Jurnal Perencanaan Pembangunan
The Indonesian Journal of Development Planning
*Volume V No. 2 October 2021/Page: 177-194*
*Doi: 10.36574/jpp.v5i2.198*

# The Earnings Difference of Workers in Indonesia: 2007 and 2014

Aryo Prabowo[1]
International University of Japan

Submitted: July 18th 2021 | Accepted: October 23rd 2021 | Published: October 31st 2021

**Abstract**

The purposes of this study are to measure the earnings difference and the factors that influence earnings difference between 2007 and 2014 using data sourced from the Indonesian Family Life Survey (IFLS) in 2007 and 2014. The income determinant analysis results found that the longer a person's education year and work experience, the higher the income. Income will be even greater if someone is a man, lives in urban areas, and works in the non-agriculture sector. Moreover, there is no evidence that religion and ethnicity affect income. Then, I used the Blinder-Oaxaca decomposition method to distinguish the factors contributing to the difference in income to be explained factors and unexplained factors. It was found that that the income gap between 2007 and 2014 was 13.2 percentage points. Endowment factor contribution is more significant than unexplained factors. Furthermore, decomposition at different income levels shows that the endowment factor's effect on earning difference is getting smaller at higher income levels.

**Keywords:** decomposition; blinder-oaxaca; RIF.

[1] Graduate Student of International University of Japan, Minamiuonuma, Japan. E-mail: aryo@iuj.ac.jp.

# The Earnings Difference of Workers in Indonesia: 2007 and 2014

Aryo Prabowo

## I. Introduction

There was an upward trend in income inequality after the Asian financial crisis in Indonesia, at the national level, urban and rural. According to research from Yusuf, Sumner, and Rum (2014), income inequality in Indonesia denoted by the Gini coefficient increased rapidly by 32% from 0.32 in 2003 to 0.43 in 2013.

According to Statistics Indonesia, the level of inequality in Indonesia worsened in the 2007-2014 period. In 2007, the level of inequality measured by the Gini ratio was at the level of 0.376. Then, inequality changed to 0.414 in 2014, the highest level recorded by Statistics Indonesia.

Inequality would be expected at some stage to provide an impetus for the economy to continue to expand faster. However, continued income inequality would influence economic growth in the future (J. E. Stiglitz 2016). Therefore, income inequality has always been spotlighted in many countries, including Indonesia.

There has been a lot of research on inequality in Indonesia, and some look at the disparities between regions. Heryanah (2017) looked at the condition of Indonesia's income gap by using three gap indexes, namely the world bank criteria, the Gini coefficient, and the Theil index. Yusuf, Sumner, and Rum (2014) estimate expenditure inequality using the Gini coefficient, decile dispersion ratio, and Theil index of interdistrict inequality and interprovincial inequality.

Some researchers focus on the factors and causes of inequality. Taniguchi and Tuwo (2014), Sukma and Kadir (2019) analyze the gender wage gap in Indonesia and decompose it into explained factors and unexplained factors by using the Oaxaca-Blinder decomposition method. Moreover, the decomposition method is focusing the research on identifying the source factors that influence inequality. To some extent, it is necessary since it can prevent the increase of income inequality.

Decomposition of income inequality can be done in several ways. Decomposition of population subgroups or factor components is the most widely used method (Bourguignon 1979, Shorrocks 1980). The use of gender, age, and race variations in decomposition analysis is an example of population subgroup decomposition. Even though it was popular, this approach cannot control the contribution of other factors. Accordingly, it spoils the other factors' contribution, such as experience and education (Shorrocks and Wan 2005). Researchers may assign income inequalities by the source of income using a factor-component decomposition. This approach, however, is unable to account for the underlying variables that lead to income disparities, such as education, wealth, and other personal or family characteristics.

The other methodological framework, regression-based decomposition, allows researchers to get around the previous method's limitation. Oaxaca (1973) and Blinder (1973) pioneered this method, which was later developed by Jann (2008) and Fortin, Lemieux, and Firpo (2011). Using this analytical framework, researchers can simultaneously monitor many factors' contributions and define fundamental aspects' role in explaining inequality. This benefit is significant because studies have looked at not only whether inequality is needed for accumulation and how income distribution varies with economic growth, but also the determinant or how the large of income difference.

Earlier research discovered several causes that play a significant role in earnings difference. Tang and Hsu (2014) looked at how men's earnings changed in urban China by using 1986 and 2006 urban household surveys. They performed Oaxaca-Blinder decompositions of the mean wage differential, as well as newly created quantile decompositions, using the regular Mincer equation. They looked at the degree to which worker characteristics and returns to these characteristics would explain earnings differences over time, both on average and over the whole earnings distribution. They discover that earnings are favorably associated with experience and education over time. They also found that the type of ownership has a substantial effect on earnings disparities.

Neog and Sahoo (2019) explored the discrimination among formal and informal workers in India based on caste and gender. They used Employment-Unemployment Survey from the National Sample Survey Office for the four main waves from 1999-2000 to 2011-2012. Using the Oaxaca-Blinder decomposition, they found that wage inequality is slightly higher in informal jobs than formal jobs, and caste-based discrimination is often found to be less common than gender-based discrimination. The findings of the quantile decomposition demonstrated that inequality varies across quantiles.

Using Vietnam's national household info, Bui and Imai (2019) analyzed the determinants of the rural-urban divide in household welfare in Vietnam from 2008 to 2012. They used recentered influence function (RIF) decomposition to conduct quantile decomposition studies across the entire distribution to determine underlying reasons for the rural-urban variance. According to the findings, Basic schooling benefits the rural poor and ethnic minorities in raising their living conditions.

Sohn (2015) used Blinder-Oaxaca decomposition to break down the gender wage gap into explained and unexplained factors using data from the 2007 Indonesia Family Life Survey. He conducted it not only at the mean but also across the entire distribution. According to his findings, in both paying jobs and self-employment, women received around 30% less than men.

Research using the decomposition method to determine the causes of inequality is undoubtedly beneficial. However, not many researchers have used the decomposition method to find the cause of inequality between two different years. This study aims to measure how large the earnings difference is between 2007 and 2014 using the Blinder-Oaxaca decomposition. This study also seeks to determine the factors that influence earnings between the two years and how much endowment and coefficient factors contribute to earnings difference in Indonesia.

