

IZA DP No. 7708

**What If You Had Been Less Fortunate:
The Effects of Poor Family Background
on Current Labor Market Outcomes**

Sungwook Cho
Almas Heshmati

October 2013

What If You Had Been Less Fortunate: The Effects of Poor Family Background on Current Labor Market Outcomes

Sungwook Cho
Sogang University

Almas Heshmati
*Sogang University
and IZA*

Discussion Paper No. 7708
October 2013

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

What If You Had Been Less Fortunate: The Effects of Poor Family Background on Current Labor Market Outcomes

This study examines the correlation between childhood poverty and its influence on adulthood wage distribution, where childhood poverty refers to experience of poverty or poor family background during one's childhood. With the data from Korean Labor Income Panel Study, KLIPS, quantile regression technique and decomposition method are conducted to identify and decompose the wage gap between low (poor) and middle class income group along the whole current wage distribution, based on a simulated counterfactual distribution. The results show that, those who had been less fortunate during their childhood likely had less opportunity to gain labor market favored characteristics such as a higher level of education, and even earn lower returns to their labor market characteristics in the current labor market. This leads to a discount of about fifteen percentage points off of the wage on average in total for those with underprivileged backgrounds during childhood compared to those with the middle class background, and that disadvantage is observed heterogeneously, greater at the lower quantiles than the higher quantiles of the current wage distribution. Then this research contributes to the literature by providing a partial understanding of poverty in Korea and its possible causes, in particular, in form of poor family background or childhood poverty, with which the implication of intergenerational effect issue is considered.

JEL Classification: C21, E24, J13, J31, J62, O15

Keywords: labor market outcome, family background, childhood poverty, wage inequality, wage distribution, decomposition, quantile regression, Korea

Corresponding author:

Almas Heshmati
Department of Economics, Room K526
Sogang University
35 Baekbeom-ro (Sinsu-dong)
Mapo-gu, Seoul 121-742
Korea
E-mail: heshmati@sogang.ac.kr

1. Introduction

Poverty is, fundamentally, an issue of the individual. It restricts people to live well with their vigor and deprives them of opportunities to fully develop their potential, thus limits for an economy to maximize the present welfare. Poverty is an issue of a unit as well. A single unit of a household simultaneously suffers from poverty, which in turn, transferred easily from one generation (parents) within household to its post-generation (children). In this sense, poverty is also a task of future welfare an economy will face. Most of previous research, therefore, has studied the index of poverty to accurately measure it and done policy evaluation based on the index, but little literature is found about the transferring mechanism of poverty between generations and about poor family background during childhood and its influence on earnings during adulthood. A number of questions can be asked in this regard. What if you had been less fortunate? If you had grown up in a poor family background, then how would that fact affect your current labor market outcome and would the effect be even heterogeneous across the wage distribution?

In this paper, we examine the correlation between childhood poverty and adulthood wage distribution. Here childhood poverty refers to experience of poverty and poor family background during one's childhood. Specifically we use a quantile regression technique and decomposition method developed by Machado and Mata (2005) so as to decompose the wage gap between two main groups along the whole wage distribution. The first group consists of those who grew up in poor family background, while the other is those grown up with middle class family. The rich family background is not considered in this study so as to more clearly capture the poverty effect.

Korea has been known as a relatively equal society without big concern of poverty in spite of its unprecedented rapid growth for decades, low taxation and redistribution and high concentration of assets. The recent research, however, indicates the Gini coefficients have been underestimated due to technical problems and shows that Korea is actually a high-inequality country among OECD (Kim & Kim, 2013). At this point, we pose a question; what about the other aspect of inequality, which could be attributed to intergenerational transmission earnings status? Lacking of Korean household panel data, there is little or no literature on the correlation of earnings status between generations. As filling a gap in the literature, our research would not only integrating poverty and inequality issues in Korean wage distribution but also is an attempt to provide a partial explanation to it. In addition it sheds light on the necessity to reduce poverty and inequality and at the same time to strengthen the human capital impact on economic growth and competitiveness through suitable policies.

This research contributes to the literature in a number of ways. First, it provides a partial understanding of poverty in Korea and its possible causes, in particular, in form of poor family background or childhood poverty, with which the implication of intergenerational effect issue is considered. Also, using the accepted decomposition methodology, we use a new grouping that is different than those used in traditional literature such as by male versus female, urban versus rural or private sector versus public sector. Lastly, the third contribution is the unique policy implication of our result in form of possibility of expanding educational opportunity for the lower income class of the work force.

The rest of this study is organized as follows. In the next section the literature on

the theories about labor market outcomes and its individual and family background as determinants is reviewed. In Section 3 the Korean household data is described and define poor family background. In Section 4 the Mincerian earning model is specified and estimated using traditional least squares and quantile regression methods. Common decomposition method is also used to decompose the income gaps. The estimation results are discussed in Section 5 and policy implications of the results are presented. The final section concludes.

2. Review of the Literature

In the literature there is voluminous evidence on the relationship between level of income or earnings and characteristics of individuals such as age, education and gender since Mincer (1974) provided the seminal framework. The researchers have also considered the parents education or intergenerational earnings correlation. Related with our research, there have been attempts which considered the family background during childhood and its influence on earnings during adulthood. For instance, Becker and Tomes (1979) devised a theoretical model and illustrated several crucial aspects of the intergenerational transmission of earnings status. That is, intergenerational influences on the child's future earnings during adulthood may depend on investment in child's human capital by parents and also depend on the child's endowed capacities. Becker and Tomes (1994) and Black and Devereux (2011) are among excellent starting references for theoretical model and empirical evidence on the research topic. In the sociology literature, there are also empirical studies on earnings intergenerational correlation conducted, based on Panel Study of Income Dynamics, PSID (Duncan & Yeung, 1998; Duncan, Ziol-Guest, & Kalil, 2010).

However, despite progress discussed above, little or no literature is found about, specifically, 'poor' family background during childhood and its influence on earning during adulthood. At this point, we examine poor family background during childhood and its influence on labor market endowments and reward structure which one may face, and thus eventually on earnings during adulthood. Moreover, we trace its heterogeneous effects across different quantiles of the earnings distribution. In order to investigate those intergenerational effect presented above in earnings, we use quantile regression and a decomposition method to shed lights on influence of poor family background on earnings and decompose the earning gaps between different family income classes during childhood.

Quantile regression was introduced firstly by Koenker and Bassett (1978). Buchinsky (1998a) is an excellent guide for the use quantile regression technique in earnings analysis. Quantile regression has been widely used in estimation of wage equations since this technique enables us to characterized relationship between earnings and labor market endowments in the entire wage distribution. Buchinsky (1994) presented comprehensive analysis of wage structure in USA using quantile regression method. In the literature, there is evidence that the returns to education tend to increase across quantiles (Hartog, Pereira, & Vieira, 2001). Mwabu and Schultz (1996) suggested an alternative explanation for quantile estimates of education in this case. Relating with the topic of intergenerational transmission (Eide and Showalter 1999) application to Korea data include Kang and Yun (2008) and Lee and Lee (2006).

Another toolkit in the earnings analysis is decomposition method. The decomposition method in labor economics started with the seminal papers of Oaxaca (1973) and Blinder (1973). Another major advancement in decomposition method literature is made by Juhn, Murphy, and Pierce (1993). In this paper we use the approach by Machado and Mata (2005), which is a decomposition method based on quantile regression. More specific overview on the decomposition method literature is found in Fortin, Lemieux, and Firpo (2011). There are applications of Machado and Mata decompositions method employed to decompose the wage gap between male and female (Albrecht et al. 2003, 2009), and between urban and rural areas (Nguyen and Albrecht 2007).

3. The Data

The data used in this study is obtained from Korean Labor and Income Panel Study (KLIPS). The data is freely and publicly available to registered researchers. The surveys are collected on an annual basis with national coverage. The annual samples are stratified and representative of the population. To investigate our question, basically we need to identify one's family background during one's childhood, but there is no exact data available; the panel data has only ten years from 1998 to 2008, which is not accumulated enough to track one's childhood family background. Therefore, we proxy it by using information provided in the survey. There is a survey question in KLIPS, stating "Which of the following did your family environment belong to when you were fourteen years old?" and five answers are available, 'very high', 'high', 'middle', 'low', and 'very low'. In this study we constitute low class group of those who answered 'low' or 'very low', and middle class group who answered 'middle'. In grouping the sample, we exclude those who answered 'high' and 'very high' so that we could more clearly capture the poverty effect with comparing poor group with the middle class group rather than 'the poor' versus 'the non-poor'. Based on that, because of some data issues, we limited the sample which consists of wage earners, but not including drop-outs from the education system. It included only those who answered 'graduated' to completion of education at one's final education level, excluding 'drop-outs', 'leave of absence from school' or 'attending'. For example, those who dropped out from middle school or those attending college were excluded.

