RIETI Discussion Paper Series 22-E-099

An *RCM* Approach to Causal Inference with Two-level Data and Unobserved Social Contextual Heterogeneity: An application for the decomposition analysis of the gender income gap and the gender gap in positional rank in Japan

YAMAGUCHI, Kazuo
RIETI

RIETI
Research Institute of Economy, Trade & Industry, IAA
The Research Institute of Economy, Trade and Industry
https://www.rieti.go.jp/en/

RIETI Discussion Paper Series 22-E-099

October 2022

An *RCM* Approach to Causal Inference with Two-level Data and Unobserved Social Contextual Heterogeneity: An application for the decomposition analysis of the gender income gap and the gender gap in positional rank in Japan[1]

Kazuo YAMAGUCHI

The University of Chicago, RIETI

Abstract

This article introduces a new RCM method based on the inverse probability of treatment weighting for the analysis of two-level data of individuals and their social contexts when we expect unobserved contextual effects on the treatment and outcome. The method is an alternative to the use of fixed effects for social contexts in the estimation of propensity score when the fixed effects cannot be included in the estimation of propensity score due to small sample sizes for a non-negligible number of social contexts. The method is based on a novel ignorability assumption that may hold in many cases and permits the elimination of confounding unobserved contextual effects under such a situation.

An application of the new method to the decomposition analysis of inequality by combining it with the DiNardo-Fortin-Lemieux method focuses on the decomposition of the gender income gap and gender gap in positional rank among white-collar regular employees in Japan when their employers are the social contexts. The application provides findings that are consistent with the hypotheses that women tend to remain employed in firms for which their relative income and relative opportunity of being promoted to supervisory positions compared with men are better than in other firms, and that the gender gap is consequently reduced among those who remain employed.

Keywords: causal inference, decomposition analysis, two-level data, unobserved contextual effects, gender inequality

JEL classification: C31, C51, D31, J71

[1]The research in this article was supported by the Research Institute of Economy, Trade, and Industry in Japan through my research on "Women's economic empowerment, low fertility rate, work-life balance, and the establishment of evidence-based policy making in Japan" as a visiting fellow there.

INTRODUCTION

This article introduces a new variation of Rubin's causal model (RCM) by extending the semiparametric causal inference method and a related decomposition method based on the inverse probability of treatment (IPT) weighting for the case where we have two-level cross-sectional data, one at the individual level and the other at the social contextual level. By social context, we refer to firms, schools, school classes, communities, and other forms of social organizations for which the population of individuals in each of mutually exclusive social contexts can be clearly defined. We also assume a situation where we expect the effects of unobserved contextual heterogeneity on the treatment and outcome, but data have such a characteristic that while observed social contexts are numerous, the sample size in each social context is small for a nonnegligible number of social contexts. Under such a situation, we cannot effectively control unobserved contextual heterogeneity by the use of a set of dummy variables for the fixed effects of social contexts together with individual-level covariates in estimating the propensity score. In other words, it is a situation where it is not likely that we can effectively attain $Z \perp (\mathbf{X}, \mathbf{D})$ in data under the strong ignorability assumption of $\{Y(1), Y(0)\} \perp Z \mid \mathbf{X}, \mathbf{D}$, where $\{Y(1), Y(0)\}$ are a pair of potential outcomes under treatment and under nontreatment, Z is the treatment variable, $\mathbf{X}$ is the set of observable individual-level covariates, and $\mathbf{D}$ is the set of dummy variables for social contexts. This is because we cannot expect an adequate balancing of the propensity-score distribution between the treatment group and the control group, because covariate distributions cannot be balanced within each social context. On the other hand, we also consider the use of observed contextual-level covariates $\mathbf{V}$ together with the individual-level covariates $\mathbf{X}$ to attain $Z \perp (\mathbf{X}, \mathbf{V})$ under an alternative ignorability assumption of

$\{Y(1), Y(0)\} \perp Z \mid \mathbf{X}, \mathbf{V}$ to be unsatisfactory because we believe that unobserved effects of social contexts exist.

In this article, I introduce a method that makes the ignorability assumption weaker than $\{Y(1), Y(0)\} \perp Z \mid \mathbf{X}, \mathbf{V}$ but stronger than $\{Y(1), Y(0)\} \perp Z \mid \mathbf{X}, \mathbf{D}$, such that the method can control the major portion, if not all, of the unobserved contextual effects. The reason it becomes a major portion is explained in this article. The new method, which is referred to in this article as the *marginal contextual control method*, can attain $Z \perp (\mathbf{X}, \mathbf{D})$ in data under a novel ignorability assumption introduced in this article. The advantage and the limitation of making this ignorability assumption are also discussed in the article. The new method is unlikely to suffer from the lack of balance in the propensity-score distribution between the treatment group and control group, as exemplified in the application. Note that the problem in the use of context-specific dummy variables and a related incidental-parameter problem can be solved if we employ the conditional likelihood estimation, which does not depend on the estimates of the fixed effects of the context-specific dummy variables. However, such a method requires a full parametric specification for the outcome equation, but this article is concerned with the semiparametric method without assuming any parametric outcome equation.

Like Rubin's causal model based on inverse probability of treatment (IPT) weighting (Rubin 1985), the method introduced in this article can be applied to the decomposition analysis of social inequality between groups, such as between men and women or among racial/ethnic groups. The use of IPT weighting for the decomposition analysis of social inequality was introduced by DiNardo, Fortin, and Lemieux (1996); their method will be referred to below as the DFL method. Methodological extensions and applications of the DFL method for segregation analysis were introduced recently by Yamaguchi (2017, 2021). There are commonalities and differences between causal analysis and decomposition analysis in the use of the IPT weights. The major commonality is that both analyses try to attain in data a counterfactual situation where a particular key variable Z becomes independent of its covariates by the IPT weighting. For example, for the estimation of the treatment effect for the treated, the covariate distribution of the control group is equated to that of the treatment group in order to estimate the counterfactual outcome of the treatment group under no treatment. For the decomposition analysis, the covariate distribution of women may be equated to that of men to estimate the outcome under this counterfactual situation. This counterfactual situation is the same as considering what would be the extent of gender inequality in the outcome if men were treated like women in the society, given the retention of their own covariate states.