## II. Methods/Methodology

This study used data sourced from the Indonesian Family Life Survey (IFLS) in 2007 and 2014. IFLS is a longitudinal survey that collects data from individual respondents, households, communities, residences, educational facilities, and health facilities. IFLS has been conducted five times, which are 1993, 1997, 2000, 2007, and 2014 as a collaboration between RAND Corporation and various research institutions. There are seven guidebooks of IFLS: book K (control), book I, book II, book III, book IV, book V, and book proxies.

**Table 1**. Descriptive Statistics

| Characteristics | | 2007 | 2014 |
|---|---|---|---|
| Average income (million) | | 11.43 | 13.23 |
| Average age (years) | | 39.45 | 40.34 |
| Education year (years) | | 8.305 | 9.115 |
| Work experience (years) | | 25.15 | 25.22 |
| Marital Status (%) | Not married | 16.07 | 15.10 |
| | Married | 83.93 | 84.90 |
| Gender (%) | Female | 17.46 | 19.72 |
| | Male | 82.54 | 80.28 |
| Religion (%) | Islam | 90.24 | 89.98 |
| | Christianity | 3.87 | 3.85 |
| | Catholic | 1.42 | 1.13 |
| | Hinduism | 4.19 | 4.85 |
| | Buddhism | 0.23 | 0.11 |
| | Others | 0.05 | 0.08 |
| Ethnicity (%) | Jawa | 42.79 | 44.16 |
| | Sunda | 12.87 | 12.81 |
| | Bali-Nusatenggara | 11.03 | 12.52 |
| | Sumatera | 13.30 | 14.56 |
| | Sulawesi | 4.83 | 5.48 |
| | Chinese | 0.65 | 0.34 |
| | Others | 14.53 | 10.13 |
| Living Place (%) | Rural | 46 | 38.99 |
| | Urban | 54 | 61.01 |
| Industry (%) | Agriculture, forestry, fishing, hunting | 29.77 | 24.06 |
| | Manufacturing-mining | 13.81 | 15.35 |
| | Electrify, gas, water, construction | 6.78 | 7.01 |
| | Retail, restaurants, hotels | 20.46 | 21.37 |
| | Finance, Insurance | 1.04 | 5.11 |
| | Social Services | 23.09 | 22.45 |
| | Others | 5.05 | 4.64 |
| Number of observations | | 8,098 | 9,081 |

**Notes:** Income is expressed at 2007 constant price in Indonesian Rupiah, where the exchange rate in 2007 was Rp 9,379/USD

**Source:** processed data

All variables used in this study, except work experience, were taken from book IIIA and the control book. The IIIA book contains information from household members aged 15 years and over. Meanwhile, the control book includes information on all household members' information. In this study, the unit of analysis is household members aged 15-64 years included in the workforce group, both men and women.

Furthermore, the work experience variable was obtained from the following equation:

$$work\ experience = age - (education\ year + 6)$$

This method of calculating work experience is in accordance with the formula proposed by Mincer (1974) that it is assumed that work experience begins immediately after completing education. So that the length of work experience is the same as the current age minus the age at completing education.

Table 1 shows the descriptive statistics of the data used in this study. From the IFLS4 sample in 2007, there were 8,098 individuals, consisting of 6,684 men and 1,414 women. Meanwhile, from the IFLS5 sample in 2014, there were 9,081 individuals, consisting of 7,290 men and 1,791 women. The number of employed men is more than employed women; this condition is perhaps because there are differences in gender roles. Women are more responsible for household chores and childcare, while men earn a living (Becker, 1985). Therefore, the number of women who are available in the labor market is less than men.

The workforce has an average age of 39 years in 2007 and 40 years in 2014. Based on their education, most of the workforce in 2007 and 2014 have low education, shown by the average length of education at 8 and 9 years. In general, it can be said that most individuals in the workforce who worked, both in 2007 and 2014, were married and lived in urban areas.

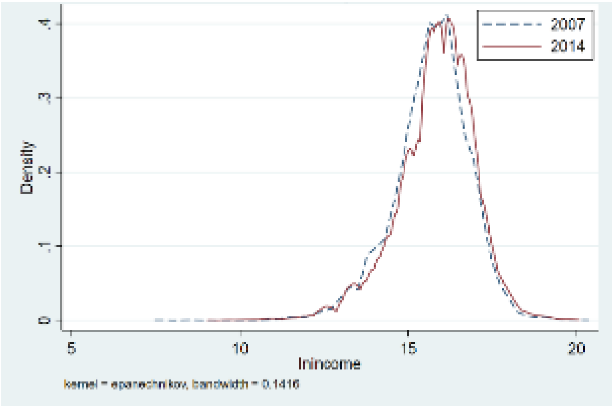


**Figure 1.** Kernel Density Plots of Annual Log Income

**Source:** processed data

Figure 1 shows kernel density plots of the log of the annual income in both 2007 and 2014. At 2007 constant price, it appears that the density of the two years is almost the same. Nevertheless, income in 2007 has a higher density than income in 2014 in the lower-income range. Otherwise, 2014 income has a higher density than 2007 income in the higher income range.

To obtain an income difference, firstly, an estimate of the income function must be made. Furthermore, the method used to estimate income refers to the Mincer earnings function (1974). The following is a form of the Mincer equation:

$$lnY = \alpha_0 + \alpha_1 EDUC + \alpha_2 EXP + \alpha_3 EXP^2 + \sum_i \gamma_i \, CONTROLS_i + \varepsilon$$

where $lnY$ is the natural logarithm of the last 12 months' earnings received by individuals, $\alpha$ and $\gamma$ are the coefficients of determining earnings, and $\varepsilon$ is the error coefficient. $EDUC$ is years of schooling, and $EXP$ is potential years of work experience.

Control variables include dummy variable marital status ("not married" =0 "married" =1), gender ("female" =0 "male" =1), living place ("rural" =0 "urban" =1). Moreover, control variables also include religion ("Christian" is the omitted category), ethnicity ("Chinese" is the omitted category), and industry types ("agriculture, forestry, fishing and hunting" is the omitted category).

