Discussion Paper Series

RIEB

Kobe University

DP2017-01

Determinants of Rural-urban Inequality in Vietnam: Detailed Decomposition Analyses Based on Unconditional Quantile Regressions*

Thanh BUI
Katsushi S. IMAI

Revised August 22, 2018

Research Institute for Economics and Business Administration
Kobe University
2-1 Rokkodai, Nada, Kobe 657-8501 JAPAN

Determinants of rural-urban inequality in Vietnam: Detailed decomposition analyses based on unconditional quantile regressions

Thanh Bui* and Katsushi S. Imai†€

*Global Development Institute, University of Manchester, the UK

†Department of Economics, University of Manchester, the UK

€ RIEB, Kobe University, Japan

First Draft: 5th February 2017

This Draft: 22nd August 2018

Abstract

This study examines determinants of the rural-urban gap of household welfare in Vietnam during 2008-2012 using national household data. We have used unconditional quantile regressions (UQR) to carry out quantile decomposition analyses to identify underlying causes for the rural-urban disparity across the entire distribution. Our analyses have overcome the limitations of Oaxaca-Blinder decomposition, namely, (i) decomposition is made only at mean and (ii) a dependent variable has a linear and parametric relationship with covariates. For these purposes, we have carried out detailed decomposition analyses and the UQR decomposition (Fortin et al, 2011) combined with the reweighting technique. Our results show that basic education is beneficial to the rural poor and ethnic minorities in improving their living standards. Remittances generally improve rural welfare, but do not reduce within- or between-inequality. Public policy should ensure easier access to education for the rural poor and support the self-employed to raise and stabilise income.

Key Words: Inequality, Rural-urban disparity, Quantile Regression, Decomposition, Vietnam

JEL Classification: C21, I31, O18

*Corresponding Author

Katsushi S. Imai (Dr.)
School of Social Sciences, University of Manchester, UK
Address: Arthur Lewis Building, Oxford Road, Manchester M13 9PL, UK
Email: katsushi.imai@manchester.ac.uk

Acknowledgements
We would like to thank Kuhu Dahiya, Chung Yi Lam, Hilary Pun, Brian Ramprasad, Chen Xin and participants in the seminar at the University of Manchester for their useful suggestions. Any errors that remain are our sole responsibility.

Determinants of rural-urban inequality in Vietnam: Detailed decomposition analyses based on unconditional quantile regressions

I. Introduction

Since 1987, the economy of Vietnam has experienced a rapid economic growth as well as the structural transformation from a centrally planned economy to a market-based economy. This process has involved rapid urbanisation and rural-to-urban migration. As a result, the poverty in Vietnam has decreased dramatically: poverty headcount ratio based on the international poverty line of $1.25 ($2.00) a day (2005 PPP) decreased from 43.6% (58.7%) in 1993 to 14.3% (36.9%) in 2008 (Begun, 2012). Although it has followed similar structural reform like other developing countries such as China or India, Vietnam achieved continuous growth with relatively lower inequality. However, the recent widening gap within both urban and rural and between-inequality has become a problem that needs to be addressed as Vietnam urbanises.

Studies on inequality in Vietnam has so far focused mainly on income or expenditure differentials by applying Oaxaca-Blinder decomposition (Oaxaca, 1973; Blinder 1976) at the mean (Van de Walle and Gunewardena 2001; Takahashi, 2007; Imai et al., 2011a, Imbert, 2013). However, this approach may mask the intra-group heterogeneity in terms of behavioural responses to exogenous changes within a group, which is important in characterising the inequality. For instance, within a rural area, there is a high degree of heterogeneity depending on the geographic characteristics (remoteness) or cultural factors (Cao and Akita, 2008). On the other hand, decomposition based on the quantile regression allows the marginal effects of covariates to vary across the entire distribution. However, in conditional quantile regression (CQR) the marginal effects of covariates cannot be interpreted straightforwardly and as we will discuss later, it cannot be easily applied to the Oaxaca-

Blinder type decompositions at different percentile points. We propose to overcome two limitations of Oaxaca-Blinder decomposition - (i) decomposition is made only at mean and (ii) an objective variable has a linear and parametric relationship with covariates. To do this we will use unconditional quantile regression (UQR) (Firpo, et al. 2009) to derive an estimate of the marginal effect of covariates (e.g. education) at each quantile across the entire distribution of expenditure. We will also carry out quantile decomposition analyses based on and Melly (2005) and the UQR decomposition (Fortin et al, 2011) combined with the reweighting technique (DiNardo et al. 1996), or the re-centered influence function RIF-OLS approach. This is in our view a more useful method for decompositions than the previous studies to identify underlying causes for the rural-urban disparity considering the heterogeneity among households across the entire distribution.

To our knowledge only two studies, Thu Le and Booth, (2014) and Nguyen et al. (2007) used quantile regression techniques to assess the expenditure inequality between urban and rural residents over the entire distribution. Drawing upon the UCR, Thu Le and Booth found that urban-rural inequality continued to increase over the years due to both covariate effects and returns to those covariate effects. On the other hand, using the Machado-Mata (2005) technique Nguyen et al. found that the welfare disparity between the two sectors was mainly explained by the impact of structural effects. Building upon these studies we have applied the techniques adopted by these studies and the reweighted regression approach to further overcome the second limitation of the Oaxaca-Blinder decomposition (i.e. an objective variable assumes to have a linear and parametric relationship with covariates) to carry out detailed decomposition analyses for the entire distribution. When the assumption of linearity is violated, the results from decomposition will provide biased estimates of structural and composition effects. In contrast, reweighted RIF-OLS decomposition could take advantage of the flexibility of an ordered probit model. Moreover, the reweighting process could yield

efficient estimates and guarantee robust results to RIF regression assumption (Ashenfelter and Card, 2010).

During the 1990s, data provided by General Statistic Office (GSO) revealed that the increase in inequality in Vietnam was due to a widening gap between urban and rural sectors. In contrast, during the past decade, inequality is mostly due to within-group disparity because of the disproportional increase in returns on human capital and the number of salaried workers in the households within each sector (Fritzen et al., 2005). However, urban-rural inequality studies in Vietnam have predominantly assessed the determinants of inter-sector disparity. To our knowledge, only Cao and Akita (2008) have investigated the within gap by carrying our inequality decomposition by urban and rural sectors. Using the household survey data in 2002 and 2004, they used the Theil indices and disaggregated household income into different income sources in each sector. Given the recent rapid economic growth of Vietnam (ranging between 5.2% and 7.6% in 2005-2015 in terms of annual GDP per capita growth), there is a need for detailed decomposition analyses of inequality using more recent dada. The present study will fill the gap by estimating expenditure differentials both between- and within-sectors using the three most recent Vietnam Household Living Standards Survey (VHLSS) longitudinal survey data from 2008 to 2012. These surveys cover the period during which Vietnam gained more access to foreign markets because of joining the World Trade Organization (WTO) in 2007, and in the meantime, the country faced significant challenges after the global financial crisis. More importantly, economic growth in this period witnesses growing unequal income distribution from several angles including regional, rural-urban, ethnic, gender inequality and especially the gap between top income earners and a majority people (World Bank, 2014), which indicates a problem for long term growth. Our analysis of micro-level determinants of interregional and intraregional consumption expenditure will be useful for identifying challenges for policy makers during this challenging period.

More specifically, this paper aims to (i) identify factors influencing between- and within-urban and rural inequality in Vietnam and (ii) assess the contributions of the characteristic effects (e.g. education) and the effects of returns to the characteristics across the entire distribution to identify the factors underlying the rural-urban inequality. The rest of the paper is structured as follows. Section II provides an overview of Vietnam's urban-rural economic trends and reviews the relevant literature. Section III provides the details of the VHLSS data. Section IV summarises the methodologies adopted in this study, including conditional and unconditional decomposition methods and econometric methods. Section V discusses details about our empirical analysis processes and the estimation results. Section VI concludes with policy implications.

II. Literature Review

In 1986, Vietnam's Communist Party made a decisive step to abandon the central planning model of socialism and to adopt a ''market-oriented socialist economy under state guidance'' - also known as Doi Moi (Renovation) (Beresford, 2008). By controlling the credit growth and reducing the subsidies to state-owned enterprises besides opening the economy to international trade, state and society have undergone dramatic transformations in the economy, which has been successful as measured by a sharp reduction in inflation from 170% to 5% within a decade and a dramatic growth in GDP per capita over the last three decades. While Doi Moi policies have been well perceived by a majority of the population as a major contributor to the high economic growth in the economy, this came with a cost. The price of growing inequality is creating a wedge in the living standards between urban and rural sectors (Mundle and Van Arkadie, 1997; Rama, 2008). Increased inequality observed between 1993 and 1998 has often been referred to as an 'urban-rural phenomenon'. The economic reforms, which intended to raise the standard of living in rural areas were not

necessarily successful and led to the unbalanced growth where growth rate was much higher in urban than in rural areas (Webster, 2004). This unbalanced growth creates a remarkable inequality between urban and rural areas in employment opportunities and living standards (Phan and Coxhead, 2010).

Nguyen et al. (2007) applied a quantile regression decomposition method to analyse the sources of urban-rural household welfare inequality. They stressed that the welfare disparity between the urban and rural sector was mainly due to the change in return of household characteristic. In particular, they claimed that the return to education, ethnicity and agricultural activities dramatically changed from 1993 to 1998. Amid those adjustments, the return to education improved the most, thus they further argued that the development policy had an urban bias. This was because urban households generally received better education and were more likely to benefit from the economic reform than rural households. Their results were consistent with Fesselmeyer and Le (2010) who applied the Theil Index decomposition and found that during the same period inequality within the rural and urban sectors remained stable, while the between inequality increased 61.9 %.

While a radical increase in economic inequality was observed during the 1990s, its evolution did not follow the same trend in the 2000s. The rural and urban inequality continuously increased after 1998, but peaked at 2002 and marginally decreased from 2002 to 2006 (Thu Le and Booth, 2014; Fritzen et al. 2005). Applying the decomposition for UQR, Thu Le and Booth (2014) found that, while the return to education and remittances is the most important factor in explaining the rural-urban gap, the remittances play a significant role in narrowing the urban–rural expenditure gap in between 2002 and 2006. However, it is unclear how and why the rural-urban gap has evolved in Vietnam in more recent years and the present study attempts to fill the gap.

III. Data

The present study makes use of the Vietnam Household Living Standards Survey (VHLSS) data in 2008, 2010 and 2012. The VHLSS is a national representative survey, carried out every two years by the General Statistics Office (GSO) with technical assistance from the World Bank. The master sample was a random sample based on the 1999 Population Census enumeration areas. This was based on a two-stage procedure where communes were selected in the first stage and then 3 enumeration areas were selected for each commune in the second stage.[1] Finally, within each commune, individual households were selected by a systematic sampling. By using this sample design on a large scale, VHLSS is deemed the best data source to obtain reliable estimates at national or regional levels given its geographical coverage of households and rigorous sampling methods.

However, a drawback of VHLSS is non-inclusion of urban migrants, some of whom were not included in the household registration system. For this reason, the studies based on VHLSS could underestimate the urban population. Also, if migrants to urban areas were systematically poorer than the average urban residents, the rural-urban disparity might be overestimated and within-urban inequality may be underestimated (Haughton, 2010). Since 2010, however, the renewed VHLSS round tried to include urban migrants who were living in the cities without a permanent resident permit, which helped to mitigate these errors. Hansen and Le (2013) carefully compared VHLSS data and census data in 1999 and 2009 and found that households residing in VHLSS communes were on average better off than those in other communes in both years. Despite this limitation, VHLSS can be generally seen as a high-quality source of data.

[1] In each round, about 3000 communes (corresponding to 30 per cent of all communes in Vietnam) were selected.

In three rounds of VHLSSs, although the total number of observations is over 40,000 households, the household-level observations which sustain throughout the questionnaires are approximately 9000 every year. In particular, there are 9,189, 9,402 and 9,399 household-level observations in 2008, 2010, and 2012 respectively. Taking into account the regional differences in prices and living costs, regional price indices are used to deflate household expenditures by using a spatial cost of living index (SCOLI) and temporal price deflators will also be applied to each expenditure respectively to derive the inflation-adjusted per capita real expenditure.[2]

IV. Methodology

Unconditional quantile regression

To address heterogeneity in the response of a dependent variable to covariates, numerous studies have used a conditional quantile regression (CQR). CQR examines distributional outcomes of observed attributes at different points in the conditional distribution of the dependent variable across its entire distribution, not at the mean (Koenker and Bassett Jr, 1978; Koenker and Hallock, 2001; Koenker, 2006). Quantile regression (QR) estimator is more robust to outliers in the dependent variables that often occur in the data of developing countries like Vietnam (Hampel et al., 2011). Therefore, a large measurement error only has a small impact on the coefficient estimates while it could cause a severe bias to OLS (Deaton, 1997). Especially, since the QR technique requires less strict assumptions about the distribution of the error term[3], it is preferred to OLS when the error term is not normally

[2] Conversions follows the CPI formula: $\frac{CPI_t}{CPI_{baseyear}} = \frac{Price_t}{Price_{baseyear}}$

[3] When CQR is estimated, some distributional assumptions are needed. For instance, the distribution of the error terms is continuous and differentiable. It also assumes a specific structure for the conditional quantiles, that is, the linear dependence of quantiles on parameters. In contrast, UQR is a

distributed.[4] Exploring the determinants of log per capita real expenditure on the whole distribution but not only at moments, this study will apply a method of unconditional quantile regression (hereafter UQR) introduced by Firpo et al. (2009).

