Generative Process of Wage Gaps: Analyses of Nation-wide Employer-Employee Matched Data

Koji TAKAHASHI (Japan Institute for Labour Policy and Training / University of California, Los Angeles)

Introduction

- This poster aims to introduce an approach to investigate wage gaps in the labor market, using employer-employee matching data.
- Previous studies on wage gaps have attributed wage gaps to several individual-level factors such as gender, education, and occupation.
- In daily life, however, two types of generative processes of wage gaps are recognized. They are:
  - Xs (males, college graduates, white-collars, etc.) tend to earn more because **they are paid more within each firm (each employer)**.
  - Xs tend to earn more because **they are more likely to work for high-wage firms (employers)**.
- The distinction above has not been paid enough attention. How can we separate these two mechanisms? What kind of data and method do we need to clarify the process through which each wage gap is generated?

Framework and Method

- We assume that some individual attributes have a correlation with the wage level of firms based on hiring and job-seeking activities. By controlling for the wage level of each firm, we can calculate the wage gaps induced by individual-level factors within each firm (Figure 1).

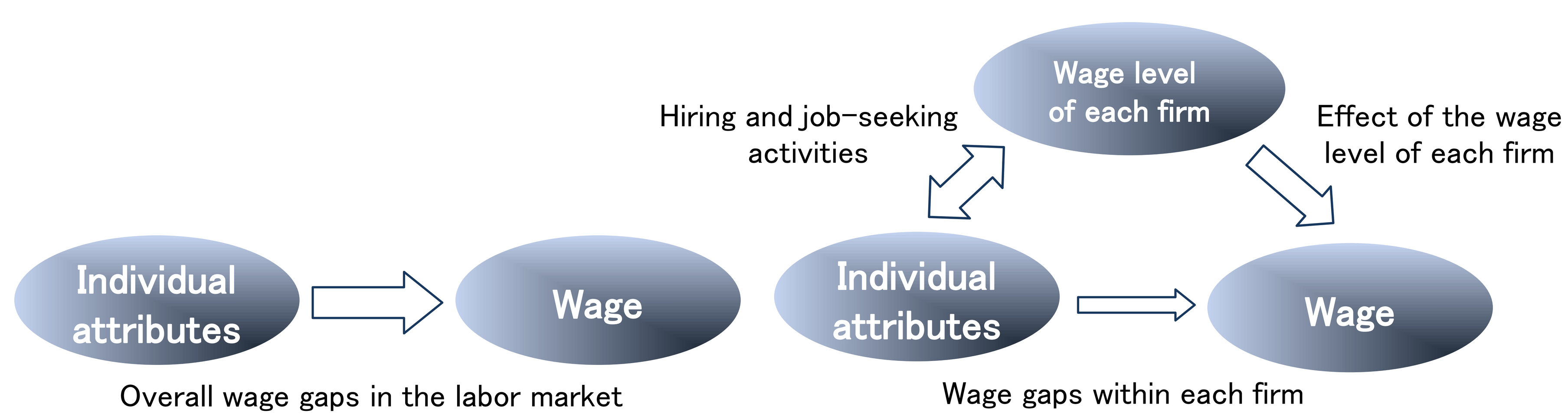

Figure 1. Framework of the Analyses

- To do this, employer-employee matching data, in which individual employees' information is nested within their employers', are required. **Wage gaps within each firm are estimated by the Fixed-effect Model (Equation (2), β_{FE})**.
- On the other hand, the results of the Pooled OLS Model are considered to correspond to the overall wage gaps in the labor market (Figure 1) (Equation (1), $\boldsymbol{\beta}_{\mathbf{Pooled}}$).
- Then, **wage gaps generated through hiring and job-seeking activities are estimated by subtracting the coefficients of Equation (2) (β_{FE}) from those of Equation (1) (β_{Pooled})**. This is because the differences between $\boldsymbol{\beta}_{\mathbf{FE}}$ and $\boldsymbol{\beta}_{\mathbf{Pooled}}$ expand when individual attributes correlate strongly with the wage level of firms.