The procedure used to estimate the Mincer equation is the Ordinary Least Square (OLS). In addition, quantile regression (QR) was used to estimate covariates' effects on income at different conditional distribution points (Koenker & Basset, 1978). Results are reported for three points: 25th quantiles, 50th quantiles, 75th quantiles.

The next step is to measure the income difference between years using the Blinder-Oaxaca decomposition, a method used to measure differences in income based on differences in endowment and coefficient (Oaxaca, 1973; Blinder, 1973). Oaxaca (1973) developed a method to see the difference in wages between two groups by decomposing the difference in wages into two parts: endowments or explained factors and unexplained factors. Explained factors are a part that explains income wages seen from each of the variables used in the income model. Explained factors can be in the form of individual demographic characteristics (age, gender, marital status, etc.), individual productivity factors (education, skills, health, etc.), or other considered influential factors such as place of residence.

The following is the Oaxaca-Blinder decomposition form for this research:

$$ln\bar{Y}_{2014} - ln\bar{Y}_{2007} = (\bar{X}_{2014} - \bar{X}_{2007})\hat{\beta}_{2014} + \bar{X}_{2007}(\hat{\beta}_{2014} - \hat{\beta}_{2007})$$

where $(\bar{X}_{2014} - \bar{X}_{2007})\hat{\beta}_{2014}$ is the income gap due to differences in endowment and $\bar{X}_{2007}(\hat{\beta}_{2014} - \hat{\beta}_{2007})$ is the income gap due to differences in the coefficient. $Y$ denotes the income in the last 12 months, $X$ is a vector of individual characteristics affecting earnings, $\beta$ is a vector of returns to these characteristics. This study adopts the assumption of conditional independence hypothesis following Fortin, Lemiux, and Firpo (2011). If the conditional independence hypothesis can be established, it will imply that the earnings during the two years may be produced under the same explanatory variables. Then, the effect of endogeneity problem or self-selection problem are the same. Therefore, even in different years, consistent estimates of the decomposition method can be obtained.

The Oaxaca-Blinder method measures the income difference between 2014 and 2007 and decomposes the causes of income difference, namely observed characteristics (endowment factors) and unobservable characteristics (coefficient). In this study, the endowment factors consist of variables education year, work experience, marital status, gender, religion, ethnicity, living place, and industry. With this method, it can be seen how much the contribution of the endowment factor and the coefficient to the income gap between 2007 and 2014.

Although Oaxaca-Blinder decomposition's original approach was designed to examine variations in result at the mean, several subsequent papers such as Decomposition Methods in Economics (Fortin, Lemieux, & Firpo, 2011) provided extensions and refinements that enabled the research to be extended to other distributional statistics. The aggregate structure effect can also be defined and viewed as a treatment effect under the assumptions of ignorability (conditional independence) and overlapping support (Rios-Avila, 2019).

This research followed the recentered influence function (RIF) decomposition method proposed by Firpo, Fortin, and Lemieux (2007). The method's thought is for the distributional statistics of interest, the procedure substitutes the dependent variable with the related recentered influence function.

## III. Results, Analysis, and Discussions

In 2007, for OLS estimates, all the independent variables together significantly affected the dependent variable at the 95 percent confidence level with an F value of 152.68. The coefficient of determination R squared is 0.3108, which means that 31.08 percent of the variation in the dependent variable's value is contributed by all independent variables used. At the same time, the rest comes from other variables that are not used in the model. Separately, each independent variable in the model significantly affects the dependent variable at the 95% confidence level.

As shown in Table 2, the variable education year has a significant effect on income. For OLS estimation, the coefficient is 0.099, which means that each additional year of schooling will increase income by 9.9 percent. For the QR method, the coefficient value varies but is still in the range of 9 percent.

**Table 2.** Estimation of Mincer Equation, 2007

| VARIABLES | OLS | Q25 | Q50 | Q75 |
|---|---|---|---|---|
| Education year | 0.099*** | 0.095*** | 0.097*** | 0.098*** |
| | (0.003) | (0.005) | (0.004) | (0.004) |
| Work experience | 0.032*** | 0.035*** | 0.029*** | 0.026*** |
| | (0.004) | (0.005) | (0.004) | (0.004) |
| Work experience2.1000 | -0.445*** | -0.584*** | -0.400*** | -0.290*** |
| | (0.063) | (0.076) | (0.075) | (0.061) |
| Marital Status | 0.066* | -0.035 | 0.063* | 0.130*** |
| | (0.038) | (0.052) | (0.037) | (0.038) |
| Gender | 0.577*** | 0.719*** | 0.490*** | 0.348*** |
| | (0.038) | (0.058) | (0.039) | (0.037) |
| Religion | | | | |
| Islam | 0.079 | 0.094 | 0.062 | 0.052 |
| | (0.059) | (0.067) | (0.062) | (0.053) |
| Catholic | -0.033 | 0.045 | 0.059 | 0.080 |
| | (0.111) | (0.164) | (0.102) | (0.087) |
| Hinduism | 0.118 | 0.032 | 0.128 | 0.198** |
| | (0.085) | (0.113) | (0.087) | (0.091) |
| Buddhism | -0.162 | -0.044 | 0.010 | 0.223 |
| | (0.371) | (0.760) | (0.361) | (0.542) |
| Others | 0.422 | 0.570 | 0.054 | 0.245 |
| | (0.357) | (0.658) | (0.581) | (0.429) |
| Ethnicity | | | | |
| Jawa | -0.630*** | -0.169 | -0.573*** | -0.782*** |
| | (0.188) | (0.281) | (0.175) | (0.154) |
| Sunda | -0.640*** | -0.117 | -0.567*** | -0.842*** |
| | (0.189) | (0.281) | (0.176) | (0.158) |
| Bali-NusaTenggara | -0.681*** | -0.174 | -0.654*** | -0.923*** |
| | (0.191) | (0.288) | (0.176) | (0.154) |
| Sumatera | -0.438** | 0.0968 | -0.366** | -0.626*** |
| | (0.189) | (0.278) | (0.180) | (0.155) |
| Sulawesi | -0.938*** | -0.646** | -0.855*** | -0.871*** |

