NEW EVIDENCE ON THE GENDER WAGE GAP IN INDONESIA

Kiyoshi Taniguchi and Alika Tuwo

NO. 404

September 2014

ADB ECONOMICS WORKING PAPER SERIES

ASIAN DEVELOPMENT BANK

ADB

ADB

ADB Economics Working Paper Series

New Evidence on the Gender Wage Gap in Indonesia

Kiyoshi Taniguchi and Alika Tuwo

No. 404 | 2014

Kiyoshi Taniguchi is Senior Economist at the Private Sector Operations Department of the Asian Development Bank (ADB) and Alika Tuwo is Economist at the Jakarta Country Office of the World Bank.

The authors are highly indebted to Charles Horioka (University of the Philippines), Yoko Niimi (ADB), Sonomi Tanaka (ADB), Sakiko Tanaka (ADB), and Manami Suga (ADB) for their insightful comments. They would also like to thank the participants of the seminar hosted by the Economics and Research Department and the Gender Equity Community of Practice at ADB. All remaining errors are our own.

ASIAN DEVELOPMENT BANK

Asian Development Bank
6 ADB Avenue, Mandaluyong City
1550 Metro Manila, Philippines
www.adb.org

September 2014
ISSN 2313-6537 (Print), 2313-6545 (e-ISSN)
Publication Stock No. WPS146823-3

Note: In this publication, "$" refers to US dollars.

The ADB Economics Working Paper Series is a forum for stimulating discussion and eliciting feedback on ongoing and recently completed research and policy studies undertaken by the Asian Development Bank (ADB) staff, consultants, or resource persons. The series deals with key economic and development problems, particularly those facing the Asia and Pacific region; as well as conceptual, analytical, or methodological issues relating to project/program economic analysis, and statistical data and measurement. The series aims to enhance the knowledge on Asia's development and policy challenges; strengthen analytical rigor and quality of ADB's country partnership strategies, and its subregional and country operations; and improve the quality and availability of statistical data and development indicators for monitoring development effectiveness.

The ADB Economics Working Paper Series is a quick-disseminating, informal publication whose titles could subsequently be revised for publication as articles in professional journals or chapters in books. The series is maintained by the Economics and Research Department.

CONTENTS

TABLES AND FIGURES

TABLES

FIGURES

ABSTRACT

Indonesia has been experiencing impressive economic growth and rapid urbanization in recent years. However, urbanization could affect income inequality through people's movement from rural to urban areas. Using the 2010 National Labor Force Survey (Sakernas) in Indonesia, this study examines how monthly wages are distributed between male and female workers and tests whether a wage gap exists between them. Regression results reveal that urbanization tends to benefit male workers more favorably, in terms of monthly wages, than female workers. The wage gap tends to be wider among younger workers, particularly among those who are underemployed and severely underemployed. It is also greater among public sector workers than those in the private sector. Gender wage gap in Indonesia is mainly due to gender discrimination. An act to equalize opportunity and wages among workers, especially in the public sector, is proposed.

Keywords: gender, wage distribution, gender wage gap, Indonesia, urbanization, inclusive growth, migration

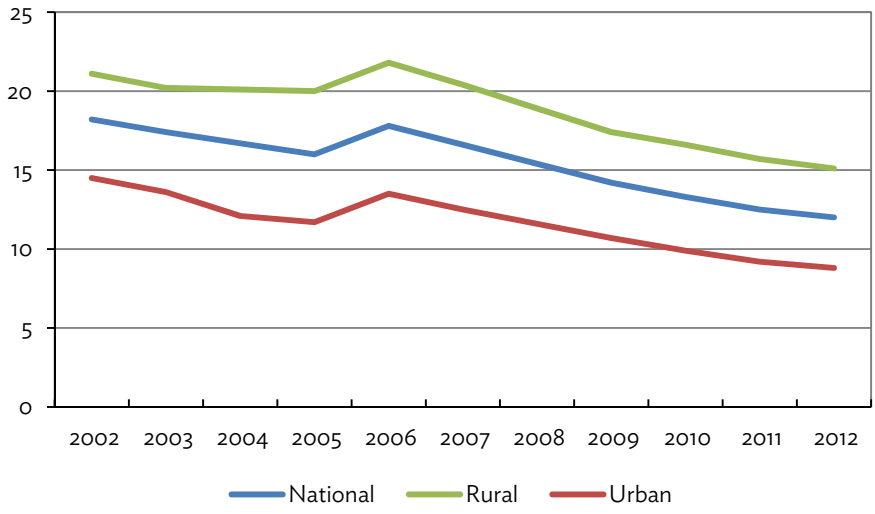
JEL Classification: E24, J16, J31, R23

I. INTRODUCTION

Indonesia has recorded impressive economic growth and poverty reduction in the last few decades. Its life expectancy at birth has lengthened to 70 years in 2011 from only 44 years in 1960 (World Bank 2014), exceeding the mean for lower middle-income countries (i.e., 66 years) in the same year. The per capita gross domestic product (GDP) at constant 2011 purchasing power parity (PPP) has more than doubled in 2012 from 1990. The last few decades also saw a substantial reduction in the proportion of Indonesian population living in absolute poverty, from 60% in 1970 (Mason and Baptist 1996) to only 12% in 2012 (World Bank 2014).

Both the increase in real income and the reduction in poverty were felt in both urban and rural areas. For the last decade, the poverty head count ratio in the national level, urban areas, and rural areas has been steadily declining, and the downward trend seems to be continuing (Figure 1).

Figure 1: Poverty Headcount Ratios at National Poverty Line
(% of population)

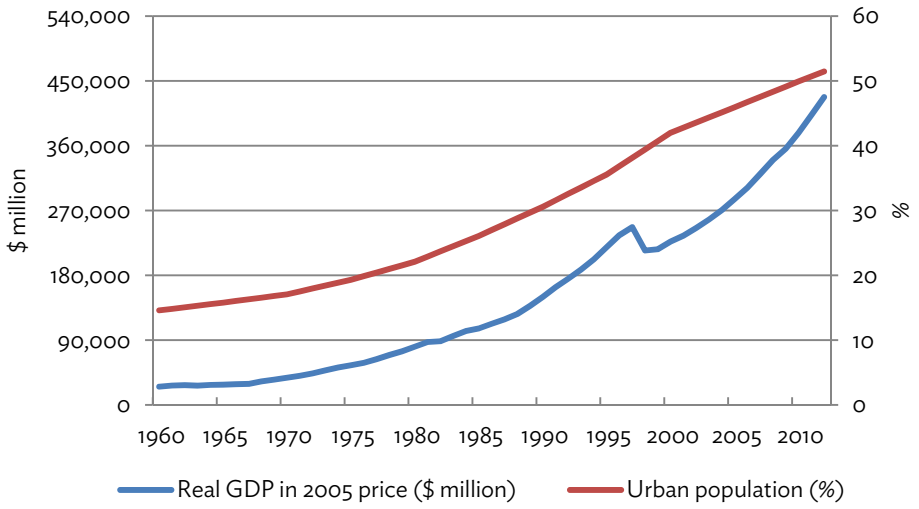


Source: World Bank. 2014. World Development Indicators.

As in most countries, Indonesia's growth was accompanied by rapid urbanization or an increase in the share of population residing in urban areas.[1] Majority of the population in Indonesia are now residing in urban areas (United Nations Population Fund [UNFPA] 2007). The proportion of urban population grew, exponentially in the past 5 decades as shown in Figure 2.

[1] Urbanization in this study simply refers to an increase in the urban share of total population that leads to a transition from a rural- to urban-centered economy.

Figure 2: Real GDP and Proportion of Urban Population



Source: World Bank. 2014. World Development Indicators.

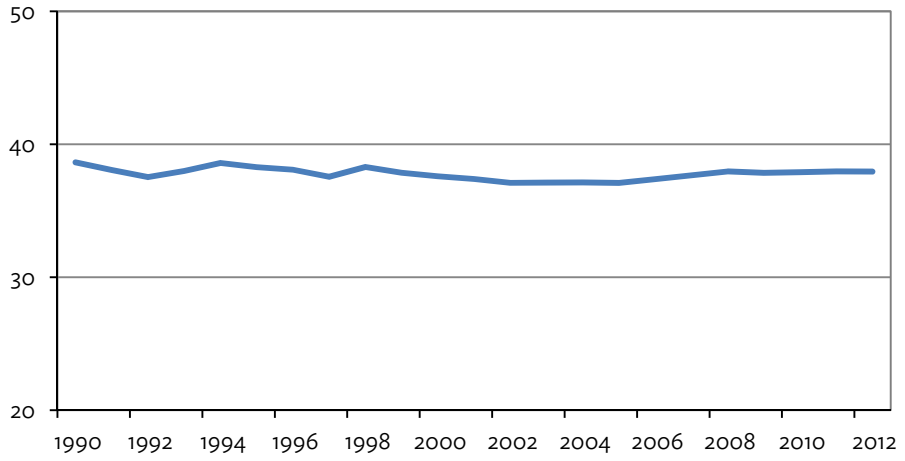
These changes were also accompanied by a structural shift in the country's industry composition. The share of agriculture in GDP declined from 42% in the latter half of 1960s to only 14% in 2012 (Hayashi 2005 and World Bank 2014).

Although the robust relationship between income growth and urbanization is well known (Spence et al. 2009), urbanization does not automatically imply an equitable distribution of wealth. Urbanization has several implications for income growth and regional inequalities. For instance, as cities grow, the urban–rural wage gap widens initially, but eventually narrows as the level of income rises to a certain level. This phenomenon is often referred as a variation of the Kuznets curve (Kuznets 1955 and Fields 2001).

Urban residents are highly dependent on cash incomes to satisfy their basic household needs (Tacoli 2012, Beall and Fox 2007). In addition, prices of goods and services in urban areas are typically higher than in rural areas. Hence, income dimensions tend to have higher weights to determine the welfare and poverty levels of urban residents. As a result, the equitable distribution of wage income opportunities is particularly more important in urban areas.

Data indicates that female labor participation rate has been steady for the last 2 decades in Indonesia (Figure 3). Badan Pusat Statistik's (BPS – the Central Agency on Statistics) definition of unemployment and its measurement went through major changes in 1994 and 2001, and more people were qualified as unemployed (Dhanani et al. 2009). For instance, in 1994, the definition was expanded from "someone who looked for a job last week" to "someone who looked for a job last week and who is currently looking for a job." These changes may account for some volatility in female labor participation rate, which remained stable within 1.5 percentage points. Acknowledging the stable rate of the female workers' share, there must be fundamental, institutional, and structural factors that affect participation of female workers in the labor market. In this study, we analyze potential causes why female workers decide not to enter the labor market.

Figure 3: Proportion of Female Workers
(% of total labor force)



Note: Labor force refers to economically active population ages 15 years old and above.
Source: World Bank. 2014. World Development Indicators.

This study focuses on measuring the changes in gender wage gap during the period of high-income growth and rapid urbanization in Indonesia. In particular, we are interested in the complex relationship between income and the gender wage gap in urban areas. We highlight the implications of income growth on the distribution of income among male and female workers and on improvements in living standards, particularly in urban areas. The livelihood and welfare of workers depend on labor income, especially in urban residents (Tacoli 2012, Beall and Fox 2007). From the social equality and egalitarian's points of view, wages should match marginal productivity of labor as well as skill level. If the gender wage gap exists in Indonesia, the government has obligations to narrow such gap through labor policy reforms from both views.

By focusing on gender wage gap in Indonesia and extending the analysis to examine its relationship with urbanization, this study contributes to the literature on urbanization and the gender wage gap, which remains limited (Tacoli 2012), and its implications, which is yet to be analyzed. It also adds to the literature on gender wage gap (the American Association of University Women [AAUW] 2014, Tijdens and van Klaveren 2012, Chevalier 2007, and Blau and Kahn 2007).

The remainder of the paper is organized as follows. Section II presents a literature review on gender wage difference. In Section III, we present data and stylized facts on gender wage differences in Indonesia. Section IV presents empirical models to estimate the gender wage gap, and its results. Section V presents a summary of findings and some policy recommendations.

II. LITERATURE REVIEW

Economic development is often associated with a structural shift from agriculture to manufacturing, and to services. Workers migrate from rural to urban areas because the expected wage is higher in urban than in rural areas (Todaro 1969 and Harris and Todaro 1970). Urbanization has several implications in the context of economic development. It is well known that there is a robust positive relationship between income growth and urbanization (Spence et al. 2009).

However, it may result in potential income distribution issues between the rural and urban areas, as well as within cities. Specifically, through the urbanization process, the rural-urban wage gap tends to widen. This is often referred as a variation of the Kuznets curve (Kuznets 1955 and Fields 2001). In addition to the rural-urban wage gap, another issue arising from this process is who benefits from the higher income within urban areas.

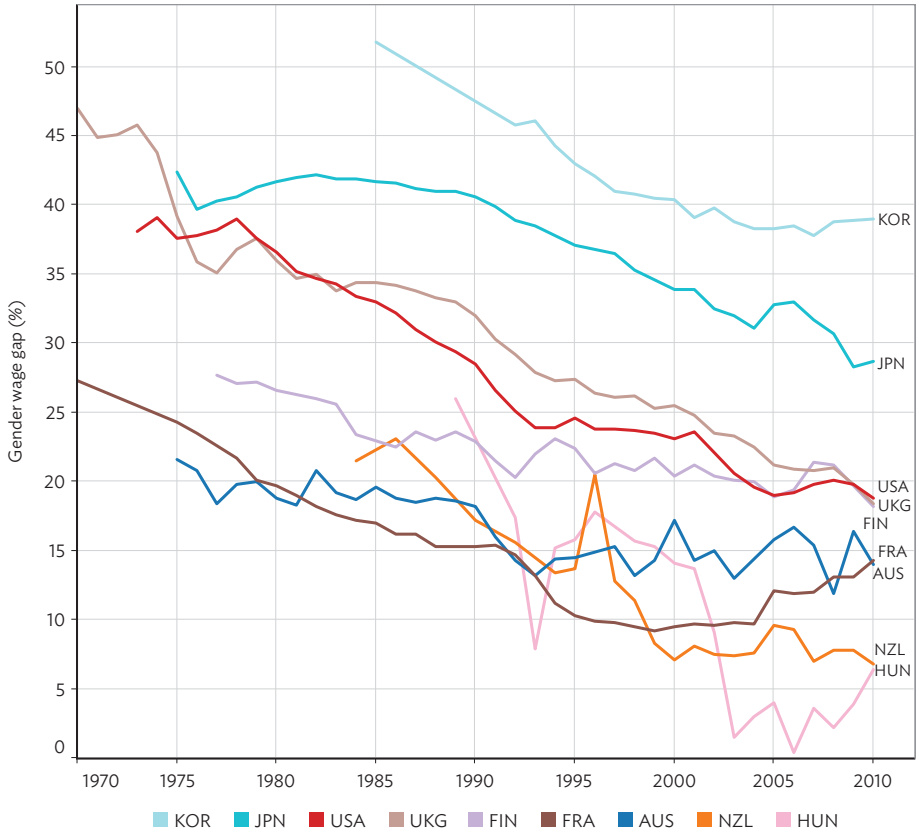
As pointed out by many studies on economic growth and income distribution, income inequality in urban areas remains a development challenge even though the average income level in urban areas tends to be higher. The Kuznetz hypothesis (Kuznets 1955) states that income inequality tends to increase in the early stages of economic development and decrease in the later stages. Fields (2001) empirically tested the Kuznets hypothesis and concluded that the Kuznets' inverted–U curve depends mainly on policies that a country adopts. Based on a cross-country study, Ravallion et al. (2007) maintained that the prevalence of poverty is increasingly becoming an urban phenomenon because the population of the urban poor is growing faster than the population as a whole.

