suspect its inclusion as an exogenous variable. The fact that the health status measure pre-dates the current wage used suggests the assumption of exogeneity could be viewed as relatively innocuous. However, the absence of good instruments to test explicitly this exogeneity proposition suggests some interpretational caution should perhaps be exercised in respect of this variable's estimates.

10. The Multinomial logit (MNL) model passes the test for the independence from irrelevant alternatives (IIA) proposition for both ethnic groups using the Small-Hsiao (1987) test in all but one case, thus broadly satisfying a necessary condition for the use of the Lee (1983) procedure in our application. The test results are available from the corresponding author on request.

11. Using data for the full sample of households from the same VHLSS 2002, the *ceteris paribus* gap in living standards between the majority and the minority, as measured by per capita household expenditure, is nearly 26 percent. This *ceteris paribus* gap is obtained from the

ethnic dummy estimate in an OLS equation of the log household per capita expenditure on a comparably defined ethnic dummy and characteristics capturing the education, age, and gender of the household head, household demographic structure, household landholding, other non-labour income sources, and location. This specification is similar to that reported in Baulch *et al.*, (2004) with the exception that the head of household education attainment levels are specified as in Van de Walle and Gunewardana (2001).

12. Albrecht, Björklund and Vroman (2003), Arulampalan, Booth and Bryan (2005), and De la Rica, Dolado and Llorens (2007) provide explanations of 'glass ceilings' and 'sticky floors' for a variety of countries in the context of gender pay gaps using quantile regression models.

13. There were 503 district effects introduced into the majority mean regression equation and 167 into the corresponding minority equation.

14. Unfortunately, the parameters of the quantile regression models are not estimable with standard econometric software when the district effects (see footnote 13) are included given the prohibitively large number of parameters this implies. Thus, the decompositions are based on models that exclude the district effects but include the standard regional controls. However, we take some comfort from the fact that mean regression estimates for treatment and endowment effects are found to be fairly insensitive to the exclusion of the district effects.

15. In order to determine whether the reported results are sensitive to the removal from the majority group of the small Chinese ethnic group, we re-estimated the wage regression models for the Kinh as the only majority group. The resultant estimates for the treatment and endowment components were found to be invariant to this exclusion using both mean and quantile regression models. The results of this exercise are available from the corresponding author on request.

16. Given reservations expressed about the use of the *ad hoc* selection term in the quantile regression models, the decompositions were also re-computed for quantile regression models estimated without the inclusion of a selection term. This exercise resulted in only modest changes to the relevant point estimates for the treatment and endowment effects and did not alter any of the inferences contained in the text. The results of this exercise are available from the corresponding author on request.

17. The weighted least squares procedure in this application is designed to give a higher weight in estimation to the more precisely estimated treatment effects within the minority sub-sample.

18. The magnitude of the ethnic wage gap in the Vietnamese labour market is modest when compared to the recent experience of a number of transitional European labour markets. For example, see Kroncke and Smith (1999) for Estonia and Giddings (2003) for Bulgaria.

References

ADB (Asian Development Bank) (2004), "The Factor Markets in Vietnam", Discussion Paper No. 5, Making Markets Work Better for the Poor, Asia Development Bank, Vietnam Resident Mission, Hanoi.

Albrecht, J., Björklund, A, and Vroman, S. (2003), Is there a Glass Ceiling in Sweden, *Journal of Labor Economics*, Vol. 21, pp.145 –177.

Arcand, J.L. and D'Hombres, B. (2004), "Racial Discrimination in the Brazilian Labour Market: Wage Employment and Segregation Effects", Journal of International Development, Vol. 16 No. 8, pp. 1053-1066.

Arias, O., Yamada, G. and Tejerina, L. (2004), "Education, Family Background and Racial Earnings Inequality in Brazil", International Journal of Manpower, Vol. 25, pp. 355-374.

Arulampalam, W., Booth, A. and. Bryan, M. (2005), "Is there a Glass Ceiling over Europe? An Exploration of Asymmetries in the Gender Pay Gap across the Wages Distribution", Institute for Social and Economic Research (ISER) Working Paper # 2005-25, Essex University.

Baulch, B., Truong, K.T.C., Haughton, H. and Haughton, J. (2004), "Ethnic Minority Development in Vietnam: A Socioeconomic Perspective", in Glewwe, P., Agrawal, N. and Dollar, D. (Ed.), Economic Growth, Poverty, and Household Welfare in Vietnam, World Bank Regional and Sectoral Studies, World Bank, Washington D.C., pp. 273-310

Brassard, C. (2004), "Wage and Labour Regulation in Vietnam within the Poverty Reduction Agenda", mimeo, Public Policy Programme, National University of Singapore.

Buchinsky, M. (1999), "The Dynamics of Changes in the Female Wage Distribution in the USA: A Quantile Regression Approach", Journal of Applied Econometrics, Vol. 13, pp. 1-30.

Chamberlain, G. (1994), "Quantile regression, censoring and the structure of wages", in Sims, C. and Laffont, J.J. (Ed.), Proceedings of the sixth world congress of the Econometrics Society, Cambridge University Press, New York, pp. 171-209.

Dang, N.V., Chu, T.S. and Luu, H. (2000), Ethnic Minorities in Vietnam, The Gioi Publisher, Hanoi.

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

De la Rica, S., Dolado J.J. and Llorens, V. (2007), "Ceilings or Floors ?: Gender Wage Gaps by Education in Spain", *Journal of Population Economics*, forthcoming.

Gallup, J.L. (2004), "The Wage Labor Market and Inequality in Vietnam", in Glewwe, P., Agrawal, N. and Dollar, D. (Ed.), Economic Growth, Poverty, and Household Welfare in Vietnam, World Bank Regional and Sectoral Studies, World Bank, Washington D.C., pp. 53-94.