| | | | | |
|---|---|---|---|---|
| | (0.195) | (0.288) | (0.185) | (0.158) |
| Others | -0.465** | 0.0643 | -0.424** | -0.678*** |
| | (0.189) | (0.285) | (0.175) | (0.156) |
| Living place | 0.164*** | 0.208*** | 0.167*** | 0.144*** |
| | (0.024) | (0.035) | (0.026) | (0.024) |
| Industry | | | | |
| Manufacturing-Mining | 0.445*** | 0.576*** | 0.466*** | 0.349*** |
| | (0.037) | (0.048) | (0.040) | (0.038) |
| Electrify, gas, water, construction | 0.265*** | 0.377*** | 0.307*** | 0.192*** |
| | (0.044) | (0.056) | (0.047) | (0.045) |
| Retail, restaurants, hotels | 0.406*** | 0.444*** | 0.394*** | 0.312*** |
| | (0.034) | (0.046) | (0.037) | (0.040) |
| Finance, Insurance | 0.194*** | 0.869*** | 0.613*** | 0.410*** |
| | (0.044) | (0.154) | (0.062) | (0.077) |
| Social Services | 0.544*** | 0.510*** | 0.486*** | 0.378*** |
| | (0.114) | (0.053) | (0.034) | (0.038) |
| Others | 0.437*** | 0.246*** | 0.216*** | 0.0702 |
| | (0.035) | (0.066) | (0.047) | (0.052) |
| Constant | 14.05*** | 13.04*** | 14.20*** | 15.04*** |
| | (0.202) | (0.291) | (0.204) | (0.168) |
| Observations | 8,098 | 8,098 | 8,098 | 8,098 |
| R-squared | 0.311 | 0.189 | 0.198 | 0.196 |

**Notes**: Asterisks report the level of significance (*** p-value <0.01, ** p-value <0.05, * p-value <0.1) and standard errors are in brackets. The reference category for religion is "Christianity". The reference category for ethnicity is "Chinese". The reference category for the industry is "Agriculture, forestry, fishing, and hunting".

**Source**: processed data

The regression results on work experience show that this variable has a significant effect on income. OLS estimates indicate that every increase of one year in work experience will increase income by 3.21 percent. The 25th quantiles regression results show that every increase of one year in work experience will increase income by 3.47 percent. Lower results were obtained in the 50th quantiles and 75th quantiles, namely 2.9 percent and 2.6 percent. Furthermore, the work experience squared shows a negative sign on both OLS and QR estimates.

Gender is also a variable that significantly affects income. The OLS estimation shows that men have a higher income by 57.7 percent than women. Using the QR method shows that the difference in income is huge at the 25th quantiles, which is 71.9 percent. The effect of gender on income is getting smaller in the 50th quantiles and 75th quantiles, 49 percent, and 34.8 percent, respectively. The detail can be seen in Figure 2.

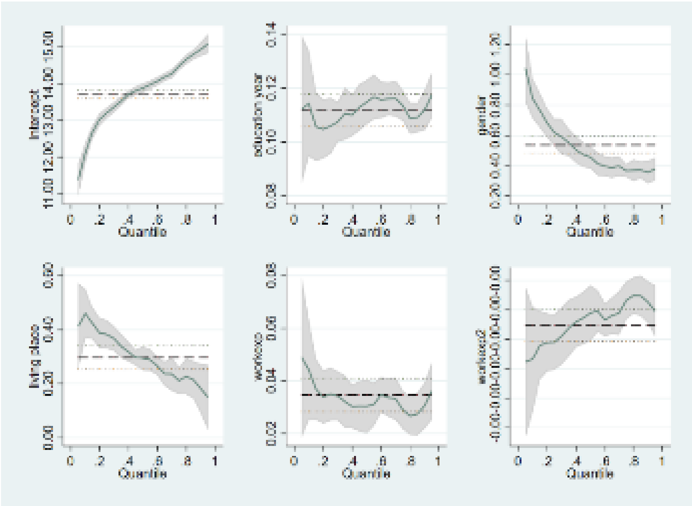


**Figure 2.** Quantile Regression Estimation, 2007

**Source**: processed data

The living place also significantly affects the annual income received. OLS estimates show that those living in urban areas have a higher income by 16.4 percent than those living in rural areas. Quantile Regression shows that in quantile 25th quantiles, the difference in income is even at 20.8 percent. The lower number is in the 75th quantiles, which is 14.4 percent.

In the industry variable, agriculture, forestry, fishing, and hunting sector became the reference categories. By using OLS estimation, all categories are significant and positive. It means that all other sectors produce a more significant income when compared to the agriculture, forestry, fishing, and hunting sector. The two sectors with the largest coefficient value are social services and manufacturing mining, which are 0.544 and 0.445.

Furthermore, the results were obtained using the quantile regression method. Almost all categories are significant and positive. The results tend to be uniform, with the smallest quantile (25th quantiles) have the largest coefficient value in each category. Then the coefficient value decreases in the following quantiles. When comparing between sectors, the largest coefficient value is in the finance and insurance sector. In the smallest quantile, the coefficient value is 0.869 or 86.9 percent higher than the agricultural sector. In the following quantiles, the finance and insurance sector's coefficient is still larger than that of other sectors.

Meanwhile, the estimation results of several variables showed less significant values. Religion variable, which used Christianity as a reference variable, was one of them. OLS estimation results showed insignificant results even for p-value <0.01. Using the Quantile Regression method, the results obtained were also not significant, except for the Hinduism category, which was positive at a confidence level of p-value <0.05.

The marital status variable shows a significant value (p-value <0.01) only at 75th quantiles, with a coefficient value of 0.130. Meanwhile, other quantiles, as well as estimation using the OLS method, showed less significant results.

The estimation results for the ethnicity variable using Chinese as the reference category showed varying results. Using the OLS method, the results obtained were negative and significant (p-value <0.01), which means that other ethnic income tends to be smaller than Chinese ethnic income. Using the QR method, the results obtained were also negative and significant at 50th quantiles and 75th quantiles at p-value <0.01. However, at the 25th quantiles, the estimation results obtained were not significant.