The major difference between the two analyses is the causal order between the key variable and covariates. As shown in Figure 1, covariates are confounding variables that affect the treatment variable Z in causal analysis. On the other hand, for the decomposition analysis of inequality between two groups, covariates are mediating variables that are affected by the group variable Z, such as gender. Hence, by IPT weighting, decomposition analysis tries to eliminate the path from Z to $\mathbf{X}$, while causal analysis tries to eliminate the path from $\mathbf{X}$ to Z, in the IPT-weighted data. The method introduced in this article can be combined with both types of analyses.

(Figure 1 about here)

A related major difference between the two analyses is that in the causal analysis, confounding variables can be endogenous; that is, they can be affected by an unobserved variable that affects the outcome, as indicated from path U to $\mathbf{X}$ in the causal analysis component of Figure 1. On the other hand, as the DFL method assumed, and as Yamaguchi (2015) showed formally, the mediating variables have to be exogenous variables; that is, they should not be affected by an unobserved variable that affects the outcome, and, therefore, the dotted line from U to $\mathbf{X}$ in the decomposition analysis component of Figure 1 must be absent, because otherwise the effects of $\mathbf{X}$ on the outcome will be biased even though they are not estimated. It follows that when $\mathbf{X}$ are endogenous, even after a path from Z to $\mathbf{X}$ is eliminated by the IPT weighting, the remaining effect of Z on the outcome cannot be interpreted as the "direct effect" of Z on the outcome, because the "indirect effect" through $\mathbf{X}$ eliminated by the IPT weighting is biased. In this article, I assume the exogeneity of mediating variables for the decomposition analysis.

In the application described below, the new method is combined with the DFL method for the decomposition analysis of the gender income gap and the gender gap in the attainment of managerial and supervisory positions, which we will refer to as the gender gap in positional rank, among regular employees in Japan, with firms as social contexts. Japan has the second largest gender wage gap, next to South Korea, among the OECD countries, and Japan and South Korea have also the highest gender gap in the attainment of managerial positions among the OECD countries (Yamaguchi 2019, Youm et al. 2021).

A distinct methodological elaboration is introduced in application for the significance tests of decomposed components of inequality when the gender gap is decomposed into the element within firms and the element due to firm selection by gender, with and without an additional control for individual-level human-capital variables. Note that, unlike the decomposition of variance into within-group and between-group components, the "between-group" component in the decomposition analysis of inequality can be opposite in direction to the within-group component of inequality. This is because, in the analysis of inequality, within-group inequality can be greater than the observed prima facie inequality, because inequality is measured by difference rather than by variance. Hence, we refer to the difference between the observed prima facie inequality and the within-context inequality as the effect of the selection of social context by the group, such as the effect of firm selection by gender, on inequality.

The application of the new method introduced in this article has a major substantive aim. It is to clarify the role that the selection of firms plays in explaining the observed gender income gap and gender gap in positional rank. In particular, the article is concerned with assessing the effects of selectivity which results from women's entry into and separation from their employers, compared with men's, among regular employees with unspecified term employment contracts, called *seishain* in Japan. Depending on cohorts, 40%–60% of women who are *seishain* leave the employer at the time of marriage or childrearing (Kenjoh 2015), and only a small percentage of women who leave the labor force or change to irregular employment can return to regular employment, when the time of rearing small children is over, due to the "lifetime employment system" of large and medium-sized Japanese firms that prioritize new and recent graduates from schools for their regular employment. Hence, we expect the selection bias of women among those who remain regularly employed to occur largely in their leaving of firms rather than in their entry into firms, which leads to specific hypotheses described in the application section.

METHODS

The Marginal Contextual Control Method

Since we apply the new method to the decomposition analysis of gender inequality in the outcome Y, we first consider a counterfactual situation where men and women are randomly assigned to social context. Let us denote men by $Z = 1$ and women by $Z = 0$.

Then, under the hypothetical situation of random assignment to social contexts, we should expect $P(D_j = 1 | Z = 1) = P(D_j = 1 | Z = 0)$ to hold, where D_j is the dummy variable for social context j. Since

$$P(D_j = 1 | Z = k) = \frac{P(D_j = 1)P(Z = k | D_j = 1)}{P(Z = k)}, \text{ for } k = 0,1, \qquad (1)$$

holds, equation $P(D_j = 1 | Z = 1) = P(D_j = 1 | Z = 0)$ indicates that random assignment realizes a state where the marginal probability of being in context j, $P(D_j = 1)$, is retained in the sample, and the conditional probabilities of group membership for each given context, $P(Z = k|D_j = 1)$, are equalized between men ($Z = 1$) and women ($Z = 0$), given that the numbers of men and women to be assigned, and thereby $P(Z = k)$ for $k = 0,1$, are fixed.

Since both the estimates of $P(D_j = 1)$ and $P(Z = k|D_j = 1)$ for $k = 0,1$ are available in the sample, it may seem easy to retain a counterfactual situation that retains observed $P(D_j = 1)$ and equalizes $P(Z = k|D_j = 1)$ between men and women in the sample. Suppose, however, that we wish to realize a counterfactual situation where the distribution of some covariates, such as human-capital variables, also need to be equalized between men and women within each social context. As I discussed in the introduction, the characteristics of data we have do not permit a realization of such a counterfactual situation in data by the use of context-specific dummy variables in the estimation of propensity scores. On the other hand, an equalization of observed individual-level and contextual-level covariates between men and women by the use of IPT weights can neither retain $P(D_j = 1)$ nor equalizes $P(Z = k|D_j = 1)$ between men and women for each context in the weighted data.