Giddings, L. (2003), "Continued Decline for Ethnic Minorities in the Transition?: Changes in Ethnic Earnings Differentials in Bulgaria, 1986, 1993 and 1997", Economics of Transition, Vol. 11, No.4, pp. 621-648.

Gyourko, J. and Tracy, J. (1988), "An Analysis of Public- and Private-Sector Wages Allowing for Both Endogenous Choice of Government and Unions", Journal of Labor Economics, Vol. 6, pp. 229-253.

Heckman, J. (1979), "Sample Selection Bias as a Specification Error", Econometrica, Vol. 47, pp.153-161.

Huber, P.J. (1967), "The behaviour of maximum likelihood estimates under non-standard conditions", in Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, CA, Vol. 1, pp. 221-223.

Koenker, R. (2005), Quantile Regression, Econometric Society Monograph Series, Cambridge University Press, Cambridge.

Koenker, R. and Bassett, G. (1978), "Regression Quantiles", Econometrica, Vol. 46 No. 1, pp. 33-50.

Kroncke, C. and Smith, K. (1999), "The wage Effects of Ethnicity in Estonia", Economics of Transition, Vol. 7, No.1, pp. 179-199.

Lee, L.F. (1983), "Generalized Econometric Models with Selectivity", Econometrica, Vol. 51 No. 2, pp. 507-512.

Liu, A.Y.C. (2004), "Gender Pay Gap in Vietnam: 1993-1998", Journal of Comparative Economics, Vol. 32 No. 3, pp. 586-596.

Machado, J. and Mata, J. (2005), "Counterfactual Decomposition of Changes in Wage Distributions Using Quantile Regression Models", Journal of Applied Econometrics, Vol. 20, pp. 445-465.

Mincer, J. (1974), Schooling, Experience, and Earnings, Columbia University, New York.

MoET (Ministry of Education and Training) (2006), "Education and Training for Ethnic Minorities at Universities and Colleges", report by the University Department, MoET (in Vietnamese).

MOLISA (Ministry of Labour, War Invalids, and Social Affairs) and UNDP (United Nations Development Program) (2004), "Taking stock, planning ahead: Evaluation of the national targeted programme on hunger eradication and poverty reduction and programme 135", mimeo, UNDP Country Office, Hanoi.

Moock, P., Patrinos, H.A. and Venkataraman, M. (1998), "Education and Earnings in a Transition Economy The Case of Vietnam", mimeo, Development Research Group, World Bank, Washington D.C.

Neuman, S. and Oaxaca, R.L. (2004), "Wage Decompositions with Selectivity-Corrected Wage Equations: A Methodological Note", Journal of Economic Inequality, Vol. 2, pp.3-10.

Newey, W.K. (1999), "Two Step Series Estimation of Sample Selection Models", mimeo, Department of Economics, Massachusetts Institute of Technology.

Oaxaca, R.L. (1973), "Male-Female Wage Differentials in Urban Labour Markets", International Economic Review, Vol. 14, pp. 693-709.

Pham, T.H. and Reilly, R. (2007), "Gender Pay Gap in Vietnam in 1993-2002: A Quantile Regression Approach", Journal of Asian Economics, Vol. 18, No.5, pp.775-806.

Phung, D.T. and Nguyen, P. (2006), "Vietnam Household Living Standard Survey (VHLSS), 2002 and 2004: Basic Information", mimeo, General Statistics Office, Hanoi.

PTF (Poverty Task Force) (2002), "Localizing MDGs for Poverty Reduction in Vietnam: Promoting Ethnic Minority Development", report by Poverty Task Force, World Bank Vietnam Office, Hanoi.

Reilly, B. (1991), "Occupational Segregation and Selectivity Bias in Occupational Wage Equations: An Empirical Analysis Using Irish Data", Applied Economics, Vol.23, No.1, pp.1–7.

Small, K.A. and Hsiao, C. (1985), "Multinomial Logit Specification Tests", International Economic Review, Vol. 26 No. 3, pp. 619-627.

SRV (Socialist Republic of Vietnam) (2001), "National Target Programme on HEPR and Employment for the Period 2001-2005", Ministry of Labour, Invalids and Social Affairs, Socialist Republic of Viet Nam, Hanoi.

Swinkels, R. and Turk, C. (2006), "Explaining Ethnic Minority Poverty in Vietnam: a summary of recent trends and current challenges", background paper for Ministry of Planning and Investment meeting on Ethnic Minority Poverty, Sept 2006, Hanoi.

Van de Walle, D. and Gunewardena, D. (2001), "Sources of Ethnic Inequality in Viet Nam", Journal of Development Economics, Vol. 65, pp. 177-207.

Winrock International (1996), "Ethnic Minorities in Vietnam: A Country Profile", report prepared for the World Bank, Winrock International, Hanoi.

World Bank (2000), Vietnam Living Standard Survey (VLSS), 1992-93: Basic Information, Poverty and Human Resources Division, World Bank, Washington, D.C.

World Bank (2001), Vietnam Living Standard Survey (VLSS), 1997-98: Basic Information, Poverty and Human Resources Division, World Bank, Washington, D.C.