Table 3 shows the 2014 estimation of the Mincer equation. In 2014, for OLS estimates, all the independent variables together significantly affected the dependent variable at the 95 percent confidence level with an F value of 148.5. The coefficient of determination R squared is 0.2893, which means that 28.93 percent of the variation in the dependent variable's value is contributed by all independent variables used. At the same time, the rest comes from other variables that are not used in the model. Separately, each independent variable in the model significantly affects the dependent variable at the 95% confidence level.

**Table 3.** Estimation of Mincer Equation, 2014

| VARIABLES | OLS | Q25 | Q50 | Q75 |
|---|---|---|---|---|
| Education year | 0.093*** | 0.086*** | 0.093*** | 0.097*** |
| | (0.003) | (0.005) | (0.004) | (0.003) |
| Work experience | 0.023*** | 0.024*** | 0.016*** | 0.015*** |
| | (0.004) | (0.005) | (0.004) | (0.003) |
| Work experience2.1000 | -0.404*** | -0.511*** | -0.285*** | -0.194*** |
| | (0.063) | (0.096) | (0.071) | (0.059) |
| Marital Status | -0.003 | -0.021 | 0.028 | 0.024 |
| | (0.037) | (0.048) | (0.034) | (0.032) |
| Gender | 0.714*** | 0.931*** | 0.606*** | 0.433*** |
| | (0.035) | (0.048) | (0.040) | (0.037) |
| Religion | | | | |
| Islam | 0.059 | 0.131* | 0.099 | 0.077 |
| | (0.058) | (0.071) | (0.066) | (0.051) |
| Catholic | 0.206* | 0.417*** | 0.213** | 0.268*** |
| | (0.117) | (0.119) | (0.108) | (0.089) |
| Hindu | 0.189** | 0.326*** | 0.200** | 0.185** |
| | (0.080) | (0.100) | (0.088) | (0.082) |
| Buddhist | 0.391 | 0.553 | 0.242 | 0.425 |
| | (0.271) | (0.569) | (0.310) | (0.290) |
| Others | 1.143*** | 1.114*** | 0.791 | 1.572*** |
| | (0.306) | (0.406) | (0.505) | (0.505) |
| Ethnicity | | | | |
| Jawa | -0.063 | -0.341 | -0.173 | 0.080 |
| | (0.178) | (0.286) | (0.156) | (0.172) |
| Sunda | -0.068 | -0.378 | -0.133 | 0.086 |
| | (0.180) | (0.283) | (0.161) | (0.176) |
| Bali-NusaTenggara | -0.103 | -0.407 | -0.208 | 0.045 |
| | (0.182) | (0.283) | (0.162) | (0.177) |
| Sumatera | 0.156 | -0.112 | 0.016 | 0.222 |
| | (0.179) | (0.283) | (0.160) | (0.174) |
| Sulawesi | 0.051 | -0.315 | -0.048 | 0.222 |

| | | | | |
|---|---|---|---|---|
| | (0.183) | (0.292) | (0.166) | (0.178) |
| Others | 0.125 | -0.153 | 0.029 | 0.285* |
| | (0.181) | (0.288) | (0.157) | (0.171) |
| Living place | 0.219*** | 0.303*** | 0.228*** | 0.115*** |
| | (0.024) | (0.033) | (0.028) | (0.022) |
| Industry | | | | |
| Manufacturing-Mining | 0.429*** | 0.579*** | 0.440*** | 0.388*** |
| | (0.035) | (0.056) | (0.046) | (0.040) |
| Electrify, gas, water, construction | 0.173*** | 0.270*** | 0.201*** | 0.058* |
| | (0.041) | (0.058) | (0.046) | (0.034) |
| Retail, restaurants, hotels | 0.349*** | 0.449*** | 0.297*** | 0.292*** |
| | (0.035) | (0.047) | (0.042) | (0.035) |
| Finance, Insurance | 0.181*** | 0.590*** | 0.393*** | 0.337*** |
| | (0.054) | (0.057) | (0.057) | (0.055) |
| Social Services | 0.446*** | 0.309*** | 0.246*** | 0.255*** |
| | (0.050) | (0.049) | (0.040) | (0.034) |
| Others | 0.236*** | 0.334*** | 0.212*** | 0.156*** |
| | (0.035) | (0.060) | (0.051) | (0.049) |
| Constant | 13.73*** | 13.25*** | 14.05*** | 14.46*** |
| | (0.196) | (0.306) | (0.186) | (0.186) |
| Observations | 9,081 | 9,081 | 9,081 | 9,081 |
| R-squared | 0.289 | 0.182 | 0.176 | 0.175 |

**Notes**: Asterisks report the level of significance (*** p-value <0.01, ** p-value <0.05, * p-value <0.1) and standard errors are in brackets. The reference category for religion is "Christian". The reference category for ethnicity is "Chinese". The reference category for the industry is "Agriculture, forestry, fishing, and hunting". QR was estimated by bootstrapping the results 200 times.

**Source**: processed data

As can be seen in Table 3, the variable education year has a significant effect on income. For OLS estimation, the coefficient is 0.093, which means that each additional year of schooling will increase income by 9.33 percent. Furthermore, the Quantile Regression result showed that in the highest quantile, education years have more effect on income than in the lowest quantile. The result of QR estimation of education also can be seen in Figure 3.

The regression results on work experience show that this variable has a significant effect on income. The coefficient value of work experience was 0.023. It means that every increase of one year in work experience will increase income by 2.3 percent. The 25th quantiles regression results show that every increase of one year in work experience will increase income by 2.4 percent. Lower results were obtained in the 50th quantiles and 75th quantiles, namely 1.6 percent and 1.5 percent. Furthermore, the work experience squared shows a negative sign on both OLS and QR estimates.

Gender is also a variable that significantly affects income. The OLS estimation shows that men have a higher income by 71.4 percent than women. Using the QR method shows that the difference in income is very large at the 25th quantiles, which is 93.1 percent. The effect of gender on income is getting smaller in the income 50th quantiles and 75th quantiles, 60.6 percent and 43.3 percent, respectively.