Let us call a group of social contexts with the same observed values of both $P(D_j = 1)$ and $P(Z = 1|D_j = 1)$ in the sample a stratum of social contexts. Note that If sampling weights need to be applied to the data in order to estimate the average treatment effect, then we need to express the observed values of both $P(D_j = 1)$ and $P(Z = 1|D_j = 1)$ by data weighted by the sampling weights. The same remark applies to the estimation of the propensity score described below.

Then, since we expect that, under the ideal counterfactual situation of $Z \perp (\mathbf{X}, \mathbf{D})$, no within-strata variability exists in values of either $P(D_j = 1)$ or $P(Z = 1|D_j = 1) = P(Z = 1)$, we may consider, as an alternative to using context-specific dummy variables, using stratum-specific dummy variables in the estimation of the propensity score and equalizing the gender composition among strata. In fact, the use of stratum-specific dummy variables together with observed covariates in the estimation of propensity scores can attain $Z \perp (\mathbf{X}, \mathbf{D})$ as described below under a modified ignorability assumption. The advantages and limitations of this modified ignorability assumption is described in detail later. This article will also demonstrate in the application that the control for the strata of social contexts is much more effective than the use of observed contextual-level covariates in capturing the variability in the propensity score.

More formally, the method for causal inference introduced in this article assumes the following besides the standard SUTVA assumption as Assumption 1.

Assumption 2

The unobserved effects of social contexts on the outcome are fixed effects and are linearly additive to the effects of the treatment variable and covariates.

It follows this assumption that, without loss of generality, the outcome equation can be specified by the following semiparametric model with individual-specific heterogenous treatment effects:

$$y_{ij} = \alpha(\mathbf{x}_i, \mathbf{v}_j \mid \boldsymbol{\theta}) + \beta_i Z_i + U(\mathbf{d}_j), \qquad (2)$$

where subscript i indicate a person, subscript j indicates a social context, $\alpha(\mathbf{x}_i, \mathbf{v}_j \mid \boldsymbol{\theta})$ is an unspecified function of observable individual-level covariates $\mathbf{x}_i$ and contextual-level covariates $\mathbf{v}_j$ with parameters $\boldsymbol{\theta}$ and can reflect the interaction effects of individual-level covariates and observed contextual-level covariates but does not reflect the main effects of contextual-level covariates due to their collinearity with $U(\boldsymbol{d}_i)$, and $U(\mathbf{d}_i) = \sum_j u_j D_j(i)$ represents the effects of social contexts, where u_j is the fixed effect of social context j. Parameter β_i indicates the treatment effect for person i and is assumed to vary with individuals.

Assumption 3

$$Y(0) \perp Z \mid \mathbf{X}, \mathbf{V}, U$$

In order to estimate the causal effect of Z on the outcome, this ignorability assumption requires us obtain a balancing score $b(\mathbf{X}, \mathbf{D})$ that satisfies $Z \perp (\mathbf{X}, \mathbf{V}, U) \mid b(\mathbf{X}, D)$ under the data constraint that we cannot use the set of context-specific dummy variables in the estimation of propensity score. Note that we do not need an additional assumption of $Y(1) \perp Z \mid \mathbf{X}, \mathbf{V}, U$ when we assume a counterfactual situation to estimate the average treatment effect for the treated, or its equivalent in the decomposition analysis of inequality.

Let a stratum be a group of social contexts within which the observed values of $P(D_j = 1)$ and $P(Z = 1 | D_j = 1)$ are identical, and let us denote the set of dummy variables to distinguish strata as **s**. We then also make the following assumption:

Assumption 4

$$U \perp (Z, \mathbf{X}) \,|\, \mathbf{s}, \mathbf{v} . \tag{3}$$

This assumption implies that, controlling for strata and contextual-level covariates (**s**,**v**), unobserved effects of social contexts are independent of the treatment variable and individual-level covariates. Assumption 4 places a constraint on Assumption 3 and may appear to be a strong assumption, and, therefore, we discuss below why it can be regarded as a reasonable assumption in many cases.

Generally, the condition of Assumption 4 can be rewritten as the combination of the following two element conditions:

$$U \perp Z \,|\, (\mathbf{s}, \mathbf{v}) \text{, and} \tag{3A}$$

$$U \perp \mathbf{X} \,|\, (\mathbf{s}, \mathbf{v}, z) . \tag{3B}$$

Under the situation where the $P(Z = 1 | D_j = 1) = P(Z = 1 | s(j))$ and is invariant across social contexts within each stratum in the sample, condition (3A) always holds in the sample from the definition of strata, and this condition is therefore not an assumption. However, it will become an assumption which is very likely to hold when each stratum clusters samples of observations for which $P(Z = 1 | D_j = 1)$ has a very small variability among social strata and is not constant within each stratum.

For condition (3B), note that the effects of U on the outcome reflect the effects of social contexts that are not explained by the effects of individual-level variables $\mathbf{X}$ on the outcome, as indicated in equation (2). Whether this condition holds or not may depend on the effective control for observable contextual-level covariates $\mathbf{V}$. Let me explain this by using an example of the decomposition analysis of gender inequality, where $\underline{Z}$ represents gender, $\mathbf{X}$ represent individual human-capital variables, and $\mathbf{s}$ indicate firm strata. Generally, firms differ in the college graduation rate and the average duration of employment in the firm among their employees, but if the effects of college graduation and employment duration on the outcome, such as income or wage, are uniform across firms, condition (3B) is not violated. However, if firms with a higher proportion of college-graduated employees or with a higher average duration of employment among employees have higher average income for employees beyond the effects of individual human capital than other firms, $U \perp \mathbf{X}$ does not hold unconditionally. The higher average income for the employees of firms with higher average human-capital characteristics, beyond the effects of individual human-capital characteristics on income, however, is likely to be a function of observable firm characteristics such as firm size and industry, which can be controlled by variables $\mathbf{V}$. Firm strata also cluster firms with the same observed value of $P(D_j = 1)$, and therefore a conditioning on firm strata will control differences in firm size efficiently. Condition (3B) indicates that, among firms with the same stratum category and observable firm-level characteristics, U becomes independent of $\mathbf{X}$. Hence, we consider that, while condition (3B) may not necessarily hold, the control for $\mathbf{V}$ as well as firm strata will significantly reduce the extent of to which it is violated. Generally, it is important to select observable contextual-level variables that will reduce the violation of condition (3B). In particular, if context-specific effect u_j is constant, random, or a mixture of the two among firms in each stratum, Assumption 4 holds. However, that is a sufficient but not a necessary condition for Assumption 4.