Studies (Suryahadi et al. 2009 and Sugiyarto et al. 2006) indicate that poverty is more prevalent in rural Indonesia. Their empirical results indicate that growth in the urban services sector has the highest impact on reducing rural, as well as urban poverty. In contrast, industrial growth has a smaller poverty reduction effect. Furthermore, since the distribution of the poor tends to be clustered around the national poverty line, even a small increase in the poverty threshold would result in a sharp increase in poverty.

Women's participation in the labor market is crucial to the prosperity of cities, as well as to poverty reduction. Urbanization could open up greater employment opportunities for men and women. However, compared to the endogenous growth process of urbanization and agglomeration, the relationship between urbanization and gender equality is less studied. The World Bank (2011) reports gender issues on the development context. Based on interviews with 2,000 women across 19 countries, the study identifies the following three main factors to empower women: (i) occupational and economic change, (ii) financial management, and (iii) education and training. While the study acknowledges that no single factor could empower women, economic and financial independence, and education are major sources of womens' empowerment. At the same time, the World Bank report identifies that women's participation in the labor force has not been translated into equal employment opportunities or equal earnings for men and women. The report concludes that persistent gender gaps in earnings stem from differences in: (i) use of time (particularly for household work), (ii) access to assets and credits, and (iii) legal and regulatory framework.

The Organisation for Economic Co-operation and Development (OECD 2014) defines the "gender wage gap" (in an unadjusted form) as the difference between male and female earnings expressed as a percentage of male earnings. Figure 4 shows selected OECD member country trends in the gender wage gap in median earnings of full-time employees from 1970 to 2010. Even though there are year-to-year variances, the general downward trend is visible. Based on the OECD data, the gender wage gaps are narrowing as the economies grow. However, it is noted that the level of the gender wage gap varies significantly across countries. The Republic of Korea and Japan have one of the highest gaps, while Hungary and New Zealand have the lowest wage gaps.

Figure 4: Gender Wage Gap in Selected OECD Countries (1970–2010)



AUS = Australia, FIN = Finland, FRA = France, HUN = Hungary, JPN = Japan, KOR = Republic of Korea, NZL = New Zealand, UKG = United Kingdom, USA = United States.
Source: Authors' calculations based on OECD (2014).

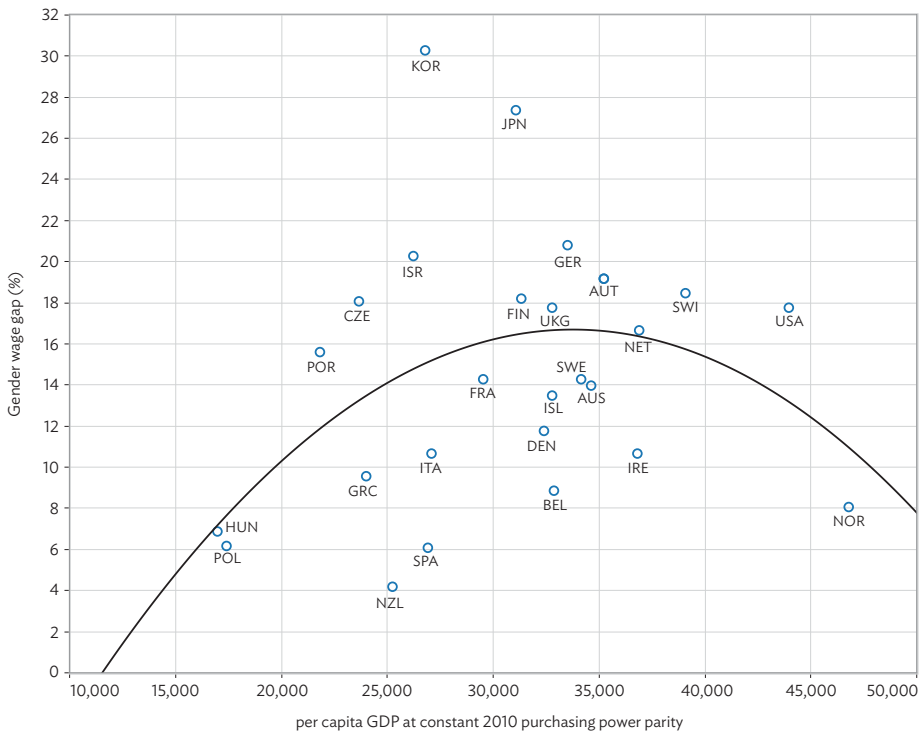
Figure 5 plots the per capita GDP at constant 2005 PPP in 2010 and gender wage gap for the latest year in 21 OECD member countries.[2] There is no simple linear relationship between the income level (i.e., per capita GDP) and the gender wage gap. This most likely stems from the labor market structure of each economy as well as the wage distributional issues within the economy. Nonetheless, the inverted U shape has a striking resemblance to the Kuznets curve. Even though there is weak empirical backing of the Kuznets hypothesis (Fields 2001), the gender gap might follow the hypothesis; as the income level rises through urbanization, the gender wage gap widens, and it starts to become narrow once a country reaches a certain income level.

Some studies summarize possible causes of the gender wage gap (Chevalier 2007, Blau and Kahn 2001). The source of the gender wage gap could be either at the labor supply (workers) side or at the labor demand side (employers). Among the "gender specific" factors that may influence gender differences in wages, the level of educational attainment is identified as a critical determinant. In particular, men tend to have higher educational attainment and specialize in career-oriented fields of study such as engineering and business, which eventually lead to relatively high earnings. Work experience may also influence the gender wage gap. Women tend to have relatively shorter work experience as they enter and exit the labor market due to family considerations. Also, women anticipate shorter or more discontinuous work lives. They tend to have lower incentives to invest in

[2] Chile is not in the data set.

education or obtain on-the-job training. There might be discriminatory exclusion of women from "male" jobs (e.g., construction work), which results in an excess supply of labor in "female" occupations and depress female wages. According to literature, the gender wage gap could be caused by: (i) hiring or employer's discrimination (Goldin and Rouse 2000); (ii) choice of entering the labor market due to social expectations (Polachek and Kim 1995); (iii) education, training, and skills (Machin and Puhani 2003, Siddique 2007); (iv) occupational choice (Baker and Fortin 2001); and (v) bearing children (Anderson, Binder, and Krause 2002).

Figure 5: Per Capita GDP (2010) and Gender Wage Gap (Latest Year)



AUS = Australia, AUT = Austria, BEL = Belgium, CZE = Czech Republic, DEN = Denmark, FIN = Finland, FRA = France, GER = Germany, GRC = Greece, HUN = Hungary, ISL = Iceland, IRE = Ireland, ISR = Israel, ITA = Italy, JPN = Japan, KOR = Republic of Korea, NET = Netherlands, NZL = New Zealand, NOR = Norway, POL = Poland, POR = Portugal, SPA = Spain, SWE = Sweden, SWI = Switzerland, UKG = United Kingdom, USA = United States.

Notes:

A polynomial trend model of degree 2 is computed as follows:

Wage Gap = -21.83 + 0.002 * pc GDP – 3.38e^-8 * pc GDP^2
(0.196) (0.04) (0.06)

R-squared: 0.18

where Wage Gap is the wage gap between male and female, pc GDP is per capita GDP in 2010, and the numbers in parentheses are P-values for coefficients.

Source: Authors' calculations based on OECD (2014) and World Bank (2014).

Various studies have empirically estimated gender wage gap across countries (Table 1). Based on the 2008 National Labor Force Survey data, van Klaveren et al. (2012) found that on average, women receive monthly wages that are 22.8% lower than men. Of this, 8.8 percentage points differential is because women work shorter hours per month, while the remaining 14 percentage points represents the pure gender wage gap due to other latent causes.

In an analysis of the impact of globalization on regional gender wage gap in Indonesia from 2001 to 2010, Fitrania (2013) found a positive correlation between regional income growth and the gender wage gap, after controlling for provincial fixed effects. He concluded that Indonesia is still at the nascent stage of industrialization and globalization.

Based on the Philippine Labor Force Survey, Cabegin (2012) found that although female workers on average have higher educational attainments than men, they still tend to receive lower compensation. This gender wage gap is attributed to gender discrimination as well as macroeconomic fluctuations, for example, in economic and productivity growth.

Table 1: Selected Studies on the Gender Wage Gap

Country (Year)	**Gender Wage Gap** (%)	**Study**
Republic of Korea (2011)	30.3*	OECD (2014)
New Zealand (2011)	4.2**	
United States (2012)	23.0	AAUW (2014)
Indonesia (2010)	20.4	Fitrania (2013)
43-country average (various years)	18.4	Tijdens and van Klaveren (2012)
Indonesia (2008)	13.7	
Zambia (2005)	46.0*	
Slovenia (2008)	4.0**	
Philippines (2005)		
Management	10.0	Cabegin (2012)
Administrative	15.0	
United States (1998)	20.3	Blau and Kahn (2001)

AAUW = American Association of University Women, OECD = Organisation for Economic Co-operation and Development,
* (**) denotes having the widest (narrowest) gap across sample countries.
Source: Authors' compilation.

Based on existing literature on gender wage gap, we test the following hypotheses using the National Labor Force Survey of Indonesia:

- A gender wage gap exists in Indonesia, as women tend to receive lower wage than men;
- The gap is narrower in urban areas where the labor market is larger and more efficient[3]; and
- Socioeconomic factors, such as marital status and educational attainment, are significant determinants of gender wage gap. These factors are more pronounced in urban areas than rural areas because the former tend to have higher weights on the income dimension of livelihood.

III. DATA AND STYLIZED FACTS

In order to analyze the distributional welfare impacts of workers in urban and rural areas, we analyze the wage differentials among workers. We utilize the 2010 labor force survey (Survei Angkatan Kerja Nasional [Sakernas]) to examine the nature of gender based wage differentials in Indonesia. Based on results, we analyze how income is distributed through workers' wage.

Sakernas follows labor force definitions set by BPS as reflected in the Standard Labor Force Concept. It defines labor force as follows:

[3] The gender wage gap is assumed to be the second degree price discrimination. Since there are more job offerings in urban areas, the magnitude of price discrimination is assumed to be less.

- Population is categorized into working age and non-working age. Those who are in the working age group are further disaggregated into labor force and non-labor force. This grouping is based on the type of activities conducted over the week prior to the survey enumeration (i.e., the reference period);
- While the non-labor force comprise those who are not engaged in an economic activity because one is a student or engaged in other personal activities such as housekeeping and social activities;
- The labor force is further divided into those who are employed and unemployed;
- In order for an individual to be qualified as being employed, one needs to be engaged in a work activity for at least one hour during the reference period with the clear objective of earning an income or profit. The referred minimum one hour of work should be uninterrupted and without any break. The employed include those who are currently working as well as those temporarily not working due to various reasons: illness, leave of absence, harvest season, and voluntary strike; and
- The unemployed refers to non-working individuals, who are currently looking for work,[4] unable to obtain work (e.g., discouraged),[5] preparing businesses,[6] or have been accepted for a job but have not yet started.

Our analytical approach slightly deviates from earlier studies in several ways. First, while many studies looked at hourly wage (e.g., the monthly total wage is divided by the actual total number of hours worked), we analyze the total monthly wage. This is because our primary focus is the welfare impact of the urbanization process on workers.

Second, we estimate real monthly wage by deflating the nominal wage by the official poverty line, with Jakarta's poverty line as the numéraire. The official poverty line is used as a proxy for the general price indicator. We use the official national poverty line because the consumer price index (CPI) dominantly reflects the price level in the urban areas, and there is no price indicator to adequately reflect the general price level in the rural areas.

Third, instead of comparing the median wages between male and female, which does not capture socioeconomic factors such as educational attainment, marital status, and number of hours worked, we analyze the gender wage gap using multivariate regression analysis. Regression analysis, in contrast, makes full use of available information, and hence, allows a more precise estimation of gender wage gap, conditional on differences in each worker's social and economic characteristics.

A. Distribution of Real Wages and Hours Worked in Urban and Rural Areas

Figures 6 and 7 show kernel density of the log of the monthly real wage by gender in both urban and rural areas. We note three key issues that have policy and social safety net implications. First, in the lower wage region, women's wage has a higher density than men's wage in both urban and rural areas. This implies that women's wage is skewed toward the lower wage range in the whole wage spectrum. Second, men's wage is denser than women's wage around mean wage in both urban and rural areas. This implies that men's wage is clustered around the mean with low variance. Third, wage distribution in the upper range is similar for both men and women. Hence, for workers who receive above-average wage, there is gender wage differential. Interestingly, this is true for both urban and rural areas.

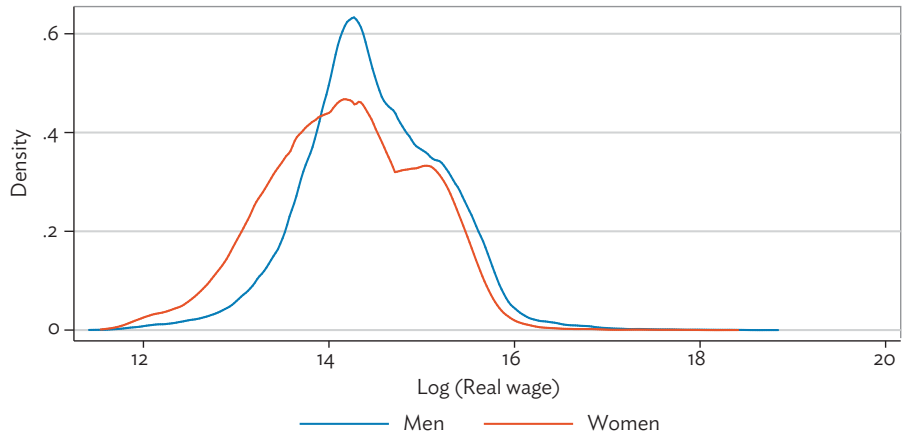
---

[4] Looking for work is defined as displaying efforts to obtain a job during the reference period.

[5] Discouraged individuals are those who have repeatedly sought work but have been unsuccessful to obtain a job and who feel that the current conditions and season do not allow them to obtain a job which they want.