Equation (1) [Pooled OLS Model]

$$\mathbf{y_{ij}} = \alpha + \mathbf{X_{ij}}\boldsymbol{\beta}_{\mathbf{Pooled}} + \varepsilon_{ij}$$

Equation (2) [Fixed-effect Model]

$$\mathbf{y_{ij}} = \alpha + \mathbf{X_{ij}}\boldsymbol{\beta}_{\mathbf{FE}} + \delta_j + \varepsilon_{ij}$$

Notation:

- The subscript j is the ID of the firm;
- The subscript i is the ID of an individual employee working for firm j;
- $\mathbf{y_{ij}}$ is the wage of the individual employee i working for firm j;
- $\mathbf{X_{ij}}$ denotes the vector of individual-level variables of individual employee i working for firm j;
- α is the constant (intercept) for all individual employees;
- $\boldsymbol{\beta}$ is the vector of the slope for individual-level variables;
- δ_j is a unique constant (intercept) for individual employees working for firm j;
- ε_{ij} is the error factor for all individual employees.

Data

- Japan's nation-wide employer-employee matched data from the Ministry of Health, Labour and Welfare, "General Survey on Diversified Types of Employment" conducted in 2010 and 2014 were used (see Table 1).

Data (cont'd)

Table 1. Data Overview (Unweighted)

			2014	2010
Overview of the survey	Employers:	Number of distribution	16,973	16,886
		Number of response	10,938 (64.4%)	10,414 (61.7%)
	Employees:	Number of distribution	52,949	51,152
		Number of response	34,511 (65.2%)	33,087 (64.7%)
Target of the analyses	Employers		6,641	7,067
	Employees		21,100	21,856
	Regular employees		7,312	7,991
	Non-regular employees		13,788	13,865
	Number of employees per employer (Max./Min./Ave.)		1 / 22 / 3.2	1 / 21 / 3.1

Note: 1) The main target of the analyses is the data from 2014. 2010 data were analyzed to confirm the robustness of the results (not for comparison).
2) Employees who were 60 years or above, those dispatched from other firms, students, and those working less than 20 hours per week were excluded from the analyses.
3) All the tables in this poster were published in the Japan Institute for Labour Policy and Training ed. (2018).

- Dependent variable: Logarithm of the hourly wage of individual employees.
- Independent variables: Employment types (regular employee dummy), gender (male dummy), age, square of age, education (college graduate dummy), occupations (white-collar dummy), and years of service.
- We weighted them back to all individual employees working for all firms in the population. Table 2 shows the descriptive statistics.

Table 2. Descriptive Statistics (Weighted)

	2014					2010				
	N	Ave.	S.D.	Min.	Max.	N	Ave.	S.D.	Min.	Max.
Ln (wage)	35,379,496	7.25	0.44	4.75	8.67	33,104,911	7.26	0.45	4.75	8.67
Regular employee	35,379,496	0.71	0.45	0	1	33,104,911	0.71	0.45	0	1
Male	35,379,496	0.56	0.50	0	1	33,104,911	0.59	0.49	0	1
Age	35,379,496	40.65	10.14	17.50	57.50	33,104,911	40.78	10.29	17.50	57.50
Square of age	35,379,496	1754.93	820.58	306.25	3306.25	33,104,911	1768.49	845.70	306.25	3306.25
College graduate	35,379,496	0.37	0.48	0	1	33,104,911	0.36	0.48	0	1
White-collar	35,379,496	0.65	0.48	0	1	33,104,911	0.65	0.48	0	1
Years of service	35,379,496	10.77	8.30	0.13	25.00	33,104,911	10.96	8.31	0.13	25.00

Results

- By focusing on the difference between the results of "Pooled OLS Model" ($\boldsymbol{\beta}_{\mathbf{Pooled}}$) and those of "Fixed-effect Model" ($\boldsymbol{\beta}_{\mathbf{FE}}$), we can judge which type of generative process is more relevant to each wage gap. For example:
  - **Gender wage gap is primarily generated within each firm**; Figure 2 demonstrates how gender has little to do with the wage level of firms.
  - **Education affects both the chance to work for high-wage firms and the compensation within each firm**; Figure 2 proves how education has much to do with the wage level of firms.
- In addition, we find that the results of 2014 and 2010 are almost the same.