Table 1: Summary Statistics for Majority and Minority Groups

	Majority (1)	Minority (2)	z-score or t-statistics (2) - (1)	Wage majority (3)	Wage Minority (4)	z-score or t-statistics (4) - (3)
Household Living standards						
− Poverty headcount (%)	23.81	65.40	52.28***	17.24	43.53	14.15***
− Poverty gap (%)	5.00	21.02	35.81***	5.00	19.02	10.62***
− Mean expenditure p.c (VND 1000)	3,697	1,911	35.11***	4,361	2,712	7.72***
− Sample size (no. of households)	25,454	4,078	*f*	10,301	575	*f*
Individual educational attainments (%)						
− Illiteracy	19.66	49.39	40.88***	13.69	18.70	12.21***
− Primary education	27.98	28.16	1.02	22.66	24.79	1.81*
− Lower secondary education	31.75	15.86	21.55***	24.46	21.32	5.93***
− Upper secondary education	12.56	4.23	11.35***	15.01	17.40	2.87***
− Technical/vocational training	4.46	1.71	5.48***	11.81	11.17	3.49***
− College/university	3.60	0.65	10.71***	12.37	6.63	3.70***

Source: calculations from the VHLSS 2002

Notes to table 1:

a. (1) represents the Kinh (or Chinese) headed households; (2) comprises the ethnic minority households ; (3) refers to the majority households that have at least one member who was wage-employed; and (4) denotes the ethnic minority households that have at least one member who was wage-employed.

b. The poverty headcount and poverty gap index is calculated using the 2002 general poverty line developed by GSO and the World Bank. This general poverty line is based on expenditure on the food basket of the 3rd quintile that is equivalent to average consumption of 2,052 calories per day plus the non-food expenditures of the 3rd quintile. The rice equivalence of the general poverty line is 51kg. The real monetary value of the general poverty line in 2002 is VND 1,920 thousand (adjusted by monthly and regional price indices), where VND denotes the Vietnamese Dong. See World Bank (2003) for more details.

c. *f* denotes 'not applicable'.

d. The individual educational attainments are calculated using the sample of individuals aged from 18 to 60 years who were working over the past 12 months.

e. The z-scores and t-tests are computed to test for statistical differences between the majority and minority groups. The test values are reported in absolute terms ***, **, and * denotes statistical significance at the 0.01, 0.05 and 0.1 levels respectively.

Table 2: Majority-Minority Differences by Selected Characteristics

	Majority	Minority	z-score or t-statistics	% Change
Labour force participation rate (%)	81.54	89.70	14.27***	-9.10
Employment structure (%)				
− Agriculture	48.26	77.40	44.13***	-37.65
− Wage employment	29.55	15.01	24.55***	96.87
− Self-employment outside agriculture	22.18	7.60	27.33***	191.84
Wage rates by sector (VND 1000)				
− Government agencies	4.44	4.41	2.72***	0.68
− State-owned enterprises	3.88	2.93	1.84*	32.42
− Formal private enterprises	3.59	2.48	0.23	44.76
− Household or informal businesses	3.23	2.64	4.98***	22.35
Wage rates by education (VND 1000)				
− Illiteracy	3.11	2.63	3.80***	18.25
− Primary education	3.41	2.97	2.71***	14.81
− Lower secondary education	3.78	3.48	1.67*	8.62
− Upper secondary education	4.48	4.33	2.57***	3.46
− Vocational and technical training	5.05	4.98	5.29***	1.41
− College/University	6.95	5.71	2.80***	21.72
Average hourly real wage rates (VND 1000)				
− All	4.19	3.76	4.46***	11.44
− Urban	5.04	4.44	3.35***	13.51
− Rural	3.54	3.24	3.14***	9.26

Source: calculations from the VHLSS 2002

Notes to table 2:

a. The labour force participation rate is calculated as percentage of those who were working during the past 12 months in the total labour force aged from 18 to 60 years old.

b. Wage rate includes salary, and other payments in cash and kind (bonuses, holidays, subsidies *etc.*). The monthly and regional prices indices were used to adjust the nominal rate to the hourly real wage rate (Jan. 2002 prices).

c. Wage and employment outcomes are defined exclusively in terms of the primary job.

d. The z-scores and t-tests are testing for the statistical differences between majority and minority outcomes. The test results are reported in absolute terms. See notes to table 1 for the starring convention used.

e. The '% change' column gives the percentage difference between the majority and minority outcomes.

Figure 1: Kernel Densities of Majority and Minority Real Wages
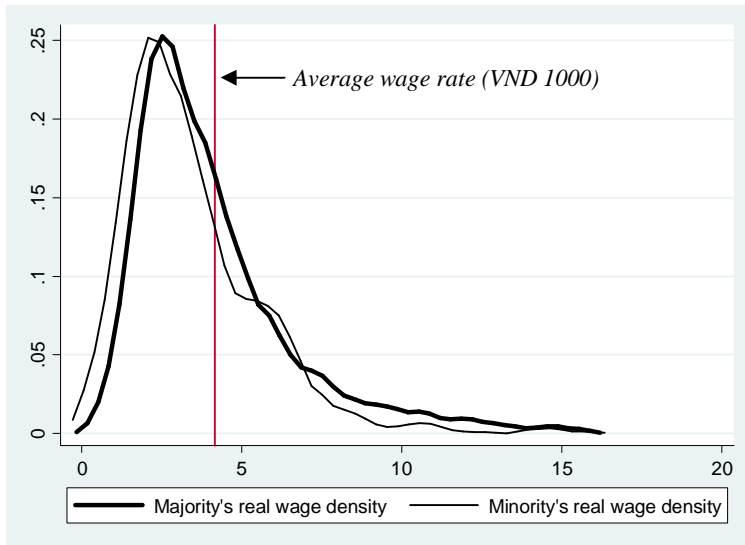
Source: VHLSS 2002