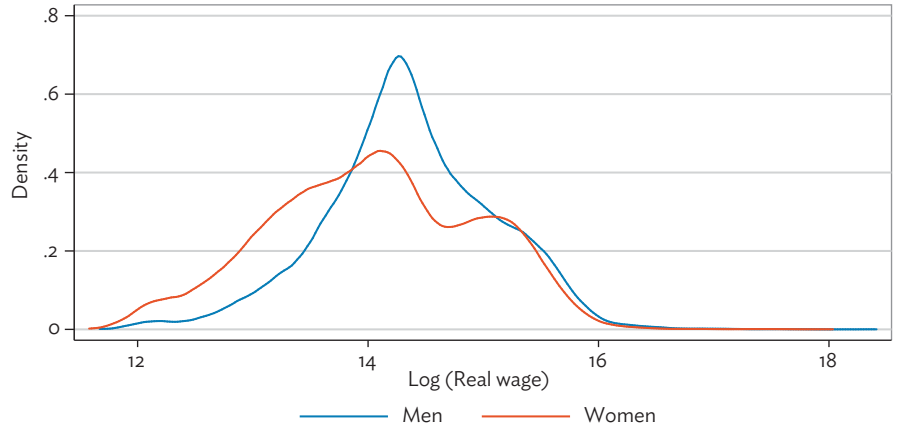
[6] Preparing businesses refers to activities done to set up as new business. This includes accumulating capital, identifying the location of business, and obtaining business licenses.

Figure 6: Kernel Density of Monthly Urban Real Wages by Gender



Source: Authors' estimates based on Sakernas (2010).

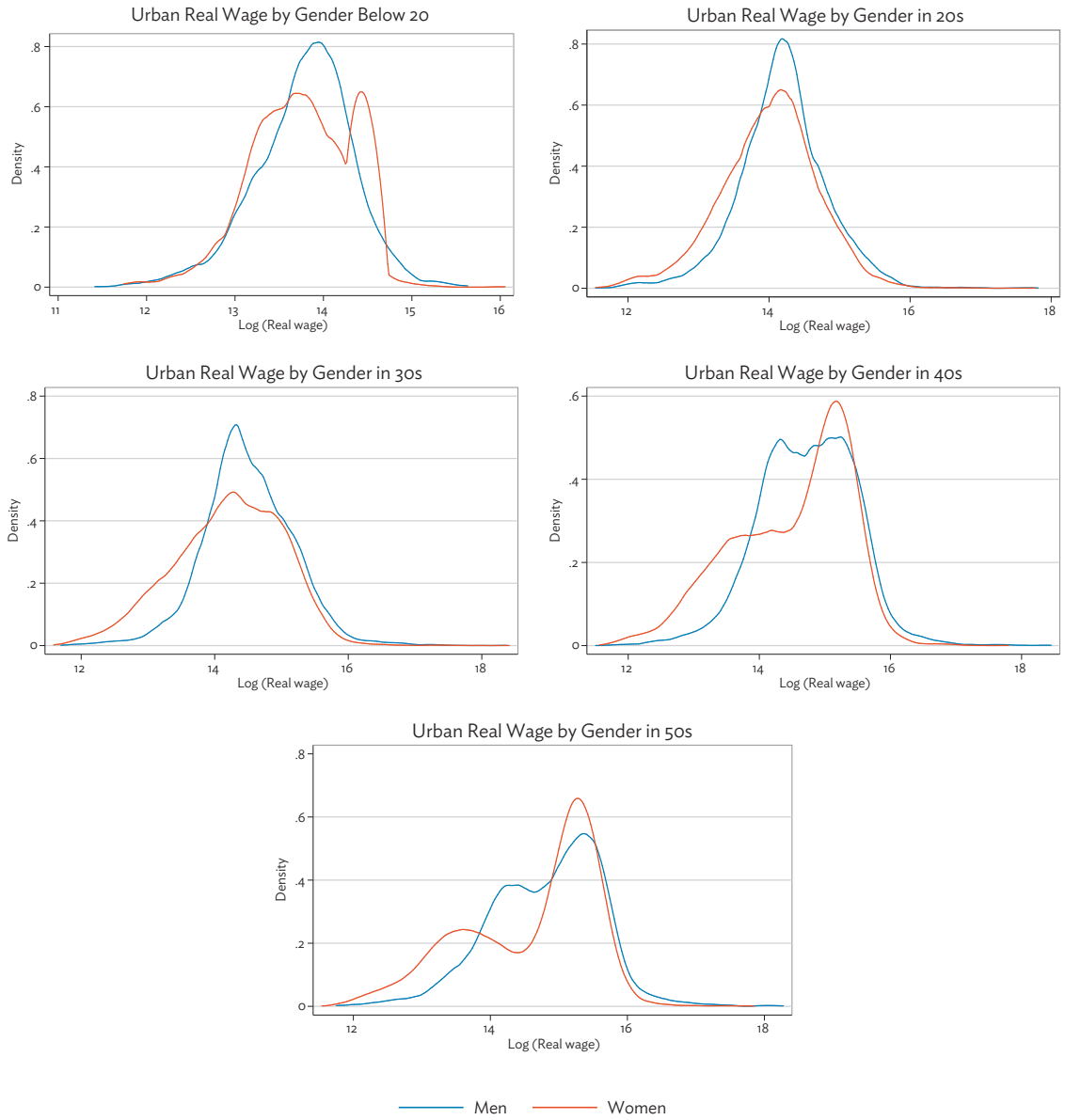
Figure 7: Kernel Density of Monthly Rural Real Wages by Gender



Source: Authors' estimates based on Sakernas (2010).

Figure 8 shows kernel density of monthly urban real wage for different age cohorts. In general, as in the total data, the women's wage has a dense distribution at the lower range of the wage distribution except for workers below 20 years old. This is probably because the absolute level of the junior level staff wages and salaries is very low that any gender wage gap is not apparent. In addition, women's life cycle events in this age cohort (e.g., giving birth or taking care of parents) limit the number of hours they can work. In their 40s and 50s, the upper range of the women's real wage is distributed similar to the one for men. From the social welfare and equity point of view, the government should focus on female workers who are in the lower wage bracket. There female workers in the lower wage range are sizable throughout the age distribution.

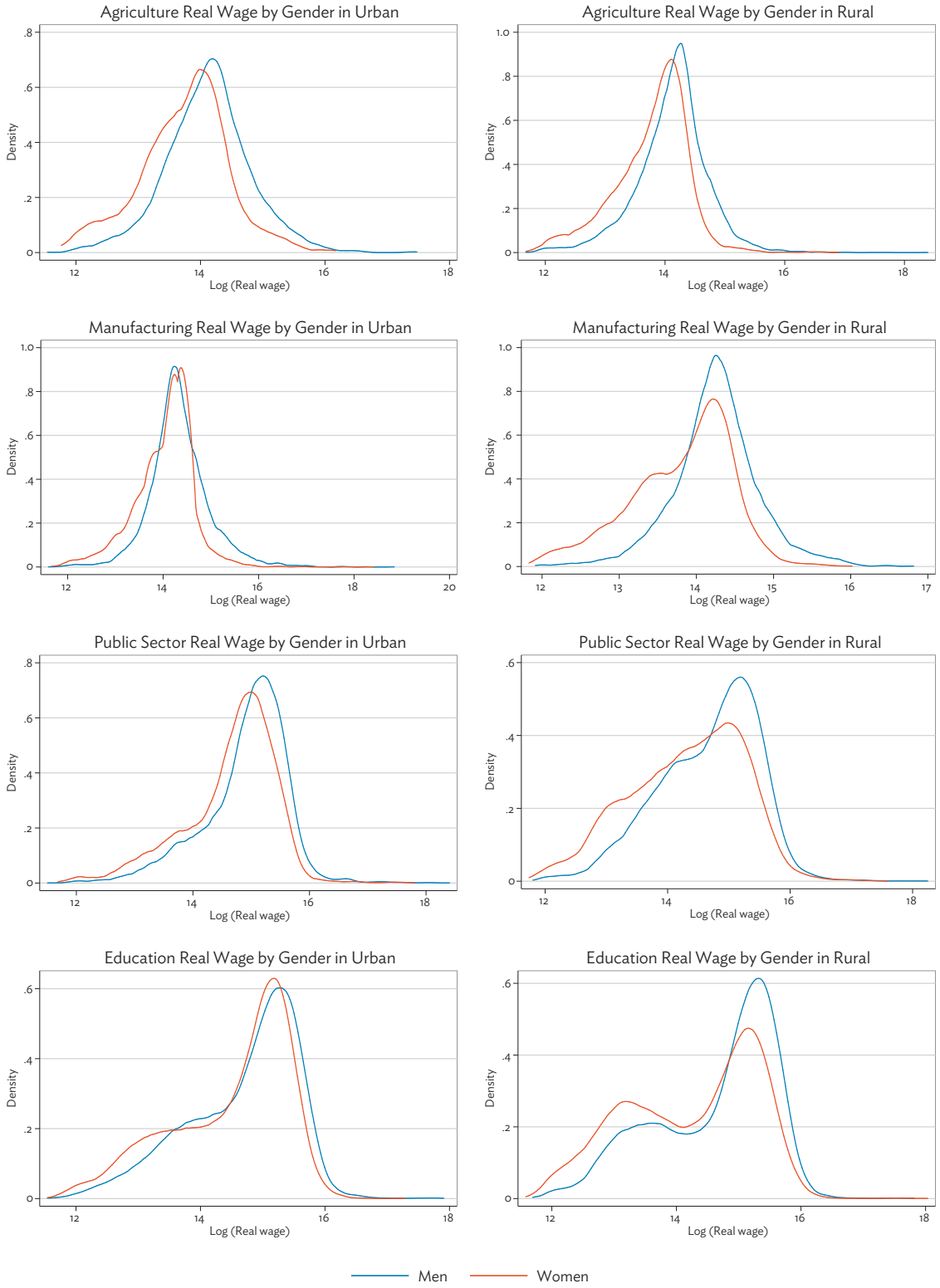
Figure 8: Kernel Density of Monthly Real Wages by Gender and Age Cohort in Urban Areas



Source: Authors' estimates based on Sakernas (2010).

Looking at the kernel density of monthly real wage in selected industries (Figure 9), women's wages are again mostly skewed to the lower wage range, except in the manufacturing sector. Interestingly, the peak of the women's wage distribution is placed higher, albeit slightly, than the men's wage distribution peak in the manufacturing industry.

Figure 9: Distribution of Real Monthly Wages by Gender, Area, and Industry

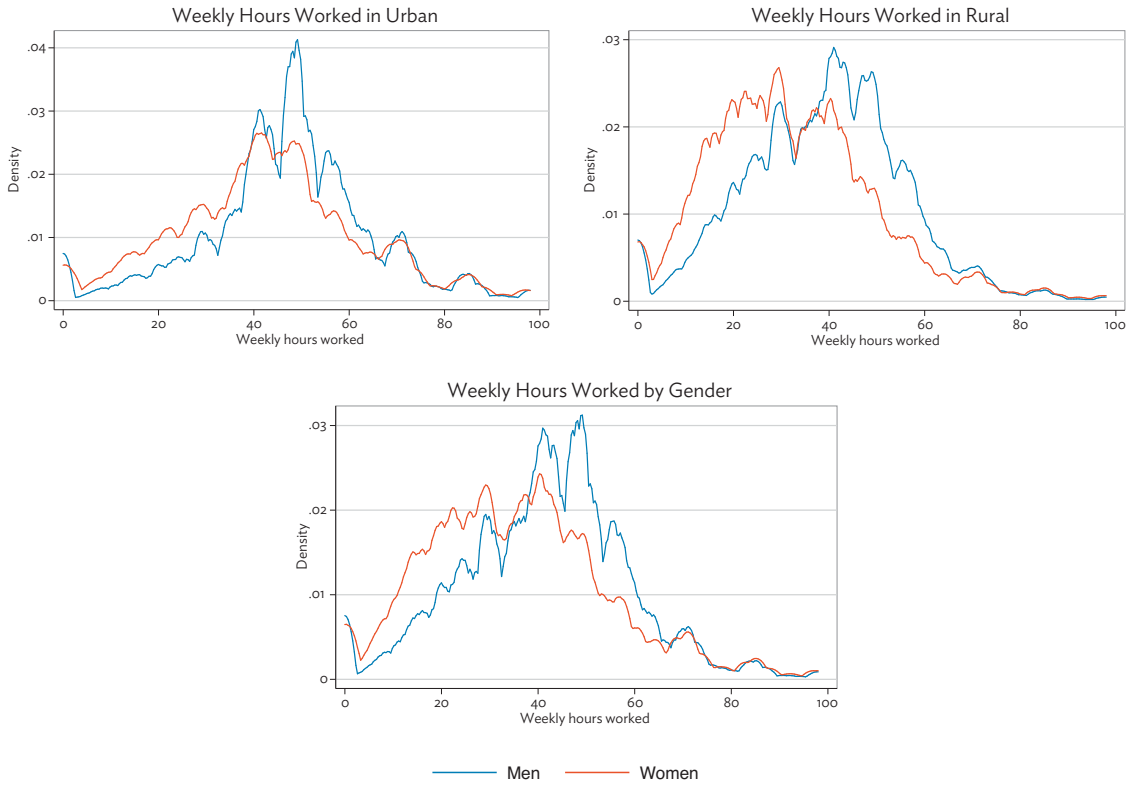


Source: Authors' illustrations based on Sakernas (2010).