Under the QR framework, conditional quantile regression (CQR) is the most commonly used method in the literature. However, when the effects of covariates vary for different quantiles or for different set of covariates, it is not straightforward to interpret the results of CQR as its estimates cannot be generalized to the entire population (Borah and Basu, 2013). That is, in CQR the estimated marginal effect at each percentile is valid only if the change in an independent variable (e.g. expenditure) does not move the observation into a different percentile. Therefore, CQR cannot be easily applied to the Oaxaca-Blinder type decomposition. Machado and Mata (2005) proposed the counterfactual simulation method, but it is not only computationally demanding but also unsuitable for detailed decomposition as it cannot provide composition effects (Fortin et al., 2011).

In contrast, UQR based on the approach of Firpo *et al*. (2009) is more generalizable and intuitive to interpret due to its straightforward implementation, especially in the presence of multiple covariates. Moreover, it is preferred to the conditional method since it provides the overall inequality as measuring directly the effects on the expenditure distribution without conditioning on the explanatory variables. This property is particularly important for policy implications when we need to draw the causal inference from observational data (Angrist and Pischke, 2009). Besides, the method is also easy to be generalised to yield other distributional statistics such as the Gini or Theil coefficients. In UQR the estimated marginal effect at each percentile expresses the marginal effect of covariate and can be used more straightforwardly

non-parametric estimator, therefore it is robust to even a misspecified set of control variables (Maclean et al., 2014). Authors thank an anonymous referee for pointing this out.

[4] The normality test (D'Agostino et al., 1990) shows that the hypothesis that log per capita household expenditure is normally distributed is rejected for all the rounds of our dataset. We have applied the same test for the residuals of OLS and found that they are not normally distributed.

for the Oaxaca-Blinder type decomposition. However, the estimate shows only a partial-equilibrium effect under the assumption that the unobserved heterogeneity is independent of the observed characteristics and no reversal causality exists (Koske et al., 2012).

The key idea of UQR is based upon the re-centered influence function (RIF) which is the sum of given distributional statistic (q_θ) and the influence function (IF) (Firpo et al. 2009, p.956).

$$RIF(Y_i; q_{\theta;}F_Y) = q_\theta + IF(Y_i; q_\theta; F_Y) \quad (1)$$

Here IF represents the influence of an individual observation Y_i (household consumption expenditure in our case) on the distributional statistic, q_θ, that is, the θ^{th} quantile of the unconditional distribution of Y_i. F_Y is the (unconditional) cumulative distribution function of Y_i. That is, RIF is the 'adjusted quantile' at θ^{th} which is equal to q_θ with the effect of Y_i on q_θ given the distribution of Y_i accounted for. More specifically, it is written as:

$$RIF(Y_i; q_{\theta;}F_Y) = q_\theta + \frac{\theta - \mathbb{I}\{Y_i \leq q_\theta\}}{f_Y(q_\theta)} \quad (1)'$$

As detailed by Firpo et al., $f_Y(q_\theta)$ is the marginal density of Y_i at the θ^{th} quantile, which could be estimated using kernel method, while $\mathbb{I}\{Y_i \leq q_\theta\}$ is an indicator function taking 1 if household expenditure is below or equal to the expenditure at the θ^{th} quantile. This is rewitten as:

$$RIF(Y_i; q_{\theta;}F_Y) = c_{1,\theta} \cdot \mathbb{I}\{Y_i > q_\theta\} + c_{2,\theta}$$

$$\text{where } c_{1,\theta} = \frac{1}{q_\theta}; c_{2,\theta} = q_\theta - c_{1,\theta} \cdot (1 - \theta)$$

Here the influence function corresponds to an observed Y (household expenditure) of a real values function, $f_Y(q_\theta)$, that is, the probability density function of the marginal distribution of Y (Firpo et al. 2009, p.958). It can be seen from this that the expectation of the

RIF equals the population quantile of the unconditional distribution.[5]

Following the law of iterated expectations, with a vector of covariate X_i, the unconditional distributional statistic q_θ, can be expressed as the conditional expectation of the re-centered influence function:

$$q_\theta = \mathbb{E}\big[RIF\big(Y_i; q_{\theta;}F_y\big)\big] = \mathbb{E}[\mathbb{E}\big[RIF\big(Y_i; q_{\theta;}F_y\big)\big|X_i\big]] \quad (2)$$

We can also integrate over the conditional mean to obtain the distributional statistics using regression methods:

$$\mathbb{E}\left[\mathbb{E}\big[RIF\big(Y_i; q_{\theta;}F_y\big)\big|X_i\big]\right] = \int \mathbb{E}\big[RIF\big(Y_i; q_{\theta;}F_y\big)\big|X_i\big]dF_X(x) \quad (3)$$

where $F_X(x)$ is the marginal distribution function of covariates, x. In the simplest case, the conditional expectation of $RIF(Y_i; q; F_y)$ can be written as the linear function of observable covariates:

$$\mathbb{E}\big[RIF\big(Y_i; q_{\theta;}F_y\big)\big|X_i\big] = X\beta_q + \varepsilon \quad (4)$$

Under the zero conditional mean assumption for error term and equations (2)-(4), we obtain:

$$q_\theta = \mathbb{E}_X(X_R)\beta_U \quad (5)$$

This expression shows that RIF provides a tool to capture the effects of explanatory variables on the distributional statistic of interest without facing the difficulty of computing the counterfactual distribution.

Following the Firpo et al.'s (2009) procedure, firstly, one can compute the sample quantile $\widehat{q_\theta}$ of marginal distribution of Y as presented in Koenker and Bassett Jr (1978) and generate the density estimate $\widehat{f_y}$ using kernel method.[6] Subsequently, an estimate of RIF could be

[5] $\mathbb{E}[RIF(Y; q_\theta; F_Y)] = \mathbb{E}[q_\theta] + \frac{\theta - \mathbb{E}[\mathbb{I}\{Y_i \le q_\theta\}]}{f_Y(q_\theta)} = q_\theta + \frac{\theta-\theta}{f_Y(q_\theta)} = q_\theta$

[6] Firpo et al.'s (2009, p.960) showed that the estimator of the θth population quantile of the marginal distribution of Y_i is $\widehat{q_\theta}$, the usual θth sample quantile, which can be represented as:

obtained by plugging each estimate back to equation (1) above. Finally, the coefficient matrix could be estimated as:

$$\widehat{\beta_\theta} = (\sum_{i=1}^{N} X_i \cdot X_1')^{-1} \sum_{i=1}^{N} X_i \widehat{RIF}(Y_i, \widehat{q_\theta}, F_Y) \quad (6)$$

which is the marginal effect of an infinitesimal location shift in the distribution of observed covariates X on the $\boldsymbol{\theta}^{th}$ unconditional quantile of Y, *ceteris paribus*.

Decomposition of differences in distribution

While the method of decomposing the Theil index into a between and within group components is useful in identifying relative importance among different sources for inequality, that method also comes with disadvantages. First, although this decomposition method can handle the simple case where the whole population is decomposed into subgroups by a single criterion (e.g. decomposing inequality by ethnic groups, as in Imai et al., 2011a), it cannot deal with more complex cases, such as multivariate decompositions with a mixture of factors (Cowell and Jenkins, 1995; Jenkins, 1995). Second, this method cannot quantify the magnitude of the effect of a certain factor on the overall inequality. Third, this method does not shed light on the causes of changes in the distribution (Fortin et al., 2011).

A regression-based approach, such as Oaxaca-Blinder (OB) decomposition will address these limitations of the Theil decomposition by groups by decomposing the contribution of the expenditure gap into differences in characteristics and differences in returns to those characteristics. The Oaxaca-Blinder ~~method~~ is easy to implement, but decomposition can be made only at the mean, rather than across the entire distribution and the dependent variable is

---

$\widehat{q_\theta} = \arg\min_q \sum_{i=1}^{N} (\theta - \mathbb{I}\{Y_i - q \leq 0\} \cdot (Y_i - q))$

Firpo et al.'s (p.961) estimated the density of Y_i, $\widehat{f_Y}(\cdot)$, using the kernel density estimator:

$$\widehat{f_Y}(\widehat{q_\theta}) = \frac{1}{N \cdot b} \cdot \sum_{i=1}^{N} \kappa_Y \left(\frac{Y_i - \widehat{q_\theta}}{b}\right)$$

where $\kappa_y(\cdot)$ is a kernel function and b is a positive scalar bandwidth.

assumed to have a linear and parametric relationship with covariates. The levels of inequality could be quite distinctive when the studies focused on different parts of the distribution (Melly, 2005). Among a number of earlier studies to address the limitations of the OB method (Juhn et al., 1993; DiNardo et al., 1996; Fortin and Lemieux, 1998; Donald et al., 2000; Machado and Mata, 2005; Melly, 2005[7]), we will use Melly (2005) using quantile regression decomposition. One advantage of this method compared with Juhn et al. (1993) is that it can avoid assuming the independence between residuals and independent variables. We will also adopt the RIF-OLS technique (Firpo et al., 2009) to avoid this limitation and provide detailed insights into the roles of each factor on the consumption expenditure inequality. Fundamentally, the distributions of expenditures between urban and rural sectors can be decomposed into (i) differences in the covariate distributions of households observable characteristics X in these sectors, (ii) differences in the conditional distribution of expenditures when both areas have the same characteristics (returns to observable characteristics), and (iii) differences in the distribution of unobservable characteristics (residual).

It is noted, however, RIF-regressions are only carried out locally and the linear assumption may not hold when there are significant changes in the explanatory variables (Inchauste et al., 2014). For instance, age appears to have a convex relationship with consumption expenditure, the coefficients under linear assumptions will tend to increase when the distribution of X shifts up even though the distribution of log consumption expenditure remains unchanged. Angrist and Pischke (2009) emphasise the issue of potential misspecification when the linear assumption may not hold. This issue could be avoided by

[7] Machado and Mata (2005) proposed a more advantageous approach based on the transformation of observed dependent variable into a counterfactual observation using conditional quantile regression. Machado and Mata (2005)'s decomposition is similar to that of Melly, but because of using simulation technique, it faces a computational demanding problem and time consuming when there is a large dataset.

combining reweighing technique (DiNardo et al., 1996) together with the RIF regression. Technically using reweighing to adjust the distribution of covariates in the first sample group to have the same distribution as in the second sample, one can estimate the counterfactual mean of covariates from the RIF-regression on the reweighed sample. Rather than mechanically comparing between rural and urban households, using reweighting function reweights the urban household sample in which the distribution of their characteristics is similar to that of rural households. As Firpo et al.'s (2009) discussion suggests, the aggregate decomposition for the distributional statistic is divided into the composition effect and the structural effect using simple reweighing. Here, the composition effect can be decomposed into the pure composition effect and the specification error component. Similarly, the structural effect could be written in terms of the pure structural effect and the reweighing error.

Model specification and estimation methods

In this study, the empirical model is estimated by applying UQR to three rounds of cross-sectional data based on VHLSS surveys in order to assess the changes in the unconditional distribution of expenditure as a consequence of changes in the distribution of the observed characteristics. First, in order to examine within-group inequality, we will estimate the re-centered influence functions for each sector. Two separate conditional RIF for urban and rural residents can be expressed as a linear approximation as follows:

$$\begin{aligned}\mathbb{E}[RIF(Y_{U,i}; q_\theta; F_Y|X_{U,i}] &= X_{U,i}\beta_{\theta,U} + \epsilon_U \\ \mathbb{E}[RIF(Y_{R,i}; q_\theta; F_Y|X_{R,i}] &= X_{R,i}\beta_{\theta,R} + \epsilon_R\end{aligned} \tag{7}$$

$X = (1, X_{i1}, X_{i2}, \dots, X_{ik})$ is a vector of $k+1$ observable characteristics including demographic and human capital variables, employment, industry, remittances and regional

dummies for household head i. Vectors of coefficients of UQR, $\beta_{\theta,U}$ and $\beta_{\theta,R}$, show the marginal effects of changes in explanatory variables at the θ^{th} quantile of expenditure distributions for urban and rural sectors respectively. RIF function will be calculated by the aforementioned procedure (Equations (1)-(6)). After reweighing and estimating the coefficients by RIF-regressions for each group, the OB type decomposition at the θ^{th} unconditional quantile point is expressed as:

$$\begin{aligned}\hat{\Delta}_{S,R}^{v} &= \bar{X}_U\left(\hat{\beta}_{U,\theta} - \hat{\beta}_{R,\theta}\right) + \left(\bar{X}_U - \bar{X}_R\right)\hat{\beta}_{R,\theta} \\ &= \hat{\Delta}_S^{\theta} + \hat{\Delta}_X^{\theta}\end{aligned} \tag{8}$$

where $\hat{\Delta}_X^{\theta}$ denotes the differences in characteristics (composition effect) using the estimated coefficients for rural households as the reference group and $\hat{\Delta}_S^{\theta}$ denotes the difference in returns to those characteristics (structural effect). There are three major methods to estimate the unconditional quantile partial effect, OLS regression, logit regression and nonparametric estimators. As Firpo et al. (2009) pointed out, OLS and logit estimators are close to the fully nonparametric estimator, and so this study will adopt the RIF-OLS.