Table 3: Ethnic Wage Gap in Vietnam: Pooled Regression Results

	Mean	10th	25th	50th	75th	90th
Kinh (and Chinese) majority	0.1099***	0.1891***	0.1699***	0.1077***	0.0682***	0.0416**
	(0.019)	(0.042)	(0.025)	(0.019)	(0.022)	(0.021)
Male worker	0.1303***	0.1278***	0.1121***	0.0871***	0.097***	0.078***
	(0.009)	(0.018)	(0.009)	(0.01)	(0.011)	(0.015)
Married	-0.0702***	-0.081***	-0.0733***	-0.0787***	-0.0747***	-0.091***
	(0.011)	(0.025)	(0.011)	(0.015)	(0.014)	(0.019)
Aged from 26 to 35	0.1182***	0.172***	0.1263***	0.1047***	0.0817***	0.0948***
	(0.012)	(0.028)	(0.012)	(0.014)	(0.017)	(0.022)
Aged from 36 to 45	0.1557***	0.1831***	0.154***	0.1424***	0.1353***	0.148***
	(0.014)	(0.03)	(0.017)	(0.016)	(0.018)	(0.024)
Aged from 46 to 55	0.1611***	0.1007***	0.1355***	0.1803***	0.1791***	0.1956***
	(0.016)	(0.029)	(0.02)	(0.018)	(0.021)	(0.027)
Aged from 56 to 60	0.1197***	0.0411	0.0454	0.1557***	0.1645***	0.235***
	(0.041)	(0.115)	(0.057)	(0.034)	(0.053)	(0.063)
Primary education	0.1048***	0.1478***	0.0801***	0.1065***	0.108***	0.1252***
	(0.015)	(0.027)	(0.012)	(0.015)	(0.016)	(0.028)
Lower secondary education	0.1103***	0.1877***	0.1267***	0.124***	0.0888***	0.0798***
	(0.016)	(0.029)	(0.017)	(0.016)	(0.018)	(0.029)
Upper secondary education	0.2049***	0.2371***	0.1479***	0.1777***	0.2146***	0.305***
	(0.022)	(0.041)	(0.029)	(0.024)	(0.028)	(0.04)
Vocational/Tech. training	0.4645***	0.5269***	0.4203***	0.4333***	0.4766***	0.4517***
	(0.024)	(0.039)	(0.031)	(0.03)	(0.027)	(0.046)
College/University	0.6749***	0.7659***	0.7126***	0.6516***	0.6307***	0.5769***
	(0.026)	(0.046)	(0.034)	(0.025)	(0.028)	(0.04)
Urban	0.1243**	0.1381***	0.1501***	0.1308***	0.0942***	0.0871**
	(0.025)	(0.045)	(0.032)	(0.032)	(0.029)	(0.045)
Urban*Primary education	0.0145	0.0179	0.0154	-0.0411	-0.0601**	-0.0373
	(0.031)	(0.062)	(0.033)	(0.037)	(0.028)	(0.053)
Urban*Lower secondary	0.0639**	0.0099	0.0123	0.0506*	0.042	0.075*
	(0.031)	(0.048)	(0.043)	(0.028)	(0.038)	(0.046)
Urban*Upper secondary	0.1543***	0.1373**	0.1513***	0.108***	0.1081***	0.0958**
	(0.035)	(0.061)	(0.039)	(0.025)	(0.038)	(0.039)
Urban*Vocational or Tech.	0.0319	0.1078*	0.1055***	0.0043	0.066**	0.0818*
	(0.034)	(0.056)	(0.041)	(0.039)	(0.03)	(0.046)
Urban*College or University	0.1276***	0.128**	0.0725**	0.1022**	0.1132**	0.226***
	(0.035)	(0.059)	(0.037)	(0.043)	(0.047)	(0.053)
Government agencies	0.053***	0.0824***	0.0306	0.0609***	0.1445***	0.1261***
	(0.015)	(0.018)	(0.022)	(0.017)	(0.023)	(0.029)
State-owned enterprises	0.1537***	0.0636***	0.1356***	0.1875***	0.1973***	0.1914***
	(0.014)	(0.023)	(0.016)	(0.018)	(0.019)	(0.024)
Formal private enterprises	0.2074***	0.1916***	0.1936***	0.2335***	0.2069***	0.16***
	(0.014)	(0.024)	(0.019)	(0.017)	(0.019)	(0.025)
Medical treatment	-0.0315*	-0.0142	-0.0226	-0.0218	-0.0399*	-0.044
	(0.019)	(0.043)	(0.026)	(0.024)	(0.021)	(0.034)
Northern uplands	-0.1295***	-0.1542***	-0.1493***	-0.1596***	-0.1545***	-0.1112***
	(0.018)	(0.027)	(0.019)	(0.022)	(0.02)	(0.027)
Red River Delta	-0.1632***	-0.223***	-0.1754***	-0.1796***	-0.1706***	-0.1326***
	(0.015)	(0.023)	(0.018)	(0.016)	(0.018)	(0.024)
North Central Coast	-0.1743***	-0.2295***	-0.1939***	-0.2155***	-0.1766***	-0.1349***
	(0.019)	(0.026)	(0.019)	(0.019)	(0.028)	(0.033)
South Central Coast	-0.0241*	-0.0195	-0.0208	-0.047**	-0.0588***	-0.0519**
	(0.015)	(0.021)	(0.017)	(0.019)	(0.016)	(0.027)
Central Highlands	0.1766***	0.078**	0.102***	0.1479***	0.2323***	0.2985***
	(0.02)	(0.035)	(0.025)	(0.026)	(0.024)	(0.035)
Mekong River Delta	0.0272*	-0.0182	0.0029	0.003	0.0169	0.0647***
	(0.014)	(0.024)	(0.015)	(0.014)	(0.014)	(0.023)
Selection bias correction term	-0.2671***	-0.2354**	-0.2729***	-0.3457***	-0.3219**	-0.2672**
	(0.078)	(0.103)	(0.097)	(0.136)	(0.155)	(0.138