Now, let us define the following propensity score:

$$\theta(\mathbf{x},\mathbf{s},\mathbf{v}) = P(Z=1 \mid \mathbf{x},\mathbf{s},\mathbf{v}). \tag{4}$$

I prove below that under the set of Assumptions 1 through 4, $(s,\theta(\mathbf{x},s,\mathbf{v}))$, where s indicates each given stratum, is a balancing score to satisfy $Z \perp (U,\mathbf{X},\mathbf{V}) \mid (s,\theta(\mathbf{x},s,\mathbf{v}))$.

Proof:

In order to prove $Z \perp (\mathbf{X},\mathbf{V},U) \mid (s,\theta(\mathbf{x},s,\mathbf{v}))$, we need to prove:

$$P(Z=1 \mid U,\mathbf{x},s,\mathbf{v},\theta(\mathbf{x},s,\mathbf{v})) = P(Z=1 \mid s,\theta(\mathbf{x},s,\mathbf{v})). \tag{5}$$

For the left-hand side of equation (5), we obtain:

$$\begin{aligned}
&P(Z=1 \mid U,\mathbf{x},s,\mathbf{v},\theta(\mathbf{x},s,\mathbf{v})) = P(Z=1 \mid U,\mathbf{x},s,\mathbf{v}) \text{ (because } \theta(\mathbf{x},s,\mathbf{v}) \text{ becomes a constant)}\\
&= \frac{P(Z=1,U,\mathbf{x},s,\mathbf{v})}{P(U,\mathbf{x},s,\mathbf{v})} = \left\{\frac{P(U,s,\mathbf{v})P(Z=1,\mathbf{x},s,\mathbf{v})}{P(s,\mathbf{v})}\right\}\left\{\frac{P(s,\mathbf{v})}{P(U,s,\mathbf{v})P(\mathbf{x},s,\mathbf{v})}\right\} \text{ (from } U \perp (Z,\mathbf{X}) \mid \mathbf{s},\mathbf{v})\\
&= \frac{P(Z=1,\mathbf{x},s,\mathbf{v})}{P(\mathbf{x},s,\mathbf{v})} = P(Z=1 \mid \mathbf{x},s,\mathbf{v}) = \theta(\mathbf{x},s,\mathbf{v}). \qquad (6)
\end{aligned}$$

From the right-hand side of the equation, we obtain:

$$\begin{aligned}
P(Z=1 \mid s,\theta(\mathbf{x},s,v)) &= E(Z \mid s,\theta(\mathbf{x},s,v))\\
&= \int_{\mathbf{x},\mathbf{v}} E(Z \mid \mathbf{x},s,\mathbf{v},\theta(\mathbf{x},s,\mathbf{v}))P(\mathbf{x},\mathbf{v} \mid s,\theta(\mathbf{x},s,\mathbf{v}))d\mathbf{x}d\mathbf{v}\\
&= \int_{\mathbf{x},\mathbf{v}} \theta(\mathbf{x},s,\mathbf{v})P(\mathbf{x},\mathbf{v} \mid s,\theta(\mathbf{x},s,\mathbf{v}))d\mathbf{x}d\mathbf{v} \text{ (from equation (6))}\\
&= \theta(\mathbf{x},s,\mathbf{v})\int_{\mathbf{x},\mathbf{v}} P(\mathbf{x},\mathbf{v} \mid s,\theta(\mathbf{x},s,\mathbf{v}))d\mathbf{x}d\mathbf{v} \text{ (because } \theta(\mathbf{x},s,\mathbf{v}) \text{ is fixed)}\\
&= \theta(\mathbf{x},s,\mathbf{v}).
\end{aligned}$$

Q.E.D.

Note that $(s,\theta(s,\mathbf{v}))$, where $\theta(s,\mathbf{v}) = P(Z=1 \mid s,\mathbf{v})$, also becomes the balancing score to attain $Z \perp (U,\mathbf{V}) \mid (s,\theta(s,\mathbf{v}))$, because Assumption 4 holds when **X** is held constant. Hence, by the use of those propensity scores, we can attain counterfactual situations where the treatment variable becomes independent of the unobserved context variable U as well as observed covariates. In other words, the use of $\theta(\mathbf{x}, s, \mathbf{v})$ and $\theta(s, \mathbf{v})$ for the IPT weights permits us to realize in data counterfactual situations where $Z \perp (\mathbf{X}, \mathbf{V}, U)$ and $Z \perp (\mathbf{V}, U)$ hold, respectively, for the IPT-weighted data. Note that if observed contextual-level covariates **V** do not have unique effects on the propensity score, which may occur when the effects of context strata are controlled, we can omit them entirely from the analysis.

Decomposition of the observed prima facie group effect on the outcome into elements through and not through social contexts and the significance test of the elements

The method is described using the case of the decomposition of gender inequality on which this article focuses in the application--although the method can be applicable for a general case of using the DFL method for decomposition analysis with two-level data. I consider the following four situations, including one observed and three counterfactual situations for the decomposition:

Situation 1A: The observed situation

Situation 1B: The counterfactual situation where the distribution of firm strata **S** and observed firm characteristics **V** for female employees becomes the same as the observed distribution of male employees

Situation 2A: The counterfactual situation where the distribution of human-capital variables **X** for female employees becomes equal to the observed distribution of male employees within each group of firms with the same firm stratum and observed firm characteristics; note that in this model, the distributions of men and women across firms with different strata and observed firm characteristics are not equalized

Situation 2B: The counterfactual situation where the joint distribution of firm strata, observed firm characteristics, and human-capital variables of female employees becomes equal to the observed distribution of male employees

I express by $Z = 1$ the group of male employees, and by $Z = 0$ the group of female employees. Human-capital variables are denoted by **X**. For all situations, the IPT weights for male employees are 1.0, because the counterfactual situation does not change their covariate distribution. For female employees, the IPT weights for situation 1A are also 1.0, that is, $\omega_{1A} = 1.0$.