The variables used in modeling income of the two groups are divided into dependent and explanatory variables. The dependent variable is logarithm of average monthly wage of individuals. The key variable here is classification of class status of household where individuals are grown up. It is represented as dummy variable, 1 for belonging to the low or less fortunate group in the childhood and 0 for the middle group counterpart based on the survey question as mentioned above. Explanatory variables other than group separator are of traditional type such as education, age, gender, marital status, regional location and occupation. The education variable is grouped into four different educational categories: below middle school graduates, middle school graduates, high school graduates and college graduates or above. Age is a continuous variable. Following the tradition, its square is used to capture non-linearity in its effect on wages. Gender is a dummy variable, 1 for male and 0 for the female counterpart. Marital status is also represented by a dummy variable where 1 if

married, while 0 unmarried including divorced and widows. Regional location is grouped into three regions, those living in Seoul, in big urban cities including Kyung-gi province and in other cities. Occupation is also a categorical variable. The first type represents professionals, technicians and associated professionals, legislators, senior officials and managers. The second type is clerks and the third category includes shop and market sales and service workers. The last occupation type includes area of production. It involves craft and related trade workers, plant and machine operators and assemblers, as well as elementary occupations.

[Insert Table 1 here]

Descriptive summary statistics of the used data are presented in Table 1. Several points are to be noted in Table 1. First, middle class group log wage is larger than low-group at specified percentiles as well as at the mean point. Second, middle class group members are likely more educated than low class. The median of educational attainment of low group is high school degree while it is college degree for the middle class. Finally, data indicate that those who belong to low class are more likely aged than middle class. It should be noted because this may reflect that each of those two groups have spent their childhood at different time periods, which could be a potential problem for the group comparative analysis. For instance, those who spent their childhood in the 70's are more likely to answer 'poor' compared with 80's because Korean economy began its high growth rate period in 80's. Possibly this may be able to mislead the result by mixing poverty effect with, say, generational effect. Therefore this sample biased problem should be addressed because the poverty effect has to be solely segregated from the generation effect.

To claim that the abovementioned poverty effect is isolated from generational effect, a few remarkable points are presented. First, the answer of the survey question is, exactly saying, a relative terms of speaking; for instance, 'very low' is actually written in the survey as 'very low compared with average'. Second, although the t-statistic of the mean difference between the groups indicates that the mean age of groups is statistically different, the 95% confidence interval for the age difference of groups is in the interval from 3.7 to 5.3 years. This relatively small estimated age difference between groups may enable us to assume that both groups spent their childhood in similar times, namely both of the two groups are of same generations in general.

4. Estimation and Testing Procedures

Our research is essentially on the wage equation, thus begins by introducing the basic theoretical framework, Mincerian earnings model to investigate the hypothesis we propose in the paper. The wage equation consists of wage as a dependent variable and explanatory variables of traditional type such as education, age, gender and in addition, we consider marital status, regional location and occupation as control variables. Our hypothesis is basically that the childhood poverty which one may have experienced during one's childhood may affect one's adulthood labor market outcome, in which two routes are assumed; one is through his or her labor market endowments (i.e. characteristic effects) such as level of education, and another is through reward to labor market endowments (i.e. compensation effects) such as returns to education. Moreover,

this poverty effect could be revealed differently across the wage distribution. Therefore, we consider two methods to investigate our hypothesis as in the following; the quantile regression to capture heterogeneous effect across the wage distribution, and a decomposition method to decompose the poverty effects.

4.1 Quantile regression method

Quantile regression was firstly introduced by Koenker and Bassett (1978), extending the notion of ordinary quantile in a location model to a more general class of linear models in which the conditional quantiles have a linear form. Buchinsky (1998) and Koenker (2005) are excellent references on quantile regression. Quantile regression has been used in a variety of research on the wage structure since Buchinsky (1998) studied changes in US wage structure during a decade. Thomas (1996) was able to reveal heterogeneous growth trends of education by ethnic groups in South Africa through quantile regression. Mwabu and Schultz (1996) gave a fresh explanation of quantile regression coefficient of education in the wage equation. The estimation procedure of quantile regression is outlined below.

The τ^{th} quantile of random variable Y with cumulative distribution function $F(y)$ is the minimum value q_τ such that:

$$(1) \quad q_\tau = F(\tau) = \inf \{y \mid F(y) > \tau\}.$$

The quantiles are simply obtained from the following relation:

$$(2) \quad \min_x E \rho_\tau(Y - \hat{y}) = \min_x (\tau - 1) \int_{-\infty}^{\hat{y}} (Y - y) dF(x) + \tau \int_{\hat{y}}^{\infty} (Y - y) dF(y),$$

where $\rho_\tau(u) = u(\tau - I(u < 0))$ is called a loss function and I is a characteristic function. This is nothing but a way of imposing asymmetric weighted residual by a loss function. After differentiating with respect to \hat{x} to obtain the optimized value, we have;

$$(3) \quad \begin{aligned} 0 &= (1 - \tau) \int_{-\infty}^{\hat{y}} dF(y) - \tau \int_{\hat{y}}^{\infty} dF(y) \\ &= \int_{\hat{y}}^{\infty} dF(y) - \tau \int_{-\infty}^{\infty} dF(y) \\ &= F(\hat{y}) - \tau \end{aligned}$$

Then, we obtain a quantile estimator $\hat{y} = F^{-1}(\tau)$ since F is monotonous. Based on the above, the sample quantile is obtained by minimizing the sample expected loss;

$$(4) \quad \arg \min_{\hat{y}_\tau} \frac{1}{n} \sum_{i=1}^n \rho_\tau(y_i - \hat{y}),$$

which yields τ^{th} sample quantile. Now in the context of a linear regression, assuming that Y_i is a linear function of x_i , the conditional τ^{th} quantile $q_\tau(y_i | x_i)$ is obtained by minimizing the following relation:

$$(5) \quad \min_{\hat{\beta}_\tau} \sum_{i=1}^n \rho_\tau(y_i - x_i' \hat{\beta}),$$

which yields τ^{th} sample quantile regression coefficient $\hat{\beta}_\tau$.

Quantile estimates $\hat{\beta}_\tau$ at distinct quantiles τ could be interpreted as the response of the dependent variable to changes in the regressors at various points in the conditional distribution of the dependent variable. And as one increase τ continuously from 0 to 1, one traces the entire conditional distribution of dependent variable, conditional on the regressors. For more properties of quantile estimates as well as asymptotics, the readers are referred to Buchinsky (1998).

4.2 Decomposition method

The decomposition method in labor economics started with the seminal papers of Oaxaca (1973) and Blinder (1973). This method decomposes the gap between groups, namely earnings gap, into two parts, differential in characteristics and differential in compensation often labeled as coefficients, namely differential in labor market endowments and heterogeneity in returns to the labor market endowments. In terms of our research, the Oaxaca-Blinder decomposition is able to decompose the wage gap between those who had experienced the poverty (hereafter, low class) and those who had not (hereafter, middle class) as specified before. This only holds, however, at the mean point, which is referred to as a limitation of this method. In other words, Oaxaca-Blinder decomposition is only able to deal with the gap between the averages of each of group. Thus, the method is not able to capture the composition of the gap at the various points of the distribution other than the mean.

In order to overcome the limitations of the method, Machado and Mata (2005) introduced a new way of decomposing the gap beyond the mean using quantile regression. In our research, we are interested in the effect of poverty during childhood on the current labor market outcome across wage distribution, thus need to compare wages between groups at various quantiles, namely, $q_\tau(Y | low\ class)$ and $q_\tau(Y | middle\ class)$. However, quantile regression estimates only yield conditional quantile specified as:

$$(6) \quad q_\tau(Y | low\ class, X) = X' \hat{\beta}_\tau^{Low}$$

$$(7) \quad q_\tau(Y | middle\ class, X) = X' \hat{\beta}_\tau^{Mid}$$

Of course, conditional quantiles and unconditional quantiles cannot be linked directly as ordinary regression does because the law of iterated expectation is not applicable in conditional quantile estimator. Here, Machado and Mata suggest

simulating the unconditional distribution based on quantile regression estimates so as to compare and decompose the quantiles of each of groups. The simulation procedure in our case is as the following steps:

(i) Draw a random sample of size S from a $uniform(0,1)$, and denote $\tau_1, \tau_2, \dots, \tau_s, \dots, \tau_S$.

(ii) Estimate a linear quantile regression for τ_s -th quantile in each groups, then denoted by $\{\hat{\beta}_{\tau_j}^{Mid}\}_{j=1}^S$ and $\{\hat{\beta}_{\tau_j}^{Low}\}_{j=1}^S$.

(iii) Draw a random sample of size S from the set of observed covariates of low class and middle class, $\{x_i^{Mid}\}_{i=1}^{N_1}$, and $\{x_i^{Low}\}_{i=1}^{N_2}$, where N_1 and N_2 are the sample sizes of the two groups, then denoted by $\{\tilde{x}_j^{Mid} \mid j = 1, 2, \dots, S\}$ and $\{\tilde{x}_j^{Low} \mid j = 1, 2, \dots, S\}$.

(iv) Generate the simulated values as $\{\tilde{Y}_j^{Mid} = \tilde{x}_j^{Mid'} \hat{\beta}_{\tau_j}^{Mid}\}_{j=1}^S$ and $\{\tilde{Y}_j^{Low} = \tilde{x}_j^{Low'} \hat{\beta}_{\tau_j}^{Low}\}_{j=1}^S$ for the unconditional distribution, and $\{\tilde{Y}_j^C = \tilde{x}_j^{Mid'} \hat{\beta}_{\tau_j}^{Low}\}_{j=1}^S$ for the unconditional counterfactual distribution.