B. Distribution of Hours Worked by Urban and Rural Workers

Since our main focus is the monthly wage, work duration (i.e., the number of hours worked) could be one of the key determinants of total monthly wage. As shown in Figure 10, in general, male workers tend to work longer hours than females, in both urban and rural areas, although to a lower extent in urban areas. Comparing duration of work suggests that rural workers tend to work shorter hours.[7]

Figure 10: Distribution of Hours Worked by Gender and Area

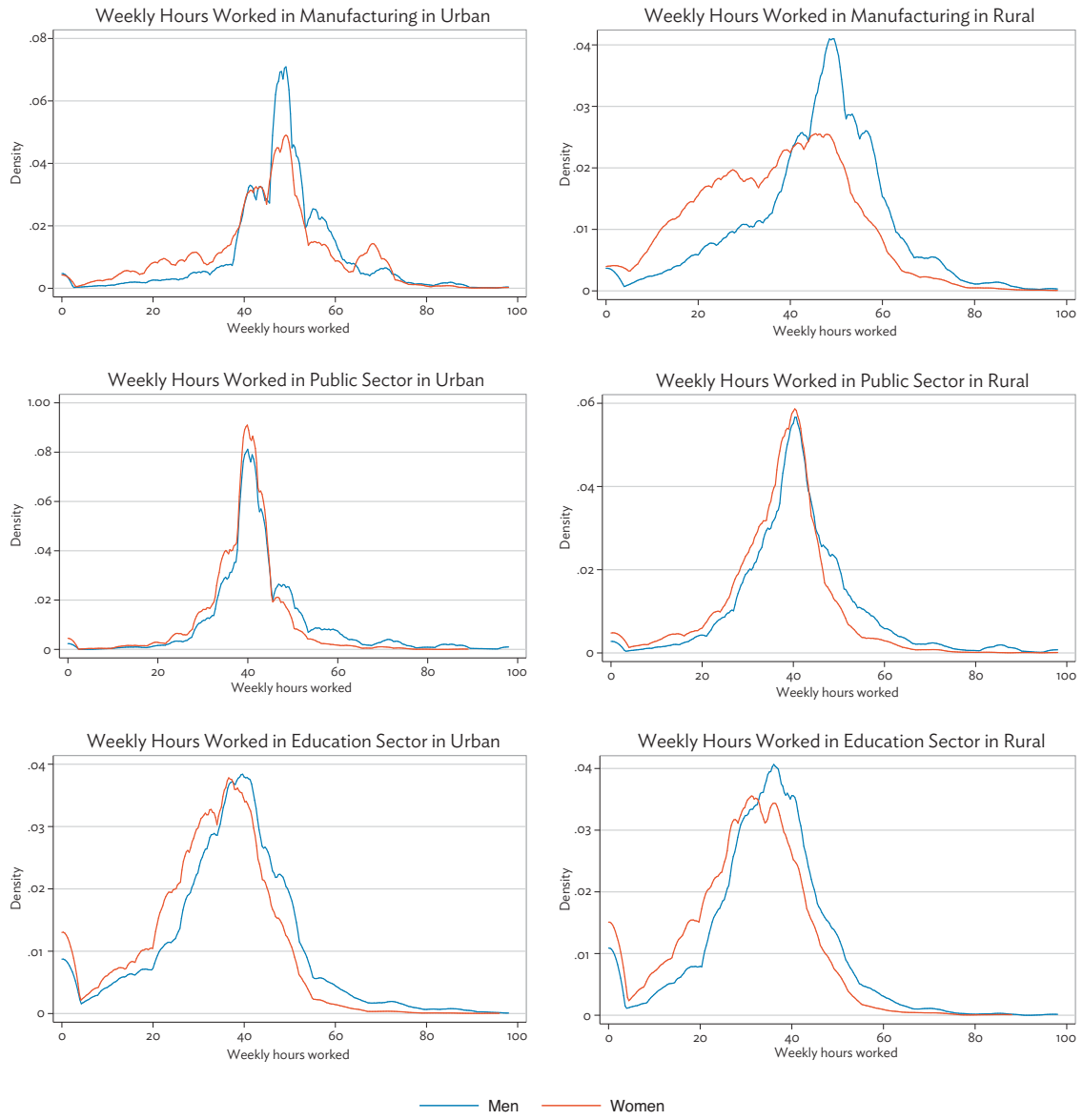


Source: Authors' illustrations based on Sakernas (2010).

Figure 11 illustrates the distribution of hours worked of those in the manufacturing sector, government administration (the public sector), and education (mainly teachers in the public sector). Compared to workers in the public sector, those in the manufacturing sector understandably tend to work longer. The distribution of hours worked is symmetric, except in the manufacturing sector in rural areas. Compared to the private sector, gender differences in hours worked is lower in the public sector, particularly in urban areas.

[7] Potentially, there is a seasonal effect on hours worked in rural areas. However, given the database, it is not possible to control potential seasonal effect.

Figure 11: Distribution of Hours Worked by Gender, Area, and Industry



Source: Authors' illustrations based on Sakernas (2010).

C. Full Employment, Underemployment, and Severe Underemployment

After entering the labor market, some female workers decide to work part time, either intentionally or unintentionally. BPS defines underemployment as working 35 hours per week or less.[8] Meanwhile, working less than 15 hours per week is defined as severe underemployment (Mason and Baptist 1996).

As shown in Table 2, on average, male workers work longer hours than female workers in both 1993 and 2010. Table 2 also shows that unemployment is more of an urban phenomenon, while severe underemployment and underemployment are more prevalent in rural areas. Underemployment rate among females in urban areas was almost halved from 1993 to 2010. However, severe

[8] In Bahasa Indonesia, underemployment is referred as *Setengah penganggur* or the half-unemployed. See http://sirusa.bps.go.id/index.php?r=indikator/view&id=43.

underemployment has worsened among male and female workers in both urban and rural areas between 1993 and 2010.

Table 2: Average Hours Worked by Type of Employment, Gender, and Area (1993 and 2010)

	1993*				2010			
	Male		Female		Male		Female	
	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural
Average hours worked	47.1	40.2	45.4	35.6	46.9	38.9	42.9	32.3
Unemployment (%)	7.8	2.9	11.5	5.3	7.2	3.7	10.9	6.3
Severe underemployment (%)	1.7	4.1	3.1	7.3	3.7	6.0	3.7	7.4
Underemployment (%)	16.7	36.2	26.7	52.2	13.6	31.2	13.4	29.9

* See Table 6 in Mason and Baptist (1996).
Notes: Severe underemployment is less than 15 hours a week worked. Underemployment is defined as less than 35 hours a week worked. These are defined solely based on the number of hours worked in a week.
Sources: Mason and Baptist (1996) and Sakernas (2010).

In order to capture the gender wage differentials in detail, we run the Heckman's two-step model over different data samples: (i) all workers, (ii) full-time workers (i.e., those who worked 35 hours or more), (iii) underemployed workers (i.e., less than 35 hours a week), and (iv) severely underemployed workers (i.e., less than 15 hours a week). These data samples were further divided into urban and rural areas.

D. Predictive Margins of Real Wages

We estimate the relationship between predictive margins of the log of real wages and worker's characteristics using the following reduced form of the regression model.

$$y_i = x_i\beta + \varepsilon_i \tag{1}$$

where y_i is the log of worker i's real wage, x_i is the set of the factors that may influence it, including worker i's characteristics (i.e., gender, age, and education level), hours worked, work type, and industry type, and β is the set of parameters associated with these factors. Also, to distinguish any possible gender differences in the effect of these variables on real wage, gender interaction terms, such as gender–education level, gender–hours worked, gender–work type, and gender–industry type, are also included. Finally, ε represents errors that are assumed to be uncorrelated with the regressors. Full regression results are omitted here but available upon request.[9]

Summary statistics used to analyze the marginal effects to the log of the real wage are shown at Table A.1 in Appendix 1. Table 3 shows the definition of dummy variables used for the reduced model.

[9] Results were omitted for brevity.

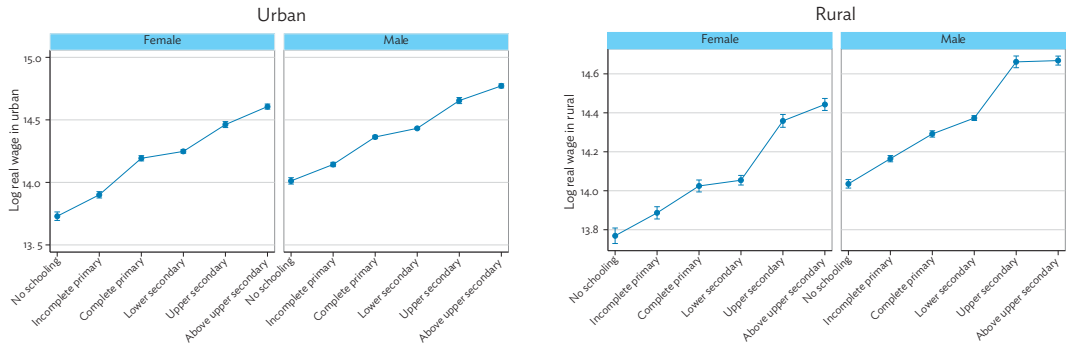
Table 3. Dummy Variables in the Reduced Model

Category	Classification
Gender	male, female
Educational level	no schooling, incomplete primary, complete primary, lower secondary, upper secondary, above upper secondary
Work type	Managers and professional, technicians, administrative staff, sales, farming, laborers, military and police
Industry category	agriculture, mining, manufacturing, construction and utilities, retail and hotel, transport and communication, finance and real estate, public administration, education, health and public services, personal and household services, international corporation

Source: Authors' classification.

Figure 12 shows predictive margins of real monthly wage, estimated by the reduced form regression model, based on educational attainment in urban and rural areas. The vertical line at each observation signifies 95% confidence intervals. Clearly, there is a positive correlation between the level of educational attainment and the higher monthly wage for both urban and rural areas. A minor exception is the upper secondary and above upper secondary levels in rural areas where the increment of the monthly wage is nearly zero. This is because of lack of job opportunities which require higher educational attainments in rural areas.

Figure 12: Predictive Margins of Urban and Rural Real Wages Based on Educational Attainment



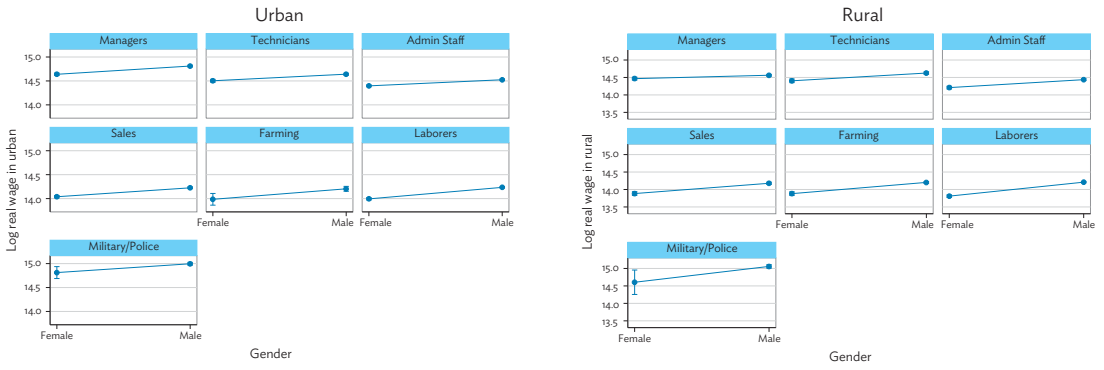
Source: Authors' calculation based on Sakernas (2010).

For each educational level attained by a worker, male workers receive higher wages in both urban and rural areas. Estimated by the reduced form regression, the results can be interpreted as the gender wage gap conditional on age, hours worked, and other variables which are in the regression model. Hence, we can conclude that a gender wage gap exists at every educational attainment in both urban and rural areas.

Figure 13 shows predictive margins of real monthly wage by type of work in urban and rural areas. These figures correspond to the type of work dummy variable, which include managers and professionals, technicians, administrative staff, sales, farming, laborers, and military and police. The vertical line at each observation signifies 95% confidence intervals. If the line is upward sloping, the female workers in the work category receive less than the male workers. The steepness of the slope indicates the magnitude of the wage gap, i.e., the steeper the line is, the wider is the wage gap between male and female.

The margins are higher across all types of work, suggesting that a gender wage gap exists across work types in favor of male workers. There is a wide margin of error for female military and police officers. This is due to lack of samples (i.e., few female military personnel and police officers). The steepness of the slope varies across types of work. Hence, even though the gender wage gap exists, the magnitude varies depending on the type of work. Except military and police, the lower wage type works (e.g., sales, farming, and laborers) tend to have steeper slopes than the higher wage type works (e.g., managers, technicians, and administrative staff).

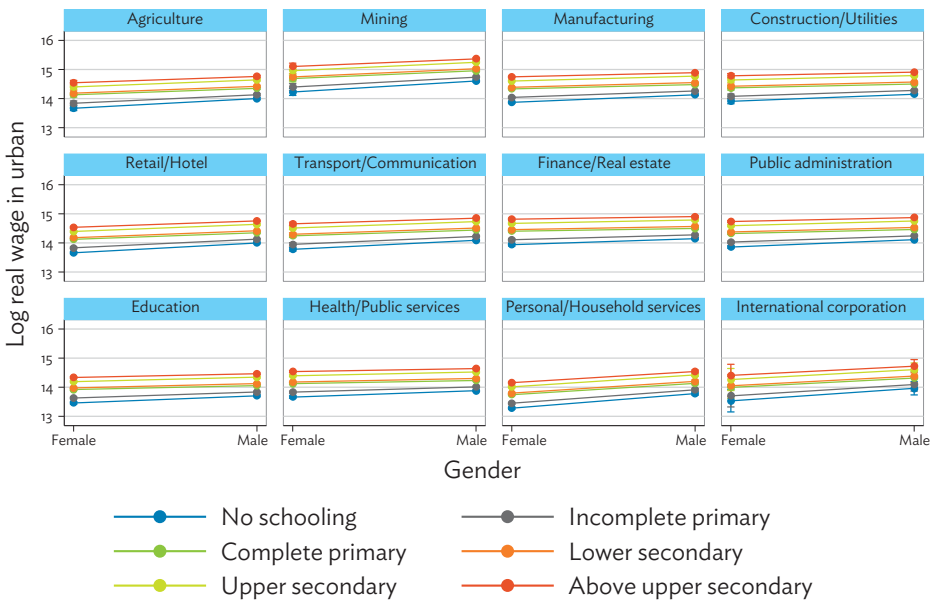
Figure 13: Predictive Margins of Urban and Rural Real Wages Based on Type of Work



Source: Authors' calculation based on Sakernas (2010).

Figures 14 and 15 show predictive margins of real wage based on industry category and educational attainment in urban and rural areas, respectively. Each line is further divided into different educational attainments of no schooling, incomplete primary, complete primary, lower secondary, upper secondary, and above upper secondary. As in the previous figures, the vertical line at each observation signifies 95% confidence intervals. If the line is upward sloping, female workers in the industry category receive less than the male workers. The slope indicates the magnitude of the gender wage gap, i.e., a steeper line represents a wider gap.

Figure 14: Predictive Margins of Urban Real Wage Based By Industry Category and Educational Attainment

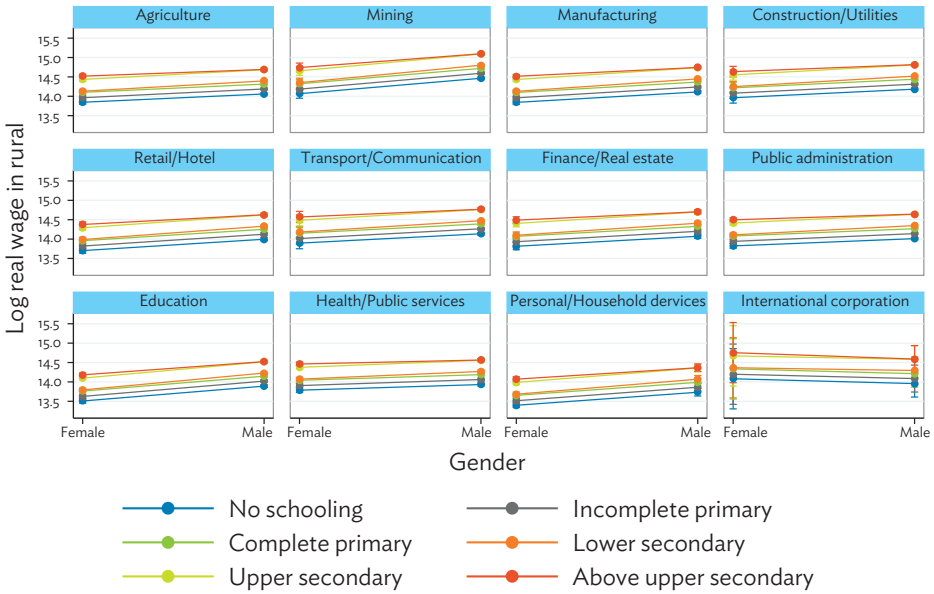


Source: Authors' calculation based on Sakernas (2010).