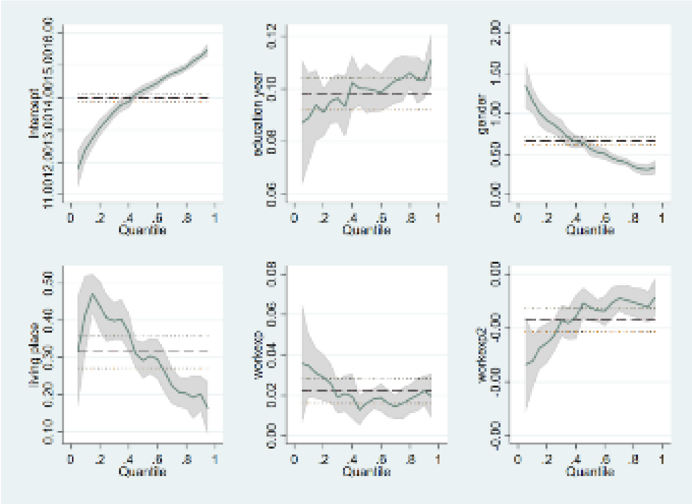


**Figure 3.** Quantile Regression Estimation, 2014

**Source**: processed data

The living place also significantly affects the annual income received. OLS estimates show that those living in urban areas have a higher income by 21.9 percent than those living in rural areas. Quantile Regression shows that in quantile 25th quantiles, the difference in income is at 30.3 percent. The lower number is in the 75th quantiles, which is 11.5 percent.

In the industry variable, agriculture, forestry, fishing, and hunting sector became the reference categories. By using OLS estimation, all categories were significant and positive. It means that all other sectors produce a larger income when compared to the agriculture, forestry, fishing, and hunting sector. The two sectors with the largest coefficient value are social services and manufacturing mining, which were 0.446 and 0.429.

Furthermore, the estimation results for industry variables obtained using the quantile regression method showed that almost all categories are significant and positive. The results tend to be uniform, with the smallest quantile (25th quantiles) have the largest coefficient value in each category. Then the coefficient value decreases in the following quantiles. When comparing between sectors, the largest coefficient value among categories in each quantile was varied. In the 25th quantiles, the highest coefficient value was finance and insurance, with a coefficient value of 0.59. In the following quantiles, the manufacturing-mining sector had the highest coefficient value, which was 0.44 and 0.388, respectively.

Meanwhile, the estimation results of several variables showed less significant values. OLS and QR estimation results for marital status variable were less significant with both negative and positive values. Less significant results also occurred in the estimation of ethnicity variables—the coefficients obtained with Chinese as the reference category varied with negative and positive signs but not significant. Moreover, there was no category for the religion variable that received significant results at p-value <0.01.

However, estimating the income function using the OLS method can lead to biased results due to self-selection problems. Therefore, the estimation of the Mincer equation will be conducted using the decomposition method.

The decomposition of the mean income differential between 2007 and 2014 is presented in Table 4. Positive variable values that reflect differences in the endowment variable will increase the income gap. Meanwhile, the negative variable value will reduce the income gap between 2007 and 2014.

Based on calculations using the Blinder-Oaxaca decomposition method, the size of the income gap between 2014 and 2007 is 0.132. This value indicates a difference in the income earned by workers in that year, where the average income in 2014 was 13.2 percentage points greater than the average income in 2007. From the 13.2 percentage points of the income gap, 9.8 percentage points or 74.24 percent of total difference were contributed by the difference in the observed characteristics (explained factor), and 3.4 percentage points or 25.76 percent of total difference from the return to the characteristics (unexplained factor). With these results, it can be concluded that the endowment factor plays a more significant role in explaining the income gap between 2007 and 2014.

**Table 4.** Decomposition of The Mean Income Differential

| Total Difference $ln\bar{Y}_{2014} - ln\bar{Y}_{2007}$ | Explained $(\bar{X}_{2014} - \bar{X}_{2007})\hat{\beta}_{2014}$ | Unexplained $\bar{X}_{2007}(\hat{\beta}_{2014} - \hat{\beta}_{2007})$ |
|---|---|---|
| 0.132 (0.017) | 0.098 (0.010) | 0.034 (0.015) |
| Education year | 0.076 | -0.048 |
| | (0.007) | (0.038) |
| Work experience | 0.002 | -0.234 |
| | (0.005) | (0.116) |
| Work experience2 | 0.001 | 0.034 |
| | (0.005) | (0.066) |
| Marital Status | -0.000 | -0.057 |
| | (0.000) | (0.040) |
| Gender | -0.016 | 0.113 |
| | (0.004) | (0.036) |
| Living place | 0.015 | 0.029 |
| | (0.002) | (0.018) |
| Industry | | |
| Manufacturing-Mining | 0.007 | -0.002 |
| | (0.002) | (0.007) |
| Electrify, gas, water, construction | 0.000 | -0.006 |
| | (0.001) | (0.004) |
| retail, restaurants, hotels | 0.003 | -0.012 |
| | (0.002) | (0.010) |
| Finance, Insurance | 0.018 | -0.001 |
| | (0.002) | (0.001) |
| Social Services | -0.002 | -0.046 |
| | (0.002) | (0.011) |
| Others | -0.001 | -0.001 |
| | (0.001) | (0.004) |
| Ethnicity | | |
| Jawa | -0.001 | 0.243 |
| | (0.003) | (0.108) |
| Sunda | 0.000 | 0.074 |
| | (0.000) | (0.033) |
| Bali-NusaTenggara | -0.002 | 0.064 |

| | | |
|---|---|---|
| | (0.003) | (0.028) |
| Sumatera | 0.002 | 0.079 |
| | (0.003) | (0.034) |
| Sulawesi | 0.000 | 0.048 |
| | (0.001) | (0.013) |
| Others | -0.005 | 0.086 |
| | (0.009) | (0.037) |
| Religion | | |
| Islam | -0.000 | -0.019 |
| | (0.000) | (0.072) |
| Catholic | -0.001 | 0.003 |
| | (0.000) | (0.002) |
| Hindu | 0.001 | 0.003 |
| | (0.001) | (0.005) |
| Buddhist | -0.000 | 0.001 |
| | (0.000) | (0.001) |
| Others | 0.000 | 0.000 |
| | (0.000) | (0.000) |
| Constant | | -0.316 |
| | | (0.269) |

**Notes**: standard error is reported in brackets.

**Source**: processed data

From the 9.8 percentage points contributed by the explained factor, 7.6 percentage points or 77.55 percent factor endowment came from the Education Year. This result looks enormous compared to the variable work experience, which only resulted in 0.2 percentage points. This result means that the education year increased the gap in workers' income between 2007 and 2014. This result also shows that the addition of school years is more dominant in determining income than work experience.