Constant	0.5539***	-0.0387	0.3559***	0.603***	0.9217***	1.1196***
	(0.026)	(0.048)	(0.03)	(0.029)	(0.027)	(0.035)
R^2/Psuedo R^2	0.2918	0.1402	0.1613	0.1901	0.2157	0.2206
Wald Test (1) $\sim \chi^2_{31}$	67.534***	78.901***	95.725***	71.029***	105.67***	86.205***
Wald Test (2) $\sim \chi^2_{32}$	71.105***	102.56***	69.744***	98.601***	85.611***	78.529***
Number of observations	16,170	16,170	16,170	16,170	16,170	16,170

Notes to table 3:

a. ***, **, and * denotes statistical significance at 0.01, 0.05 and 0.1 levels respectively.

b. Standard errors are reported in parentheses. The OLS standard errors are based on Huber (1967) and the quantile regression estimates are based on bootstrapping with 200 replications.

c. The specifications also include three controls for the quarters within which the interview was conducted.

d. The Wald statistics test for common parameters across the two ethnic groups. Wald Test (1) provides an overall test for ethnic differences across all parameters other than the constant term. Wald Test (2) provides an overall test for ethnic differences for all parameters including the constant term.

Table 4: Ethnic Wage Decomposition Analysis at the Mean

	Due to Endowment Differences	Due to Treatment Differences	Due to Selection Differences	Due to District Differences	Total Difference
Using equation (11)					
Without District Effects	0.0355***	0.0712**	0.0019*	Not applicable	0.1087***
	(0.0139)	(0.0301)	(0.0012)		(0.0202)
With District Effects	0.0323**	0.0652***	0.0018*	0.0094	0.1087***
	(0.0162)	(0.0258)	(0.0011)	(0.0069)	(0.0211)
Using equation (12)					
Without District Effects	0.0387**	0.0681***	0.0019*	Not applicable	0.1087***
	(0.0179)	(0.0251)	(0.0012)		(0.0202)
With District Effects	0.0298**	0.0677**	0.0018*	0.0094	0.1087***
	(0.0126)	(0.0275)	(0.0011)	(0.0069)	(0.0211)

Notes to table 4:

a. ***, **, and * denotes statistical significance at the 0.01, 0.05 and 0.1 levels respectively;

b. Standard errors are reported in parentheses;

Table 5: Ethnic Wage Decomposition Analysis at Selected Quantiles

	Due to Endowment Differences	Due to Treatment Differences	Due to Selection Differences	Due to Unobservable Differences	Total Difference
Using equation (13)					
Ethnic wage gap at 10th	0.0198**	0.1316***	0.0213	0.0509	0.2236
	(0.0103)	(0.0482)	(0.0149)		
Ethnic wage gap at 25th	0.0328**	0.1072***	0.0215	0.0179	0.1794
	(0.0149)	(0.0357)	(0.0159)		
Ethnic wage gap at 50th	0.0338***	0.0749***	0.0042*	-0.0009	0.1121
	(0.0135)	(0.0266)	(0.0024)		
Ethnic wage gap at 75th	0.0279**	0.0495**	0.0069	-0.0085	0.0758
	(0.0142)	(0.0223)	(0.0051)		
Ethnic wage gap at 90th	0.0139**	0.0156**	0.0087	-0.0055	0.0327
	(0.0068)	(0.0075)	(0.0062)		
Using equation (14)					
Ethnic wage gap at 10th	0.0402***	0.1112***	0.0213	0.0509	0.2236
	(0.0147)	(0.0442)	(0.0149)		
Ethnic wage gap at 25th	0.0441**	0.0959**	0.0215	0.0179	0.1794
	(0.0225)	(0.0461)	(0.0159)		
Ethnic wage gap at 50th	0.0397***	0.069***	0.0042*	-0.0009	0.1121
	(0.0135)	(0.0257)	(0.0024)		
Ethnic wage gap at 75th	0.0357**	0.0417**	0.0069	-0.0085	0.0758
	(0.0169)	(0.0187)	(0.0051)		
Ethnic wage gap at 90th	0.0097*	0.0198***	0.0087	-0.0055	0.0327
	(0.0058)	(0.0077)	(0.0062)		

Notes to table 5:

a. ***, **, and * denotes statistical significance at the 0.01, 0.05 and 0.1 levels respectively.

b. Standard errors are reported in parentheses;

Table 6: WLS Regression of Treatment Effects on Ethnic Groups

	Mean	10th	25th	Median	75th	90th
Tay, Thai, Muong, Nung	-0.1062***	-0.0984***	-0.1842***	-0.149***	-0.0739***	-0.1076***
	(0.023)	(0.03)	(0.03)	(0.014)	(0.016)	(0.024)
Other Northern Uplands	-0.0399*	-0.0788	-0.1481**	-0.0817***	-0.0204	-0.0778*
	(0.022)	(0.063)	(0.06)	(0.029)	(0.035)	(0.048)
Central Highland Minorities	0.3422***	0.725***	0.4005***	0.2034***	0.2236***	0.1568***
	(0.027)	(0.042)	(0.039)	(0.019)	(0.023)	(0.033)
Other Minority groups	0.0683	0.0944	0.1747	0.0068	0.0459	0.023
	(0.079)	(0.16)	(0.153)	(0.073)	(0.085)	(0.134)
Constant	0.1143***	0.1006***	0.1298***	0.1223***	0.103***	0.1481***
	(0.02)	(0.018)	(0.018)	(0.008)	(0.01)	(0.014)
Unadjusted R^2	0.2693	0.2248	0.2979	0.2666	0.1782	0.1008
Number of observations	968	968	968	968	968	968