V. Results

Description of the key variables

The outcome variable of interest is the logarithm of real per capita expenditure of households, which is used as a measure of economic well-being.[8] [9]

[8] The aggregate consumption expenditure consists of food expenditure (including regular and holiday food consumption), non-food expenditure (items purchased daily and annually), durable goods (e.g. T.V, motorcycles), expenditure on health and education, and the imputed rent. Lending, debt repayments, investment, costs for buying gold, gemstone for savings, expenses for building a new house and expenditures for production purposes are not included in the consumption expenditure.

[9] Following Thu Le and Booth (2014), we took the logarithm of real per capita household expenditure to facilitate the interpretation of parameter estimates so that they represent percentage changes in per capita household expenditure corresponding to one unit or one percentage change in explanatory

The choice of covariates included in the model follows the economic theory as well as the previous empirical studies (e.g. economic theory assumes that households will allocate their resources including financial assets, natural resources to generate consumption such that they could maximise their utility or well-being). Demographic characteristics include age, age squared, variables for marriage, ethnic minority, gender and household size. Education as a proxy for human capital is one of the key factors included in the model. Variables on education attainment reflect completion of primary, lower secondary, higher secondary school, and the degree level. Employment characteristics of household heads contain variables indicating self-employment (farm and non-farm), the number of working members and the number of people working.[10] Overall, there are 99 occupations classified into 21 economic industries.[11] The regional dummies are created for eight regions from 64 provinces in total.[12] Although the percentage of the population who work in the non-farm sector increased over time, there was still about 65 % involved in agricultural activities. Thus, land area is an important variable indicating their income and expenditure.

International and domestic remittances are found to increase both income and consumption expenditure so it is crucial to include these as in Thu Le and Booth (2014).

---

variables. Given the nature of the national currency, without taking the logarithm of the outcome variable, the coefficients would be too small and make it harder to interpret. On the other hand, it would be more comparable to the OLS regressions reported in the results. Moreover, as quantile regression has a property of monotonic transformation, this monotone transform would not affect the results.

[10] Including working for wage/salary and self-employment.

[11] Industry categories including agriculture and aquaculture, mining, processing, utility production and distribution, construction, trading and reparation vehicles, hotel and restaurant, transportation and communication, finance and credit, science and technology activities, asset business and consulting services, administration, governmental social insurance, education and training, health and social relief, cultural and sport activities, housework and other services, activities for foreign organizations. Only six dummies for industry having significant impact on expenditure will be included in the regressions.

[12] In 2008, Ha Tay province located in the South-West of Hanoi was merged with Hanoi municipality. Therefore, the total number of provinces changed from 64 to 63. Eight regions includes Red River Delta, North East, North West, North Central Coast, South Central Coast, Central Highlands and South East, Mekong River Delta

Over the whole sample periods, the number of households receiving remittances was relatively stable. The proportions of households receiving international transfer in urban areas increased and remained much larger than those in rural areas, from 5.2% to 7.3% in urban areas and 2.6% to 4.1% in rural areas. While domestic remittances have a much higher percentage of recipients, on average roughly at 82% and 84% for urban and rural sectors respectively, the amount of international remittances received per household is much larger. In addition, we also include interaction terms of male and industry, ethnic minority and industry in order to examine whether the gender effect is more pronounced in some industries than others, or for ethnic majorities than minorities.[13]

We report definitions of the variables in Table A.1 and their descriptive statistics of variables by regions in Table A.2. Mean differences in observable characteristics between rural and urban samples are statistically significant. On average, household heads living in urban areas are 1.56 years older than the rural counterparts, have 2 more years of schooling, have a higher rate of self-employment in business or trading and a lower proportion of people working in the agricultural sector. Urban residents are also less likely to be married, prevalently get more international and internal remittances although they are more likely to be unemployed.

Surprisingly, female-headed households seem to be better off, though it is not statistically significant. This could be due to larger transfers for female-headed households. The mean household size is ranged from 3.8 to 4.2, with two-thirds of households living in rural area. Statistically, the percentages of household heads that finished primary and secondary schools are higher in rural than in urban areas, while urban household heads' education levels at upper secondary or higher education are higher than those of rural household heads.

[13]Other interactions (e.g. male*agriculture, male*hospitality, male*administration, minority*agriculture, minority*political) have also been included.

(Figure 1 to be inserted)

Figure 1 plots the rural-urban expenditure gap across different quantiles in 2008, 2010 and 2012. The figure on the left-hand side shows that the inequality in 2010 is generally lower than that in 2008, but at the bottom and the top end percentiles, the expenditure gaps in 2010 are significantly higher than those in 2008. The gap between the two sectors became smaller across different quantiles from 2010 to 2012 in particular at the middle and higher quantiles, with a distinctive pattern of the distribution (the figure on the right-hand side). As the rural-urban expenditure gap varies across different percentiles, it is necessary to extend the traditional OB decomposition using the quantile regression decomposition.

(Table 1 to be inserted)

Furthermore, Table 1 reports the Gini index and general entropy measures of expenditure inequality, namely the Theil-L and Theil-T indices, for both urban and rural areas. The Gini index increased from 0.36 to 0.39 from 2008 to 2010 and then decreased to 0.36 in 2012. The Theil-L and Theil-T indices show similar trends. Decomposition of the Theil indices into within and between components suggest that the former (inequality within rural areas or urban areas) remained not only dominant in absolute terms (0.74 to 0.83), but also expanded over time, while the latter (the rural-urban disparity) shrank from 0.22 to 0.19 in 2008-2012.

Determinants of urban and rural households inequality

We have run Firpo et al.'s (2009) unconditional quantile regression (UQR) method at 10^{th}, 50^{th} and 90^{th} percentiles. Sampling weight will be applied for both regressions and computation of the weights in decomposition to make our results broadly nationally-representative.[14] Tables 2, 3 and 4 report the results of URQ together with standard OLS for

[14] The weight in our data set is defined as a sampling weight, which is the inverse of the probability that the observation is include, like in most of the LSMS. Therefore, in estimating OLS, 'pweight' is

both urban and rural sectors for 2008, 2010 and 2012. The parameter estimates show marginal effects of explanatory variables on household consumption expenditure.[15]

(Tables 2, 3 and 4 to be inserted)

In general, the parameter estimates show expected signs in line with earlier studies. We will report the results below selectively. The standard errors of estimated coefficients are generally smaller at the mean (OLS, Column 1) and at the median (Columns 3 and 7) than at 10^{th} or 90^{th} quantiles, implying that there is a larger degree of uncertainty about the estimates at the tails. This will not affect our estimates since analytical weights ('*aweight*' in Stata) are employed in the quantile regressions. In terms of demographic characteristics, the estimated coefficients for the household head belonging to an ethnic minority group are negative and statistically highly significant. This shows that ethnic minority groups have generally significantly lower consumption expenditure than majority groups, notably in rural areas for all the three rounds - across the entire distribution, except at the end tail in 2008 and at the top tail in 2012. In urban areas, the majority-minority gap is observed only at the middle of distributions in 2008 (at the median) and in 2012 (at the mean). The results are consistent with Imai et al.'s (2011a) finding that ethnic minority households in Vietnam are commonly poorer and have lower living standards than the majority. Glewwe et al. (2002) also showed that the ethnic minority group found it hard to escape from poverty and continue to fall

employed. In estimating CQR, we use 'fweight' (after Stata command qreg2), or frequency weights, the weights which indicate the number of duplicated observations as this is only a possible option for CQR. On the other hand, in estimating UQR, we apply 'aweight', or analytical weights, the weights which denote the proportional to the inverse of the observation's variance (Stata command rifreg does not allow 'pweight'), given that UQR is primarily based on OLS.

[15] Online Appendix compares the estimates from UQR and CQR at the 10th, 50th and 90th percentiles of the consumption expenditure distribution in both rural and urban areas in 2008, 2010 and 2012. While the pattern of the results is mostly similar (e.g. the impact of having a degree is consistently higher at the lower point percentiles in rural area in both CQR and UQR), the heterogeneity across different quantiles of CQR is less dispersal than UQR. It is noted that the advantage of UQR over CQR is that the parameter estimates show the marginal effects and the results do not depend on a particular set of covariates, which is more useful for deriving policy implications for a sub-group of the population.

behind the majority. As the size of a household gets larger, per capita expenditure tends to decrease across different quantiles, possibly due to the decreased relative share of food expenditure for the larger household (Deaton, 1997). A household with a married household head, or an older head, spends more only in rural areas. Marginal effects of age are higher at the 10^{th} quantile than at the top end of the distribution in rural areas, implying that the average returns to experience are greater for poorer households.

Having health insurance does not influence consumption expenditure of urban residents significantly, apart from rural households at the lowest quantile of the distribution. It is conjectured that the access to health insurance or free treatment will reduce the burden of health expenditure of poor and spend more on improving the living standard.[16] Given the impacts of insurance and remittances on household welfare, there may arise the problems of endogeneity due to the causal effects (i.e. households with higher income are often more able to buy the health insurance as well as more mobility to migrate to other places such that they can obtain the remittances). In the previous literature in Vietnam, both these covariates were not considered to have potential endogeneity problem (Nguyen et al., 2009). In Mitra et al.'s (2015) study about the impact of health shocks in Vietnam, it has tried a number possible instruments and concluded that the results between before and after using IV were similar, which showed no evidence of endogeneity of health insurance in the welfare model. Likewise, Pham (2008) found that the current international remittance situation in Vietnam is unlikely related to the changes in household welfare.[17]

[16] According to GSO's calculation, the high proportion of health care visits are found among the poorest.

[17] We have included health insurance and remittances as determinants of household expenditure in UQR because the earlier poverty studies on Vietnam have pointed to the importance of these variables in changing household welfare and vulnerability (e.g. Imai et al., 2011b). As we have noted, we cannot completely deny the possibility that both health insurance and remittances are endogenous variables if individuals with unobserved characteristics self-select into their insurance programme. When these variables were removed, we have found that the results of UQR and decomposition are

In Vietnam, remittances play an important role in improving household welfare as it not only serves as a risk-coping device but also an income source for investment in agriculture or in non-agricultural business (e.g. Niimi and Reilly, 2011). Both domestic and international remittances significantly increase the household expenditure of rural households over the years, with the effect larger and more significant as they become more affluent. In urban areas, domestic remittances are significant across all the quantiles and foreign remittances are not significant in 2008. In 2010 foreign remittances are positive and significant for the rural poor only, while the domestic remittances are positive and significant at the mean. In 2012 only domestic remittances are positive and significant in increasing the expenditure of the rural households at the median and the top quantile.

On education, educational attainment is found to be positively correlated with consumption expenditure across different points in the distribution for both rural and urban households (except in 2008 for urban areas). The effects of educational attainment increase with the household heads' educational level. Interestingly, changes in the highest level of household head's education also have a different impact on the distribution of outcome variable. The returns to education in compulsory and basic education is greater at the lower end of the distribution.[18] However, at higher education level, the impact of education at the top of the expenditure distribution is much higher than lower quantile, especially in urban areas. The results in 2010 and 2012 imply that basic education helps reduce the within-group inequality, while higher education widens this within group gap.

On employment, over a half of household heads in our sample worked in the agricultural

broadly similar. URQ with an instrumental variable (IV) is not easily implemented and this will be left for a topic for future research.

[18] In Vietnam, primary education is compulsory and free from tuition fees. Basic education consists of 12 years of schooling and divided into 3 levels: 5 years of primary, 4 years of lower secondary and 3 years of upper secondary.

sector in rural areas, and they are likely to be self-employed farmers.[19] Empirically, the effects of being self-employed in the agricultural sector are negative and statistically significant. About 21% of self-employed household heads in the sample were engaged in businesses as well as services and their expenditure was overall higher. Our estimations are consistent with earlier studies, which claimed that non-farm self-employment helps alleviate poverty as well as increase the rural household expenditure in Vietnam (Hoang et al., 2014).

In terms of spatial disparity, a number of regional dummies have found to be significant with expected signs. Urban residents in South East and Red River Delta have higher expenditure level compared with other regions across the whole distribution. It is not surprising since two biggest urban centres, Hanoi and Ho Chi Minh City with better prospects in the labour market and incomes are located in these regions. The expenditure gap between the top quantiles and lower quantiles tend to decrease throughout the years, indicating lower within-urban inequality trend. North Central Coast still remains to have relatively low living standards.[20]

Machado-Mata and Melly (2005) decomposition

CQR is applied to decompose the expenditure gap between urban and rural area into different components. Following Machado and Mata (2005) and Melly (2005), the welfare disparity is decomposed using rural household heads as a reference group.