Based on Table 4, the living place was also one of the most significant factors contributing to income inequality. The Living place coefficient was 1.5 percentage point or 15.31 percent of the explained factor. This value means that workers who live in cities earn a higher income than workers who live in rural areas.

From the industrial sector, the Finance and Insurance industry contributed 1.8 percentage points or 18.37 percent endowment gap. This sector is followed by manufacturing which has a coefficient value of 0.7 percentage point. This result is reasonable because the financial and the manufacturing sector are efficient sectors. Therefore, companies in the financial and manufacturing sectors can provide higher returns for workers in this sector when compared to other business sectors.

Meanwhile, the gender variable has a coefficient value of 1.6 percentage points and is marked negative. This result means that a female worker will get better return in 2014 when compared to the return obtained in 2007.

Table 5 shows the result of Recentred Influence Function (RIF) decomposition in three income levels: 25th quantiles, 50th quantiles, 75th quantiles. The earnings gap shows a lower gap in lower quantiles compares to upper quantiles. At 25th quantiles, the earnings gap accounts for 8.8 percent, which increases to 12.7 percent at 50th quantiles, and 23.1 percent at 75th quantiles.

**Table 5.** Decomposition in Different Income Level

| | Q25 | Q50 | Q75 |
|---|---|---|---|
| Overall | | | |
| group_1 | 15.225 | 15.951 | 16.618 |
| group_2 | 15.136 | 15.824 | 16.388 |
| difference | 0.088 | 0.127 | 0.231 |
| explained | 0.109 | 0.101 | 0.117 |
| unexplained | -0.020 | 0.026 | 0.114 |
| | | | |
| Explained | | | |
| Education year | 0.066 | 0.075 | 0.097 |
| Work experience | 0.002 | 0.001 | 0.001 |
| Work experience2 | 0.001 | 0.000 | 0.000 |
| Marital Status | -0.001 | -0.000 | 0.000 |
| Gender | -0.024 | -0.013 | -0.007 |
| Living place | 0.022 | 0.016 | 0.012 |
| Industry | | | |
| Manufacturing-Mining | 0.011 | 0.008 | 0.006 |
| Electrify, gas, water, construction | 0.001 | 0.000 | -0.000 |
| retail, restaurants, hotels | 0.006 | 0.004 | 0.002 |
| Finance, Insurance | 0.028 | 0.022 | 0.017 |
| Social Services | -0.003 | -0.001 | -0.001 |
| Others | -0.002 | -0.001 | -0.000 |
| Ethnicity | 0.000 | -0.009 | -0.010 |
| Religion | 0.001 | -0.001 | 0.000 |
| | | | |
| Unexplained | | | |
| Education year | -0.040 | 0.061 | -0.025 |
| Work experience | -0.170 | 0.195 | -0.338 |
| Work experience2 | 0.009 | 0.117 | 0.132 |
| Marital Status | -0.102 | 0.069 | -0.118 |
| Gender | 0.107 | 0.063 | 0.105 |
| Living place | 0.003 | 0.030 | 0.029 |
| Industry | | | |
| Manufacturing-Mining | -0.004 | 0.012 | 0.019 |
| Electrify, gas, water, construction | -0.016 | 0.007 | -0.005 |
| retail, restaurants, hotels | 0.002 | 0.016 | 0.021 |
| Finance, Insurance | -0.001 | 0.001 | -0.003 |
| Social Services | -0.047 | 0.019 | -0.043 |
| Others | 0.002 | 0.006 | 0.011 |
| Ethnicity | 0.420 | 0.390 | 0.679 |
| Religion | -0.034 | 0.125 | -0.060 |
| Constant | -0.102 | 0.420 | -0.290 |

**Source**: processed data

The endowment's role at 25th quantiles is very large, but it got a negative effect from the unexplained variable. At 50th quantiles, the effect of endowment on earning differences was 79.53 percent. Then, at 75th quantiles, explained variable has 50.65 percent effect on the earning difference.

In the Explained part, positive sign variables reflect differences in the endowment variable will increase the income gap. Meanwhile, the negative sign variables will reduce the income gap between 2007 and 2014.

In the three quantiles, the role of education towards income difference remains the most dominant. In the 25th quantiles, the education year contributed 60.55 percent to the endowment. Meanwhile, at 50th quantiles was 74.25 and at 75th quantiles was 82.91 percent. This result shows that in the larger quantile, education's role on the earning difference is getting bigger.

The living place factor is also the most significant contributor to the earning difference. For 25th quantiles, this variable has a value of 20.18 percent, 50th quantiles of 15.84 percent, and 75th quantiles of 10.26 percent. These results indicate that residence is still a dominant factor affecting a person's earnings. However, that effect gets smaller and smaller at the larger quantile.

From the industrial sector, Finance and Manufacture are the biggest variables contributing to the earning difference. At 25th quantiles, the financial sector contributed 25.69 percent endowments, which then decreased to 21.78 percent and 14.53 percent in 50th quantiles and 75th quantiles. Meanwhile, manufacturing sector contributes 10.09 percent endowment at 25th quantiles, 7.92 percent at 50th quantiles, and 5.13 percent at 75th quantiles.

Gender became a variable that contributes to reducing the earning gap. At 25th quantiles, it contributed 22.02 percent to reduce the endowment factor. However, the percentage reduction becomes smaller in the larger quantile. At 50th quantiles, it values 12.87 percent, and 5.98 percent at 75th quantiles.

In general, earnings gaps across all earnings levels are caused due to explained factors. These factors are dominant to determine the earning difference. Furthermore, from several explained factors included in this study, education is the most significant factor in determining the earning gap between 2007 and 2014.