Except for the international corporation category in rural areas, all lines are upward sloping. This implies that the gender wage gap exists in all industry categories except the international corporation category in rural areas, and the wage gap is favorable for male workers. The downward sloping international corporation category in rural areas (i.e., the gender wage gap is favorable for female workers) is due to lack of observations for female workers in this industry in rural areas. Due to the large margin of errors, the gender wage gap for this category is inconclusive.

For all categories, the higher educational attainment workers have, the higher real wage workers receive. In rural areas, the workers with upper secondary and above upper secondary education tend to receive distinctively higher wages than other educational categories. This phenomenon is not visible in urban areas, and the wage increment vis-à-vis educational attainment in urban areas seems to have higher positive correlation.

Figure 15: Predictive Margins of Rural Real Wage Based By Industry Category and Educational Attainment



Source: Authors' calculation based on Sakernas (2010).

We can conclude again that the gender wage gap exists in the majority of the industry category. The male workers receive higher wages than the female workers in all industry categories in both urban and rural areas, except the inconclusive case for the international corporate category in rural areas.

IV. EMPIRICAL MODEL AND RESULTS

In the previous section, we saw that female workers tend to receive lower wages conditional on several socioeconomic characteristics across types of work and industry categories. A pertinent question is what the key determinants of the gender wage gap are. In order to answer this question, several issues need to be addressed.

A. Labor Force Participation Selection Bias

In the survey data, we only observe wages for those who have decided to participate in the labor market. For instance, a housewife might have a latent reservation wage to determine if she would join the labor force (Heckman 1974). With regard to women who have decided to work at home, this is an incidental truncation, instead of censoring, because data distribution is truncated at a certain reservation wage (Maddala 1983). Since incidental truncation refers to a sample that is not randomly selected, the selection bias needs to be corrected. In this regard, this paper applies for the Heckman selection model (Heckman 1976 and Greene 2000).

We are interested in testing whether there is any significant gender difference in the expected value of the real wage. In order to incorporate incidental truncation of the labor market participation, this study applies the Heckman's two-step model as follows:

Regression equation: $y_i = x_i\beta + \varepsilon_i$;
Selection equation: $w_i^* = z_i + u_i$, if $w_i^* > 0$, and $w_i = 0$ otherwise;
$Prob(w_i = 1|z_i) = \Phi(z_i\gamma)$ and $Prob(w_i = 0|z_i) = 1 - \Phi(z_i\gamma)$. (2)

For the regression equation, y is the log of the real wage, and x is a vector of age, a gender dummy (Table 3), an education level dummy (Table 3), a work type dummy (Table 3), an industry category dummy (Table 3), hours worked per week, a dummy if one has a secondary work, the years of experience at the current work, and 33 province dummies. For the selection equation, independent variables include a marital status dummy (i.e., single, married, and divorced and widowed), the number of children who is 10 years old and below, and the education level dummy (Table 3). Φ is the standard normal cumulative function, and ϕ is the standard normal density function. Since we are using the survey data, we applied sampling weights for regression estimates to conservatively handle standard errors. Summary statistics of variables are shown in Tables A.1 and A.2 in the Appendix.

In Heckman's two-step model, we include a selection bias variable denoted as *lambda*. *Lambda* is the inverse Mills ratio, which is calculated as the ratio of the probability density function to the cumulative distribution function for each wage worker. In Heckman's model, since sample selection can be defined as omitted-variable bias, inclusion of lambda will make an unbiased inference of all sample population.

B. Selection Bias Correction Based on the Multinomial Logit Model

Heckman's two-step model appropriately corrects selection bias and serves as the base model. However, one of the limitations of Heckman's two-step model includes the bivariate selection correction. For our study, the bivariate model for the full sample between the working and non-working corrects the possible bias appropriately by Heckman's model. Once the choice becomes multivariate like full-time, underemployment, severely underemployment, and unemployment, Heckman's two-step model with the bivariate probit procedure presents restrictions on the structure of the error terms. Hence, Heckman's two-step model with the bivariate selection correction could be a suboptimal application to correct selection bias in polytomous cases.

In order to overcome restrictions on the error terms for the bivariate selection model, we apply linear regression models where selectivity is modelled as a multinomial logit for some selected cases. Results from Heckman's two-step model where selectivity is modeled as univariate probit are compared with a multinomial logit for full-employment, underemployment, and severe underemployment cases.

For the regression, we follow a correction model for multivariate cases developed by Lee (1983) and Dubin and McFadden (1984). The mathematical exposition below closely follows Bourguignon, Fournier, and Gurgand (2007).

First, we set up the following model:

$$y_s = x_s\beta_s + u_s$$
$$y_s^* = z_s\gamma_s + \eta_s \quad s = 1 \ldots M \qquad (3)$$

where x_s and z_s are exogenous, and the error term, u_s, verifies $E(u_s|x,z) = 0$ and $V(u_s|x,z) = \sigma_s^2$. S is a categorical variable among M alternatives. The outcome variable y_s is observed if and only if the category s is chosen. β and γ are parameters, and η is an error term for the selection equation.

Following Heckman (1979) and Lee (1983), a consistent estimate of β_s would be obtained by least squares with:

$$y_s = x_s\beta_s - \sigma_s\rho_s \frac{\phi(J_s(z_s\gamma_s;\Gamma))}{F_s(z_s\gamma_s;\Gamma)} + \nu_s \tag{4}$$

where the cumulative distribution of ε_s for a given Γ and $\sum_{j\neq s} exp(z_j\gamma_j)$ is denoted as $F_s(\varepsilon;\Gamma) = \frac{exp(\varepsilon)}{exp(\varepsilon)+\Gamma}$ and $J_s(\varepsilon;\Gamma) = \Phi^{-1}(F_s(\varepsilon;\Gamma))$. Φ is the standard normal cumulative function, and ϕ is the standard normal density function. ρ_s is the correction between u_s and $J_s(\varepsilon;\Gamma)$. ν_s is an independent random term.

C. Oaxaca-Blinder Decomposition

In order to analyze the sources of the gender wage gap, this paper applies the Blinder-Oaxaca decomposition (Blinder 1973 and Oaxaca 1973). The Blinder-Oaxaca decomposition separates the wage differentials from explained to unexplained variations. Explained variations can be differences in socioeconomic characteristics like educational attainments, while unexplained variations can be derived from latent variables like gender discrimination. In this paper, we follow exposition by Jann (2008), which explains the decomposition technique in detail.

D. Empirical Results

Tables 4 to 6 show regression results from Heckman's two-step model to analyze the gender wage gap.[10] As expected, the gender dummy variables (i.e., one for male, while zero for female) in the regression results are significant and positive for all models we analyzed. These results imply that the real wage for the male workers under full employment, underemployment, and severe underemployment in both urban and rural areas is consistently higher than the female worker's wage on average.

Tables 4 to 6 also show that the inverse Mills ratio (lambda) under the all data category is significant for all, urban, and rural areas. This implies that there is a binding selection bias to enter the labor market. In general, marital status, the number of children below 10 years old, and educational attainments matter in the choice whether to enter the labor market or not. The most estimated coefficients of lambda are negative, except for underemployment, indicating that the potential real wage for the whole population is higher than the estimated wage for wage workers. This implies that observed skill sets for larger wages are underutilized in both urban and rural areas. There is a potential labor force who could earn larger wages if one decides to enter the labor market. The absolute value of lambda is larger in rural areas, compared to urban areas; hence, the magnitude of labor force underutilization is more severe in rural areas.

[10] Results from the multinominal logit model will be provided upon request.

Table 4: Regression Results from Heckman's Two-Step Model by Type of Employment (Full Data)

	All Data		Full Employment 35 Hours per Week or More		Underemployment Less than 35 Hours per Week		Severe Underemployment Less than 15 Hours per Week	
Dependent Variable: ln (Real Monthly Wage)	Regression Equation	Selection Equation	Regression Equation	Selection Equation	Regression Equation	Selection Equation	Regression Equation	Selection Equation
Explanatory variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age	0.00737***		0.00697***		0.00698***		0.00768***	
	(0.000325)		(0.000329)		(0.000733)		(0.00117)	
Male	0.178***		0.157***		0.281***		0.297***	
	(0.00767)		(0.00836)		(0.0126)		(0.0245)	
Technicians	-0.103***		-0.154***		0.0963***		0.0606	
(work type)	(0.0113)		(0.0116)		(0.0280)		(0.0669)	
Admin staff	-0.201***		-0.257***		-0.115***		-0.188***	
(work type)	(0.0101)		(0.0107)		(0.0208)		(0.0592)	
Sales	-0.504***		-0.550***		-0.359***		-0.396***	
(work type)	(0.0133)		(0.0138)		(0.0330)		(0.0712)	
Farming	-0.493***		-0.526***		-0.504***		-0.414***	
(work type)	(0.0166)		(0.0177)		(0.0391)		(0.0827)	
Laborers	-0.499***		-0.536***		-0.446***		-0.450***	
(work type)	(0.0123)		(0.0130)		(0.0287)		(0.0634)	
Military and police	0.147***		0.118***		0.341***		0.485***	
(work type)	(0.0150)		(0.0154)		(0.0475)		(0.132)	
Mining	0.474***		0.484***		0.460***		0.588***	
(industry type)	(0.0207)		(0.0226)		(0.0442)		(0.0904)	
Manufacturing	0.0388***		0.0234*		-0.0669**		0.134**	
(industry type)	(0.0134)		(0.0136)		(0.0305)		(0.0587)	
Construction and utilities	0.113***		0.0992***		0.132***		0.254***	
(industry type)	(0.0123)		(0.0123)		(0.0301)		(0.0548)	
Retail and hotel	-0.0761***		-0.0922***		0.0632*		0.302***	
(industry type)	(0.0134)		(0.0136)		(0.0345)		(0.0691)	
Transport and communication	0.0563***		0.0493***		0.173***		0.394***	

continued on next page

Table 4 continued

	All Data		Full Employment 35 Hours per Week or More		Underemployment Less than 35 Hours per Week		Severe Underemployment Less than 15 Hours per Week	
Dependent Variable: ln (Real Monthly Wage)	Regression Equation	Selection Equation	Regression Equation	Selection Equation	Regression Equation	Selection Equation	Regression Equation	Selection Equation
Explanatory variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(industry type)	(0.0139)		(0.0139)		(0.0326)		(0.0661)	
Finance and real estate	0.107***		0.0950***		0.0937**		0.321***	
(industry type)	(0.0145)		(0.0146)		(0.0460)		(0.0993)	
Public administration	0.000787		-0.00238		-0.0607*		0.0381	
(industry type)	(0.0132)		(0.0137)		(0.0320)		(0.0740)	
Education	-0.296***		-0.290***		-0.332***		-0.246***	
(industry type)	(0.0158)		(0.0157)		(0.0347)		(0.0713)	
Health and public services	-0.114***		-0.108***		-0.191***		-0.0968	
(industry type)	(0.0125)		(0.0127)		(0.0307)		(0.0645)	
Personal and household services	-0.424***		-0.379***		-0.466***		-0.169***	
(industry type)	(0.0158)		(0.0173)		(0.0307)		(0.0568)	
International corporation	-0.0289		-0.00240		-0.108		-0.252	
(industry type)	(0.0954)		(0.106)		(0.192)		(0.195)	
Incomplete primary	-0.0230**	0.164***	-0.0301***	0.188***	0.166***	0.0582***	0.124***	0.0515*
(educational attainment)	(0.0101)	(0.00862)	(0.0104)	(0.0099)	(0.0191)	(0.0133)	(0.0433)	(0.0269)
Complete primary	0.00285	0.348***	0.0307***	0.347***	0.341***	0.261***	0.0794*	0.218***
(educational attainment)	(0.0112)	(0.0109)	(0.0110)	(0.0118)	(0.0228)	(0.0147)	(0.0465)	(0.0283)
Lower secondary	-0.346***	0.897***	-0.228***	0.851***	0.496***	1.005***	-0.327***	0.867***
(educational attainment)	(0.0138)	(0.0121)	(0.0137)	(0.0133)	(0.0364)	(0.0172)	(0.0584)	(0.0260)
Upper secondary	-0.655***	1.739***	-0.424***	1.522***	1.050***	2.451***	-0.694***	2.130***
(educational attainment)	(0.0249)	(0.0167)	(0.0233)	(0.0173)	(0.0645)	(0.0269)	(0.102)	(0.0496)
Above upper secondary	0.0421**	0.891***	0.116***	0.800***	1.043***	1.106***	0.185***	1.000***
(educational attainment)	(0.0179)	(0.0155)	(0.0175)	(0.0164)	(0.0426)	(0.0223)	(0.0697)	(0.0360)
Hours worked	0.0066***		0.00216***		0.00986***		-0.0338***	
	(0.0002)		(0.0004)		(0.0005)		(0.0021)	

continued on next page

Table 4 continued

Dependent Variable: ln (Real Monthly Wage)	All Data		Full Employment 35 Hours per Week or More		Underemployment Less than 35 Hours per Week		Severe Underemployment Less than 15 Hours per Week	
	Regression Equation	Selection Equation	Regression Equation	Selection Equation	Regression Equation	Selection Equation	Regression Equation	Selection Equation
Explanatory variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Other job dummy	0.0134**		0.0117*		0.0328***		-0.00847	
	(0.0058)		(0.0065)		(0.0103)		(0.0208)	
Experience	0.0201***		0.0166***		0.0328***		0.0281***	
	(0.000384)		(0.000374)		(0.000907)		(0.00173)	
Married		0.0711***		0.104***		-0.161***		-0.0276
		(0.0093)		(0.0111)		(0.0151)		(0.0231)
Divorced and widowed		-0.271***		-0.294***		-0.403***		-0.319***
		(0.0145)		(0.0168)		(0.0218)		(0.0384)
Number of children below 10		0.0070***		0.0010		0.0192***		0.0383***
		(0.0026)		(0.0031)		(0.0057)		(0.0071)
Constant	15.18***	-1.505***	15.28***	-1.478***	12.59***	-1.436***	14.92***	-1.483***
	(0.0368)	(0.0129)	(0.0399)	(0.0148)	(0.0898)	(0.0182)	(0.148)	(0.0316)
Lambda	-0.906***		-0.781***		0.219***		-0.962***	
	(0.0103)		(0.0106)		(0.0370)		(0.0543)	
Provincial dummy	Yes	No	Yes	No	Yes	No	Yes	No
Number of observations	826,715	826,715	624,695	624,695	202,020	202,020	46,535	46,535

*** p<0.01, ** p<0.05, * p<0.1.
Notes: Standard errors in parentheses. Estimates of provincial dummy coefficients are omitted but available upon request.
Source: Authors' calculation based on Sakernas (2010).