(v) Compare the simulated distributions across quantiles, and then decompose the difference between groups as following:

$$\begin{aligned} & q_\tau(Y \mid low\ class) - q_\tau(Y \mid middle\ class) \\ &= q_\tau(Y \mid low\ class) - q_\tau(Y \mid middle\ class\ chars., low\ class\ coeffs.) \\ &+ q_\tau(Y \mid middle\ class\ chars., low\ class\ coeffs.) - q_\tau(Y \mid middle\ class). \end{aligned}$$

Albrecht et al. (2009) proved that this procedure yields consistent and asymptotically normal estimates of the quantiles of the counterfactual distribution that it is designed to simulate.

Applying this to our research, we are able to give explanations of the wage gap between groups by decomposing the gap and conjecture that the wage gap is originated from two factors; one is labor market endowments and another is reward structures of those. In other words, the decomposition method enable us to capture the constitution of the poverty effect, assuming what if one group, middle class, had been another, low class, so as to have knowledge on how the gap would be shaped or formed.

5. Quantile Estimation Results

In this section, we follow three steps to systematically analyze the log wage gap between the two groups. First, we identify the log wage gap between ‘poor family

background during childhood’ group (low class), and ‘ordinary family background during childhood’ group (middle class). We present a series of pooled quantile regressions to investigate the extent to which the log wage gap between the two groups can be attributed to the differences in labor market characteristics at various percentiles of wage distribution. In this specification, we assume that both groups are provided with homogenous reward structure of labor market characteristics.

Second, after relaxing the restriction we imposed just before, we estimate separate quantile regressions for each of the two groups to examine the extent to which the returns to the labor market characteristics at different percentiles of the distribution differ by groups, low and middle class. In doing so, we are able to identify the reward structures that each group may face, which can function as the other contributor to the gap between groups.

Finally, we carry out a decomposition analysis to identify the extent to which the log wage gap at various percentiles of each group can be explained by differentials between the groups in their common characteristics versus differences of rewards to those characteristics in labor market. We use a decomposition method developed by Machado and Mata (2005). The method decomposes the log wage gap into compensations effects and characteristics effects by considering a counterfactual situation, which is the case of the middle class group with low class’s labor market reward structure; then, this formulates a counterfactual situation that those with middle class’s endowments face low class’s reward structure so that we could provide an answer to the question we posed at the very beginning, ‘what if you had been less fortunate?’, with considering two contributors jointly to explain log wage gap between groups.

5.1 Pooled Quantile Regression with Group Dummies

We run pooled quantile regression to investigate the extent to which the difference of the log wage distributions between groups can be attributed to differences in the labor characteristics. To examine the effects of different composition of labor market characteristics between groups on the log wage gap, we carry out a series of quantile regressions on the pooled dataset with a group dummy variable and other explanatory variables, allowing for heterogeneous intercept but homogenous slopes; these pooled quantile regressions are to impose the restriction that the returns to the labor market characteristics are the same for both groups. The estimated group dummy coefficients in these regressions thus indicate the extent to which the gap between groups remains unexplained at the various quantiles when we control for group differences in various combinations of labor market characteristics.

[Insert Table 2 here]

Table 2 present the estimated group dummy coefficients at the 5th, 10th, 25th, 50th, 75th, 90th, and 95th percentiles using the pooled data with various specifications, (See Table A1-A3 in the appendix for the whole results). We also present the corresponding estimated group dummy coefficients from OLS regressions as benchmark for comparison. The first row shows the estimated log wage gap between groups across quantiles without any control variables (Table A1). This shows that the log wage

differential between groups is observed at overall distribution of wage from 8% to 21% while the OLS estimates indicate 16%. In fact, the magnitude of the differential is more significant at the first quartile rather than the last quartile, which suggests the left skewed wage distribution with high dispersion of the low class group, compared to the middle class's.

The second row presents the result, controlling for basic explanatory variables, Spec2, including education, age, age squared, and gender (Table A1), and the third row presents the result with more variables controlled, Spec3, such as marital status, regional location, and type of occupation (Table A2). The group dummy in these regressions represent the effects of group on log wage at different percentiles once we control for any different composition in these labor market characteristics between groups. In other words, the estimated group dummy is interpreted as the log wage gap between groups unexplained by labor market characteristics. As a result, 4% to 16% of the log wage differential remains unexplained with basic covariates across quantiles, and 1% to 6% is unexplained with additional control variables. OLS estimates in each specification indicate 6% and 5% of unexplained log wage differential between groups, respectively. Also, that unexplained gap tends to be smaller at the top of the distribution than the bottom as is noted in the Table 2.

The result of some estimated effects of covariates are notable as well. In both Table A2 and Table A3 in the appendix, high school diploma and college degree effects are strengthened at the upper wage distribution even though the trend is mixed below the median. This may be because we pooled the groups and imposed the homogenous wage structure between the two groups in this section. We try to address this issue in the following section more specifically. And gender effect is estimated to decrease across the whole distribution in both specifications.

In short, we identified the negative effect of poor family background during childhood on the adulthood wage and saw that the wage gap between groups can be attributed to differential composition of labor market characteristics. Yet, we also identified that there still remained about 5% of unexplained log wage gap between groups even after controlling for all covariates available as shown Table 2. As expected, this result suggests that the experience of childhood poverty still contributes negatively, beyond the path through the labor market endowments, to current log wage in the whole wage distribution.

5.2 Quantile Regressions by Groups

In the previous section, we saw the effect of differential composition of labor market characteristics on log wage gap but also identified that there still remained some fraction of the unexplained log wage gap between groups. To trace the unexplained gap, in this section, we separate the groups into low and middle class and show the extent to which returns to basic control, demographic and marital status variables differ between groups at the different quantiles in each of respective distributions. Before investigating the wage gap by groups, we carry out F-test, where the null hypothesis allows for a heterogeneous intercept but no difference at all slope coefficients between groups. The test rejects the null hypothesis and suggests prevalence of the heterogeneous slope coefficients by groups.

[Insert Figure 1 here]

Based on Spec2, Table A4 and Table A5 in the appendix present the result of quantile regressions by each group, which shows estimated effects of explanatory variables on log wage in separate groups, low and middle class at the various percentiles. Figure 1 graphically summarizes the results of Tables A4 and A5. Firstly, we need to take a look at educational explanatory variables, which are estimated with increasing trend as moving upward after the median of the wage distributions in both classes as observed in the previous literature. It suggests that education makes the dispersion in conditional wage distribution larger. Alternatively, we can take an alternative explanation of quantile estimates of education by Mwabu and Schultz (1996). Suppose that unobserved ability is reflected in wage residuals and increases across the wage distribution, and then this result can be interpreted as the evidence of complementary relation between educational attainment and unobserved ability, namely productivity. In fact, in overall, quantile estimates of educational explanatory variables show presence of strong complementary relation, in particular, above median.

The differential of the estimates between groups is shown in the third column in Figure 1. The return to high school education of the low class is larger at the very bottom of the wage distribution while the return of the middle class is larger at the top of the distribution. At 5th quantile, the return to high school education of the low class is larger by about 12% than the middle class, while it is lower by about 7% at 95th percentile. Assume each of the groups experiences identical educational system, then this heterogeneous return to each group may originate solely from the differential unobserved productivity at the same quantile. This may imply, according to the alternative explanation of quantile estimates of returns to education, the fact that at the bottom of the distribution, people of the low class are relatively more productive than the middle class counterparts. That is, it suggests the case that at the bottom of the wage distribution, relatively productive people in the low class group, compared with their counterparts in the middle class, are crowded out for some reasons, although they have more productive unobserved ability than those at the same quantile in the middle class; this would be an evidence of discrimination that the low class may suffer and we would call it crowding-out effect. Meanwhile, people of the low class at the top seem to be less productive than their counterparts in the middle class group. Then, this may tell us an evidence that the low class group people possesses less developed unobserved ability at the top of the wage distribution than their counterparts in the middle class group, which may be interpreted as disadvantage caused from poor family background during one's childhood.

It is worth to mention that, the observations made from high school education do not seem to be the case of college education. Again, highly complementary between college education and unobserved productivity is observed in both groups. However, there seems to be no systematic difference between groups. In other words, both of groups share same reward structures for college education. No more crowding-out effect at the bottom or disadvantage of low class at the top of the wage distribution is observed. To explain this identically distributed unobserved ability for college graduates, we would claim that the college education enable one to get over the unobserved productivity disadvantages caused by the poverty experience at childhood. Putting differently, unobserved ability of the low class might be developed during their

college life, or achieving college degree may mean their effort, ability, or improvement of family background at the time. That's why no heterogeneous reward structures are observed among the two groups.

Table A6 and Table A7 in the appendix are the results obtained from the model with more control variables (Spec3). Figure A1 in the appendix summarizes the results, but shows no trends of the difference in educational variables are observed. This may be because unobserved ability is just captured by additional covariates such as occupation or marital status.