(Figures 4-6 to be inserted)

As shown in Figure 4, in 2008 the total difference in living standards between rural and urban areas displays a concave shape with a sharply increase at the top 20% of the

[19] The average value for dummy of agricultural self-employment is 0.57.
[20] The results are not shown in Tables, but will be available on request.

expenditure distribution. This gap was mainly due to the structural effects or difference in returns to characteristics (or covariates), especially around the median of the distribution (0.4-0.6). However, towards the top quantile of the distribution, the expenditure gap tends to be explained more by the difference in characteristics between two groups.

On the contrary, Figure 5 shows a different pattern of the expenditure gap across different quantiles in 2010. The welfare gap is lowest at the median and highest at the tails of distribution. The decomposition results show that the difference in expenditure between two groups is again mostly explained by the difference in returns to covariates (the structural effect or 'the discrimination effect' in the literature of O-B decomposition). Interestingly, the structural effects also follow a U-shape pattern parallel to the overall expenditure gap, which means that the differences at the bottom and at the top end of the distribution are primarily due to the difference in characteristics between the two groups. This phenomenon at the end of distributions could be considered as glass-ceiling and sticky floor phenomenon (e.g. Chi and Li, 2008) .[21] In practice, there exists discrimination in employment in Vietnam through a household registration system called *ho khau*[22]. Since the 1980s, although economic growth brought about more and more employment opportunities, *ho khau* still limits labour mobility of people from rural areas to come and work in big cities such as Hanoi or Ho Chi Minh City. This explains, to some extent, why rural residents were unable to take certain jobs in the urban area, which may help enhance their living standards.

In 2012, Figure 6 shows the overall expenditure gap between two groups is lower and

[21] Glass-ceiling refers to situation where one group was preventing from getting high-paid jobs, and so they cannot enjoy higher living standards at the top of the distribution. In contrast, sticky floor refers to an opposite scenario, where the poorest household head belonging to one group makes it harder for them to get out of their situation to have better living standards.

[22] Household registration (ho khau) system was first implemented in urban Vietnam during the 1950s and then extended over the country in the 1960s. This is a compulsory documents required in any basic administrative formalities involving marriage, birth, employment. Its main purpose was to control the population movement in order to avoid population concentration in particular regions (Labbe, 2014).

remains at less than 0.24 log point. In contrast to 2008 and 2010, effects of characteristics contribute mainly explain the rural-urban expenditure gap at 10^{th} and 25^{th} quantiles, while at the higher quantile of the distribution, the difference in returns to covariates mainly contributes to the urban-rural gap.

Empirical results from reweighed RIF-OLS decomposition

In order to carry out detailed decomposition for each covariate, we have applied the RIF decomposition. The reweighting error for decomposition in 2012 is considerably high, thus, using reweighting procedure may not provide better results as non-reweighted RIF decomposition. On the other hand, the reweighting errors in our decomposition are insignificant and relatively small compared to the specification errors for all the quantiles for the period 2008-2010, which indicates that our reweighting factor is consistent and the assumption of linearity might be violated. Therefore, combining both the RIF regression and the semi-parametric reweighting method developed by DiNardo et al. (1996) is more suitable for the detailed decomposition analysis than either of these two methods to identify the underlying factors of the urban-rural gap in Vietnam. By using reweighting approach to obtain the counterfactual before breaking down the total effects into the contribution of each covariate, it is possible to avoid inaccurately specifying the welfare model. The results of both reweighting and non-reweighting model are reported in Tables 5, 6 and 7 and as expected, it shows a very different story. Here the covariates are grouped into the main categories including demographic, education (i.e. low and high), industry, employment, remittances, interaction terms and regions.[23] [24] These tables display twofold decomposition

[23] The differences of composition effects between two models are equivalent to the reweighing error, which is due to the difference between the estimated counterfactual distribution and the real one. The reweighting error should converge to zero with the appropriate reweighting. On the other hand, the differences of structural effects between two models show the specific error. Empirically, there will

results of the RIF-regression at the main 10^{th}, 50^{th} and 90^{th} percentiles for the rural-urban expenditure gap in 2008, 2010 and 2012 with the rural household heads as a base group. Overall, the difference between urban and rural sectors is significantly positive and it is increasing across quantiles. We will selectively discuss key variables that significantly affect the sectoral gap in turns.

The RIF decomposition shows that the importance of composition effects have increased over the period. In 2008, the inequality between two sectors is mainly driven by the differences in returns to those characteristics (the structural effect) across the whole distribution and it is highest at the 10^{th} percentile (explaining 82.67 % of the total welfare gap, with the statistical significance at the 1% level) (the first panel of Table 5). This is in sharp contrast with Thu Le and Booth (2014) who found that in 2006 the rural-urban inequality was shared approximately the same proportion between both composition effects and structural effects.

However, this pattern started to change in 2010. On one side, at the low quantile, the composition effects dominate the structural effects to explain about 99.4 per cent of the difference in living standards in both areas although its influence on the sectoral gap has decreased along the welfare distribution (Table 6). On the other side, at the middle and top quantiles, the change in the returns to covariates (the structure effect) still plays important parts in explaining the difference in welfare between two areas and is larger in magnitude. In other words, the discrimination to rural households became more severe in middle and high quantiles, while it did not affect the poorest very much. This shows that even rural households hold the same qualifications; they may not be able to find jobs, which have equal

---

be no specification error and reweighting error under the reweighing model if the assumption of linearity is accurately specified and the reweighting factor is consistently estimated.

[24] In our model, the specification error is far from zero for most quantiles, which shows that the RIF-regression model alone does not give good estimates of the composition effects. This will justify our use of the reweighting approach which provides more consistent estimates.

pay as those living in urban area. In 2012, the dominance of composition and structural effects remains significant in the low and higher quantile respectively (Table 7). However, instead of the structural effect, the composition effect regains its dominance (61.2%) in the middle quantile. Overall, since 2010 the discrimination effects were gradually losing its importance along the distribution and were observed mainly at the higher quantiles in 2012. This could be considered a positive signal to the labour market, especially to the poor and middle-income households. The decomposition results imply that the returns to the characteristics have been equalising between the poor and the rich.

The second and the third panels of Tables 5-7 present the detailed results of further decompositions of composition and structural effects by household characteristics. In 2008, the basic education qualifications of household head is considered to contribute significantly to the composition effects of the welfare gap between urban and rural areas (Table 5). The positive and significant coefficients for 'Basic education' are increasing across the whole distribution, which imply that the difference in basic educational achievement of rural and urban household heads causes more welfare gap for the rich more than the poor. For instance, at the 90^{th} percentile 11.23% of the gap is explained by the difference in the attainment of basic education. On the other hand, qualifications (holding a degree, for 'High education') show an opposite trend of its impact on the welfare gap and it is only statistically significant at 10^{th} percentile (explaining 6.1% of the gap). These results imply that composition effect attributed to both higher and basic education explains a significant portion of the total composition effect in explaining the rural-urban disparity in 2008. Meanwhile, there is no significant difference in return to education across different quantiles (the third panel of Table 5).

In 2010, the composition effect associated with both basic education and high education displays an opposite trend and basic education's influence on the inequality in living

standards is no longer significant at the top quantile (the second panel of Table 6). These trends remain in 2012 and its magnitudes are larger in absolute values. The negative coefficients of basic education suggest that rural residents are overqualified in comparison with urban counterparts at the same percentiles. In contrast, higher education explains a major part of difference in characteristics between two groups although it is monotonically decreasing along quantiles. In 2010, the impacts of qualification attainment contributes 14.7% on average to the gap, ranging from 18.69% at the 10^{th} percentile to 8.14% at the top quantile while in 2012, the coefficients associated with high education are almost doubled, ranging from 32.83% to 16.4%.

The increasing trend of this impact throughout the period also suggests that providing more opportunities for people in the rural area to go to high school or enter university should be the key to close the sectoral disparity. It is concluded that the composition effects associated with different educational attainments in both basic and higher education explained much of the rural-urban disparity in the period. It is also worth noting that although the difference in returns to education (the structural effect) did not significantly explain the rural-urban disparity, in 2012 the higher education was significant at the bottom and top quantiles in 2012 (Table 7). Intuitively, the opposite signs of the composition effect associated with low and higher education suggest that basic education significantly reduces the urban-rural expenditure inequality, and its influence has become stronger for the poorest group in recent year (changing from -8% in 2010 to -22.16% in 2012, all significant at 1% level). In contrast, higher education significantly contributes to the urban and rural inequality. Although the magnitudes of higher education are decreasing along the distribution, its influence has increased at all main quantiles in the distribution recently. Statistically, its impact on the urban-rural gap of people at the top quantile (90^{th}) increased by half with 16.4% of contribution rate to the composition effects. On the other hand, the statistically

non-significant results in structural effect imply that the education attainment generally does not affect discrimination before 2012. However, in 2012, the effect of higher education reduced the gap at the 90^{th} quantile, but increased the gap at the 10^{th} quantile. Its structural effect at the 10th quantile contributes 185.6% to the welfare gap while it turns significantly negative at the 90^{th} quantile. This result indicates that discrimination is the most severe at the bottom quantiles and returns to high education is lower for the rural poor given their qualifications while at the top quantile, rural rich seems to get better returns for graduating high school or holding a degree. Our results suggest that to poor households, it is necessary but not sufficient to equalise the living standards between two sectors by merely supporting them to enter the university. Policymakers should also focus on creating more jobs in rural area that have good pays as much as in urban areas. If the rural poor could obtain the jobs, which receive the same rate as urban ones giving them having the same qualifications, the urban-rural gap would reduce significantly.

In the meantime, the results show that the composition effect attributed to being self-employed (*'Employment'*) significantly explains the rural-urban gap across the entire distribution in 2008-2012. This result accounts for much higher share of agricultural self-employment in rural areas than their urban counterparts during the period. The structural effects associated with being self-employed are statistically insignificant over the years except for a few cases (the top quantile in 2012). The composition effect is strongly significant in median and top quantiles, while only weakly significant in low quantile. This could be due to the heterogeneity in the type of self-employment at different points in the expenditure distribution.

Interestingly, in contrast to the conventional argument, remittances do not consistently appear to be a factor explaining the rural-urban gap in terms of both composition and structural effects across different quantiles. This result is consistent with the finding about the

behaviour of household's consumption expenditure with higher remittances. In 2008 the structural effect attributed to remittances is significant at the median and the top quantile, while the composition effect is significant only at the top quantile (Table 5). In 2010 and 2012, the former is significant only at the top quantile, while the latter is significant at the median at the top quantile (Tables 6 and 7). While remittances positively contribute to the inequality in both composition and structural parts (Tables 5-7), they are also found to improve overall rural living standard (Tables 2-4). In other words, remittances benefit the rich more than the poor. Although a few works suggest that remittances help smooth out the income distribution within rural areas (WouTerSe, 2010), the effects of remittances on the between-inequality are more complex. Earlier, Adams (1991) and Adams et al. (2008) found that the remittances increase household welfare, while worsening the inequality. Our decomposition results show that the difference in received amount of remittances between rural and urban households could explain the main results of Adams and Adams et al.

VI. Conclusion

Vietnam has experienced high economic growth in conjunction with rapid urbanisation, which has improved household living standards, in particular, the poor living in rural areas. During the period 1993-2012, average income of the poorest 40 per cent of the population achieved one of the largest improvements with an annual increase rate of 9 per cent (World Bank, 2014). Although the economy-wide inequality level is relatively modest compared to other emerging economies like India or China, many concerns were raised about inequality regarding to gaps between those at the top of the income distribution and the majority of Vietnamese people. In particular, the problem of growing inequality has emerged in the form of widening the gap between and within urban and rural residents. We find that the rural-urban gap is increasing in 2010 then decreasing afterwards.

Firstly, the results from unconditional quantile regressions (UQR) on three most recent rounds of VHLSS data have revealed that household heads' socio-demographic characteristics have significant impacts on living standards and vary across the expenditure distribution. Second, the estimates from the two quantile decomposition techniques developed by Machado and Mata (2005) and Firpo et al. (2009) show generally consistent results and provide a number of policy recommendations. The covariate effects, that is, the effects associated with the difference in characteristics, were increasing across the welfare distribution, while the impacts of returns to those covariates were overall decreasing. This is both an opportunity and a challenge for the government to support the disadvantaged people in rural areas, as they need to obtain the skills and knowledge required to catch up with urban residents. This reinforces the policy to promote both basic and higher education in rural areas. Moreover, since within rural inequality is the main contributor to the overall inequality in Vietnam, a policy supporting disadvantaged people, an ethnic minority in less developed areas is necessary to reduce the inequality of opportunity in the context of increasingly fast economic integration in Vietnam. In particular, a high proportion of households with an uneducated head in rural sector suggest that the government should increase the public expenditure to provide more support in such a way that poor households can access basic education.