For situations 1B, 2A, and 2B, the IPT weights for female employees are given respectively as

$$\omega_{1B} = \frac{P(\mathbf{s}, \mathbf{v} \mid Z = 1)}{P(\mathbf{s}, \mathbf{v} \mid Z = 0)} = \frac{P(Z = 0)P(Z = 1 \mid \mathbf{s}, \mathbf{v})}{P(Z = 1)P(P = 0 \mid \mathbf{s}, \mathbf{v})},$$

$$\omega_{2A} = \frac{P(\mathbf{x} \mid Z = 1, \mathbf{s}, \mathbf{v})}{P(\mathbf{x} \mid Z = 0, \mathbf{s}, \mathbf{v})} = \frac{P(Z = 0 \mid \mathbf{s}, \mathbf{v})P(Z = 1 \mid \mathbf{x}, \mathbf{s}, \mathbf{v})}{P(Z = 1 \mid \mathbf{s}, \mathbf{v})P(Z = 0 \mid \mathbf{x}, \mathbf{s}, \mathbf{v})}, \text{ and} \tag{4}$$

$$\omega_{2B} = \frac{P(\mathbf{x}, \mathbf{s}, \mathbf{v} \mid Z = 1)}{P(\mathbf{x}, \mathbf{s}, \mathbf{v} \mid Z = 0)} = \frac{P(Z = 0)P(Z = 1 \mid \mathbf{x}, \mathbf{s}, \mathbf{v})}{P(Z = 1)P(Z = 0 \mid \mathbf{x}, \mathbf{s}, \mathbf{v})}.$$

Note that since $\omega_{2A} = \omega_{2B} / \omega_{1B}$, we need to estimate only two propensity score sets, $P(Z = 1 \mid \mathbf{x}, \mathbf{s}, \mathbf{v})$ and $P(Z = 1 \mid s, \mathbf{v})$, for the following decomposition analysis.

Table 1 indicates the decomposed elements of the observed gender inequality in the outcome. The components of inequality are (1) within-firm gender inequality and (2) inequality due to firm selection effect by gender, with and without a control for human-capital variables. The estimators of these components and their differences are also presented.

(Table 1 about here)

The standard errors of Models 1A through 2B are obtained by the standard errors of the difference between the two means. The significance test of differences in the outcome between models can be made as the significance test of the covariance between the weight difference and the outcome Y among women for whom $Z = 0$. For example, the significance level of the difference in the outcome between Models 2B and 1B is equal to the significance level of covariance between $\omega_{2B} - \omega_{1B}$ and the outcome Y among women, because the weight differences sum to 0 among women.

APPLICATION

Data, Variables, and the Stratification of Firms

Data are taken from the Japan component of the 2009 International Comparative Survey on Work-Life Balance conducted by the Research Institute of Economy, Trade, and Industry in Japan. The data are linked between the employer survey and employee survey. The employee survey collected data of white-collar regular employees employed by sample firms. Regular employees are largely (more than 99%) full-time workers, and paid by monthly salary. Hence, I focus on the gender gap in annual income from employment rather than the gander gap in hourly wage. The data set analyzed here consists of 6,177 sample men and 2,666 sample women aged 23–59 without missing income data, who are employed in the 1,677 nationwide sample firms in the private sector with 100 or more employees. The average number of sample persons in each firm, which is approximately proportional to the number of regular employees of the firm, is small: 5.25. The analysis is restricted to employees ages 23-59 in order to exclude ages in which the proportion of students or retirees is high. This restriction reduces the problem of sample selection bias in the effect of age.

Table 2 shows the cross-classified number of sample subjects surveyed for firms with *NM* sample men and *NF* sample women. As shown in Table 2, 2,045 sample males are employed in firms with only male sample employees (for which $NF = 0$), and 282 female samples are employed in firms with only female sample employees (for which $NM = 0$). The data of these 2,327 sample subjects cannot be used in the analysis because the gender ratio in the firm cannot be equalized by the IPT weights in the sample for the employees of those firms. The remaining samples used in the analysis presented in this article consists of 4,100 men and 2,384 women employed by 1,139 distinct firms. There is no serious sample selection bias in gender income gap due to the exclusion of those samples from the analysis. The observed prima facie gender gap in annual individual earning for the entire 8,811 sample is 1,897 thousand yen (with the standard error of 45 thousand yen), and the gap after the elimination of 2,327 samples is 1,830 thousand yen (with the standard error of 51 thousand yen), and the difference is not significant at the 5% level.

(Table 2 about here)

Regarding the stratification of firms, where a stratum has a distinct combination of the *NM* and *NF* values, strata have 42 patterns, after the exclusion of the first row and the first column of Table 2 and the cells with 0 observed frequency. However, there are some strata for which we cannot expect an adequate balancing of human-capital variables between men and women because sample sizes in those strata are too small. Hence, while I retained all distinct values of *NM* from 1 to 9 for the stratum distinction, I combined samples with 5 to 8 for *NF* values, thereby leading to the distinction of 35 strata. The distinction of these 35 strata explains 99.2% of the variance of sample firm proportion $P(D_j = 1)$ and 99.9% of the variance in the propensity score of being a male employee in the firm $P(Z = 1 | D_j = 1)$ in the distinction of all 42 strata. Generally, if a particular sample has much more variability in both $P(D_j = 1)$ and $P(Z = 1 | D_j = 1)$, we can stratify the sample by combining cases similar in both $P(D_j = 1)$ and $P(Z = 1 | D_j = 1)$ values to force each stratum to have an adequate sample size by retaining a high proportion of explanation for the variance of both $P(D_j = 1)$ and $P(Z = 1 | D_j = 1)$, especially for the latter, as in the single-dimensional stratification method of the propensity scores (Hong 2015).