However, the role of the unexplained factors in determining the earning gap is more pronounced at the bottom and top quantiles. At the bottom quantiles, the unexplained factors reduced the gap in workers' income between 2007 and 2014. On the other hand, the unexplained factors cause the gap to be wider in the upper quantiles.

The variable that contributed most to increasing the income gap between 2007 and 2014 was education. This finding is consistent with Todaro and Smith (2012) that stated that one of the adverse effects of formal education on income distribution is a positive correlation between education and life-long income. Moreover, the increase in the wage gap for workers is likely due to the changing trends in the industrial needs of types of workers because of technological developments and trade openness. The technological developments require a more skilled workforce. Therefore, this condition makes workers who have higher education will be rewarded with higher income.

The decomposition of the differences in residence also shows that the wages of workers in cities are higher than those of workers in villages. Higher wages may be due to employers providing more incentives for workers to work in less livable cities than rural areas. This result may also be related to the fact that most industries are built in urban areas so that urban people have a higher income than rural communities. The inequality of wages between urban and rural areas can lead to labor migration from rural to urban areas.

## IV. Conclussion and Recommendation

The income determinant analysis results found that the longer a person's education year and work experience, the higher the income. Income will be even greater if someone is a man, lives in urban areas, and works in the non-agriculture sector. Moreover, there is no evidence that religion and ethnicity affect income.

From the Blinder-Oaxaca decomposition analysis results, it was found that the income gap between 2007 and 2014 was 13.2 percentage points. Endowment factors and unexplained factors contributed 74.24 percent and 25.76 percent. This result shows that the contribution of the endowment factor is more significant than the unexplained factor. Furthermore, decomposition at different income levels shows that the endowment factor's effect on earning difference is getting smaller at higher income levels.

From the overall decomposition results, educational characteristics and educational returns in the two stages of decomposition contributed the most to the income gap. The following characteristic that is no less important is the almost equal proportion of residence and industrial sector in widening the gap. In connection with these findings, the policy implications that may be made to increase the endowment of workers are through education, living place, and industrial sectors.

Education is a significant contributor to creating an income gap between 2007 and 2014. With the rapid development of technology, education and skills will be very important for workers in the future. Thus, increasing human capital is urgently needed, mainly by providing easy access to education for everyone. Another step that can be taken is to expand compulsory education from 9 years to 12 years. With higher quality education, workers can face challenges in the future.

Rural development can be done by providing adequate infrastructure to facilitate the accessibility of rural residents. With good access, rural communities are not left behind from urban communities to increase the rural economy's capacity. Finally, increased economic capacity will help rural communities to earn better income.

It cannot be denied that a more efficient business sector will provide higher returns to workers in that sector. For this reason, it is necessary to develop technology so that it can improve the efficiency of the sector, which is the mainstay of the Indonesian economy. The development of the manufacturing industry may be prioritized because it can absorb a lot of labor. In addition, agricultural mechanization can also be carried out to increase the efficiency of the agricultural sector.

**References**

Becker, G. (1985, January). Human Capital, Effort, and The Sexual Division of Labor. *Journal of Labor: Trends in Women's Work, Education, and Family Building Economics, 3*(I Part 2).

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

Bui, T. P., & Imai, K. S. (2019). Determinants of Rural-Urban Inequality in Vietnam: Detailed Decomposition Analyses Based on Unconditional Quantile Regressions. *The Journal of Development Studies, 55*(12), 2610-2625.

Bourguignon, F. (1979). Decomposable income inequality measures. *Econometrica: Journal of the Econometric Society*, 901–920.

Fortin, N., Lemieux, T., & Firpo, S. (2011). Decomposition Methods in Economics. *Handbook of Labor Economics, 4A*, 1-102.

Heryanah. (2017). Kesenjangan Pendapatan di Indonesia: Studi empiris Berdasarkan Susenas 2008, 2011, dan 2013. *Jurnal BPPK, 10*(2), 43-58.

Jann, B. (2008). The Blinder-Oaxaca decomposition for linear regression models. *The Stata Journal*, 453-479.

Neog, B. J., & Sahoo, B. K. (2019). Wage Discrimination in India's Formal and Informal Labor Markets. *Singapore Economic Review*, 1-21.

Oaxaca, R. (1973). Male–Female Wage Differentials in Urban Labor Markets. *International Economic Review, 14*.

Shorrocks, A. (1980). The Class of Additively Decomposable Inequality Measures. *Econometrica*, 48(3), 613-625.

Shorrocks, A., & Wan, G. (2005, January). Spatial decomposition of inequality. *Journal of Economic Geography, 5*(1), 59-81.

Sohn, K. (2015). Gender Discrimination in Earnings in Indonesia: A Fuller Picture. *Bulletin of Indonesian Economic Studies, 51*(1), 95-121.

Sukma, W. L., & Kadir. (2019). Decomposition of the Gender Wage Gap in Indonesia: Analysis from Sakernas Data. *MPRA Paper, 94930*.

Tang, C.-t., & Hsu, S.-h. (2014). The Earnings Difference of Male Workers in Urban China between 1993 and 2006. *Journal of Social Sciences and Philosophy*, 293-331.

Taniguchi, K., & Tuwo, A. (2014). New Evidence on the Gender Wage Gap in Indonesia. *ADB Economics Working Paper Series.*

Yusuf, Arief Anshory, Andy Sumner & Irlan Adiyatma Rum (2014). Twenty Years of Expenditure Inequality in Indonesia, 1993–2013. *Bulletin of Indonesian Economic Studies*, 50:2, 243-254.

Mincer, J. (1974). *Schooling, Experience, and Earnings.* New York: National Bureau of Economic Research.

Stiglitz, J. E. (2016). Inequality and economic growth. In M. Jacobs, & M. Mazzucato (ed.), Rethinking capitalism: economics and policy for sustainable and inclusive (pp. 134-155). Oxford: John Wiley & Sons.

Todaro, M., & Smith, S. (2012). *Economic Development 11th.* Boston: Pearson.

Firpo, S., Fortin, N. M., & Lemieux, T. (2007). Decomposing Wage Distributions using Recentered. *In Mimeo.*