Notably, the effect of primary and secondary education on expenditure has become more positive in all quantiles in the rural sector in recent years. This shows that income/expenditure inequality in Vietnam results from inequality in opportunity to improve human capital through education. Therefore, the policy and programs aimed at facilitating access to basic education will narrow the gap within the rural sector as well as between urban and rural areas in the future. On the other hand, higher education is found to widen the gaps both between and within sectors. This is expected since there are relatively fewer jobs

requiring higher education in the rural areas than in urban areas. The rural poor with higher qualifications are still more disadvantaged in attaining employment than their urban counterparts who can more easily find jobs. In addition, the institutional restriction like household book has also imposed a borderline between two areas, which makes it more difficult for the rural-urban labour transfer. Affluent households in rural areas may not encounter this problem as, with a better endowment, they could migrate to urban areas and improve their chance of getting a higher paid job. It would also be easier for them to get involved in self-employed activities. This indicates the lack of social mobility among the rural poor since they do not get the same social insurance as urban residents (e.g. equal rights in the public education system). Some approach like the new practice of household registration should be extended to increase the coverage of social protection of rural-urban migrants so that a more equitable environment is guaranteed for migrants to access to the same education, health and economic resources like urban households. Although migration is considered to increase GDP and reduce inequality at the same time (Mundle and Arkadie 1997), in long term, the problem of overconcentration in urban area like housing deficiency and air pollution urges the need for rural industrialization, which could help to generate equal opportunities for rural residences to improve their income. By focusing on developing better infrastructures in rural area, the government can attract more investment in rural area and obtain a more sustainable growth. Aforementioned above, having health insurance has a significant impact on improving living standards of rural households at the bottom of the distribution while this does not influence consumption expenditure of urban counterparts.[25] This is due to the inefficiency of current policy in providing free medical-aid to rural poor while out-of-pocket health expenditure accounts for a significant proportion in their

[25] According to GSO's statistics, the high proportion of health care visits are found among the poorest.

expenditure. This result would be important for policy implication if the government wanted to reduce the burden of health expenditure for the poor and lift them out of poverty trap.

On the other hand, in line with previous studies in Vietnam, the estimates of the effectiveness of remittances show that it increases not only the overall household living standards but also within rural inequality, while its effect on the gap between two groups is rather complex. On average, urban residents receive more remittances than their rural counterparts since poor household heads in rural areas may not have enough endowments to afford the cost of migrating and remittance behaviour. In our analysis, self-employment in non-farm activities is considered a key factor to increase the welfare of a developing country like Vietnam. It is expected that the government's effort in sending workers abroad could help to improve their family situation by remittances from those exported workers. However, in the long-term, it is important to encourage investment in rural areas by supporting labour-intensive manufacturers and new firms, as suggested by World Bank (2013). As a result, highly educated poor people would be encouraged to work in the rural area instead of migrating to urban area.

In addition, self-employment activities in agriculture are found to close the gap between urban and rural sectors across the entire distribution. This is due to low productivity and efficiency of farming in Vietnam. The costs of production inputs, such as fuel and fertilisers are rising, while it is hard to find a place in the market for their products. Moreover, they also face many challenges such as having poor nutrients in soils and being vulnerable to extreme weather events. Given a very high proportion of households working in the agricultural sector, the government will need to provide more subsidies (e.g. facilitate the access to loan, microcredit programs) to farmers to fight against poverty and achieve a sustainable positive agricultural growth trajectory. Furthermore, households with self-employed heads in non-agricultural sectors are found to have relatively better living standards than those with heads

working in agricultural sectors or unemployed. Thus, diversifying non-farm activities among poor households would be a key strategy to eliminate poverty. However, the gap among self-employed workers is widening within rural areas in recent years. So it is important for the government to support the rural poor at the bottom of the distribution so that their income or expenditure growth is sustainable and matching the higher growth of other sections of the nation. It is also noted that given a large amount of government investment is spent on industrial and service sectors, its contribution to welfare improvement of poorest households are relatively minor. Rural households working in the industrial sectors or service sector do not have significant improvement in their living standards, which indicates that the government spending and economic growth does not extensively reach the poor.

References

Adams, R., Lopez-Feldman, A., Mora, J., Taylor, J., 2008. Remittances, inequality and poverty: Evidence from rural Mexico. Migration and development within and across borders: Research and policy perspectives on internal and international migration, pages 101-130.

Adams, R. H., 1991. The effects of international remittances on poverty, inequality, and development in rural Egypt, volume 86. Intl Food Policy Res Inst.

Angrist, J.D., and Pischke, J.S., 2009. *Mostly harmless econometrics: an empiricist's companion*. Princeton, Princeton University Press.

Begun, W., 2012. *Not Yet Done: Vietnam's Remarkable Progress on Poverty Reduction and the Emerging Challenges.* World Bank, Hanoi.

Blinder, A. S., 1973. Wage Discrimination: Reduced Form and Structural Estimates**.** *Journal of Human Resources* 8(4), 436–455.

Borah, B. J., Basu, A., 2013. Highlighting differences between conditional and unconditional

quantile regression approaches through an application to assess medication adherence. *Health economics 22*(9), 1052-1070

Beresford, M. (2008). Doi Moi in review: The challenges of building market socialism in Vietnam. *Journal of Contemporary Asia* 38(2), 221-243.

Cao, T.C.V., Akita, T. (2008). Urban and rural dimensions of income inequality in Vietnam, GSIR Working Papers, Economic Development & Policy Series EDP08-2, GSIR, Niigata, Japan.

Chi, W., Li, B., 2008. Glass ceiling or sticky floor? examining the gender earnings differential across the earnings distribution in urban china, 1987-2004. *Journal of Comparative Economics* 36(2), 243-263.

D'Agostino, R. B., A. J. Belanger, R. B. D'Agostino Jr., 1990. A suggestion for using powerful and informative tests of normality. *American Statistician* 44, 316-321.

Deaton, A., 1997. *The analysis of household surveys: a micro econometric approach to development policy*. World Bank, Washington D.C.

DiNardo, J., Fortin, N. M., Lemieux, T., 1996. Labour market institutions and the distribution of wages, 1973-1992: A semi parametric approach. *Econometrica* 64(5), pp. 1001-1044.

Donald, S. G., Green, D. A., Paarsch, H. J., 2000. Differences in wage distributions between Canada and the United States: An application of a flexible estimator of distribution functions in the presence of covariates. *The Review of Economic Studies* 67(4), 609-633.

Fesselmeyer, E., Le, K. T., 2010. Urban-biased policies and the increasing rural-urban expenditure gap in Vietnam in the 1990s. *Asian Economic Journal* 24(2), 161-178.

Firpo, S., Fortin, N. M., Lemieux, T., 2009. Unconditional quantile regressions. *Econometrica* 77(3), 953-973.

Fortin, N. M. and Lemieux, T., 1998. Rank regressions, wage distributions, and the gender gap. *Journal of Human Resources* 33(3), 610-643.

Fortin, N., Lemieux, T., Firpo, S., 2011. Decomposition methods in economics. In Ashenfelter, O. and Card, D. (Eds), *Handbook of Labor Economics,* Vol. 4, Part A of Handbook of Labor Economics, pages 1–102. Elsevier Science Publishers, Amsterdam.

Fritzen, S., Brassard, C., Bui, T., 2005. Vietnam inequality report 2005: Assessment and policy choices. Mekong Economics Ltd Synthesis Paper of the DFID Drivers of Inequality in Vietnam Project, Mekong Economics Ltd, Hanoi.

Glewwe, P., Gragnolati, M., and Zaman, H., 2002. Who gained from Vietnam's boom in the 1990s? *Economic Development and Cultural Change* 50(4),773-792.

Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J., Stahel, W. A., 2011. *Robust statistics: the approach based on influence functions,* Vol. 114. John Wiley & Sons., Hoboken, NJ.

Hansen, H., Le, T., 2013. The importance of being surveyed: The representativeness and impact of the Vietnam household living standard surveys, Working Paper at Centre for Applied Microeconmetrics (CAM), Available at http://www.econ.ku.dk/cam/calendar/seminars/cam01102013/VHLSS_Impacts_V7.pdf, CAM, Copenhagen.

Hao, L., and Naiman, D. Q., 2007. *Quantile Regression.* Sage Publications, London.

Haughton, J., 2010. *Urban poverty assessment in ha noi and ho chi minh city*, UNDP in Vietnam, Hanoi.

Hoang, T. X., Pham, C. S., Ulubasoglu, M. A., 2014. Non-farm activity, house-hold expenditure, and poverty reduction in rural Vietnam: 2002-2008. *World Development* 64, 554-568.

Imai, K. S., Gaiha, R., Kang, W., 2011a. Poverty, inequality and ethnic minorities in Vietnam. *International Review of Applied Economics* 25(3), 249-282.

Imai, K., Gaiha, R., Kang, W., 2011b. Poverty Dynamics and Vulnerability in Vietnam, *Applied Economics* 43(25), 3603-3618.

Imbert, C., 2013. Decomposing wage inequality: public and private sectors in Vietnam 1993-2006. *World Bank Economic Review* 27(1), 55-79.

Inchauste, G., Azevedo, J. P., Essama-Nssah, B., Olivieri, S., Van Nguyen, T., Saavedra-Chanduvi, J., Winkler, H., 2014. *Understanding Changes in Poverty.* World Bank, Washington D.C.

Juhn, C., Murphy, K. M., Pierce, B., 1993. Wage inequality and the rise in returns to skill. *Journal of Political Economy* 101(3), 410-442.

Koenker, R., 2005. *Quantile Regression.* Cambridge University Press, New York.

Koenker, R., 2006. Quantile regression. *Encyclopaedia of Environmetrics Vol.4*. Available at: DOI: 10.1002/9780470057339.vnn091.

Koenker, R., Bassett Jr, G., 1978. Regression quantiles. *Econometrica* 46(1), 33-50.

Koenker, R., Hallock, K., 2001. Quantile regression: An introduction. *Journal of Economic Perspectives* 15(4), 43-56.

Koske, I., Fournier, J.-M., Wanner, I., 2012. Less income inequality and more growth-are they compatible? Part 2. The distribution of labour income. OECD Economics Department Working Papers 925, Paris, OECD.

Labbe, D., 2014. Land Politics and Livelihoods on the Margins of Hanoi, 1920-2010. UBC Press, Vancouver.

Lanjouw, J. O., Lanjouw, P., 1995. Rural nonfarm employment: A survey. Policy Research Working Paper, Number 1463, World Bank, Wahshington D.C.

Machado, J. A., Mata, J., 2005. Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of Applied Econometrics* 20(4), 445- 465.

Maclean, J.C., Webber, D.A. and Marti, J., 2014. An application of unconditional quantile regression to cigarette taxes. *Journal of Policy Analysis and management*, *33*(1), pp.188-210.

Melly, B., 2005. Decomposition of differences in distribution using quantile regression. *Labour Economics* 12(4), 577-590.

Mitra, S., Palmer, M., Mont, D. and Groce, N. (2015). Can households cope with health shocks in Vietnam. *Health Economics*, 25(7), 888-907.

Mundle, S., Van Arkadie, B., 1997. The rural-urban transition in Vietnam: some selected issues., ADB publication, Available at https://www.adb.org/publications/rural-urban-transition-viet-nam-some-selected-issues, Manilla, ADB.

Nguyen, B. T., Albrecht, J. W., Vroman, S. B., Westbrook, M. D., 2007. A quantile regression decomposition of urban-rural inequality in Vietnam. *Journal of Development Economics* 83(2), 466-490.

Nguyen, C., Berg, M. and Lensink, R., 2009. The impact of international remittances on income, work efforts, poverty and inequality: Evidence from Vietnam. Available at https://mpra.ub.uni-muenchen.de/50313/.

Niimi, Y. and Reilly, B., 2011. Gender differences in remittance behavior: Evidence from Vietnam. *The Singapore Economic Review*, 56(2), 215-237.

Oaxaca, R., 1973. Male-Female Differentials in Urban Labor Markets. *International Economic Review* 14, 693–709.

Pham, T. L. T., 2008. Access to Credit, Remittances, and Household welfare: the case of Vietnam. ProQuest. Available at: http://gradworks.umi.com/33/24/3324437.html, University of Minnesota, Minneapolis MN.

Phan, D., Coxhead, I., 2010. Inter-provincial migration and inequality during Vietnam's transition. *Journal of Development Economics* 91(1),100-112.

Rama, M., 2008. *Moving to middle income status: Donors and Vietnam beyond 2010*. The World Bank, Hanoi, Vietnam.

Takahashi, K., 2007. Sources of regional income disparity in rural Vietnam: Oaxaca-blinder

decomposition, Discussion Papers No.095, IDE-JETRO, Tokyo.

Thu Le, H., Booth, A. L., 2014. Inequality in Vietnamese urban-rural living standards, 1993-2006. *Review of Income and Wealth*, 60(4), 862-886.

Van de Walle, D., Gunewardena, D., 2001. Sources of ethnic inequality in Vietnam. Journal of Development Economics, 65(1):177-207.