All in all, these results indicate that the returns to labor market variables are heterogeneous between two groups. In the next section, we investigate whether the log wage gap between groups at various points in the wage distribution is due to differences in labor market characteristics by group or due to differences in the returns to these characteristics as the gap at each point of the distribution is further decomposed into its main constituent components.

5.3 Decomposition analysis

In this section, we use an approach developed by Machado and Mata (2005) to decompose the difference between the group log wage distributions into a component that is due to differences in labor market characteristics between low and middle class groups, and the other component that is due to differences in the rewards that each of two groups receives for their labor market characteristics.

[Insert Table 3 here]

Table 3 shows the results with basic covariates (Spec2) and with more additional variables (Spec3), respectively. Figure 2 shows the results of Spec2 graphically (See Figure A2 in the appendix for Spec3). Both of those specifications indicate similar results. The predicted overall gap shows a decreasing trend in the absolute term across quantiles from 23% to 7% (Spec2) and 22% to 8% (Spec3). This suggests the log wage gap caused by the poor family background decreases as we move up along the wage distribution. Such a shape of a treatment effect shifts the distribution to the left and increases its dispersion (Koenker & Bassett, 1978). In other words, the past poverty experience during childhood shifts the current wage distribution to the left and increases its dispersion, accordingly exacerbating inequality. It implies, for instance, that an economic crisis shock that had brought about growing number of poor household may have a sustained, long-run negative effect on the inequality, which may have accounted for the highly sustained Gini coefficient in Korea after the economic crisis in 1997.

[Insert Figure 2 here]

This poverty effect, then, can be decomposed into two factors, characteristics effect and compensation effect. In both specifications, the decreasing trend of characteristics effects in absolute term is observed across quantiles, from 13% to 6% (Spec2) and from 16% to 5% (Spec3). Thus the characteristic differential seems to make up a major part and main driving force - could have been greater because we excluded drop-outs in the

sample who may be greater portion in the low class than the middle class - of the trend of overall wage gap. This may imply that past poor family background limited opportunities to cultivate the favored labor market characteristics during one's childhood. In the specification of this paper, education level may be the only selectable variable by individual which is directly correlated with childhood family background. Accordingly, the education seems to be a main factor contributing the most in the wage gap even if we are not able to carry out a detailed decomposition through this approach. Then, we would claim that the less fortunate environments at their childhood limit their opportunities to cultivate their human capital thus lead to discount the current wage. In sum, the result points to the fact that education may serve as a facilitator of intergenerational transition effect.

The coefficient effects are estimated significantly across the whole wage distribution from 10% to 1% (Spec2) and from 7% to 3% (Spec3). The observed trends in both specifications show that the gap caused by coefficients is more influential below median than above. The estimated gap by coefficients may be explained by considering two cases. First, suppose that the employer in the labor market is able to observe and distinguish which groups of individual applicants belonged to through some other characteristics such as appearance and way of speaking. In this case, the differential by coefficients between groups can be interpreted as discrimination. Second, more likely to be realistic, suppose that one's childhood family background is not identified by employers. Then, the wage gap from the different returns to the labor market characteristics may imply that the past poor family background would have negatively affected accumulating non-cognitive skills and unobserved ability such as confidence, sincerity or sociability which could be complementary to labor market endowments. This, in turn, made those less productive and thus is reflected in the current labor market outcome. The relatively small gap observed above the median may explain fact that the discrimination is somewhat improved at the top of the distribution, or based on the other way of interpretation, we may conjecture that unobserved ability or non-cognitive skill is similarly developed at least among the top of quantiles regardless of the group membership.

6. Conclusion and Policy Recommendations

In this study, we presented a series of quantile regressions with a group dummy variable to investigate the extent to which the log wage gap between groups can be explained by the differences in labor characteristics. In doing this, we imposed upon the restriction that both groups are paid the same rewards for their labor characteristics. Then, we found out that unexplained log wage gap remained even after one control for labor market characteristics. Next, in a second step we estimated separate quantile regressions for those groups so as to examine the extent to which the returns to the labor market characteristics differ by groups. Then, we found the evidence that suggests the returns to the labor market characteristics differ by groups, seeking potential explanations of the heterogeneity. Finally we carry out a decomposition analysis to identify the extent to which the log wage gap at various percentiles of the wage distribution can be explained by differences between the groups in characteristics versus differences in labor-market rewards to those characteristics. As a result, we

found out that characteristic differential between groups was the driving force causing the log wage gap between groups, and the returns to the characteristics differed by group across overall wage distribution so as to contribute to the wage gap as well. In total, this poverty effect works heterogeneously, shown more markedly at the lower quantiles rather than the higher.

Now let's get back to the question we proposed at the beginning. What if you had been less fortunate? If you had been less fortunate in your childhood, then you might have gotten less opportunity to gain a higher level of education and even lower returns to your educational characteristics in adulthood. Our estimations indicate that this leads to a discount of about fifteen percentage points off of your wage on average, otherwise you were at higher quantile of the current wage distribution, and this disadvantage is observed greater at the lower quantiles than the higher. Such a shape of heterogeneous poverty effect, as a kind of a treatment effect, shifts the current wage distribution to the left and increases its dispersion (Koenker & Bassett, 1978), accordingly exacerbating inequality.

There is a potential caveat that needs to be taken into account in the result of this paper. It is reliability of the proxy for the past family background. In our research, dictated by data availability, we just depend on the respondent's choice among multiple choice questions to find the data of past family background. However, the respondent's current status in wage distribution may affect the answer in the survey; lower income class may put the blame their current situation on their less fortunate past so that they are likely to choose 'poor family background' in the survey. This potential endogenous problem, however, is expected to be neutralized by the opposite case; higher income class is also likely to choose 'poor family background' as well in order to overstate their current success. In addition, this potential limitation is also not expected to occur in the sense that respondents just honestly answer the question because this is only a survey which has no benefit to them no matter how they answer.

Lastly, this research provides some policy implications. First, expanding educational opportunity may contribute to reducing the wage gap originated in childhood poverty. In practice, the compulsory education course in Korea was set at the middle school since 1991. According to our study, this policy would be in effect although more data needs to be collected for proper policy evaluation of the program effect. However, it should be recognized that the education could just function as a signaling device, so that this aspect of education may weaken the policy effect of expanding educational opportunity; this compulsory school attendance policy may increase the educational attainment of those who are not directly affected so as to render the policy ineffective. Secondly, our results point out the gap by compensation effects which imply the necessity to take a look at factors beyond expanding formal education. Because this gap may be from non-cognitive skills, individual productivity or genetic influence, it is demanding to consider and implement a proper policy to reduce or desirably to eliminate the gap. However, through the active policy such as introduction of child allowance linked to family income, we may support them not to lose their potential sociability during childhood because of parent's financial problems at least.

Table 1. Summary statistics of the data, 2008.

Variable	Definition		Total	Low class	Middle class
wage (Std. Dev.)	Average monthly wage in terms of million Won(₩)		196.50 (125.92)	183.03 (128.96)	208.38 (121.99)
5th			60.00	50.00	62.00
10th			80.00	70.00	84.00
25th			100.00	100.00	120.00
50th	Monthly wage at the given quantile		170.00	150.00	180.00
75th			250.00	240.00	260.00
90th			350.00	330.00	380.00
95th			430.00	400.00	438.00
log wage (Std. Dev.)	Logarithm of monthly wage		5.10 (0.61)	5.02 (0.62)	5.17 (0.59)
midedu	Dummy	Middle school graduates or below	0.19	0.28	0.11
hiedu		High school graduates	0.37	0.40	0.35
coledu		College graduates or above	0.44	0.32	0.55
gender	Dummy	Male=1, Female=0	0.60	0.62	0.58
age (Std. Dev.)	Individual current age		40.91 (11.15)	43.30 (11.54)	38.81 (10.35)
region1	Dummy	Seoul	0.23	0.26	0.21
region2		Big urban and Kyung-gi province	0.53	0.51	0.55
region3		Others	0.24	0.23	0.24
occ1	Dummy		0.28	0.22	0.33
occ2			0.18	0.14	0.21
occ3			0.14	0.15	0.14
occ4			0.40	0.49	0.32
middle	Dummy	The answer of the survey question on which of your family environment belonged to when you were fourteen years old?	1,545	-	1,545
low			1,032	1,032	-
very low			330	330	-
obs.			2,907	1,362	1,545

95% confidence interval of age-difference between groups = (3.7, 5.3)

Table 2. The estimated group dummy coefficients using the pooled data with various specifications.

VARIABLES	(1) OLS	(2) 5th	(3) 10th	(4) 25th	(5) 50th	(6) 75th	(7) 90th	(8) 95th
low (Spec1)	-0.158 ^a (0.023)	-0.215 ^b (0.098)	-0.182 ^a (0.043)	-0.182 ^a (0.014)	-0.182 ^a (0.028)	-0.080 (0.052)	-0.141 ^a (0.050)	-0.091 ^c (0.049)
low (Spec2)	-0.059 ^a (0.018)	-0.160 ^a (0.050)	-0.084 ^b (0.037)	-0.056 ^a (0.021)	-0.038 ^c (0.022)	-0.061 ^b (0.030)	-0.047 (0.028)	-0.042 (0.038)
low (Spec3)	-0.043 ^b (0.018)	-0.058 (0.047)	-0.044 (0.031)	-0.035 (0.023)	-0.043 ^b (0.021)	-0.056 ^a (0.021)	-0.035 (0.024)	-0.011 (0.043)
Observations	2,907	2,907	2,907	2,907	2,907	2,907	2,907	2,907
R-squared	0.017							

Notes: Standard errors in parentheses. For OLS estimates, robust standard errors are used, and for Quantile regression estimates, standard errors are calculated using the bootstrap method with 300 replications.

a: Significant at 1% level; b: Significant at 5% level; c: Significant at 10% level.