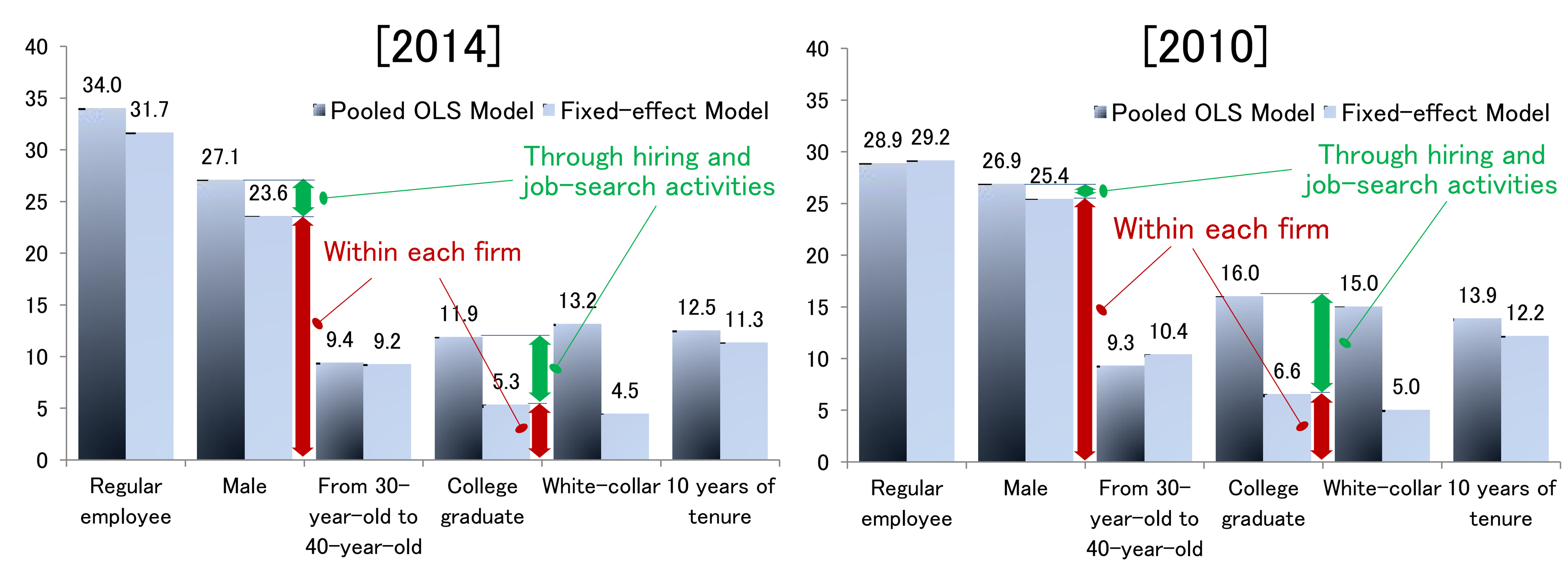

Figure 2. Rates of difference of wages (%)

Conclusion and Implication

- **Difference in Generative Process**: Studies of gender wage gap and policies to deal with it should focus on job allocation and promotion within each firm. On the other hand, as far as education is concerned, selection and matching process should also be paid enough attention.
- **Reliability of the Method**: Although the number of employees nested in each firm was not necessarily large, the two data sets (2014 and 2010) produced similar results. This method can be applied to data in other countries. International comparisons of the generative process of wage gaps can substantially contribute to the study of social stratification.

Reference: Japan Institute for Labour Policy and Training ed. (2018) *Results of the Secondary Analyses of Ministry of Health, Labour and Welfare "General Survey on Diversified Types of Employment": Transition of the Diversified Employment No.4, Using Data of 2003, 2007, 2010, and 2014*, Japan Institute for Labour Policy and Training. (in Japanese).

July 18 (Wed), 2018, RC 28 Poster Session "Cutting-Edge Research in Social Stratification," XIX ISA World Congress of Sociology, Toronto, Canada