Table 5: Regression Results from Heckman's Two-Step Model (Urban Data Only)

	All Data		Full Employment 35 Hours per Week or More		Underemployment Less than 35 Hours per Week		Severe Underemployment Less than 15 Hours per Week	
Dependent Variable: ln (Real Monthly Wage)	Regression Equation	Selection Equation	Regression Equation	Selection Equation	Regression Equation	Selection Equation	Regression Equation	Selection Equation
Explanatory variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age	0.00820***		0.00828***		0.00674***		0.00310*	
	(0.000412)		(0.000424)		(0.00107)		(0.00173)	
Male	0.155***		0.141***		0.240***		0.284***	
	(0.0102)		(0.0109)		(0.0187)		(0.0381)	
Technicians	-0.167***		-0.202***		0.0241		-0.0109	
(work type)	(0.0135)		(0.0137)		(0.0380)		(0.0841)	
Admin staff	-0.249***		-0.294***		-0.108***		-0.216***	
(work type)	(0.0120)		(0.0121)		(0.0308)		(0.0735)	
Sales	-0.556***		-0.586***		-0.393***		-0.524***	
(work type)	(0.0152)		(0.0152)		(0.0448)		(0.0938)	
Farming	-0.573***		-0.586***		-0.655***		-0.573***	
(work type)	(0.0343)		(0.0351)		(0.0893)		(0.160)	
Laborers	-0.561***		-0.582***		-0.521***		-0.521***	
(work type)	(0.0150)		(0.0152)		(0.0419)		(0.0873)	
Military and police	0.0691***		0.0515***		0.335***		0.440***	
(work type)	(0.0186)		(0.0187)		(0.0640)		(0.122)	
Mining	0.627***		0.622***		0.580***		0.714***	
(industry type)	(0.0402)		(0.0385)		(0.0949)		(0.146)	
Manufacturing	0.0964***		0.0655***		1.63e-05		0.216**	
(industry type)	(0.0252)		(0.0241)		(0.0689)		(0.104)	
Construction and utilities	0.149***		0.118***		0.211***		0.416***	
(industry type)	(0.0245)		(0.0231)		(0.0689)		(0.100)	
Retail and hotel	-0.0248		-0.0526**		0.0939		0.442***	
(industry type)	(0.0242)		(0.0224)		(0.0714)		(0.113)	
Transport and communication	0.0901***		0.0739***		0.186***		0.483***	

continued on next page

Table 5 continued

	All Data		Full Employment 35 Hours per Week or More		Underemployment Less than 35 Hours per Week		Severe Underemployment Less than 15 Hours per Week	
Dependent Variable: ln (Real Monthly Wage)	Regression Equation	Selection Equation	Regression Equation	Selection Equation	Regression Equation	Selection Equation	Regression Equation	Selection Equation
Explanatory variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(industry type)	(0.0252)		(0.0235)		(0.0689)		(0.112)	
Finance and real estate	0.178***		0.151***		0.175**		0.557***	
(industry type)	(0.0255)		(0.0248)		(0.0807)		(0.133)	
Public administration	0.0886***		0.0685***		0.0150		0.210*	
(industry type)	(0.0250)		(0.0248)		(0.0636)		(0.117)	
Education	-0.292***		-0.287***		-0.343***		-0.195*	
(industry type)	(0.0273)		(0.0264)		(0.0666)		(0.113)	
Health and public services	-0.0799***		-0.0888***		-0.158**		-0.0613	
(industry type)	(0.0242)		(0.0228)		(0.0638)		(0.109)	
Personal and household services	-0.367***		-0.329***		-0.423***		-0.105	
(industry type)	(0.0271)		(0.0267)		(0.0655)		(0.0995)	
International corporation	0.0380		0.0579		-0.748***		-0.563	
(industry type)	(0.113)		(0.107)		(0.262)		(0.387)	
Incomplete primary	0.0166	0.151***	0.00375	0.172***	0.193***	0.0454*	0.181***	-0.0158
(educational attainment)	(0.0142)	(0.0126)	(0.0145)	(0.0140)	(0.0338)	(0.0274)	(0.0619)	(0.0469)
Complete primary	0.185***	0.257***	0.173***	0.266***	0.448***	0.192***	0.312***	0.0210
(educational attainment)	(0.0152)	(0.0150)	(0.0157)	(0.0163)	(0.0358)	(0.0269)	(0.0636)	(0.0501)
Lower secondary	-0.0114	0.676***	0.0187	0.676***	0.534***	0.704***	0.535***	0.454***
(educational attainment)	(0.0168)	(0.0166)	(0.0170)	(0.0185)	(0.0414)	(0.0257)	(0.0720)	(0.0483)
Upper secondary	-0.132***	1.312***	-0.0540**	1.202***	1.027***	1.851***	1.138***	1.425***
(educational attainment)	(0.0260)	(0.0199)	(0.0256)	(0.0214)	(0.0642)	(0.0410)	(0.117)	(0.0702)
Above upper secondary	0.120***	1.112***	0.168***	1.021***	1.183***	1.513***	1.275***	1.217***
(educational attainment)	(0.0236)	(0.0181)	(0.0232)	(0.0192)	(0.0590)	(0.0336)	(0.113)	(0.0599)
Hours worked	0.0054***		0.0014***		0.0068***		-0.0368***	
	(0.0003)		(0.0005)		(0.0007)		(0.0030)	

continued on next page

Table 5 continued

	All Data		Full Employment 35 Hours per Week or More		Underemployment Less than 35 Hours per Week		Severe Underemployment Less than 15 Hours per Week	
Dependent Variable: ln (Real Monthly Wage)	**Regression Equation**	**Selection Equation**	**Regression Equation**	**Selection Equation**	**Regression Equation**	**Selection Equation**	**Regression Equation**	**Selection Equation**
Explanatory variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Other job dummy	0.0185**		0.0192**		0.0271		-0.0322	
	(0.0083)		(0.0093)		(0.0169)		(0.0321)	
Experience	0.0193***		0.0164***		0.0325***		0.0295***	
	(0.0005)		(0.0005)		(0.0013)		(0.0023)	
Married		0.0653***		0.0778***		-0.199***		-0.220***
		(0.0151)		(0.0175)		(0.0232)		(0.0428)
Divorced and widowed		-0.311***		-0.351***		-0.435***		-0.568***
		(0.0221)		(0.0250)		(0.0381)		(0.0613)
Number of children below 10		0.0198***		0.0164***		0.0395***		0.0735***
		(0.00438)		(0.00511)		(0.00976)		(0.0156)
Constant	14.66***	-1.169***	14.83***	-1.183***	12.75***	-0.923***	12.82***	-0.806***
	(0.0511)	(0.0186)	(0.0579)	(0.0209)	(0.113)	(0.0306)	(0.206)	(0.0547)
Lambda	-0.754***		-0.656***		0.173***		0.259***	
	(0.0120)		(0.0123)		(0.0360)		(0.0904)	
Provincial dummy	Yes	No	Yes	No	Yes	No	Yes	No
Number of observations	295,141	295,141	255,305	255,305	39,836	39,836	10,854	10,854

*** p<0.01, ** p<0.05, * p<0.1.
Notes: Standard errors in parentheses. Estimates of provincial dummy coefficients are omitted but available upon request.
Source: Authors' calculation based on Sakernas (2010).

Table 6: Regression Results from Heckman's Two-Step Model (Rural Data Only)

	All Data		Full Employment 35 Hours per Week and More		Underemployment Less than 35 Hours per Week		Severe Underemployment Less than 15 Hours per Week	
Dependent Variable: ln (Real Monthly Wage)	Regression Equation	Selection Equation	Regression Equation	Selection Equation	Regression Equation	Selection Equation	Regression Equation	Selection Equation
Explanatory variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Age	0.0062***		0.0051***		0.0074***		0.0087***	
	(0.0005)		(0.0005)		(0.0010)		(0.0017)	
Male	0.209***		0.184***		0.309***		0.347***	
	(0.0090)		(0.0095)		(0.0163)		(0.0330)	
Technicians	0.0106		-0.0541***		0.167***		0.169*	
(work type)	(0.0174)		(0.0190)		(0.0403)		(0.0982)	
Admin staff	-0.135***		-0.187***		-0.132***		-0.194**	
(work type)	(0.0160)		(0.0179)		(0.0297)		(0.0829)	
Sales	-0.389***		-0.451***		-0.330***		-0.262***	
(work type)	(0.0230)		(0.0249)		(0.0493)		(0.0974)	
Farming	-0.363***		-0.413***		-0.393***		-0.302***	
(work type)	(0.0215)		(0.0243)		(0.0488)		(0.104)	
Laborers	-0.362***		-0.419***		-0.333***		-0.336***	
(work type)	(0.0181)		(0.0204)		(0.0406)		(0.0840)	
Military and police	0.288***		0.254***		0.317***		0.521*	
(work type)	(0.0229)		(0.0242)		(0.0741)		(0.304)	
Mining	0.405***		0.415***		0.429***		0.574***	
(industry type)	(0.0217)		(0.0237)		(0.0510)		(0.120)	
Manufacturing	0.0597***		0.0543***		-0.0566		0.166**	
(industry type)	(0.0146)		(0.0146)		(0.0356)		(0.0758)	
Construction and utilities	0.144***		0.140***		0.125***		0.212***	
(industry type)	(0.0138)		(0.0143)		(0.0351)		(0.0725)	
Retail and hotel	-0.0169		-0.0354*		0.122***		0.251**	
(industry type)	(0.0174)		(0.0182)		(0.0449)		(0.0987)	
Transport and communication	0.107***		0.0974***		0.235***		0.396***	

continued on next page

Table 6 continued

	All Data		Full Employment 35 Hours per Week and More		Underemployment Less than 35 Hours per Week		Severe Underemployment Less than 15 Hours per Week	
Dependent Variable: ln (Real Monthly Wage)	Regression Equation	Selection Equation	Regression Equation	Selection Equation	Regression Equation	Selection Equation	Regression Equation	Selection Equation
Explanatory variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(industry type)	(0.0167)		(0.0176)		(0.0419)		(0.0878)	
Finance and real estate	0.0934***		0.0897***		0.0731		0.110	
(industry type)	(0.0253)		(0.0268)		(0.0638)		(0.156)	
Public administration	0.0167		0.0212		-0.0197		0.0775	
(industry type)	(0.0173)		(0.0180)		(0.0403)		(0.0964)	
Education	-0.143***		-0.135***		-0.206***		-0.103	
(industry type)	(0.0208)		(0.0217)		(0.0440)		(0.0949)	
Health and public services	-0.0313*		-0.00787		-0.123***		0.0102	
	(0.0179)		(0.0184)		(0.0405)		(0.0884)	
Personal and household services	-0.312***		-0.289***		-0.362***		-0.0826	
(industry type)	(0.0227)		(0.0220)		(0.0535)		(0.0871)	
International corporation	0.0275		0.0232		0.181		0.133	
(industry type)	(0.153)		(0.276)		(0.227)		(0.316)	
Incomplete primary	-0.0462***	0.155***	-0.0632***	0.183***	0.153***	0.0626***	0.0996*	0.0761**
(educational attainment)	(0.0147)	(0.0113)	(0.0157)	(0.0133)	(0.0220)	(0.0156)	(0.0557)	(0.0336)
Complete primary	-0.0639***	0.297***	-0.0430***	0.301***	0.273***	0.240***	0.0460	0.242***
(educational attainment)	(0.0164)	(0.0137)	(0.0161)	(0.0151)	(0.0282)	(0.0181)	(0.0641)	(0.0369)
Lower secondary	-0.521***	0.883***	-0.380***	0.818***	0.428***	1.013***	-0.417***	0.930***
(educational attainment)	(0.0188)	(0.0146)	(0.0186)	(0.0152)	(0.0579)	(0.0220)	(0.0944)	(0.0336)
Upper secondary	-1.131***	1.980***	-0.833***	1.711***	0.990***	2.646***	-0.857***	2.484***
(educational attainment)	(0.0337)	(0.0204)	(0.0318)	(0.0233)	(0.117)	(0.0338)	(0.186)	(0.0670)
Above upper secondary	0.0942***	0.615***	0.182***	0.466***	0.885***	0.869***	0.193**	0.840***
(educational attainment)	(0.0231)	(0.0196)	(0.0240)	(0.0216)	(0.0563)	(0.0261)	(0.0959)	(0.0439)
Hours worked	0.00796***		0.00377***		0.0120***		-0.0315***	
	(0.000223)		(0.000316)		(0.000660)		(0.00268)	

continued on next page

Table 6 continued

	All Data		Full Employment 35 Hours per Week and More		Underemployment Less than 35 Hours per Week		Severe Underemployment Less than 15 Hours per Week	
Dependent Variable: ln (Real Monthly Wage)	**Regression Equation**	**Selection Equation**	**Regression Equation**	**Selection Equation**	**Regression Equation**	**Selection Equation**	**Regression Equation**	**Selection Equation**
Explanatory variables	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Other job dummy	0.00516		-0.000402		0.0406***		0.0268	
	(0.00637)		(0.00645)		(0.0135)		(0.0282)	
Experience	0.0213***		0.0170***		0.0328***		0.0286***	
	(0.000559)		(0.000577)		(0.00115)		(0.00258)	
Married		0.0916***		0.141***		-0.158***		-0.0538*
		(0.00887)		(0.0104)		(0.0181)		(0.0293)
Divorced and widowed		-0.214***		-0.201***		-0.392***		-0.343***
		(0.0151)		(0.0174)		(0.0268)		(0.0487)
Number of children below 10		0.00911***		0.000622		0.0326***		0.0399***
		(0.00275)		(0.00313)		(0.00604)		(0.00865)
Constant	15.47***	-1.663***	15.59***	-1.648***	12.46***	-1.572***	14.71***	-1.635***
	(0.0489)	(0.0146)	(0.0452)	(0.0164)	(0.143)	(0.0225)	(0.255)	(0.0414)
Lambda	-1.068***		-0.951***		0.211***		-0.870***	
	(0.0138)		(0.0137)		(0.0646)		(0.0984)	
Provincial dummy	Yes	No	Yes	No	Yes	No	Yes	No
Number of observations	531,574	531,574	369,390	369,390	162,184	162,184	35,681	35,681

*** p<0.01, ** p<0.05, * p<0.1.
Notes: Standard errors in parentheses. Estimates of provincial dummy coefficients are omitted but available upon request.
Source: Authors' calculation based on Sakernas (2010).

Since our model is a semi-logarithmic regression model (i.e., the dependent variable is in a log form, while the gender variable is a dummy variable), we need to transform the coefficient in order to interpret the gender dummy. The gender wage gap in percentage form is represented as $100 * [e^{-\beta} - 1]$ where e is the base of the natural logarithm (i.e., Napier's constant), and β is a coefficient of the gender dummy variable (Giles 2011). Using the regression results, Table 7 shows the gender wage difference for full-time employment, underemployment, and severe underemployment.