Notes to table 6:

a. The weighted least square regressions are estimated using as weights the square root of the sampling variances of the treatment effects estimated for each ethnic minority-headed household.

b. The unadjusted R^2 is computed as the squared correlation coefficient between the actual treatment effects and their predicted values from the relevant regression model.

c. Standard errors are reported in parentheses.

d. ***, **, and * denotes statistical significance at the 0.01, 0.05 and 0.1 levels respectively.

e. The sample proportions for the following five ethnic groups are in parentheses: the Khmer and Cham (0.39); the Tay, Thai, Muong, Nung (0.38); Other Northern Uplands (0.07); Central Highlands minorities (0.13); Other Minority Groups (0.03).

Appendix

Table A1: Variables and Summary Statistics

Variables	Variable Description	Majority	Minority
Hourly real wage rate	Hourly real wage rate including all payment in cash and kind, given in Vietnamese Dong (VND) thousand (Jan 2002 price)	4.1923 (2.958)	3.7587 (2.446)
Male worker	= 1 if male worker, = 0 if female	0.6244	0.5702
Married	= 1 if married, = 0 otherwise	0.6470	0.6736
Aged from 18 to 25	= 1 if aged less than 25, = 0 otherwise	0.2785	0.2769
Aged from 26 to 35	= 1 if aged from 26 to 35, = 0 otherwise	0.3063	0.2707
Aged from 36 to 45	= 1 if aged from 36 to 45, = 0 otherwise	0.2733	0.2965
Aged from 46 to 55	= 1 if aged from 46 to 55, = 0 otherwise	0.1271	0.1426
Aged from 56 to 60	= 1 if aged from 56 and over, = 0 otherwise	0.0149	0.0134
Illiteracy	= 1 if no schooling, = 0 otherwise	0.1364	0.2789
Primary education	= 1 if having primary education, = 0 otherwise	0.2179	0.1932
Lower secondary education	= 1 if having lower secondary education, = 0 otherwise	0.2440	0.1601
Upper secondary education	= 1 if having upper secondary education, = 0 otherwise	0.1370	0.1043
Vocational/Tech. training	= 1 if having vocational/technical training, = 0 otherwise	0.1339	0.1736
College/University	= 1 if having college/university education, = 0 otherwise	0.1310	0.0899
Government agencies	= 1 if working for government agencies, = 0 otherwise	0.1058	0.1973
State-owned enterprises	= 1 if working for state-owned enterprises, = 0 otherwise	0.0644	0.0341
Formal private enterprises	= 1 if working for formal private enterprises, = 0 otherwise	0.0557	0.0279
Household enterprises	= 1 if working for other household enterprises	0.7741	0.7407
Medical treatment	= 1 if treated in hospital recently, = 0 otherwise	0.0470	0.0589
Urban	= 1 if residing in urban areas, = 0 otherwise	0.4359	0.4318
Northern Uplands	=1 if Northern Uplands, = 0 otherwise	0.0992	0.4628
Red River Delta	= 1 if Red River Delta, = 0 otherwise	0.2475	0.0238
North Central Coast	= 1 if North Central Coast, = 0 otherwise	0.0857	0.0176
South Central Coast	= 1 if South Central Coast, = 0 otherwise	0.1535	0.0196
Central Highlands	= 1 if Central Highlands, = 0 otherwise	0.0811	0.1250
Southeast	= 1 if Southeast, = 0 otherwise	0.1222	0.0558
Mekong River Delta	=1 if Mekong River Delta, = 0 otherwise	0.2108	0.2955
Interviewed in the 1st quarter	= 1 if being interviewed in quarter 1, = 0 otherwise	0.2404	0.2851
Interviewed in the 2nd quarter	= 1 if being interviewed in quarter 2, = 0 otherwise	0.2698	0.2459
Interviewed in the 3rd quarter	= 1 if being interviewed in quarter 3, = 0 otherwise	0.2544	0.2531
Interviewed in the 4th quarter	= 1 if being interviewed in quarter 4, = 0 otherwise	0.2354	0.2159
Number of observations		15202	968

Source: calculations from the VHLSS 2002

Notes to table A1:

a. The standard deviations for continuous variables only are reported in parentheses.