Figure 1. The estimated effects of a variety of explanatory variables on log wage in separate groups, low and middle classes at the various percentiles (Spec2).

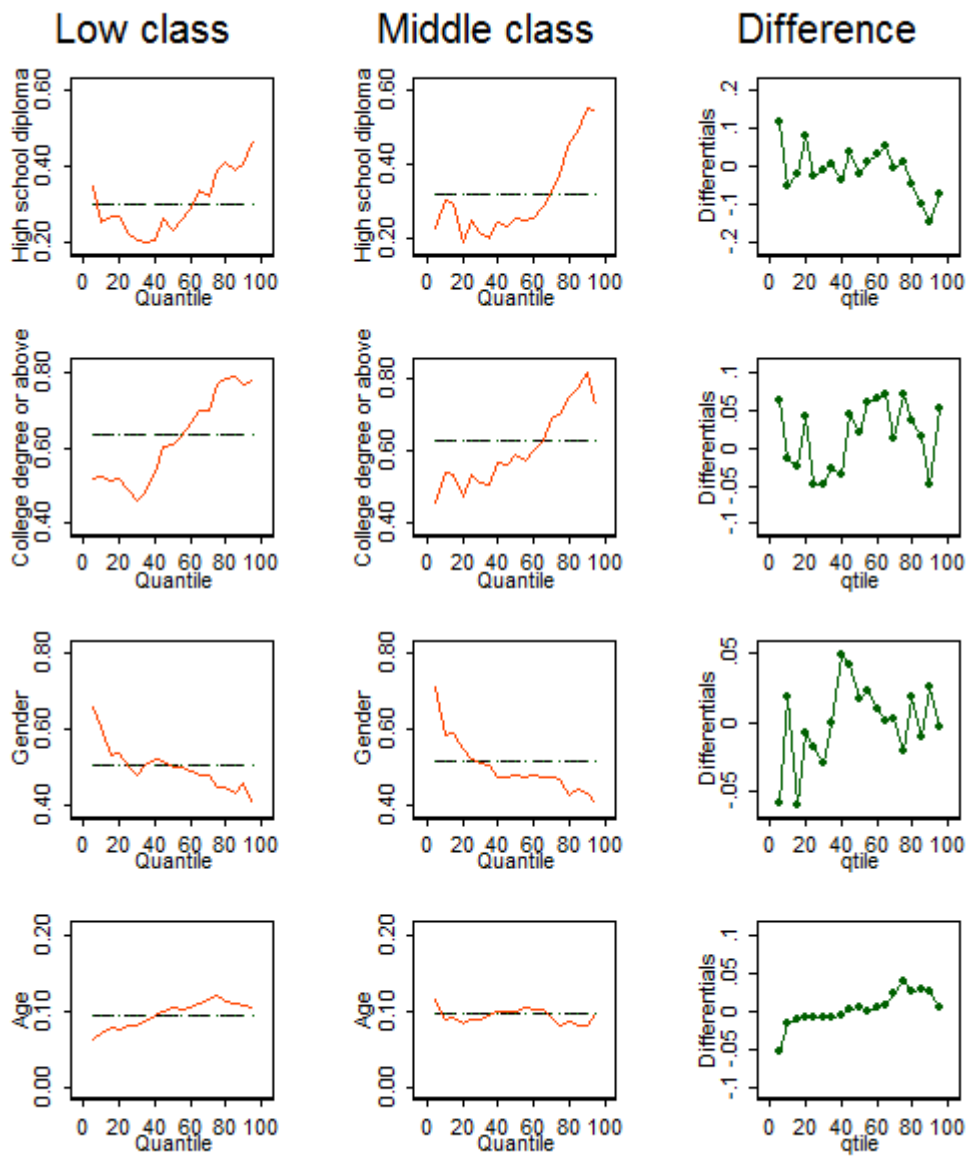
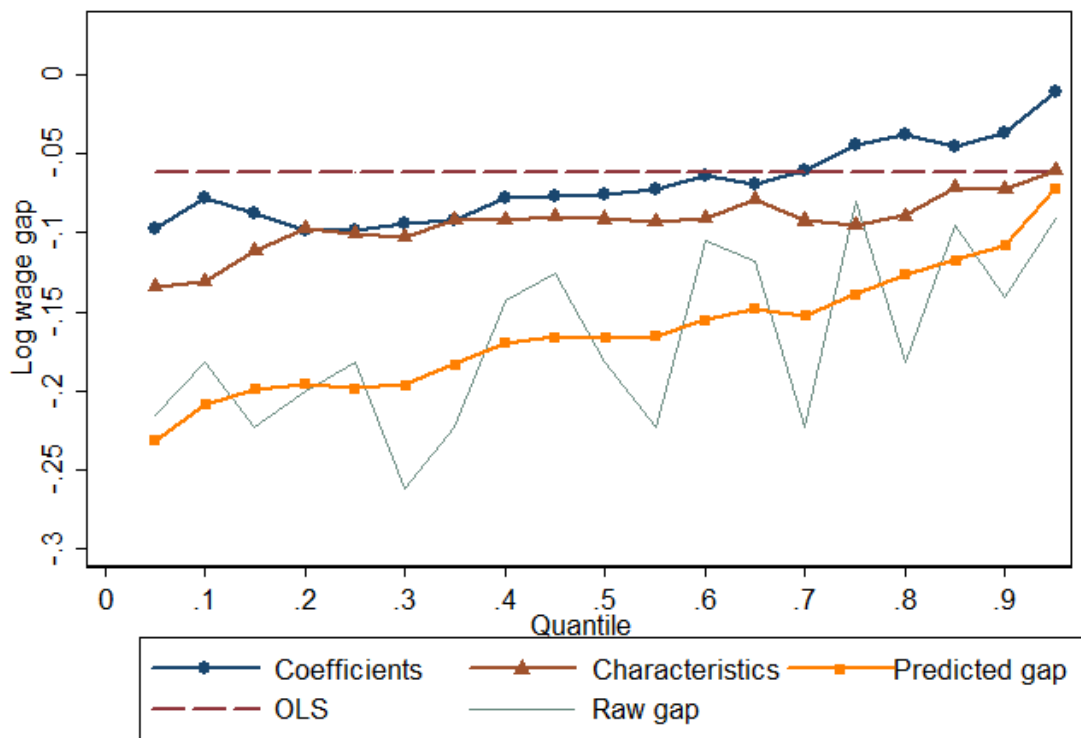


Table 2. Summary of the results from decomposition analysis based on basic explanatory variables (Spec2) and on more general specification (Spec3).

		5th	10th	25th	50th	75th	90th	95th
Spec2	Predicted gap	-0.232 (0.031)	-0.209 (0.023)	-0.199 (0.014)	-0.167 (0.014)	-0.139 (0.017)	-0.108 (0.018)	-0.071 (0.021)
	Char. gap	-0.134 (0.028)	-0.131 (0.021)	-0.100 (0.016)	-0.091 (0.014)	-0.095 (0.017)	-0.072 (0.020)	-0.060 (0.022)
	Coeff. Gap	-0.097 (0.029)	-0.078 (0.020)	-0.098 (0.016)	-0.075 (0.015)	-0.044 (0.017)	-0.036 (0.019)	-0.011 (0.020)
Spec3	Predicted gap	-0.220 (0.029)	-0.210 (0.023)	-0.202 (0.016)	-0.162 (0.014)	-0.134 (0.017)	-0.107 (0.019)	-0.079 (0.022)
	Char. gap	-0.162 (0.027)	-0.140 (0.020)	-0.126 (0.016)	-0.105 (0.014)	-0.102 (0.016)	-0.075 (0.019)	-0.047 (0.021)
	Coeff. Gap	-0.057 (0.028)	-0.070 (0.021)	-0.077 (0.016)	-0.057 (0.014)	-0.032 (0.016)	-0.032 (0.019)	-0.032 (0.019)

Notes: Standard errors in parentheses.

Figure 2. Decomposition of log wage gap between groups into compensation and characteristics effects (Spec2).