Table 7: Gender Wage Differentials By Employment Category
(%)

	Overall	Full Employment (>= 35 hours per week)		Underemployment (< 35 hours per week)		Severe Underemployment (< 15 hours per week)	
	HTR	HTR	MLR	HTR	MLR	HTR	MLR
Overall	30.8	31.4	32.9	27.8	31.1	27.3	23.3
Urban	31.5	32.0	31.8	28.9	30.9	27.7	26.3
Rural	29.9	30.6	32.6	27.0	31.3	26.0	25.8

HTR = Heckman two-step regression, MLR = multinomial logit regression.
Notes: HTR with bivariate selection correction results. MLR with selection bias correction results (available upon request). 30.8 means that the female worker's monthly real wage is 30.8% percent lower than the male worker's wage in the same category.
Source: Authors' calculation based on Sakernas (2010).

For the estimation method, Heckman two-step represents Heckman's two-step with the bivariate selection correction. Multi logit represents the selection bias correction based on the multinomial logit model.

Overall, the gender wage difference is prevalent for any worker in the urban or rural areas. Estimation results are, in general, similar for full employment and severe underemployment between two different estimation methods, while there are some differences for underemployment. Depending on the estimation method, the wage gap varies from 23.3% (in severe underemployment in rural areas by the multinomial logit) to 32.9% (in full employment in urban areas by the multinomial logit). These results are generally consistent with previous findings on the Indonesia gender wage gap based on the 2008 labor survey data (Tijdens and van Klaveren 2012). It is rather surprising to see that the gender wage gap is wider for full employment compared to underemployment and severe underemployment. Since wages for fully employed workers are higher, there are wider margins for the gender wage gap compared to underemployed and severely underemployed workers.

We further applied the same Heckman's two-step and multinominal logit model for different age cohorts.[11] Table 8 summarizes the gender wage differentials. Overall, the gender wage differential varies from 39.3% (full employment and in their 50's in rural by the multi-logit nominal model) to 17.9% (severely underemployed and below 20's in rural by the multi-logit model). In urban areas, there is tendency, albeit weakly, that the gender wage gap becomes less as the age cohort goes up. In rural areas, there is no such tendency, and the wage gap seems to be random to the age cohort.

[11] Regression results will be provided upon request.

Table 8: Gender Wage Differentials By Employment Category and Age Cohort
(%)

Age Cohort	Full Employment (>= 35 hours per week)				Underemployment (< 35 hours per week)				Severe Underemployment (< 15 hours per week)			
	Urban		Rural		Urban		Rural		Urban		Rural	
	HTR	MLR	HTR	MLR	HTR	MLR	HTR	MLR	HTR	MLR	HTR	MLR
Below 19 years old	34.1	35.0	32.1	31.8	36.2	38.1	29.5	29.1	38.8	N/A[a]	26.2	17.9
20–29 years old	33.3	33.1	32.2	33.2	30.7	29.8	29.9	30.3	33.7	28.4	30.0	31.2
30–39 years old	30.9	31.2	29.6	30.1	27.6	31.1	26.3	32.2	22.6	20.3	23.7	22.1
40–49 years old	30.3	31.2	30.1	33.7	27.3	28.7	28.4	33.6	26.1	25.0	25.0	20.6
50–59 years old	30.7	29.6	31.0	39.3	29.0	32.0	29.5	35.5	25.9	28.0	28.7	38.7

HTR = Heckman two-step regression, MLR = multinomial logit regression, N/A = not available.
Notes:
[a] The multinomial logit model does not converge. A result of 34.1 means that the female worker's monthly real wage is 34.1% lower than the male worker's wage on average in the category.
Source: Authors' calculation based on Sakernas (2010).

Table 9 summarizes the gender wage gap in different industry categories, based on Heckman's two-step regression results.[12] The widest gender wage gap turns out to be finance and real estate, which is followed by public administration. In fact, the public sector categories (i.e., public administration, education, and health and public services) tend to have wider gender wage gap compared to other industry categories. Results are consistent with previous findings that the gender wage gap is wider in urban areas than in rural areas.

Table 9: The Gender Wage Differentials Based on Industry Category
(%)

Industry Category	All	Urban	Rural
Agriculture	28.1	26.5	28.3
Mining	26.1	28.0	25.5
Manufacturing	28.4	29.9	26.0
Construction and utilities	30.6	32.3	28.2
Retail and hotel	29.5	30.1	27.9
Transport and communication	31.8	32.4	29.5
Finance and Real estate	33.1	33.4	31.1
Public administration	33.0	32.9	32.5[a]
Education	32.9	32.7	33.3
Health and public services	32.3	32.3	32.6
Personal and household services	22.6	22.5	22.8
International corporation[b]	N/A	N/A	N/A

N/A = not available.
Notes:
[a] The linear regression of equation 2 is used instead of Heckman's two-step model because the model did not converge for rural areas.
[b] Results are insignificant.
28.1 means that the female worker's monthly real wage is 28.1% lower than the male worker's wage on average in the category.
Source: Authors' calculation based on Sakernas (2010).

Table 10 shows the regression results of the Oaxaca-Blinder decomposition for the monthly real wage differentials for male and female workers. The base regression model is the Heckman's two-step

[12] Regression results will be provided upon request.

model (Equation 2). The expected log wage for female and male is estimated based on the regressors associated with female and male, respectively. The difference value is the outcome differential of the log gender wage gap. The difference value (i.e., the gender wage gap) is further decomposed to an explained and unexplained part. The explained part is due to socioeconomic characteristics (like age or educational attainment) which each worker has. The unexplained part is due to latent variables. Hence, we define the unexplained part of the wage gap as the discrimination effect. The degree of gender discrimination is the share of the unexplained part to the wage difference value. For instance, 93.3% means that 93.3% of the wage differential cannot be explained by the socioeconomic variables in the regression model; hence, this is the share of gender discrimination in the wage differential.

The Oaxaca-Blinder decomposition results indicate that 93.3% of the wage differential for the real monthly wage is not explained by socioeconomic characteristics. Hence, the gender wage gap is largely due to gender discrimination, which is not related to any worker's attributes. Wage gender discrimination is more severe in rural areas than urban areas. In urban areas, gender discrimination to wage differentials is 87.5%, while it registers 96.7% in rural areas.

Table 10: Oaxaca-Blinder Decomposition of Gender Wage Gap

	All	Urban	Rural
Predicted log wage (Female)	14.13***	14.18***	14.05***
	(0.0098)	(0.0129)	(0.0134404)
Predicted log wage (Male)	15.41***	15.13***	15.71***
	(0.0178)	(0.0188)	(0.0224)
Difference	-1.274***	-0.942***	-1.661***
	(0.0197)	(0.0223)	(0.0261)
Explained	-0.0848***	-0.118***	-0.0551***
	(0.0088)	(0.0134)	(0.0107)
Unexplained	-1.189***	-0.824***	-1.606***
	(0.0183)	(0.0188)	(0.0249)
Degree of gender discrimination (% share of unexplained to difference)	93.3	87.5	96.7

*** p<0.01, ** p<0.05, * p<0.1.
Notes: Standard errors in parentheses. Estimates of provincial dummy coefficients are omitted but available upon request. The Oaxaca-Blinder decomposition is applied for Heckman's two-step model. For the degree of gender discrimination, 93.3% means 93.3% of the wage differential is unexplained by socioeconomic variables in the regression model; hence, 93.3% of the wage differential is due to gender discrimination.
Source: Authors' calculation based on Sakernas (2010).

Finally, we multiply the gender wage gap by the share of gender discrimination, which represents the gender wage gap attributable to gender discrimination. Even though the gender wage gap is wider in urban areas, the gender gap attributable to discrimination is lower because the proportion that is due to discrimination is lower in urban areas. In general, the gender wage gap attributable to discrimination is very similar for both urban and rural areas.

Table 11: Gender Wage Gap Attributable to Discrimination
(%)

	Gender Wage Gap	Gender Discrimination	Gender Wage Gap Attributable to Discrimination
All	30.8	93.3	28.7
Urban	31.5	87.5	27.6
Rural	29.9	96.7	28.9

Source: Authors' calculation based on Sakernas (2010).

V. SUMMARY FINDINGS AND POLICY RECOMMENDATION

Urbanization, a ubiquitous phenomenon in a developing world, can be a dynamic component of the national poverty reduction process as long as the right conditions are met, and the appropriate policies are placed (UNFPA 2007). The historical process of economic development is usually characterized by an economic shift from the rural areas toward urban areas. In the process, the productivity of manufacturing and services sectors in urban areas rises, and accelerates income growth further. The rise of the middle class in the urban areas reinforces the endogenous growth process.

This study analyzes the real gender wage gap, with focus on urban workers. On average, workers in the urban areas receive higher real wages than those in the rural areas. However, conditional on age, hours worked, educational attainment, work type, industry type, and geographical location men's real wage is consistently higher than female's. On average, female's real wage is 30.8% lower than male's. The gender wage gap exists for any age cohort.

We can summarize the regression results according to the following hypotheses concerning gender wage gap in Indonesia.

Hypothesis 1: Gender Wage Gap Exists in Indonesia

Regression results consistently indicate that the gender wage gap exists in Indonesia. Women do receive lower wages than men conditional on socioeconomic characteristics in both urban and rural areas.

Hypothesis 2: Urban Gender Wage Gap is Less than the Gap in Rural Areas

Contrary to our initial hypothesis, the gender wage gap is wider in urban areas, after controlling for socioeconomic variables. However, the gender discrimination part of the gender wage gap (i.e., unexplained portion of the wage differential) is larger in rural areas than in urban areas. This might imply that the labor market is more efficient in urban areas than in rural areas.

Unlike in rural areas, urban gender wage gap is wider among younger age cohorts. It is also wider in the public sector than in the private sector. These results are robust to model specification (i.e., use of Heckman's two-step model is applied).

Hypothesis 3: Socioeconomic Characteristics are Key Determinants of Gender Wage Gap

The Blinder-Oaxaca decomposition reveals that the critical determinant of the gender wage gap is gender discrimination, rather than socioeconomic characteristics. Educational attainment, the number of children below 10 years old, and the marital status are all significant when one decides to enter the labor market. Significant inverse Mills ratios in Heckman's two-step model confirm that the selection bias exists. However, education attainment dummy variables are not always significant to determine the level of the monthly wage. Other characteristics like hours worked or years of experience at the current position significantly determine the monthly wage level.

In sum, even though the average wage is higher in urban areas, the benefits of urbanization, in terms of monthly wages, tend to fall more on male workers rather than female workers. Based on the age cohort analysis, results suggest that public policy needs to focus on female workers in urban areas

who are trapped in the lower real wage range. Also, the gender wage gap in public sectors is more prevalent than private sectors, which can be fixed by the policy change immediately.

We conclude that the labor market in Indonesia could be imperfect. There is significant gender wage gap by job type, industry category, work duration, age cohort, and educational attainment. The female worker's wage is consistently and significantly lower than the male worker's wage due to non-market reasons.

This study proposes that the government implement an equal opportunity act for male and female workers who should be given equal access to work and receive equal wage given the same credentials. In particular, the government can immediately fix the gender wage gap in the public sector. It is neither fair nor efficient that equally qualified female workers receive lower wages than male workers. The Indonesian government needs to aim for more equitable growth in view of urbanization. The economy might exhibit high but unequal growth in the future, and this is something policy makers need to address.

APPENDIX

Appendix Table 1: Summary Statistics of Variables for the Margin Plot *(continued)*

		All									
		Female					Male				
		Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
Log real wage		49,242	14.13	0.86	11.54	18.41	90,783	14.42	0.75	11.41	18.84
Age		415,228	38.47	16.27	15	98	411,487	38.19	16.17	15	98
Hours worked (week)		205,512	35.77	18.87	0	98	327,739	41.52	17.35	0	98
Education level											
1	No schooling	415,228	0.19	0.39	0	1	411,487	0.18	0.38	0	1
2	Incomplete primary	415,228	0.26	0.44	0	1	411,487	0.27	0.44	0	1
3	Complete primary	415,228	0.21	0.41	0	1	411,487	0.22	0.41	0	1
4	Lower secondary	415,228	0.19	0.39	0	1	411,487	0.24	0.43	0	1
5	Upper secondary	415,228	0.03	0.16	0	1	411,487	0.02	0.14	0	1
6	Above upper secondary	415,228	0.11	0.32	0	1	411,487	0.08	0.27	0	1
Work type											
1	Managers and professionals	205,512	0.09	0.28	0	1	327,739	0.06	0.23	0	1
2	Technicians	205,512	0.02	0.13	0	1	327,739	0.02	0.14	0	1
3	Administrative Staff	205,512	0.04	0.20	0	1	327,739	0.04	0.19	0	1
4	Sales	205,512	0.22	0.42	0	1	327,739	0.12	0.32	0	1
5	Farming	205,512	0.45	0.50	0	1	327,739	0.46	0.50	0	1
6	Laborers	205,512	0.18	0.38	0	1	327,739	0.30	0.46	0	1
7	Military and police	205,512	0.00	0.02	0	1	327,739	0.01	0.10	0	1
Industry type											
1	Agriculture	205,512	0.46	0.50	0	1	327,739	0.48	0.50	0	1
2	Mining	205,512	0.00	0.07	0	1	327,739	0.02	0.15	0	1
3	Manufacturing	205,512	0.10	0.30	0	1	327,739	0.08	0.26	0	1
4	Construction and utilities	205,512	0.00	0.06	0	1	327,739	0.07	0.26	0	1
5	Retail and hotel	205,512	0.24	0.43	0	1	327,739	0.13	0.34	0	1
6	Transport and Communication	205,512	0.01	0.09	0	1	327,739	0.07	0.25	0	1
7	Finance and real Estate	205,512	0.01	0.09	0	1	327,739	0.01	0.11	0	1
8	Public Administration	205,512	0.03	0.16	0	1	327,739	0.05	0.21	0	1
9	Education	205,512	0.07	0.26	0	1	327,739	0.03	0.18	0	1
10	Health and public Services	205,512	0.04	0.20	0	1	327,739	0.04	0.20	0	1
11	Personal and household Services	205,512	0.03	0.17	0	1	327,739	0.01	0.07	0	1
12	International Corp	205,512	0.00	0.02	0	1	327,739	0.00	0.02	0	1

Obs = observations, Std. Dev. = standard deviation.
Source: Authors' calculation based on Sakernas (2010).