Table A2: Kinh (and Chinese) Majority Wage Equation

	Mean	10th	25th	50th	75th	90th
Male worker	0.1087***	0.1485***	0.123***	0.1024***	0.1093***	0.0908***
	(0.01)	(0.02)	(0.013)	(0.011)	(0.012)	(0.018)
Married	-0.0737***	-0.0841***	-0.0655***	-0.082***	-0.0787***	-0.0921***
	(0.011)	(0.025)	(0.016)	(0.012)	(0.015)	(0.021)
Aged from 26 to 35	0.1129***	0.1603***	0.1162***	0.0981***	0.0821***	0.086***
	(0.012)	(0.025)	(0.013)	(0.013)	(0.017)	(0.021)
Aged from 36 to 45	0.1453***	0.1564***	0.1372***	0.1354***	0.1417***	0.1456***
	(0.014)	(0.026)	(0.016)	(0.015)	(0.018)	(0.03)
Aged from 46 to 55	0.1505***	0.0671**	0.105***	0.1807***	0.1904***	0.2088***
	(0.018)	(0.032)	(0.023)	(0.025)	(0.022)	(0.033)
Aged from 56 to 60	0.0814*	0.0925	0.0238	0.1426***	0.1942***	0.2297***
	(0.044)	(0.124)	(0.056)	(0.049)	(0.057)	(0.064)
Primary education	0.1033***	0.1391***	0.0828***	0.1024***	0.1089***	0.1259***
	(0.015)	(0.027)	(0.015)	(0.017)	(0.017)	(0.029)
Lower secondary education	0.0997***	0.1761***	0.1174***	0.1102***	0.083***	0.0652**
	(0.016)	(0.024)	(0.017)	(0.017)	(0.015)	(0.034)
Upper secondary education	0.1964***	0.2268***	0.1597***	0.1659***	0.1921***	0.297***
	(0.023)	(0.026)	(0.021)	(0.029)	(0.03)	(0.052)
Vocational/Tech. training	0.4551***	0.5453***	0.446***	0.4117***	0.4396***	0.4177***
	(0.026)	(0.047)	(0.03)	(0.04)	(0.032)	(0.052)
College/University	0.6746***	0.7724***	0.7288***	0.6404***	0.6084***	0.5559***
	(0.029)	(0.06)	(0.044)	(0.03)	(0.041)	(0.051)
Urban	0.1208***	0.1024**	0.1349***	0.1163***	0.0675**	0.0701**
	(0.027)	(0.047)	(0.037)	(0.032)	(0.033)	(0.035)
Urban*Primary education	-0.0356	-0.0215	-0.0084	-0.0338	-0.0599*	-0.0336
	(0.032)	(0.058)	(0.04)	(0.034)	(0.036)	(0.056)
Urban*Lower secondary	0.0372	0.0608	0.0206	0.0242	0.053	0.0801
	(0.032)	(0.058)	(0.042)	(0.041)	(0.035)	(0.059)
Urban*Upper secondary	0.1317***	0.0676**	0.0996**	0.1186**	0.1459***	0.0973
	(0.037)	(0.035)	(0.046)	(0.048)	(0.047)	(0.081)
Urban*Vocational or Tech.	0.0833**	0.0459	0.0674*	0.0727**	0.0137	0.0643*
	(0.037)	(0.062)	(0.039)	(0.038)	(0.044)	(0.037)
Urban*College or University	0.1168***	0.0836**	0.0916**	0.0977**	0.1584***	0.2586***
	(0.037)	(0.042)	(0.046)	(0.045)	(0.051)	(0.067)
Government agencies	0.0429**	0.0353	0.0974***	0.075***	0.0925***	0.0875***
	(0.02)	(0.036)	(0.021)	(0.021)	(0.028)	(0.03)
State-owned enterprises	0.1576***	0.1166***	0.1748***	0.1793***	0.1675***	0.1633***
	(0.018)	(0.031)	(0.021)	(0.02)	(0.02)	(0.032)
Formal private enterprises	0.2015***	0.2375***	0.2282***	0.2155***	0.1723***	0.1204***
	(0.019)	(0.032)	(0.026)	(0.02)	(0.022)	(0.031)
Medical treatment	-0.034*	0.0031	-0.0182	-0.0304	-0.044**	-0.0581
	(0.021)	(0.037)	(0.023)	(0.021)	(0.019)	(0.049)
Northern Uplands	-0.1716***	-0.2099***	-0.211***	-0.1899***	-0.1513***	-0.1309***
	(0.018)	(0.027)	(0.024)	(0.02)	(0.022)	(0.034)
Red River Delta	-0.1803***	-0.2355***	-0.1917***	-0.1924***	-0.1684***	-0.1519***
	(0.015)	(0.028)	(0.018)	(0.018)	(0.017)	(0.023)
North Central Coast	-0.1935***	-0.2626***	-0.2251***	-0.2201***	-0.1713***	-0.1247***
	(0.019)	(0.042)	(0.023)	(0.024)	(0.024)	(0.033)
South Central Coast	-0.0404***	-0.0329	-0.0409***	-0.058***	-0.057***	-0.0674***
	(0.015)	(0.027)	(0.016)	(0.017)	(0.019)	(0.025)
Central Highlands	0.1783***	0.0745**	0.0968***	0.1634***	0.2419***	0.3169***
	(0.02)	(0.034)	(0.024)	(0.027)	(0.026)	(0.036)
Mekong River Delta	-0.0015	-0.0687***	-0.0283	-0.0148	0.0262	0.0605**
	(0.014)	(0.025)	(0.018)	(0.017)	(0.017)	(0.024)
Selection bias correction term	0.0098	0.0716***	0.052**	-0.0482***	-0.0433**	-0.0275
	(0.015)	(0.027)	(0.021)	(0.018)	(0.021)	(0.026)
Constant	0.6902***	0.0128	0.3927***	0.7203***	1.053***	1.2699***
	(0.026)	(0.055)	(0.036)	(0.033)	(0.035)	(0.045)
R^2/Psuedo R^2	0.2967	0.1354	0.1507	0.1913	0.2085	0.1989
Number of observations	15,202	15,202	15,202	15,202	15,202	15,202

Notes to table A2:

a. ***, **, and * refers to the variables of which the estimated coefficients are statistically significant at level of 0.01; 0.05; and 0.1 respectively.

b. Standard errors are in parentheses. The OLS standard errors are based on Huber (1967) and the quantile regression model estimates are based on bootstrapping with 200 replications.

c. The specifications also include three controls for the quarters within which the interview was conducted.