References

- Albrecht, J., Björklund, A., and Vroman, S. (2003), “Is There a Glass Ceiling in Sweden?” *Journal of Labor Economics*, Vol. 21 No. 1, pp. 145–177.
- Albrecht, J., Van Vuuren, A., and Vroman, S. (2009), “Counterfactual distributions with sample selection adjustments: Econometric theory and an application to the Netherlands”. *Labour Economics*, Vol. 16, No. 4, pp.383-396.
- Becker, G. and Tomes, N. (1979), “An equilibrium theory of the distribution of income and intergenerational mobility”. *The Journal of Political Economy*, Vol. 87, No. 6, pp. 1153-1189.
- Becker, G. and Tomes, N. (1994), “Human capital and the rise and fall of families”, in Becker, G. (Ed.), *Human Capital: A Theoretical and Empirical Analysis with Special Reference to Education*, The University of Chicago Press, pp. 257-298.
- Black, S. and Devereux, P. (2011), “Recent developments in intergenerational mobility”, in Card. D. and Ashenfelter. O. (Ed.), *Handbook of labor economics*, Vol. 4, Elsevier, pp. 1487-1541.

- Blinder, A.S. (1973), "Wage Discrimination: Reduced Form and Structural Estimates". *Journal of Human Resources*, Vol. 8 No. 4, pp. 436–455.
- Buchinsky, M. (1994), "Changes in the U.S. Wage Structure 1963-1987: Application of Quantile Regression". *Econometrica* Vol. 62 No. 2, pp. 405–458.
- Buchinsky, M. (1998), "Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research". *The Journal of Human Resources*, Vol. 33, No. 1, pp. 88–126.
- Duncan, G. and Yeung, W. (1998), "How much does childhood poverty affect the life chances of children?" *American sociological Review*, Vol. 63 No. 3, pp. 406-423.
- Duncan, G.J., Ziol-Guest, K.M., and Kalil, A. (2010), "Early-childhood poverty and adult attainment, behavior, and health." *Child development* Vol. 81 No. 1, pp. 306–325.
- Eide, E. and Showalter, M. (1999), "Factors affecting the transmission of earnings across generations: A quantile regression approach". *Journal of Human Resources*, Vol. 34 No. 2, pp.253–267.
- Fortin, N., Lemieux, T., and Firpo, S. (2011), "Decomposition methods in economics", in Card. D. and Ashenfelter. O. (Ed.), *Handbook of labor economics*, Vol. 4, Elsevier, pp. 1-102.
- Hartog, J., Pereira, P., and Vieira, J. (2001), "Changing returns to education in Portugal during the 1980s and early 1990s: OLS and quantile regression estimators". *Applied Economics*, Vol. 33 No. 8, pp.1021–1037.
- Juhn, C., Murphy, K., and Pierce, B. (1993), "Wage inequality and the rise in returns to skill". *Journal of political Economy*, Vol. 101 No. 3, pp. 410–442.
- Kang, B. G. and Yun, M. S. (2008), "Changes in Korean Wage Inequality, 1980–2005", working paper, IZA Discussion Papers, Bonn, Germany, October.
- Kim, N. N., and Kim, J. (2013), "Income inequality in Korea, 1933-2010: Evidence from income tax statistics", working paper, Naksungdae Institute of Economic Research, Seoul, Korea, Feb.
- Koenker, R. and Bassett, G.J. (1978), "Regression quantiles". *Econometrica: journal of the Econometric Society*, Vol. 46 No. 1, pp. 33–50.
- Lee, B. J. and Lee, M. J. (2006), "Quantile regression analysis of wage determinants in the Korean Labor Market". *The Journal of the Korean Economy*, Vol. 7, No. 1, pp. 1–31.

- Machado, J. and Mata, J. (2005), “Counterfactual decomposition of changes in wage distributions using quantile regression”. *Journal of applied Econometrics*, Vol. 20 No. 4, pp. 445–465.
- Mincer, J. (1974), *Schooling, Experience, and Earnings*. *Human Behavior & Social Institutions* No. 2. Columbia University Press
- Mwabu, G. and Schultz, T.P. (1996), “Education Returns Across Quantiles of the Wage Function: Alternative Explanations for Returns to Education by Race in South Africa”. *American Economic Review*, Vol. 86 No. 2, pp. 335–339.
- Nguyen, B. and Albrecht, J. (2007), “A quantile regression decomposition of urban–rural inequality in Vietnam”. *Journal of Development Economics*, Vol. 83 No. 2, pp. 466–490.
- Oaxaca, R. (1973), “Male-Female Wage Differentials in Urban Labor Markets”. *International Economic Review*, Vol. 14 No. 3, pp. 693–709.

Appendix

Table A1. The estimated group dummy coefficients based on least squares and quantile regression methods at different percentiles using the pooled data without any control variables.

VARIABLES	(1) OLS	(2) 5th	(3) 10th	(4) 25th	(5) 50th	(6) 75th	(7) 90th	(8) 95th
low	-0.158 ^a (0.023)	-0.215 ^b (0.098)	-0.182 ^a (0.043)	-0.182 ^a (0.014)	-0.182 ^a (0.028)	-0.080 (0.052)	-0.141 ^a (0.050)	-0.091 ^c (0.049)
Constant	5.174 ^a (0.015)	4.127 ^a (0.070)	4.431 ^a (0.040)	4.787 ^a (0.014)	5.193 ^a (0.027)	5.561 ^a (0.036)	5.940 ^a (0.037)	6.082 ^a (0.030)
Observations	2,907	2,907	2,907	2,907	2,907	2,907	2,907	2,907
R-squared	0.017							

Notes: Standard errors in parentheses. For OLS estimates, robust standard errors are used, and for Quantile regression estimates, standard errors are calculated using the bootstrap method with 300 replications.

a: Significant at 1% level; b: Significant at 5% level; c: Significant at 10% level.

Table A2. The estimated group dummy coefficients based on least squares and quantile regression methods at different percentiles using the pooled data, controlling for basic explanatory variables (Spec2).

VARIABLES	(1) OLS	(2) 5th	(3) 10th	(4) 25th	(5) 50th	(6) 75th	(7) 90th	(8) 95th
low	-0.059 ^a (0.018)	-0.160 ^a (0.050)	-0.084 ^b (0.037)	-0.056 ^a (0.021)	-0.038 ^c (0.022)	-0.061 ^b (0.030)	-0.047 (0.028)	-0.042 (0.038)
hiedu	0.305 ^a (0.028)	0.288 ^a (0.072)	0.257 ^a (0.058)	0.200 ^a (0.049)	0.234 ^a (0.029)	0.394 ^b (0.198)	0.482 ^a (0.043)	0.530 ^a (0.053)
coledu	0.627 ^a (0.033)	0.531 ^a (0.083)	0.503 ^a (0.066)	0.481 ^a (0.050)	0.602 ^a (0.036)	0.741 ^a (0.042)	0.783 ^a (0.036)	0.755 ^a (0.052)
gender	0.511 ^a (0.019)	0.674 ^a (0.048)	0.600 ^a (0.038)	0.526 ^a (0.024)	0.473 ^a (0.020)	0.457 ^a (0.034)	0.433 ^a (0.035)	0.412 ^a (0.036)
age	0.093 ^a (0.005)	0.081 ^a (0.012)	0.067 ^a (0.010)	0.079 ^a (0.006)	0.099 ^a (0.004)	0.106 ^a (0.010)	0.093 ^a (0.009)	0.106 ^a (0.011)
agesq	-0.105 ^a (0.006)	-0.098 ^a (0.014)	-0.084 ^a (0.011)	-0.094 ^a (0.008)	-0.112 ^a (0.005)	-0.114 ^a (0.013)	-0.095 ^a (0.010)	-0.111 ^a (0.013)
Constant	2.508 ^a (0.113)	2.149 ^a (0.276)	2.696 ^a (0.234)	2.711 ^a (0.133)	2.437 ^a (0.102)	2.372 ^a (0.264)	2.831 ^a (0.193)	2.731 ^a (0.225)
Observations	2,907	2,907	2,907	2,907	2,907	2,907	2,907	2,907
R-squared	0.443							

Notes: Standard errors in parentheses. For OLS estimates, robust standard errors are used, and for Quantile regression estimates, standard errors are calculated using the bootstrap method with 300 replications.

a: Significant at 1% level; b: Significant at 5% level; c: Significant at 10% level.

Table A3. The estimated group dummy coefficients based on least squares and quantile regression methods at different percentiles using the pooled data, controlling for explanatory variables (Spec3).