Appendix Table 1: Summary Statistics of Variables for the Margin Plot *(continued)*

		Urban									
		Female					Male				
		Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
Log real wage		29,591	14.18	0.83	11.54	18.41	49,067	14.48	0.75	11.41	18.84
Age		149,730	37.91	16.03	15	98	145,559	37.77	15.78	15	98
Hours worked (week)		66,808	42.93	19.75	0	98	107,875	46.94	17.55	0	98
Education level											
1	No schooling	149,730	0.13	0.34	0	1	145,559	0.10	0.30	0	1
2	Incomplete primary	149,730	0.19	0.39	0	1	145,559	0.18	0.38	0	1
3	Complete primary	149,730	0.22	0.42	0	1	145,559	0.24	0.42	0	1
4	Lower secondary	149,730	0.31	0.46	0	1	145,559	0.36	0.48	0	1
5	Upper secondary	149,730	0.04	0.20	0	1	145,559	0.03	0.17	0	1
6	Above upper secondary	149,730	0.10	0.30	0	1	145,559	0.09	0.29	0	1
Work type											
1	Managers and professionals	66,808	0.14	0.34	0	1	107,875	0.09	0.29	0	1
2	Technicians	66,808	0.04	0.19	0	1	107,875	0.04	0.20	0	1
3	Administrative Staff	66,808	0.10	0.30	0	1	107,875	0.08	0.27	0	1
4	Sales	66,808	0.37	0.48	0	1	107,875	0.22	0.41	0	1
5	Farming	66,808	0.08	0.28	0	1	107,875	0.12	0.33	0	1
6	Laborers	66,808	0.27	0.45	0	1	107,875	0.43	0.50	0	1
7	Military and police	66,808	0.00	0.04	0	1	107,875	0.02	0.14	0	1
Industry type											
1	Agriculture	66,808	0.09	0.28	0	1	107,875	0.14	0.34	0	1
2	Mining	66,808	0.00	0.06	0	1	107,875	0.02	0.15	0	1
3	Manufacturing	66,808	0.15	0.36	0	1	107,875	0.12	0.33	0	1
4	Construction and utilities	66,808	0.01	0.08	0	1	107,875	0.11	0.31	0	1
5	Retail and hotel	66,808	0.41	0.49	0	1	107,875	0.25	0.43	0	1
6	Transport and communication	66,808	0.02	0.13	0	1	107,875	0.11	0.31	0	1
7	Finance and real Estate	66,808	0.02	0.14	0	1	107,875	0.03	0.17	0	1
8	Public Administration	66,808	0.05	0.23	0	1	107,875	0.09	0.28	0	1
9	Education	66,808	0.11	0.31	0	1	107,875	0.04	0.20	0	1
10	Health and public Services	66,808	0.07	0.26	0	1	107,875	0.08	0.27	0	1
11	Personal and household Services	66,808	0.07	0.26	0	1	107,875	0.01	0.11	0	1
12	International Corp	66,808	0.00	0.02	0	1	107,875	0.00	0.03	0	1

Obs = observations, Std. Dev. = standard deviation.
Source: Authors' calculation based on Sakernas (2010).

Appendix Table 1: Summary Statistics of Variables for the Margin Plot

		Rural									
		Female					Male				
		Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
Log real wage		19,651	14.05	0.90	11.58	18.04	41,716	14.36	0.74	11.67	18.41
Age		265,498	38.78	16.40	15	98	265,928	38.42	16.37	15	98
Hours worked (week)		138,704	32.32	17.41	0	98	219,864	38.87	16.62	0	98
Education level											
1	No schooling	265,498	0.23	0.42	0	1	265,928	0.22	0.41	0	1
2	Incomplete primary	265,498	0.30	0.46	0	1	265,928	0.31	0.46	0	1
3	Complete primary	265,498	0.21	0.41	0	1	265,928	0.21	0.41	0	1
4	Lower secondary	265,498	0.13	0.33	0	1	265,928	0.17	0.38	0	1
5	Upper secondary	265,498	0.02	0.14	0	1	265,928	0.01	0.12	0	1
6	Above upper secondary	265,498	0.12	0.33	0	1	265,928	0.08	0.27	0	1
Work type											
1	Managers and professionals	138,704	0.06	0.24	0	1	219,864	0.04	0.19	0	1
2	Technicians	138,704	0.01	0.10	0	1	219,864	0.01	0.10	0	1
3	Administrative Staff	138,704	0.02	0.13	0	1	219,864	0.02	0.14	0	1
4	Sales	138,704	0.15	0.36	0	1	219,864	0.07	0.26	0	1
5	Farming	138,704	0.63	0.48	0	1	219,864	0.62	0.48	0	1
6	Laborers	138,704	0.13	0.33	0	1	219,864	0.23	0.42	0	1
7	Military and police	138,704	0.00	0.01	0	1	219,864	0.00	0.06	0	1
Industry type											
1	Agriculture	138,704	0.65	0.48	0	1	219,864	0.65	0.48	0	1
2	Mining	138,704	0.01	0.07	0	1	219,864	0.02	0.16	0	1
3	Manufacturing	138,704	0.08	0.27	0	1	219,864	0.05	0.22	0	1
4	Construction and utilities	138,704	0.00	0.04	0	1	219,864	0.06	0.23	0	1
5	Retail and hotel	138,704	0.16	0.36	0	1	219,864	0.08	0.27	0	1
6	Transport and communication	138,704	0.00	0.06	0	1	219,864	0.04	0.21	0	1
7	Finance and real Estate	138,704	0.00	0.05	0	1	219,864	0.01	0.07	0	1
8	Public Administration	138,704	0.01	0.12	0	1	219,864	0.03	0.17	0	1
9	Education	138,704	0.06	0.23	0	1	219,864	0.03	0.16	0	1
10	Health and public Services	138,704	0.02	0.15	0	1	219,864	0.02	0.15	0	1
11	Personal and household Services	138,704	0.01	0.11	0	1	219,864	0.00	0.05	0	1
12	International Corp	138,704	0.00	0.02	0	1	219,864	0.00	0.02	0	1

Obs = observations, Std. Dev. = standard deviation.
Source: Authors' calculation based on Sakernas (2010).

Appendix Table 2: Summary Statistics of Additional Variables for Heckman's Two-Step Model
(continued)

	All									
	Female					Male				
	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
Marital status										
1 Single	415,376	0.207	0.405499	0	1	411,339	0.301	0.458653	0	1
2 Married	415,376	0.658	0.474238	0	1	411,339	0.657	0.474755	0	1
3 Divorced and widowed	415,376	0.134	0.340758	0	1	411,339	0.042	0.201098	0	1
Number of children below 10 years old	415,376	0.813	0.954184	0	11	411,339	0.795	0.944225	0	11

Obs = observations, Std. Dev. = standard deviation.
Source: Authors' calculation based on Sakernas (2010).

	Urban									
	Female					Male				
	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
Marital status										
1 Single	149,878	0.252	0.434386	0	1	145,411	0.331	0.470523	0	1
2 Married	149,878	0.617	0.486164	0	1	145,411	0.63	0.48268	0	1
3 Divorced and widowed	149,878	0.131	0.337175	0	1	145,411	0.039	0.192824	0	1
Number of children below 10 years old	149,878	0.745	0.907702	0	8	145,411	0.727	0.89619	0	8

Obs = observations, Std. Dev. = standard deviation.
Source: Authors' calculation based on Sakernas (2010).

	Rural									
	Female					Male				
	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
Marital status										
1 Single	265,498	0.182	0.385943	0	1	265,928	0.285	0.451192	0	1
2 Married	265,498	0.682	0.465732	0	1	265,928	0.671	0.46974	0	1
3 Divorced and widowed	265,498	0.136	0.34275	0	1	265,928	0.044	0.205456	0	1
Number of children below 10 years old	265,498	0.852	0.977353	0	11	265,928	0.831	0.967504	0	11

Obs = observations, Std. Dev. = standard deviation.
Source: Authors' calculation based on Sakernas (2010).

REFERENCES

American Association of University Women (AAUW). 2014. *The Simple Truth about the Gender Pay Gap*. Washington DC: AAUW.

Anderson, D., M. Binder, and K. Krause. 2002. Women, Children and the Labor Market. *American Economic Review.* 92 (2). pp. 354–358.

Badan Pusat Statistik (BPS). http://www.bps.go.id

Baker, M. and N. Fortin. 2001. Occupational Gender Composition and Wages in Canada: 1987–88.*Canadian Journal of Economics.* 34 (2). pp. 345–376.

Beall, J. and S. Fox. 2007. Urban Poverty and Development in the 21st Century: Towards an Inclusive and Sustainable World. *Oxfam Research Report*. Oxford, United Kingdom.

Blau, F. D. and L. M. Kahn. 2001. Understanding International Differences in the Gender Pay Gap. *NBER Working Paper* No. 8200. Massachusetts, USA: National Bureau of Economic Research.

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

Bourguignon, F., M. Fournier, and M. Gurgand. 2007. Selection Bias Corrections Based on the Multinomial Logit Model: Monte Carlo Comparisons. *Journal of Economic Surveys.* 21 (1). pp. 174–205.

Cabegin, E. C. A. 2012. Widening Gender Wage Gap in Economic Slowdown: The Philippine Case. School of Labor and Industrial Relations. Quezon City: University of the Philippines. Unpublished.

Chevalier, A. 2007. Education, Occupation and Career Expectations: Determinants of the Gender Pay Gap for UK Graduates. *Oxford Bulletin of Economics and Statistics*. 69 (6). pp. 819-842.

Dhanani, S., I. Islam, and A. Chowdhury. 2009. *Industry and Labour in Indonesia: Changes and Challenges.* New York: Routledge Studies in the Modern World Economy.

Dubin, J. F. and D. L. McFadden. 1984. An Econometric Analysis of Residential Electric Appliance Holdings and Consumption. *Econometrica*. 52 (2). pp. 345–362.

Fields, G. S. 2001. *Distribution and Development*. New York: Russell Sage Foundation.

Fitrania, S. N. 2013. *T*he Impact of Globalization on Gender Wage Gap in Indonesia: A Sub-National Level Analysis. Master's Thesis. Tilburg University.

Giles, David E. 2011. *Interpreting Dummy Variables in Semi-logarithmic Regression Models: Exact Distributional Results*. *Econometrics Working Paper* No. 1101. Victoria BC, Canada: University of Victoria.

Goldin, C. and C. Rouse. 2000. Orchestrating Impartiality: the Impact of "Blind" Auditions on Female Musicians. *American Economic Review.* 90 (4). pp. 715–740.

Greene, W. H. 2000. *Econometric Analysis 4th Edition*. Prentice Hall.

Harris, J. R. and M. P. Todaro. 1970. Migration, Unemployment and Development: a Two-Sector Analysis. *American Economic Review.* 60 (1). pp. 126–142.

Hayashi, Mitsuhiro. 2005. Structural Changes in Indonesian Industry and Trade: An Input-Output Analysis. *The Developing Economies*. 63 (1). pp. 39–71.

Heckman, J. 1974. Shadow Prices, Market Wages and Labor Supply. *Econometrica.* 42 (4). pp. 670–694.

Heckman, J. 1976. The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models. *Annals of Economic and Social Measurement.* 5 (4). pp. 475–492.

Heckman, J. 1979. Sample Selection Bias as a Specification Error. *Econometrica*. 47 (1). pp. 153–161.

Kuznets, S. 1955. Economic Growth and Income Inequality. *The American Economic Review.* 45 (1). pp. 1–28.

Lee, L. 1983. Generalized Econometric Models with Selectivity. *Econometrica*. 51 (2). pp. 507–512.

Machin, S. and P. Puhani. 2003. Subject of Degree and the Gender Wage Differential: Evidence from the UK and Germany. *Economic Letters*. 79 (3). pp. 393–400.

Maddala, G. S. 1983. *Limited Dependent and Qualitative Variables in Econometrics.* New York: Cambridge University Press.

Mason, A. D. and J. Baptist. 1996. How Important Are Labor Markets to the Welfare of Indonesia's Poor? *World Bank Policy Research Working Paper* No. 1665. Washington DC: World Bank.

Oaxaca, R. 1973. Male-Female Wage Differentials in Urban Labor Markets. *International Economic Review.* 14 (3). pp. 693–709.

OECD. 2014. OECD Family Database. www.oecd.org/social/family/database

Polachek, S. and M. K. Kim. 1995. Panel Estimates of the Gender Earnings Gap. *Journal of Econometrics*. 61 (1). pp. 23–42.

Ravallion, M., S. Chen, and P. Sangraula, 2007. New Evidence on the Urbanization of Global Poverty. *Population and Development Review*. 33 (4). pp. 667–701.

Siddique, M. A. B. 2007. Gender Wage Differentials and Export-Oriented Development in Indonesia. In K. C. Roy and S. Chatterjee, eds. *Growth, Development and Poverty Alleviation in the Asia-Pacific*. Nova Science Publishers, Inc.

Spence, M., P. C. Annez, and R.M. Buckley, eds. 2009. *Urbanization and Growth*. Washington DC: World Bank.

Sugiyarto, G., M. Oey-Gardiner, and N. Triaswati. 2006. Labor Markets in Indonesia: Key Challenges and Policy Issues. In J. Felipe and R. Hasan, eds. *Labor Markets in Asia.* New York: Palgrave Macmillan.

Suryahadi, A., D. Suryadarma, and S. Sumarto. 2009. The Effects of Location and Sectoral Components of Economic Growth on Poverty: Evidence from Indonesia. *Journal of Development Economics.* 89 (1). pp. 109–117.

Tacoli, C. 2012. Urbanization, Gender and Urban Poverty: Paid Work and Unpaid Carework in the City. *Urbanization and Emerging Population Issues Working Paper 7*. London: International Institute for Environment and Development.

Tijdens, K. G. and M. van Klaveren. 2012. *Frozen in Time: Gender Pay Gap Unchanged for 10 Years.* Brussels: International Trade Union Confederation.

Todaro, M. P. 1969. A Model of Labor Migration and Urban Unemployment in Less Developed Countries. *The American Economic Review.* 59 (1). pp. 138–148.

van Klaveren, M., K. Tijdens, M. Hughie-Williams, and N. R. Martin. 2012. An Overview of Women's Work and Employment in Indonesia. Decisions for Life MDG3 Project Country Report No. 14. Amsterdam, Netherlands: University of Amsterdam / Amsterdam Institute for Advanced Labour Studies.

United Nations Population Fund (UNFPA). 2007. *State of World Population 2007: Unleashing the Potential of Urban Growth.* New York: UNFPA.

World Bank. 2011. *World Development Report 2012: Gender Equality and Development.* Washington DC: World Bank.

World Bank. 2014. World Development Indicators. http://data.worldbank.org/data-catalog/world-development-indicators

New Evidence on the Gender Wage Gap in Indonesia

Even though Indonesia has been experiencing impressive economic growth, urbanization could affect income inequality among workers. Using the 2010 National Labor Force Survey (Sakernas) in Indonesia, this paper examines how monthly wages are distributed between male and female workers. Empirical results indicate that urbanization tends to benefit male workers more favorably than female workers. It also shows that the gender wage gap in Indonesia is mainly due to gender discrimination. Thus, the paper proposes the implementation of laws to equalize opportunities and wages among workers, especially in the public sector.

About the Asian Development Bank

ADB's vision is an Asia and Pacific region free of poverty. Its mission is to help its developing member countries reduce poverty and improve the quality of life of their people. Despite the region's many successes, it remains home to approximately two-thirds of the world's poor: 1.6 billion people who live on less than $2 a day, with 733 million struggling on less than $1.25 a day. ADB is committed to reducing poverty through inclusive economic growth, environmentally sustainable growth, and regional integration.

Based in Manila, ADB is owned by 67 members, including 48 from the region. Its main instruments for helping its developing member countries are policy dialogue, loans, equity investments, guarantees, grants, and technical assistance.

ADB

ASIAN DEVELOPMENT BANK
6 ADB Avenue, Mandaluyong City
1550 Metro Manila, Philippines
www.adb.org