Webster, D., 004). Urbanization dynamics and policy frameworks in developing East Asia. East Asia Infrastructure Department Working Paper, 8., World Bank, Washington D.C..

World Bank, 2013. *Poverty reduction in Vietnam: Remarkable progress, emerging,* Available at http://www.worldbank.org/en/news/feature/2013/01/24/, World Bank, Washington D.C..

World Bank, 2014. *Inequality in Vietnam: A Special Focus of the Taking Stock Report July 2014 - Key Findings.* Available at http://www.worldbank.org/en/news/feature/2014/07/08/inequality-in-vietnam-a-special-focus-of-the-taking-stock-report-july-2014, World Bank, Washington D.C.

WouTerSe, F., 2010. Remittances, poverty, inequality and welfare: Evidence from the Phan, D., Coxhead, I., 2010. Inter-provincial migration and inequality during Vietnam's transition. *Journal of Development Economics* 91(1), 100-112.

Table 1. Inequality Index in Urban and Rural Area, 2008-2012

	Theil-L Index ($\theta = 0$)			Theil-T Index ($\theta = 1$)			Gini Index		
	2008	2010	2012	2008	2010	2012	2008	2010	2012
Overall	0.21	0.26	0.21	0.23	0.29	0.23	0.36	0.39	0.36
Urban	0.19	0.25	0.20	0.20	0.28	0.21	0.34	0.39	0.34
Rural	0.14	0.19	0.17	0.14	0.20	0.17	0.29	0.33	0.31
				Relative Contribution					
Within	0.74	0.79	0.83	0.74	0.80	0.83	0.44	0.45	0.48
Between	0.22	0.23	0.19	0.21	0.22	0.18	0.45	0.41	0.36

Source: VHLSS2008; VHLSS2010; VHLSS2012, own calculation

Notes: Samples are weighted using VHLSS weight.

Table 2. Unconditional quantile regressions for urban and rural sectors in 2008

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Urban				**Rural**			
	OLS	10th Quantile	50th Quantile	90thQuantile	OLS	10th Quantile	50th Quantile	90th Quantile
Ethnic Minority	-0.284	-0.261	**-0.435****	-0.312	**-0.239*****	-0.323*	**-0.258*****	**-0.210****
(Not Chinese or Kinh)	(0.153)	(0.445)	(0.153)	(0.231)	**(0.0506)**	(0.153)	**(0.0709)**	**(0.0664)**
Household Size	**-0.115*****	**-0.0841***	**-0.111*****	-0.173*	**-0.117*****	**-0.121*****	**-0.119*****	**-0.133*****
	(0.0298)	**(0.0346)**	**(0.0313)**	(0.0782)	**(0.00610)**	**(0.0128)**	**(0.00833)**	**(0.0130)**
Number of household members who are working	-0.00480	-0.0158	0.00667	0.0600	**0.0203***	0.0246	0.0165	0.0303
	(0.0313)	(0.0463)	(0.0439)	(0.0753)	**(0.00825)**	(0.0166)	(0.0119)	(0.0172)
Male dummy (1 if male, 0 otherwise)	-0.172	-0.0561	0.0552	-0.551*	0.0120	0.0760	-0.0329	0.0855
	(0.114)	(0.0996)	(0.124)	(0.252)	(0.0475)	(0.0739)	(0.0617)	(0.119)
Age	0.0108	0.0291	0.0183	0.00328	**0.0352*****	**0.0384*****	**0.0425*****	**0.0233***
	(0.0201)	(0.0297)	(0.0214)	(0.0509)	**(0.00461)**	**(0.00888)**	**(0.00613)**	**(0.0103)**
Age squared	-0.0000687	-0.000226	-0.000140	-0.000112	**-0.000290*****	**-0.000303*****	**-0.000350*****	-0.000194
	(0.000195)	(0.000285)	(0.000210)	(0.000480)	**(0.0000463)**	**(0.0000882)**	**(0.0000616)**	(0.000104)
Married (whether married)	0.0543	0.217	0.0111	-0.0251	**0.0668***	**0.117***	**0.133****	-0.0228
	(0.106)	(0.140)	(0.135)	(0.229)	**(0.0337)**	**(0.0531)**	**(0.0427)**	(0.0791)
Primary (whether head completed primary education)	-0.170	-0.150	0.0609	-0.444	**-0.205*****	-0.0116	**-0.150***	**-0.513****
	(0.192)	(0.209)	(0.220)	(0.570)	**(0.0538)**	(0.0664)	**(0.0621)**	**(0.175)**
Secondary (whether head completed secondary education)	-0.0960	0.0816	0.114	-0.515	**-0.109***	0.120	-0.0580	**-0.439***
	(0.186)	(0.187)	(0.216)	(0.553)	**(0.0537)**	(0.0647)	(0.0618)	**(0.177)**
Secondary (whether head completed high school)	0.0956	0.227	0.275	-0.0344	-0.00727	**0.182****	0.0484	-0.263
	(0.182)	(0.180)	(0.214)	(0.558)	(0.0551)	**(0.0660)**	(0.0639)	(0.181)
Bachelor (whether head obtained bachelor degree)	0.248	0.122	0.263	0.800	**0.192***	**0.218****	0.183	0.106
	(0.204)	(0.173)	(0.241)	(0.609)	**(0.0897)**	**(0.0769)**	(0.109)	(0.318)
Health insurance (whether head covered by the insurance)	-0.107	0.0278	**-0.182***	-0.294	0.0276	**0.186*****	0.000738	-0.0426
	(0.0610)	(0.0894)	**(0.0773)**	(0.153)	(0.0145)	**(0.0281)**	(0.0206)	(0.0320)
Self-employed (whether head self-employed in business)	0.146*	**0.305*****	0.135	0.122	**0.123*****	**0.0925****	**0.170*****	0.0768
	(0.0638)	**(0.0859)**	(0.0943)	(0.169)	**(0.0189)**	**(0.0291)**	**(0.0277)**	(0.0426)
Self-employed (whether head self-employed in agriculture)	**-0.155***	-0.0515	-0.126	**-0.281***	**-0.149*****	**-0.0706***	**-0.144*****	**-0.251*****
	(0.0648)	(0.0820)	(0.0858)	**(0.151)**	**(0.0244)**	**(0.0360)**	**(0.0328)**	**(0.0631)**
Log of land area (in hectare)	**0.0704*****	0.0132	0.0450	**0.163****	**0.0652*****	**0.0788*****	**0.0602*****	**0.0946*****
	(0.0200)	(0.0244)	(0.0233)	**(0.0528)**	**(0.00702)**	**(0.0127)**	**(0.00956)**	**(0.0164)**

Foreign remittances ($1*10^9$*VND)	5.34	-6.56	5.96	21.2	**4.29*** **	**1.66** **	**4.47*** **	**6.17***
	(6.60)	(8.15)	(8.97)	(20.3)	**(0.844)**	**(0.601)**	**(1.16)**	**(2.77)**
Domestic remittances ($1*10^9$*VND)	**13.7*** **	**7/48** **	**13.7*** **	**29.5** **	**10.8** **	**2.93***	**7.86** **	**25.2*** **
	(3.27)	**(2.71)**	**(3.87)**	**(9.97)**	-0.149***	**(1.29)**	**(2.97)**	**(6.65)**
Constant	9.305***	7.983***	8.891***	10.67***	8.037***	6.665***	7.840***	9.269***
	(0.540)	(0.857)	(0.626)	(1.384)	(0.139)	(0.271)	(0.185)	(0.339)
Regional Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry Dummy (occupation by industries)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Interactions (Industry Dummies x Ethnic Minority or Male Dummy)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	448	448	448	448	4069	4069	4069	4069
R^2	0.387	0.197	0.272	0.264	0.379	0.184	0.241	0.161
adj. R^2	0.334	0.127	0.208	0.200	0.373	0.176	0.234	0.154

se in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Bold figures show statistically significant cases.

Table 3. Unconditional quantile regressions for urban and rural sectors in 2010

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Urban				**Rural**			
	OLS	10th Quantile	50th Quantile	90th Quantile	OLS	10th Quantile	50th Quantile	90th Quantile
Ethnic Minority	-0.0778	-0.232	0.0146	-0.0802	**-0.258*****	**-0.445*****	**-0.256*****	-0.119
(Not Chinese or Kinh)	(0.134)	(0.303)	(0.161)	(0.270)	**(0.0406)**	**(0.112)**	**(0.0501)**	(0.0777)
Household Size	**-0.136*****	**-0.164*****	**-0.115*****	**-0.173*****	**-0.110*****	**-0.137*****	**-0.110*****	**-0.0856*****
	(0.0215)	**(0.0404)**	**(0.0247)**	**(0.0475)**	**(0.00654)**	**(0.0132)**	**(0.00847)**	**(0.0130)**
Number of household members who are working	**0.0772****	**0.100***	0.0640	**0.143***	**0.0316*****	**0.0624*****	**0.0318****	0.0194
	(0.0275)	**(0.0480)**	(0.0350)	**(0.0599)**	**(0.00867)**	**(0.0177)**	**(0.0117)**	(0.0170)
Male dummy (1 if male, 0 otherwise)	**0.123**	**0.0841**	0.0323	0.302	-0.0678	-0.00742	-0.0715	-0.191
	(0.0782)	(0.136)	(0.0920)	(0.226)	(0.0459)	(0.0571)	(0.0516)	(0.107)
Age	0.0159	0.00636	0.0123	0.00145	**0.0344*****	**0.0354*****	**0.0334*****	**0.0276*****
	(0.0117)	(0.0168)	(0.0143)	(0.0259)	**(0.00423)**	**(0.00839)**	**(0.00506)**	**(0.00804)**
Age squared	-0.000114	0.0000212	-0.0000861	0.0000472	**-0.000299*****	**-0.000313*****	**-0.000282*****	**-0.000249****
	(0.000112)	(0.000160)	(0.000140)	(0.000253)	**(0.0000433)**	**(0.0000845)**	**(0.0000515)**	(0.0000815)
Married (whether married)	-0.0144	**0.423****	-0.00862	-0.181	**0.154*****	**0.160****	**0.132****	0.134
	(0.0953)	**(0.143)**	(0.0886)	(0.224)	**(0.0374)**	**(0.0520)**	**(0.0409)**	(0.0763)
Primary (whether head completed primary education)	0.0105	0.0576	0.0594	-0.0339	**0.105*****	**0.201*****	**0.117*****	0.0256
	(0.0613)	(0.124)	(0.0772)	(0.114)	**(0.0203)**	**(0.0435)**	**(0.0252)**	(0.0371)
Secondary (whether head completed secondary education)	**0.175***	**0.286***	**0.197***	0.150	**0.186*****	**0.268*****	**0.215*****	**0.105***
	(0.0684)	**(0.125)**	**(0.0807)**	(0.136)	**(0.0209)**	**(0.0435)**	**(0.0269)**	**(0.0423)**
Secondary (whether head completed high school)	**0.391*****	0.220	**0.493*****	**0.576****	**0.304*****	**0.352*****	**0.328*****	**0.207*****
	(0.0848)	(0.138)	**(0.0866)**	**(0.199)**	**(0.0280)**	**(0.0482)**	**(0.0346)**	**(0.0614)**
Bachelor (whether head obtained bachelor degree)	**0.623*****	0.189	**0.620*****	**1.406****	**0.577*****	**0.373*****	**0.578*****	**0.785****
	(0.117)	(0.125)	**(0.107)**	**(0.444)**	**(0.0691)**	**(0.0550)**	**(0.0637)**	**(0.240)**
Health insurance (whether head covered by the insurance)	-0.0483	0.0766	-0.0742	-0.213	**0.0376***	**0.179*****	**0.0431***	**-0.0974****
	(0.0511)	(0.0858)	(0.0601)	(0.111)	**(0.0163)**	**(0.0285)**	**(0.0211)**	**(0.0345)**
Self-employed (whether head self-employed in business)	0.0931	-0.0384	**0.164***	-0.0200	**0.213*****	**0.164*****	**0.197*****	**0.281*****
	(0.0566)	(0.0998)	**(0.0670)**	(0.143)	**(0.0222)**	**(0.0284)**	**(0.0274)**	**(0.0523)**
Self-employed (whether head self-employed in agriculture)	**-0.200****	-0.185	**-0.183***	-0.356*	**-0.160*****	-0.00636	**-0.147*****	**-0.315*****
	(0.0680)	(0.100)	**(0.0747)**	(0.160)	**(0.0269)**	(0.0379)	**(0.0315)**	**(0.0651)**
Log of land area (in hectare)	**0.0427***	**0.0522***	**0.0502****	0.0557	**0.0416*****	**0.0538*****	**0.0417*****	0.0254
	(0.0166)	**(0.0228)**	**(0.0172)**	(0.0434)	**(0.00872)**	**(0.0118)**	**(0.00820)**	(0.0164)
Foreign remittances (1*10^{9}*VND)	**3.95***	**4.69****	3.19	-0.699	**4.82*****	0.845	**3.65****	**10.4*****

	(1.60)	**(1.63)**	(3.40)	(2.13)	**(0.912)**	(0.503)	**(1.13)**	**(2.12)**
Domestic remittances (1*10^9*VND)	**7.10***	**7.06***	4.46	11.4	**4.33*****	**0.989***	**2.66***	**6.92*****
	(3.00)	**(2.76)**	(3.36)	(7.10)	**(0.541)**	**(0.449)**	**(1.09)**	**(1.47)**
Constant	8.849***	7.748***	8.795***	10.50***	8.019***	6.668***	8.050***	9.599***
	(0.361)	(0.496)	(0.391)	(0.807)	(0.127)	(0.242)	(0.141)	(0.236)
Regional Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry Dummies (occupation by industries)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Interactions (Industry Dummies x Ethnic Minority or Male Dummy)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	606	606	606	606	4757	4757	4757	4757
R^2	0.322	0.219	0.244	0.197	0.367	0.236	0.232	0.120
adj. R^2	0.279	0.169	0.196	0.146	0.363	0.230	0.226	0.113

se in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Bold figures show statistically significant cases.