VARIABLES	(1) OLS	(2) 5th	(3) 10th	(4) 25th	(5) 50th	(6) 75th	(7) 90th	(8) 95th
low	-0.043 ^b (0.018)	-0.058 (0.047)	-0.044 (0.031)	-0.035 (0.023)	-0.043 ^b (0.021)	-0.056 ^a (0.021)	-0.035 (0.024)	-0.011 (0.043)
hiedu	0.230 ^a (0.028)	0.242 ^a (0.088)	0.216 ^a (0.053)	0.157 ^a (0.042)	0.171 ^a (0.028)	0.269 ^a (0.037)	0.360 ^a (0.043)	0.459 ^a (0.058)
coledu	0.417 ^a (0.035)	0.367 ^a (0.109)	0.373 ^a (0.063)	0.354 ^a (0.054)	0.394 ^a (0.038)	0.456 ^a (0.042)	0.512 ^a (0.051)	0.545 ^a (0.064)
gender	0.534 ^a (0.019)	0.668 ^a (0.047)	0.624 ^a (0.037)	0.537 ^a (0.030)	0.516 ^a (0.020)	0.482 ^a (0.026)	0.433 ^a (0.031)	0.458 ^a (0.036)
age	0.078 ^a (0.005)	0.043 ^a (0.015)	0.063 ^a (0.010)	0.067 ^a (0.008)	0.084 ^a (0.006)	0.090 ^a (0.009)	0.084 ^a (0.010)	0.089 ^a (0.013)
agesq	-0.090 ^a (0.006)	-0.057 ^a (0.016)	-0.081 ^a (0.011)	-0.082 ^a (0.008)	-0.097 ^a (0.007)	-0.100 ^a (0.010)	-0.089 ^a (0.011)	-0.093 ^a (0.014)
marr	0.135 ^a (0.021)	0.229 ^a (0.069)	0.180 ^a (0.045)	0.137 ^a (0.027)	0.133 ^a (0.022)	0.134 ^a (0.029)	0.120 ^a (0.035)	0.109 ^b (0.043)
region1	0.019 (0.025)	-0.016 (0.063)	0.029 (0.053)	0.007 (0.033)	0.031 (0.026)	0.041 (0.034)	0.017 (0.036)	0.013 (0.065)
region2	0.005 (0.020)	0.022 (0.049)	0.035 (0.038)	0.007 (0.025)	0.021 (0.021)	-0.013 (0.029)	0.001 (0.031)	-0.007 (0.049)
oc1	0.310 ^a (0.024)	0.264 ^a (0.073)	0.269 ^a (0.054)	0.253 ^a (0.033)	0.291 ^a (0.030)	0.360 ^a (0.032)	0.283 ^a (0.044)	0.260 ^a (0.060)
oc2	0.238 ^a (0.026)	0.245 ^a (0.071)	0.164 ^a (0.051)	0.167 ^a (0.032)	0.230 ^a (0.026)	0.247 ^a (0.033)	0.216 ^a (0.053)	0.245 ^a (0.070)
oc3	0.008 (0.027)	-0.033 (0.074)	-0.059 (0.060)	-0.013 (0.042)	0.026 (0.026)	0.024 (0.035)	-0.037 (0.041)	-0.045 (0.065)
Constant	2.722 ^a (0.118)	2.741 ^a (0.333)	2.590 ^a (0.230)	2.856 ^a (0.176)	2.647 ^a (0.121)	2.697 ^a (0.182)	2.996 ^a (0.193)	2.970 ^a (0.263)
Observations	2,907	2,907	2,907	2,907	2,907	2,907	2,907	2,907
R-squared	0.487							

Notes: Standard errors in parentheses. For OLS estimates, robust standard errors are used, and for Quantile regression estimates, standard errors are calculated using the bootstrap method with 300 replications.

a: Significant at 1% level; b: Significant at 5% level; c: Significant at 10% level.

Table A4. The estimated effects of explanatory variables (Spec2) on log wage of low class at the various percentiles.

VARIABLES	(1) OLS	(2) 5th	(3) 10th	(4) 25th	(5) 50th	(6) 75th	(7) 90th	(8) 95th
hiedu	0.301 ^a (0.036)	0.346 ^a (0.112)	0.253 ^a (0.085)	0.223 ^a (0.055)	0.231 ^a (0.039)	0.389 ^a (0.060)	0.409 ^a (0.063)	0.468 ^a (0.065)
coledu	0.637 ^a (0.045)	0.521 ^a (0.141)	0.527 ^a (0.115)	0.485 ^a (0.056)	0.608 ^a (0.048)	0.771 ^a (0.061)	0.766 ^a (0.053)	0.779 ^a (0.069)
gender	0.505 ^a (0.028)	0.658 ^a (0.075)	0.604 ^a (0.062)	0.502 ^a (0.041)	0.498 ^a (0.030)	0.448 ^a (0.047)	0.457 ^a (0.046)	0.405 ^a (0.065)
age	0.093 ^a (0.007)	0.061 ^a (0.020)	0.072 ^a (0.017)	0.080 ^a (0.008)	0.104 ^a (0.006)	0.122 ^a (0.011)	0.107 ^a (0.012)	0.104 ^a (0.016)
agesq	-0.103 ^a (0.008)	-0.074 ^a (0.022)	-0.086 ^a (0.018)	-0.093 ^a (0.010)	-0.115 ^a (0.006)	-0.131 ^a (0.012)	-0.113 ^a (0.012)	-0.111 ^a (0.018)
Constant	2.430 ^a (0.161)	2.331 ^a (0.477)	2.445 ^a (0.380)	2.588 ^a (0.174)	2.266 ^a (0.150)	1.982 ^a (0.244)	2.549 ^a (0.280)	2.798 ^a (0.343)
Observations	1,362	1,362	1,362	1,362	1,362	1,362	1,362	1,362
R-squared	0.444							

Notes: Standard errors in parentheses. For OLS estimates, robust standard errors are used, and for Quantile regression estimates, standard errors are calculated using the bootstrap method with 300 replications.

a: Significant at 1% level; b: Significant at 5% level; c: Significant at 10% level.

Table A5. The estimated effects of explanatory variables (Spec2) on log wage of middle class at the various percentiles.

VARIABLES	(1) OLS	(2) 5th	(3) 10th	(4) 25th	(5) 50th	(6) 75th	(7) 90th	(8) 95th
hiedu	0.321 ^a (0.049)	0.227 (0.155)	0.303 ^a (0.082)	0.249 ^a (0.083)	0.251 (4.008)	0.375 ^a (0.062)	0.555 ^a (0.055)	0.542 ^a (0.109)
coledu	0.629 ^a (0.053)	0.459 ^a (0.155)	0.540 ^a (0.087)	0.533 ^a (0.088)	0.588 (4.035)	0.701 ^a (0.056)	0.815 ^a (0.046)	0.726 ^a (0.116)
gender	0.517 ^a (0.025)	0.715 ^a (0.081)	0.585 ^a (0.050)	0.520 ^a (0.034)	0.480 ^a (0.053)	0.469 ^a (0.044)	0.431 ^a (0.041)	0.408 ^a (0.042)
age	0.097 ^a (0.008)	0.115 ^a (0.021)	0.088 ^a (0.017)	0.089 ^a (0.012)	0.100 (0.085)	0.082 ^a (0.018)	0.082 ^a (0.011)	0.098 ^a (0.016)
agesq	-0.112 ^a (0.010)	-0.145 ^a (0.027)	-0.112 ^a (0.020)	-0.109 ^a (0.014)	-0.115 (0.118)	-0.084 ^a (0.022)	-0.081 ^a (0.014)	-0.099 ^a (0.021)
Constant	2.439 ^a (0.165)	1.584 ^a (0.453)	2.276 ^a (0.365)	2.493 ^a (0.232)	2.461 (2.594)	2.866 ^a (0.333)	2.972 ^a (0.212)	2.847 ^a (0.290)
Observations	1,545	1,545	1,545	1,545	1,545	1,545	1,545	1,545
R-squared	0.426							

Notes: Standard errors in parentheses. For OLS estimates, robust standard errors are used, and for Quantile regression estimates, standard errors are calculated using the bootstrap method with 300 replications.

a: Significant at 1% level; b: Significant at 5% level; c: Significant at 10% level.

Table A6. The estimated effects of a variety of explanatory variables (Spec3) on log wage of low class at the various percentiles.

VARIABLES	(1) OLS	(2) 5th	(3) 10th	(4) 25th	(5) 50th	(6) 75th	(7) 90th	(8) 95th
hiedu	0.238 ^a (0.036)	0.236 ^b (0.108)	0.231 ^a (0.069)	0.199 ^a (0.047)	0.189 ^a (0.037)	0.272 ^a (0.041)	0.333 ^a (0.068)	0.430 ^a (0.076)
coledu	0.416 ^a (0.049)	0.300 ^c (0.157)	0.359 ^a (0.098)	0.403 ^a (0.070)	0.349 ^a (0.056)	0.493 ^a (0.052)	0.516 ^a (0.084)	0.544 ^a (0.109)
gender	0.513 ^a (0.028)	0.620 ^a (0.075)	0.549 ^a (0.054)	0.511 ^a (0.041)	0.528 ^a (0.033)	0.479 ^a (0.036)	0.416 ^a (0.054)	0.382 ^a (0.074)
age	0.079 ^a (0.008)	0.025 (0.018)	0.065 ^a (0.014)	0.080 ^a (0.012)	0.088 ^a (0.008)	0.102 ^a (0.010)	0.091 ^a (0.016)	0.100 ^a (0.020)
agesq	-0.090 ^a (0.008)	-0.038 ^b (0.018)	-0.080 ^a (0.014)	-0.093 ^a (0.012)	-0.100 ^a (0.009)	-0.113 ^a (0.011)	-0.098 ^a (0.017)	-0.110 ^a (0.021)
marr	0.143 ^a (0.032)	0.292 ^a (0.093)	0.187 ^a (0.063)	0.142 ^a (0.045)	0.122 ^a (0.030)	0.136 ^a (0.039)	0.114 ^c (0.063)	0.104 (0.083)
region1	0.040 (0.036)	0.015 (0.102)	0.039 (0.076)	0.020 (0.047)	0.034 (0.039)	0.043 (0.052)	0.036 (0.070)	0.001 (0.100)
region2	-0.007 (0.030)	0.003 (0.086)	-0.010 (0.057)	0.016 (0.043)	0.025 (0.035)	-0.026 (0.045)	-0.023 (0.053)	-0.122 (0.079)
oc1	0.325 ^a (0.040)	0.218 ^c (0.114)	0.328 ^a (0.080)	0.278 ^a (0.056)	0.345 ^a (0.046)	0.346 ^a (0.044)	0.268 ^a (0.072)	0.224 ^b (0.101)
oc2	0.225 ^a (0.039)	0.231 ^b (0.103)	0.171 ^b (0.070)	0.137 ^a (0.050)	0.221 ^a (0.045)	0.253 ^a (0.044)	0.168 ^b (0.083)	0.162 (0.143)
oc3	-0.016 (0.040)	-0.114 (0.117)	-0.133 (0.085)	-0.087 (0.061)	0.031 (0.040)	0.024 (0.048)	-0.045 (0.066)	-0.098 (0.093)
Constant	2.647 ^a (0.170)	3.105 ^a (0.426)	2.505 ^a (0.294)	2.459 ^a (0.270)	2.504 ^a (0.158)	2.368 ^a (0.218)	2.872 ^a (0.330)	2.951 ^a (0.413)
Observations	1,362	1,362	1,362	1,362	1,362	1,362	1,362	1,362
R-squared	0.489							