As for human-capital variables, I employed categorical variables for (1) educational attainment (a distinction between college graduates and those with less education), (2) age (seven categories of 23-29, 30-34, 35-39, 45-49, 50-54, 55-59), and (3) duration of employment by the current employer (7 categories for 5-year intervals plus an additional 8th category for missing data). Table 3 presents the mean of the three human-capital variables for women and men in the sample. The differences are all significant at the 0.1% level. Figure 2 also shows the historical trends in college attendance rate in Japan. Unlike many other OECD countries, Japan has a college attendance rate for women that is still lower than that of men, which places a handicap on women in their income and status attainment. Our sample, which is restricted to white-collar regular employees employed by private-sector firms with 100 or more regular employees, also shows in Table 3 a large gender gap in college graduation rate.

(Table 3 and Figure 2 about here)

As for the observed firm-level variables, I employed (4) firm size in terms of the number of regular employees (4 categories), (5) firm size in terms of the number of female regular employees (6 categories including one category for missing cases, and (6) industries of firms (3 categories of manufacturing, wholesale and retail sales, and other). These three categories are chosen to represent major industry distinctions most strongly associated with the average individual income for the data set analyzed here.

Assumed Causal Structures and Hypotheses

An interpretation of the results from the decomposition of inequality described in Table 2 requires a further assumption of the causal structure. Since a control for age and the duration of employment by the current employer will be identical, if an interval scale for measuring these two variables is employed, to control for age at entry into the current employment and the duration of that employment, I may assume that while both education and age at entry into current employment affect firm selection, firm selection affects duration of employment. Thus, I assume the following causal structure with expected signs for the direction of influence.

The direct path from gender (G) to firm strata (FS) and a path from **X** to FS do not specify a sign for the direction of influence, but they are related to the hypotheses described below. Figure 3 indicates, however, that we have the following six distinct direct and indirect paths of influence for the effects of gender on the individual outcome: ① G → Y; ② G → **X** → Y; ③ G → FS → Y; ④ G → FS → **X** → Y; ⑤ G → **X** → FS → Y; and ⑥ G → **X** → FS → **X** → Y.

(Figure 3 about here)

Since we have only four measurable elements in the decomposition analysis while we have six distinct paths, we cannot assess the contribution of each path to the observed prima facie effect of gender on the outcome. Table 4 shows the correspondence of the entries of decomposition in Table 1 to the six paths distinguished above. The two right columns and the two bottom rows of tables 1 and 4 are the four elements of the observed gender inequality. Table 4 shows that paths ③ and ④ and paths ⑤ and ⑥ cannot be separated from data because of the incompletely recursive causal structure with respect to **X** assumed in Figure 3.

(Table 4 about here)

Generally, the gender effect on the outcome through selection of firms has two components. One of them is an indirect effect, as gender affects human capital and the latter affects the selection of firms. This is a component represented by paths ⑤ and ⑥, which involve path G → **X** → FS, and can be measured by the quantity in the right-bottom cell of Tables 1 and 4. We can expect this effect to be positive when the outcome is income because higher human capital tends to lead to firms with higher average income beyond the effects of individual-level human capital. In other words, we can expect a positive sign for path X → FS. On the other hand, when the outcome is the attainment of managerial and supervisory positions in the firm, we may not be certain, because firms with employees with higher human capital may not necessarily have a greater promotion opportunity to managerial or supervisory positions in the firm. Hence, I hypothesize

Hypothesis 1: The indirect effect of gender (men vs. women) on individual income among regular employees, through the attainment of human capital and the selection of firms as a result of human-capital attainment, is positive and thereby increases the gender income gap.

The other component is the effect of gender on the outcome through the direct gender effect on the selection of firms without the mediation of human-capital attainment and is represented by paths ③ and ④, which involve path G → FS, and can be identified as the difference in the outcome between Models 2A and 2B. As noted before, a large proportion of women quit their regular employment at the time of marriage or childrearing. Hewlett et al. (2011) report that while the major reason American women quit their jobs at the time of childrearing is childrearing itself, the major reason Japanese women quite their jobs at the time of childrearing is job dissatisfaction or poor prospects for a career in the firm, and the time of childrearing just triggers leaving the firm as an expression of such dissatisfaction. Osawa (2015) reports that the major cause of women's job leaving at the time of marriage and childrearing has gradually shifted from family-related reasons to work-related reasons. Hence, for this path, we expect that, when rationally selecting whether to exit from less satisfactory firms, women will, on average, tend to remain in firms that are more advantageous for them than other firms. More concretely, I hypothesize

Hypothesis 2: Women tend to remain employed in firms which are relatively more advantageous for attaining higher income than other firms, compared with men. It follows that while the gender effect, which contrasts men to women, on the income gap is positive, the direct gender effect on the selection of firms without the mediation of human-capital attainment is negative and thereby reduces the gender income gap.

Hypothesis 3: Women tend to remain employed in firms which are relatively more advantageous for attaining managerial or supervisory positions than other firms, compared with men. It follows that while the gender effect, which contrasts men to women, on the attainment of higher rank in position is positive, the direct gender effect on the selection of firms without the mediation of human-capital attainment is negative and thereby reduces the gender gap in positional rank.

The fourth hypothesis to be tested is concerned with the usefulness of the method introduced in this article.

Hypothesis 4: The distinction of firm strata has much more explanatory power for the variability of propensity scores than the observed firm-level covariates.

ANALYSIS

Preliminary Analyses

Preliminary analyses are concerned with the explanatory power of the firm strata relative to observed firm-level variables, and a question whether the control for observed firm-level variables is necessary for the present analysis. Tables 5 and 6 present the results from the logistic regression equations to predict the gender distinction for the following models: As the predictors of gender, Model 1 includes only individual human-capital variables; Model 2 adds to Model 1 the three observed firm-level variables as predictors; Model 3 adds to Model 1 the firm strata dummy variables; and Model 4 adds to Model 1 both the three observed firm-level variables and the firm strata dummy variables. Table 5 shows how the significance of the three observed firm variables and the firm strata variable changes when each of these two groups of covariates is introduced separately (Models 2 and 3) and they are introduced jointly (Model 4) to Model 1 by the Wald statistics.