Table 4. Unconditional quantile regressions for urban and rural sectors in 2012

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Urban				**Rural**			
	OLS	10th Quantile	50th Quantile	90th Quantile	OLS	10th Quantile	50th Quantile	90th Quantile
Ethnic Minority	**-0.305****	-0.495	-0.162	-0.173	**-0.239*****	**-0.492*****	**-0.328*****	0.0493
(Not Chinese or Kinh)	**(0.106)**	(0.314)	(0.129)	(0.185)	**(0.0505)**	**(0.129)**	**(0.0585)**	(0.0944)
Household Size	**-0.122*****	**-0.170*****	**-0.0946*****	**-0.175*****	**-0.0981*****	**-0.0883*****	**-0.111*****	**-0.0877*****
	(0.0182)	**(0.0381)**	**(0.0233)**	**(0.0370)**	**(0.00663)**	**(0.0127)**	**(0.00909)**	**(0.0136)**
Number of household members who are working	0.0460	0.0819	0.0212	0.104*	0.0155	0.0149	**0.0274***	0.00117
	(0.0251)	(0.0508)	(0.0340)	(0.0526)	(0.00912)	(0.0189)	**(0.0128)**	(0.0183)
Male dummy (1 if male, 0 otherwise)	-0.0971	-0.0242	-0.0390	-0.153	0.0289	0.0714	0.0199	0.0890
	(0.0728)	(0.129)	(0.0959)	(0.167)	(0.0462)	(0.0755)	(0.0555)	(0.109)
Age	**0.0399*****	0.0218	**0.0462*****	0.0336	**0.0497*****	**0.0610*****	**0.0473*****	**0.0427*****
	(0.00997)	(0.0248)	**(0.0119)**	(0.0192)	**(0.00380)**	**(0.00868)**	**(0.00510)**	**(0.00749)**
Age squared	**-0.000340*****	-0.000139	**-0.000425*****	-0.000292	**-0.000455*****	**-0.000534*****	**-0.000435*****	**-0.000416*****
	(0.0000933)	(0.000244)	**(0.000114)**	(0.000170)	**(0.0000375)**	**(0.0000844)**	**(0.0000508)**	**(0.0000746)**
Married (whether married)	**0.193***	0.117	0.153	**0.401***	**0.113*****	**0.144****	**0.142*****	0.00835
	(0.0845)	(0.157)	(0.110)	**(0.163)**	**(0.0326)**	**(0.0552)**	**(0.0413)**	(0.0787)
Primary (whether head completed primary education)	0.0797	0.439**	0.0000823	0.0336	**0.104*****	**0.183*****	**0.112*****	**0.0861***
	(0.0611)	(0.163)	(0.0759)	(0.0978)	**(0.0199)**	**(0.0461)**	**(0.0261)**	**(0.0369)**
Secondary (whether head completed secondary education)	**0.225*****	**0.598*****	**0.168***	0.165	**0.171*****	**0.237*****	**0.175*****	**0.142*****
	(0.0656)	**(0.168)**	**(0.0818)**	(0.114)	**(0.0211)**	**(0.0465)**	**(0.0282)**	**(0.0410)**
Secondary (whether head completed high school)	**0.338*****	**0.642*****	**0.233****	0.242	**0.260*****	**0.323*****	**0.290*****	**0.196****
	(0.0675)	**(0.165)**	**(0.0856)**	(0.141)	**(0.0261)**	**(0.0487)**	**(0.0365)**	**(0.0603)**
Bachelor (whether head obtained bachelor degree)	**0.679*****	**0.750*****	**0.466*****	**1.247*****	**0.707*****	**0.411*****	**0.533*****	**1.257*****
	(0.122)	**(0.165)**	**(0.0995)**	**(0.374)**	**(0.0831)**	**(0.0630)**	**(0.0822)**	**(0.249)**
Health insurance (whether head covered by the insurance)	0.0618	**0.173**	**-0.0344**	0.175	**0.0375***	**0.205*****	**0.00357**	-0.0230
	(0.0459)	(0.0955)	(0.0594)	(0.0959)	(0.0162)	(0.0286)	(0.0224)	(0.0355)
Self-employed (whether head self-employed in business)	**0.174****	0.177	**0.189****	**0.251***	**0.191*****	**0.0544**	**0.236*****	**0.276*****
	(0.0532)	(0.109)	**(0.0726)**	**(0.121)**	**(0.0220)**	**(0.0329)**	**(0.0302)**	**(0.0559)**
Self-employed (whether head self-employed in agriculture)	-0.119	0.0946	**-0.194****	**-0.238***	**-0.113*****	**-0.0473**	**-0.111****	**-0.226*****
	(0.0612)	(0.133)	**(0.0723)**	**(0.119)**	**(0.0248)**	**(0.0349)**	**(0.0344)**	**(0.0635)**
Log of land area (in hectare)	0.000839	-0.00460	-0.00673	0.0000155	**0.0567*****	**0.0615*****	**0.0667*****	**0.0350***
	(0.0147)	(0.0254)	(0.0194)	(0.0324)	**(0.00671)**	**(0.0127)**	**(0.00869)**	**(0.0155)**

Foreign remittances (1*10^9*VND)	2.32	1.42	2.63	8.13	**3.33*****	1.28	**3.43*****	3.97
	(1.67)	(2.19)	(2.88)	(6.28)	**(0.799)**	(0.715)	**(1.01)**	(2.25)
Domestic remittances (1*10^9*VND)	**7.47*****	2.41	**8.80*****	**13.8***	**4.00****	**3.08*****	**4.64*****	**6.69***
	(2.23)	(3.20)	**(2.31)**	**(6.06)**	**(1.23)**	**(0.918)**	**(1.23)**	**(2.64)**
Constant	8.295***	7.249***	8.480***	9.084***	7.538***	6.040***	7.631***	8.718***
	(0.311)	(0.710)	(0.346)	(0.668)	(0.109)	(0.261)	(0.146)	(0.207)
Regional Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry Dummies (occupation by industries)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Interactions (Industry Dummies x Ethnic Minority or Male Dummy)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	634	634	634	634	4717	4717	4717	4717
R^2	0.357	0.232	0.237	0.170	0.377	0.247	0.245	0.106
adj. R^2	0.318	0.185	0.191	0.120	0.372	0.242	0.239	0.099

se in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Bold figures show statistically significant cases.

Table 5. RIF-OLS Decomposition results for urban-rural gap 2008[26]

	10th quantile		50th quantile		90th quantile	
	Not reweighed	Reweighed	No reweighed	Reweighed	No reweighed	Reweighed
Overall						
Urban Households(A)	8.581***	8.581***	9.277***	9.277***	10.19***	10.19***
Rural Households (B)	8.431***	8.431***	8.974***	8.974***	9.639***	9.639***
Difference (A)-(B)	0.150*	0.150*	0.303***	0.303***	0.554***	0.554***`
%Total explained (Composition Effect)	**175.33*** **	**35.4*** **	**34.32** **	**45.87*** **	1.29	**22.38*** **
%Total unexplained (Structural Effect)	-75.33	**82.67***	**65.68*** **	**53.47*** **	**98.73*** **	**74.73*** **
Constant (%)		-18.07		0.66		2.91
Composition Effect (%)						
Demographics	15.8	**12.87** **	-6.4	2.85	-4.39	**-12.44*** **
Basic education	12.33	0.59	5.97	**5.02***	6.46	**11.23*** **
High education	6.24	**6.1***	3.83	1.73	-1.8	-2.13
Industry	-11.6	-3.78	12.97	-3.3	6.7	-2.58
Employment	**44***	**8.2** **	8.12	**11.95*** **	**9.13** **	**11.66 *** **
Remittances	9.67	0.022	6.01	0.55	1.56	**5.07*** **
Industry*Minority	1.95	-0.61	-1.06	2.07	-2	-1.56
Industry*Male	24.87	4.45	-6.73	**8.09** **	**-11.66***	4
Regions	**72*** **	**7.53** **	**11.75***	**17*** **	-2.69	**9.1*** **
Structural Effect (%)						
Demographics	-892	-216.67	-263.37	**-417.49***	-145.126	-76.53
Basic education	-117.33	-56.6	-31.68	-16.53	27.62	28.51
High education	-6.44	5.4	-2.58	5.51	1.7	4.53
Industry	126	126	-67.33	-54.46	**-83.57*** **	**-67.15*** **
Employment	-73.33	-35.07	6.96	-15.71	-6.06	-19.13
Remittances	15	27.6	**11** **	**20.69*** **	**-9.06*** **	**-5.2***
Industry*Minority	19.4	0.54	4.42	8.02	1.63	0.5
Industry*Male	-29.8	-37.4	32.74	31.09	**50.9** **	**34.84** **
Regions	-144	-90.67	-33.66	**-62.05** **	-14.87	**-44.95*** **

[26] Demographic group refers to age, age squared, marriage status of the household head, the land area and whether or not having a health insurance. The low education group includes the household heads having finished primary school, secondary school, and high school. The high education includes the household head having obtained a bachelor degree. The industry group includes those working in agriculture; processing; trading, repairation of motor vehicles and motorbikes; hotel and restaurant; government administration; education and training. The employment group includes indicators for self-employment and self-employment in agricultural sector. The regional group refers to 7 regions listed above.

Constance	1027.33	359.33	**408.91***	**554.46****	**275.63***	219.31
Reweighting error	**0.2099**		**-0.035**		**-0.11683**	
Specification error	**0.237**		**-0.19059**		**-0.5227**	

* p < 0:05, ** p < 0:01, *** p < 0:001

Table 6 RIF-OLS Decomposition results for urban-rural gap 2010

	10th quantile		50th quantile		90th quantile	
	Not reweigh	Reweighed	No reweigh	Reweighed	No reweigh	Reweighed
Overall						
Urban Households(A)	8.543***	8.543***	9.347***	9.347***	10.30***	10.30***
Rural Households (B)	8.444***	8.444***	9.116***	9.116***	9.812***	9.812***
Difference (A)-(B)	0.0995	0.0995	0.230***	0.230***	0.488***	0.488***
%Total explained (Composition Effect)	**194.6*** **	**99.4*** **	**37.2*** **	**40.26*** **	3.89	**16.13*** **
%Total unexplained (Structural Effect)	-94.6	23.92	**63.04*** **	**61.74*** **	**96.1*** **	**76.02*** **
Constant (%)		-23.32		-1.76		7.84
Composition Effect (%)						
Demographics	-22.61	**29.04*****	0.89	**7.7***	-7.33	**6.07****
Basic education	-7.4	**-8.03*** **	-1.83	**-4.14*** **	0.2	-0.11
High education	**57.89***	**18.69*** **	**19.87*** **	**17.17*** **	**4.86***	**8.14*** **
Industry	17.09	4.49	-5.21	**6.65*** **	-4.73	2.23
Employment	23.72	**11.16***	**17.57****	**17.3*** **	**5.68***	**7.19*** **
Remmittances	-5.44	-1.03	-1.01	**-2.03***	-0.2	**-1.39***
Industry*Minority	63.82	**28.34*****	1.11	-1.5	0.95	-1.24
Industry*Male	4.95	8.13	0.39	-0.76	2.38	**-3.59****
Regions	**62.6***	**8.6***	5.44	-0.11	2.1	-1.17
Structural Effect (%)						
Demographics	-415.1	-85.63	-258.7	-245.65	-74.8	-95.49
Basic education	-48.54	11.56	-24.17	-25.78	-6.17	-2.77
High education	37.69	91.26	4.61	4.43	-0.45	**-5.55**
Industry	271.36	278.39	23.65	8.57	39.55	37.09
Employment	-57.1	-44.12	-15.78	-32.78	-10.18	-18.46
Remmittances	31.96	24.02	2.3	-2.87	-3.83	**-5.5***
Industry*Minority	-80.3	-29.65	-1.2	-4.74	-0.58	-2.54
Industry*Male	-1.38	-34.07	7.91	10.3	-24.59	-7.07
Regions	-125.63	-115.58	-49.1	**-64.78***	5.39	-2.4