Notes: Standard errors in parentheses. For OLS estimates, robust standard errors are used, and for Quantile regression estimates, standard errors are calculated using the bootstrap method with 300 replications.

a: Significant at 1% level; b: Significant at 5% level; c: Significant at 10% level.

Table A7. The estimated effects of a variety of explanatory variables (Spec3) on log wage of middle class at the various percentiles.

VARIABLES	(1) OLS	(2) 5th	(3) 10th	(4) 25th	(5) 50th	(6) 75th	(7) 90th	(8) 95th
hiedu	0.236 ^a (0.048)	0.183 (0.123)	0.218 ^a (0.081)	0.135 ^c (0.070)	0.178 ^a (0.045)	0.237 ^a (0.062)	0.371 ^a (0.059)	0.482 ^a (0.088)
coledu	0.423 ^a (0.054)	0.327 ^b (0.135)	0.375 ^a (0.090)	0.335 ^a (0.076)	0.422 ^a (0.056)	0.411 ^a (0.066)	0.517 ^a (0.065)	0.572 ^a (0.087)
gender	0.550 ^a (0.025)	0.747 ^a (0.067)	0.683 ^a (0.046)	0.510 ^a (0.039)	0.498 ^a (0.026)	0.479 ^a (0.037)	0.454 ^a (0.039)	0.484 ^a (0.047)
age	0.080 ^a (0.008)	0.101 ^a (0.028)	0.079 ^a (0.017)	0.071 ^a (0.012)	0.082 ^a (0.010)	0.076 ^a (0.016)	0.072 ^a (0.015)	0.079 ^a (0.019)
agesq	-0.093 ^a (0.010)	-0.128 ^a (0.032)	-0.101 ^a (0.020)	-0.091 ^a (0.014)	-0.096 ^a (0.011)	-0.082 ^a (0.020)	-0.073 ^a (0.018)	-0.079 ^a (0.024)
marr	0.123 ^a (0.028)	0.133 (0.100)	0.096 ^c (0.056)	0.140 ^a (0.034)	0.135 ^a (0.034)	0.131 ^a (0.039)	0.135 ^a (0.049)	0.128 ^b (0.053)
region1	-0.005 (0.034)	-0.083 (0.081)	-0.010 (0.075)	-0.000 (0.046)	0.037 (0.036)	0.032 (0.042)	0.001 (0.046)	-0.036 (0.069)
region2	0.014 (0.027)	0.010 (0.070)	0.047 (0.054)	0.027 (0.034)	0.026 (0.026)	-0.003 (0.038)	0.024 (0.037)	-0.012 (0.062)
oc1	0.302 ^a (0.032)	0.312 ^a (0.088)	0.240 ^a (0.059)	0.231 ^a (0.040)	0.279 ^a (0.035)	0.349 ^a (0.044)	0.311 ^a (0.059)	0.239 ^a (0.075)
oc2	0.248 ^a (0.034)	0.295 ^a (0.093)	0.185 ^a (0.061)	0.181 ^a (0.042)	0.224 ^a (0.032)	0.252 ^a (0.046)	0.276 ^a (0.068)	0.231 ^a (0.075)
oc3	0.029 (0.037)	0.143 (0.103)	0.030 (0.074)	0.001 (0.051)	0.026 (0.035)	0.020 (0.051)	0.014 (0.054)	-0.057 (0.084)
Constant	2.699 ^a (0.172)	1.661 ^a (0.585)	2.303 ^a (0.336)	2.849 ^a (0.236)	2.698 ^a (0.199)	3.000 ^a (0.313)	3.144 ^a (0.277)	3.120 ^a (0.348)
Observations	1,545	1,545	1,545	1,545	1,545	1,545	1,545	1,545
R-squared	0.470							

Notes: Standard errors in parentheses. For OLS estimates, robust standard errors are used, and for Quantile regression estimates, standard errors are calculated using the bootstrap method with 300 replications.

a: Significant at 1% level; b: Significant at 5% level; c: Significant at 10% level.

Figure A1. The estimated effects of a variety of explanatory variables on log wage in separate groups, low and middle classes at the various percentiles (Spec3).

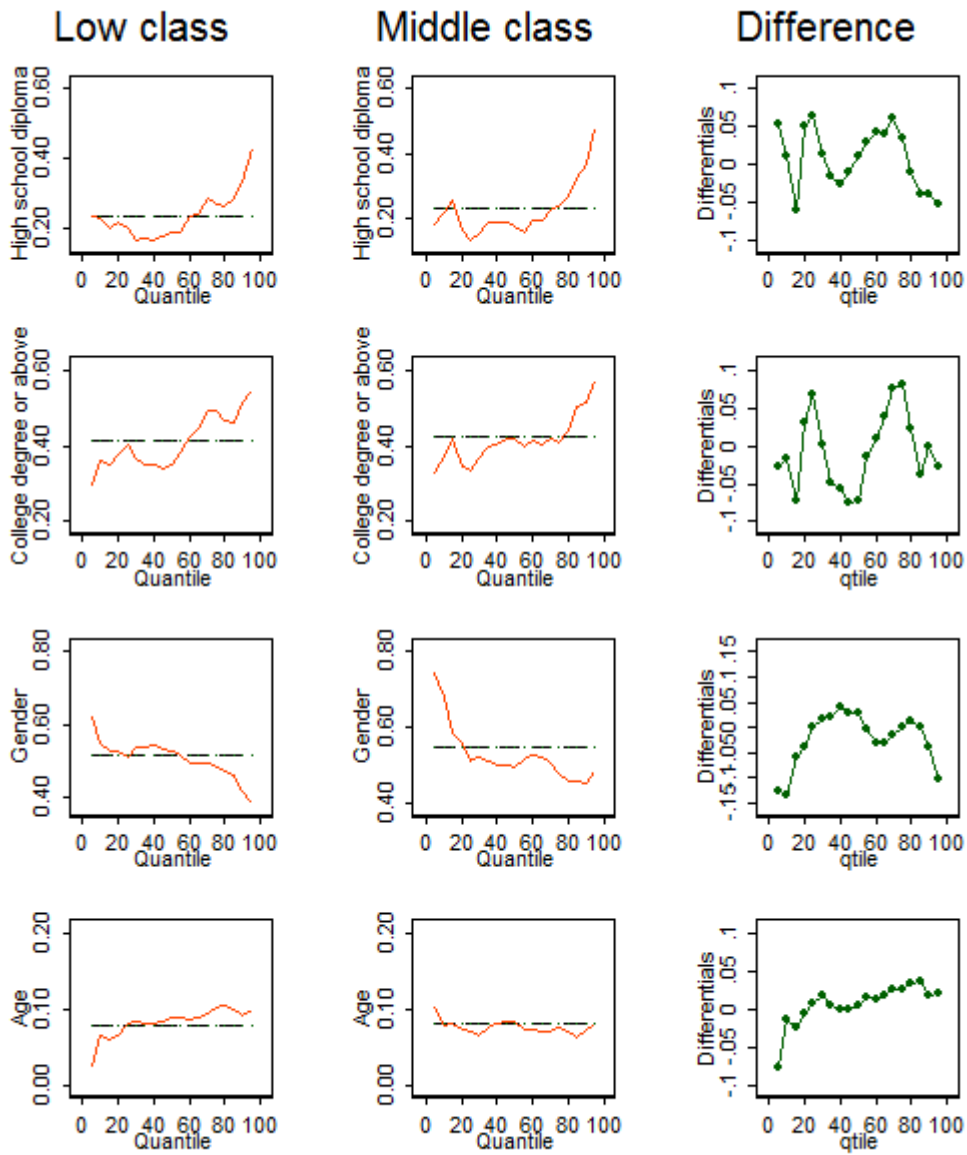


Figure A2. Decomposition of log wage gap between groups into Compensation and characteristics effects (Spec3).