(Table 5 about here)

The results of Model 2 in Table 5 show that both employee size and female employee size affect the gender ratio of employees significantly when firm strata are not controlled. The model's results show that while industry affects income (results not reported), it does not affect the gender ratio of employees significantly when two other firm-level variables are controlled. The results of Model 4 indicate, however, that neither employee size nor female employee size affects the gender ratio of employees when firm strata are controlled. Hence, I will omit those three observed firm-level variables from the main analysis. In other words, I employ Model 3 for the main analysis as Model 2B

Table 6 shows the mean and the variance of propensity scores for each of the four models. The results of Table 6 show that, compared with the variance in propensity scores for Model 4, which includes all variables in the prediction of the propensity score, Model 1, which includes only human-capital variables, explains 58.0% of that variance. The addition of observed firm-level variables in Model 2 increases the explained variance by only 4.5%, while the addition of firm-strata variables in Model 3 increases the variance by 41.2%, thereby reaching 99.2% explanation of variance in Model 4. As shown in Table 5, the further addition of observed firm-level variables to Model 3 in Model 4 does not increase the explanatory power of the propensity score significantly. A comparison of each of Model 2 and 3 against Model 1 demonstrates that the added explanatory power of variance in propensity scores by the strata dummy variables has about 9 times as much as that by the observed firm-level variables.

(Table 6 about here)

Main Analysis

Table 7 presents the balance test results for Model 2B. Note that since Model 1B with only strata dummy variables is a saturated model, its balance is perfect and does not require any diagnostic analysis. We do not need an additional balance test for Model 2A, because its IPT weights are derived as the ratio of the IPT weights of Models 1B and 2B. Model 2B includes in its logistic regression the main categorical effects of each of the four covariates plus the category-by-category interaction effects of age and college graduation, which are significant because of the higher gender gap in college attendance rate for older cohort members.

The results indicate that although the four covariates are strongly associated with gender in the observed data, none of them retain significant association with gender when the IPT weights of Model 2B are applied. Hence, we may consider that variable *Z* becomes independent of all the four covariates in the IPT-weighted data.

(Table 7 about here)

Tables 8 through 10 show the main analytical results related to Hypotheses 1, 2, and 3. Hypothesis 3 is tested separately for the attainment of managerial positions and the attainment of supervisory positions. Results of Table 8 on gender income gap show that the gender difference in human-capital characteristics explains 34.7% of the observed prima facie gender income gap, and a smaller amount, 29.5%, of the within-firm gender income gap, leaving 65.3% and 70.5% of the gender income gap unexplained, respectively. Hypotheses 1 and 2 are both supported by the data. The gender effect on the outcome through the selection of firms mediated by the effects of human-capital attainment, which involves path G → X → FS, is positive and increases the gender gap in income. On the other hand, the gender effect on the outcome through the direct selection of firms by gender, which involves path G → FS, is negative and reduces the gender gap in income. The latter finding is consistent with the hypothesis that women tend to remain employed in firms which are relatively more advantageous for them in income attainment than other firms. While the magnitude of this effect may look small, this mechanism reduces by 5.1% the observed prima facie gender income gap. Since the two selection effects, with and without mediation by human-capital attainment, are opposite in direction, they offset each other, which leads to a nonsignificant result for the firm selection effect for the situation where human-capital variables are not equalized between men and women.

(Table 8 about here)

Table 9 presents results on the decomposition of gender inequality in the attainment of the level of a manager, called *Kacho* (department head) in Japanese, or higher. It is rather striking that among regular employees of the private-sector firms with 100 or more regular employees analyzed here, gender inequality is substantial: only 3.9% of women, but 33.9% of men, have attained managerial positions. The extent to which this gender gap is explained by the gender difference in human capital is also smaller than in the case of the gender income gap: only 20.3% of gender gap in the attainment of a managerial or higher position and 18.5% for the within-firm gender gap are explained by the gender difference in human capital. The results of Table 9 also show that firm selection by gender does not play any role in explaining this gender gap, thereby rejecting Hypothesis 3 for the attainment of a managerial or higher position. Perhaps the opportunity for women to become managers is too small to be taken into consideration when they decide whether to quit their jobs at the time of marriage and childrearing.

(Table 9 about here)

Table 10 presents results on the decomposition of gender inequality in the attainment of the level of a supervisor, called *Kakaricho* (task-unit chief) in Japanese, or higher. In Japanese work organizations, the position of *Kakaricho* (supervisor) is always lower in rank than the position of *Kacho* (manager) because task-unit teams (*Kakari*) are formed within each department. In this case, the observed proportion of being a supervisor is 67.6% for men and 23.0% for women. Hence, while the gender gap is large, women have many more opportunities to attain this positional rank than a managerial position. Note that while the percentage of men in this positional rank may seem too large, about 20% of men and 50% of women were irregularly employed around the time of this survey, and the large majority of them were rank-and-file workers, and this generates a high proportion of employees with a supervisory position among regular workers.[2] Results show that the gender difference in human-capital characteristics explains 29.6% of the observed prima facie gender gap in the attainment of a supervisory or higher position, and 25.5% of the within-firm gender gap. Hypothesis 3 does not hold in the case of the attainment of managerial positions, but it does hold for the case of attainment of supervisory positions, indicated by the negative and significant direct firm selection effect, which involves path G → FS, and thereby reduces gender gap by 0.027 points in the proportion. This finding is consistent with the hypothesis that women tend to remain in firms which are relatively more beneficial to them in the attainment of supervisory positions than other firms. This mechanism reduces by 6.1% the observed prima facie gender gap in the attainment of a supervisory or higher position.