Constance	293.47	-71.86	373.48	414.78	171.52	178.89
Reweighting error	**0.0951**		**-0.007**		**-0.0597**	
Specification error	**0.1179**		**-0.003**		**-0.098**	

* p < 0:05, ** p < 0:01, *** p < 0:001

Table 7 RIF-OLS Decomposition results for urban-rural gap 2012

	10th quantile		50th quantile		90th quantile	
	Not reweighted	Reweighed	No reweigh	Reweighed	No reweigh	Reweighed
Overall						
Urban Households(A)	8.652***	8.652***	9.417***	9.417***	10.33***	10.33***
Rural Households (B)	8.566***	8.566***	9.252***	9.252***	9.916***	9.916***
Difference (A)-(B)	0.0862	0.0862	0.165***	0.165***	0.419***	0.419***
% Total explained (Composition Effect)	**337.59*** **	**128.77*** **	**40.61** **	**61.21*** **	2.48	**28.88*** **
% Total unexplained (Structural Effect)	**-237.59** **	-16.59	**59.45** **	**47.03** **	**97.37*** **	**82.34*** **
Constant (%)		-12.41		-8.18		-11.36
Composition Effect						
Demographics	91.76	**22.62***	10.42	-2.99	5.27	**-5.94***
Basic education	**-51.74***	**-22.16*** **	-8	**-10.79*** **	-1.12	**-2.46***
High education	**113.1*** **	**32.83** **	**22.97*** **	**29.64*** **	**5.63***	**16.4*** **
Industry	30.74	1.06	-8.6	-6.67	3.1	-3.5
Employment	39.68	**11.83***	**16.24***	**26.6*** **	2.89	**18.42*** **
Remmittances	2.01	0.92	0.47	**1.9***	0.24	**2.63** **
Industry*Minority	67.87	**65.2*** **	-6.54	8.36	-4.39	4.25
Industry*Male	-11.1	9.5	4.27	**11.21***	-2.43	0.06
Regions	**55.1***	7.13	9.39	3.6	**-6.7***	-0.96
Structural Effect						
Demographics	**-2219.26***	-995.36	**-665.46** **	**-846.7** **	-48.93	-183.29
Basic education	142.7	150.8	-21.76	-29.1	-12.72	-10.21
High education	**87.7** **	**185.6*** **	-3.38	-14.3	-4.91	**-16.5** **
Industry	37.35	104.06	-7.39	-23.45	**-55.85***	-11.67
Employment	-3	25.99	-2.05	-26.3	30.1	**38.19***
Remmittances	38.86	48.14	3.71	6.3	0.31	**-7.37***
Industry*Minority	-42.58	-6.1	13.76	8.48	11.63	4.8

Industry*Male	222.74	93.27	19.76	14.24	21.84	5.76
Regions	-54.18	-108.24	**-123*** **	**-187.27*** **	0.66	-23.22
Constance	1552.2	486.08	**846.06** **	**1144.8*** **	155.13	**285.92***
Reweighting error	**0.18**		**-0.034**		**-0.1106**	
Specification error	**0.1907**		**-0.0205**		**-0.063**	

* p < 0:05, ** p < 0:01, *** p < 0:001

Figure 1. Expenditure disparity by quantiles Source: VHLSS20082012; Own calculations

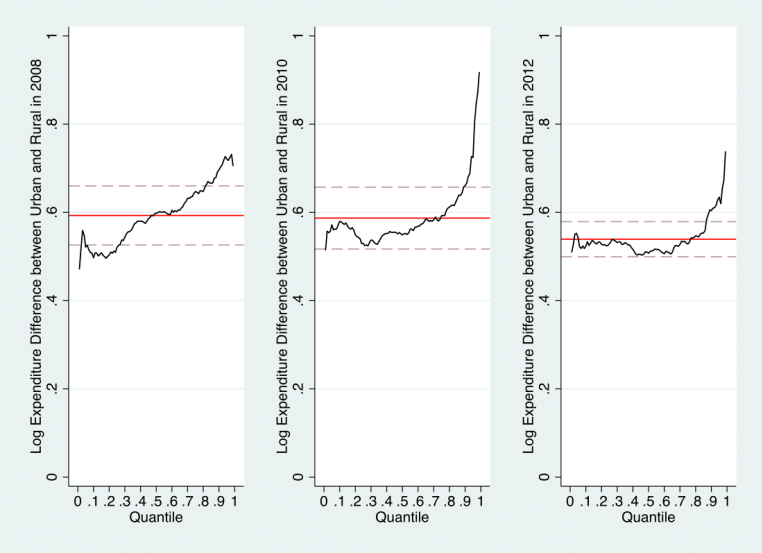


Figure 2. Expenditure disparity comparison over years

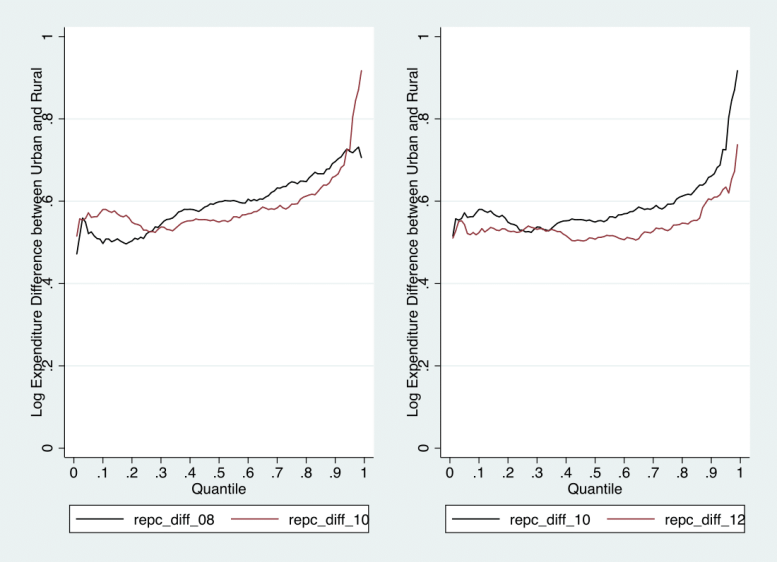


Figure 4.. Conditional quantile decomposition results of urban-rural gap in 2008(Melly, 2005)

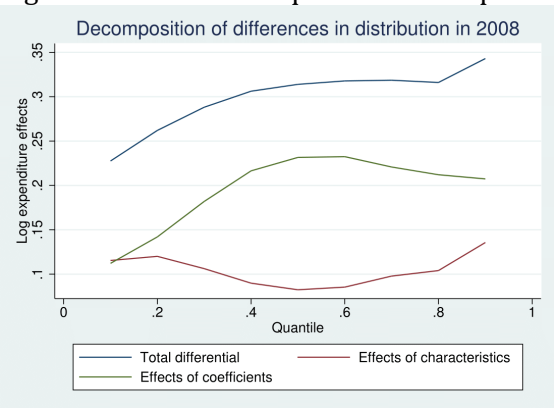


Figure 5. Conditional quantile decomposition results of urban-rural gap in 2010(Melly, 2005)

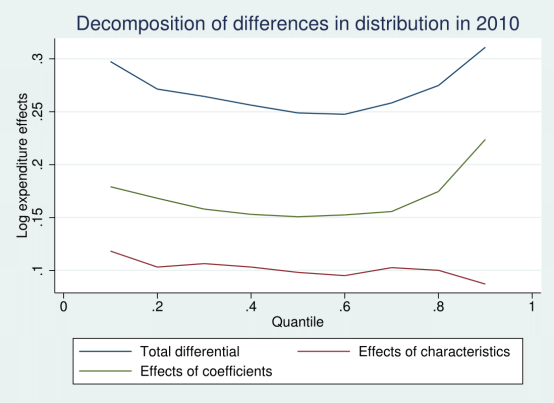
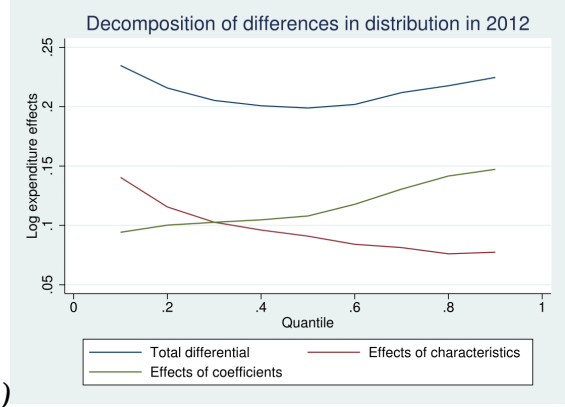


Figure 6. Conditional quantile decomposition results of urban-rural gap in 2012(Melly, 2005)



Appendix: Socioeconomic characteristics of households

	2008			2010			2012		
	Urban	Rural	difference	Urban	Rural	difference	Urban	Rural	Difference
Demographic									
Adjusted RPCE	17804.4	9593.6	8210.9***	20770.3	11154.9	9615.3***	21162.3	12261.8	8900.4***
Age	49.41	47.45	1.953***	49.31	47.31	2.003***	50.25	48.98	1.267***
Male	0.672	0.854	0.182***	0.661	0.807	0.145***	0.664	0.805	0.141***
Household Size	4.052	4.188	0.136***	3.819	3.964	0.145***	3.811	3.903	0.0923**
Minority ethnic groups (Non-Chinese/Kinh)	0.0381	0.145	0.106***	0.0556	0.183	0.127***	0.0538	0.184	0.130***
Married	0.844	0.889	0.0449***	0.800	0.851	0.0504***	0.800	0.838	0.0383***
Education Grade completed in 12	9.761	8.419	1.342***	8.910	7.152	1.758***	8.874	7.160	1.714***
Self-employed	0.332	0.211	0.121***	0.290	0.194	0.0963***	0.284	0.172	0.112***
Unemployed	0.200	0.0705	0.129***	0.207	0.0919	0.115***	0.212	0.0984	0.114***
Health Insurance Holder	0.555	0.450	0.105***	0.560	0.484	0.0761***	0.618	0.535	0.0835***
Number of HH member: salary employed	1.202	0.945	0.257***	1.178	0.990	0.188***	1.214	0.981	0.234***
Number of HH member: self-employ agriculture	0.414	1.960	1.546***	0.467	1.806	1.340***	0.463	1.785	1.323***
Number of HH member: self-employ business	0.805	0.529	0.276***	0.718	0.480	0.238***	0.682	0.438	0.245***
Number of working HH member	2.150	2.534	0.384***	2.088	2.447	0.359***	2.098	2.428	0.330***
Education									
No qualification	0.00751	0.00969	0.00218	0.127	0.237	0.109***	0.132	0.234	0.103***
Primary	0.237	0.396	0.159***	0.216	0.306	0.0895***	0.212	0.301	0.0886***
Secondary	0.311	0.431	0.120***	0.256	0.316	0.0595***	0.249	0.322	0.0737***
High school	0.288	0.141	0.146***	0.232	0.115	0.117***	0.241	0.120	0.121***
College	0.0175	0.00948	0.00804**	0.0254	0.00951	0.0159***	0.0229	0.00649	0.0164***
Bachelor	0.130	0.0115	0.118***	0.132	0.0163	0.115***	0.131	0.0156	0.115***
Master	0.00551	0.000412	0.00510***	0.00861	0.000483	0.00813***	0.00764	0.000162	0.00747***
PhD	0.00250	0	0.00250***	0.00157	0	0.00157**	0.00420	0	0.00420***
Other Education	0.00100	0.000619	0.000383	0.000783	0.000645	0.000138	0.00115	0.000324	0.000821
Region									
Red river delta	0.208	0.271	0.0626***	0.187	0.213	0.0253**	0.185	0.213	0.0282**
North east	0.138	0.160	0.0223*	0.127	0.150	0.0227**	0.125	0.154	0.0287***
North west	0.0325	0.0414	0.00889	0.0313	0.0458	0.0145**	0.0309	0.0449	0.0140**
North central	0.0751	0.148	0.0725***	0.0740	0.120	0.0464***	0.0752	0.121	0.0462***
South central coastal	0.122	0.0808	0.0409***	0.146	0.104	0.0423***	0.144	0.105	0.0385***
Central highland	0.0701	0.0584	0.0118	0.0759	0.0640	0.0119*	0.0729	0.0634	0.00954
South east	0.218	0.0860	0.132***	0.196	0.0891	0.107***	0.207	0.0848	0.122***
Mekong river delta	0.137	0.155	0.0181	0.162	0.214	0.0523***	0.161	0.214	0.0527***
Remittance									
Foreign remittance	2772.5	988.7	1783.8***	2493.6	1056.9	1436.7***	2139.9	1243.1	896.8**
Domestic remittance	4670.6	2207.1	2463.5***	4414.3	2987.5	1426.9***	5896.9	4018.3	1878.6***
* p<0.05 ** p<0.01 *** p<0.001									

Notes: Samples are weighted using VHLSS weight. The significance level of mean difference are calculated by t-tests