Table A3: Ethnic Minority Wage Equation

	Mean	10th	25th	50th	75th	90th
Male worker	0.1201*	0.0585	0.0254	0.102**	0.1048*	0.0211
	(0.063)	(0.065)	(0.045)	(0.041)	(0.059)	(0.039)
Married	-0.0488	-0.0611	-0.0057	-0.0813*	-0.0475	-0.0851
	(0.044)	(0.101)	(0.062)	(0.049)	(0.063)	(0.064)
Aged from 26 to 35	0.1299***	-0.0036	0.1352**	0.1622***	0.1394**	0.1678**
	(0.05)	(0.121)	(0.062)	(0.059)	(0.065)	(0.076)
Aged from 36 to 45	0.1845***	0.0061	0.1553**	0.1746***	0.2212***	0.264***
	(0.051)	(0.112)	(0.061)	(0.066)	(0.076)	(0.079)
Aged from 46 to 55	0.1974***	0.0044	0.1172	0.2077***	0.2473***	0.3266***
	(0.057)	(0.13)	(0.088)	(0.059)	(0.076)	(0.086)
Aged from 56 to 60	0.6317***	0.4338*	0.6195***	0.577***	0.5164***	0.38
	(0.109)	(0.25)	(0.134)	(0.088)	(0.177)	(0.253)
Primary education	-0.0206	0.0098	-0.0372	0.0251	-0.028	0.1151
	(0.06)	(0.155)	(0.062)	(0.051)	(0.065)	(0.083)
Lower secondary education	0.0423	0.0313	0.0044	0.1101	0.0279	0.1116
	(0.089)	(0.159)	(0.082)	(0.078)	(0.072)	(0.15)
Upper secondary education	0.2187**	0.241*	0.0266	0.2002**	0.171	0.1082
	(0.112)	(0.139)	(0.204)	(0.084)	(0.137)	(0.183)
Vocational/Tech. training	0.4365***	0.4983**	0.3741**	0.5224***	0.3596***	0.3217**
	(0.103)	(0.243)	(0.172)	(0.094)	(0.118)	(0.167)
College/University	0.6485***	0.6068**	0.8103***	0.6194***	0.3981***	0.3585**
	(0.112)	(0.284)	(0.194)	(0.086)	(0.125)	(0.182)
Urban	0.0997**	0.0996	0.1465	0.0838**	0.1615**	0.2141**
	(0.048)	(0.252)	(0.147)	(0.038)	(0.068)	(0.108)
Urban*Primary education	0.2807**	0.1701	0.3094*	0.2194*	0.4203***	0.1029
	(0.128)	(0.348)	(0.18)	(0.123)	(0.141)	(0.177)
Urban*Lower secondary	0.5398***	0.418	0.5182***	0.345**	0.5733***	0.335
	(0.138)	(0.294)	(0.166)	(0.141)	(0.185)	(0.246)
Urban*Upper secondary	0.4556***	0.3807	0.4851*	0.3427*	0.5207**	0.7405**
	(0.163)	(0.33)	(0.266)	(0.184)	(0.223)	(0.298)
Urban*Vocational or Tech.	0.2862**	0.0431	0.3127	0.0935	0.2827*	0.2156**
	(0.129)	(0.363)	(0.216)	(0.133)	(0.16)	(0.108)
Urban*College or University	0.1693	0.0954	0.0384	0.0666	0.2956*	0.3901***
	(0.135)	(0.362)	(0.217)	(0.151)	(0.164)	(0.146)
Government agencies	0.2524***	0.3911***	0.2703**	0.2468***	0.1717*	0.0099
	(0.072)	(0.141)	(0.133)	(0.076)	(0.095)	(0.105)
State-owned enterprises	0.2047**	0.2393	0.118	0.1058	0.1476	0.2803**
	(0.089)	(0.152)	(0.126)	(0.089)	(0.138)	(0.135)
Formal private enterprises	0.4255***	0.5716**	0.4239***	0.3803***	0.2083	0.3882**
	(0.086)	(0.21)	(0.137)	(0.085)	(0.143)	(0.172)
Medical treatment	-0.0039	-0.0795	-0.0364	-0.0453	-0.0617	0.0744
	(0.066)	(0.142)	(0.085)	(0.063)	(0.082)	(0.15)
Northern uplands	-0.2143**	1.0282**	-0.5535*	-0.2098**	-0.0609	-0.0711
	(0.095)	(-0.166)	(0.297)	(0.086)	(0.074)	(0.133)
Red River Delta	0.3043*	1.1223***	0.4878	0.1441	0.2264	0.4541
	(0.177)	(0.322)	(0.352)	(0.282)	(0.278)	(0.308)
North Central Coast	-0.3045*	1.0122***	-0.5132	-0.1829*	-0.0343	-0.1241
	(0.159)	(-0.329)	(0.388)	(0.11)	(0.153)	(0.373)
South Central Coast	0.1795	0.6398	0.4524	0.0198	-0.0281	-0.0166
	(0.154)	(-0.52)	(0.289)	(0.115)	(0.151)	(0.209)
Central Highlands	-0.3807***	1.0368***	0.6597**	-0.1317	0.118	0.1821
	(0.1)	(0.207)	(0.297)	(0.104)	(0.083)	(0.158)
Mekong River Delta	-0.4671***	1.3562***	-0.2079	-0.215**	-0.1395*	-0.0431
	(0.094)	(-0.179)	(0.266)	(0.087)	(0.086)	(0.146)
Selection bias correction term	-0.2371**	-0.0779	-0.2032**	-0.2619*	-0.1597***	-0.2595***
	(0.108)	(0.104)	(0.089)	(0.139)	(0.058)	(0.086)
Constant	0.1383	-1.3836***	-0.3445	0.32***	0.9666***	1.3954***
	(0.129)	(0.24)	(0.267)	(0.117)	(0.141)	(0.22)
R^2/Psuedo R^2	0.3178	0.2012	0.2155	0.2307	0.2316	0.2005
Number of observations	968	968	968	968	968	968

Notes to table A3: see notes to table A2