(Table 10 about here)

CONCLUSION

[2] Together with the gender gap in the attainment of managerial positions, the gender gap in the proportion of irregular workers is one of the major causes of the gender wage gap in Japan (Yamaguchi 2019).

This article has introduced a new RCM method based on the adjustment of propensity-score distribution for the control of observed contextual effects as a confounder of the key variable, which is the treatment variable in causal analysis and the group variable for the decomposition analysis of inequality between groups, in the case where we may not be able to use the set of context-specific dummy variables to control their fixed effects on the propensity score. Although the new method may not eliminate the effects of unobserved context-level variables completely unless the novel ignorability assumption introduced in this article holds, the article also explains why the novel ignorability assumption is likely to hold in many cases and is much more effective than the conventional strong ignorability assumption that relies only on the control of observable covariates of social contexts when their fixed effects cannot be controlled.

The application of the new method to the decomposition of the gender income gap and the gender gap in positional rank in Japan shows that hypotheses on firm selectivity by gender are largely consistent with the analytical findings and can provide indirect tests on hypotheses regarding women's firm-selection behavior in the labor market.

REFERENCES

DiNardo, J., N. Fortin, and T. Lemieux 1996. "Labor Market Institution and the Distribution of Wages." *Econometrica* 64:1001–44.

Hewlett, S.A., L. Sherbin, C. Fredman. C. Ho, and K. Sumberg. 2011. *Off-Ramps and On-Ramps Japan - Keeping the Talented Women on the Road to Success.* Center for Work-Life Policy: New York.

Hong, Guanglei. 2015. *Causality in Social World: Moderation, Mediation, and Spill Over*. New York: Wiley.

Kenjoh, Eiko. 2016. "Josei no Keizoku Shugyo no Goukou to Kadai (Trends and issues of women's continuation in employment)". Conference document 2-1, Conference on the Coordination and Promotion of Work-Life Balance among Ministries of Japanese Government on 11/17/2016.

Osawa. Machiko. 2015. Josei wa Naze Katsuyaku Dekinainoka? (Why Does't Women's Empowerment Occur?) Toyo Keizai Shinpou Publication.

Rubin, Donald B. 1985. "The Use of Propensity Scores in Applied Bayesian Inference". Pp. 463-72 in J. M. Bernardo, M. H. De Groot, D. V. Lindley and A. F. M. Smith (eds.), *Bayesian Statistics,* vol. 2. North-Holland: Elsevier.

Yamaguchi, Kazuo. 2015. "Decomposition of Gender or Racial Inequality with Endogenous Intervening Covariates: An Extension of the DiNardo-Fortin-Lemieux Method." *Sociological Methodology* 45:388-428.

Yamaguchi, Kazuo. 2017. "Decomposition Analysis of Segregation." *Sociological Methodology* 47:264-273.

Yamaguchi, Kazuo. 2019. *Gender Inequalities in the Japanese Workplace and Employment*. New York: Springer

Yamaguchi, Kazuo. 2021. "Multigroup Segregation Analyses with Covariates." *Sociological Methodology* 51(2): 224-252.

Youm, Youm, Yoosik, Kazuo Yamaguchi, and Kiho Sung 2021. "The tale of two inequalities: Gender Inequality Based on an Analysis of Within-Occupation and Between-Occupation Inequality）." *Korean Journal of Sociology* 55-4:161-203.

Table 1. Estimators of components of gender inequality

	Gender inequality	Gender inequality within firms	Firm selection effect by gender
Human capital not equalized	Model 1A $E(Y \mid Z=1)-E(Y \mid Z=0)$	Model 1B $E(Y \mid Z=1)-E(\omega_{1B}Y \mid Z=0)$	Model 1A-Model 1B $E((\omega_{1B}-1)Y \mid Z=0)$
Human capital equalized within firms	Model 2A $E(Y \mid Z=1)-E(\omega_{2A}Y \mid Z=0)$	Model 2B $E(Y \mid Z=1)-E(\omega_{2B}Y \mid Z=0)$	Model 2A-Model 2B $E((\omega_{2B}-\omega_{2A})Y \mid Z=0)$
Inequality due to gender difference in human capital	Model 1A-Model 2A $E(\omega_{2A}-1)Y \mid Z=0)$	Model 1B – Model 2B $E((\omega_{2B}-\omega_{1B})Y \mid Z=0)$	$E((\omega_{1B}+\omega_{2A}-\omega_{2B}-1)$ $Y \mid Z=0)$

Table 2. Sample subjects by numbers of men (*NM*) and women (*NF*) in firms

		NF									
		0	1	2	3	4	5	6	7	8	Total
NM	0	0	60	54	57	44	35	18	14	0	282
	1	141	122	153	144	50	30	7	8	18	673
	2	158	261	232	195	66	35	48	0	0	995
	3	252	300	385	210	140	72	27	20	0	1406
	4	240	455	348	273	80	81	30	0	0	1507
	5	385	264	315	264	108	70	0	0	0	1406
	6	180	266	240	180	70	0	0	0	0	936
	7	245	184	216	180	0	0	0	0	0	825
	8	160	117	140	0	0	0	0	0	0	417
	9	144	80	0	0	0	0	0	0	0	224
	10	140	0	0	0	0	0	0	0	0	140
Total		2045	2109	2083	1503	558	323	130	42	18	8811

Table 3. Gender difference in the mean of human-capital variables

	Mean by gender		
	Z=1 (Men)	*Z*=0 (Women)	Difference
College graduation (in %)	62.7	34.7	28.0***
Age (in years)	40.6	36.0	4.0***
Employment duration (in 5-year category)	3.639	3.024	0.615***

***p<0.001.

Table 4. Correspondence between decomposed elements and paths of influence

	Gender inequality	Gender inequality within firms	Firm selection effect by gender
Human capital not equalized	Model 1A ①②③④⑤⑥	Model 1B ①②	Model 1A – Model 1B ③④⑤⑥
Human capital equalized within firms	Model 2A ①③④	Model 2B ①	Model 2A – Model 2B ③④
Inequality due to gender difference in human capital	Model 1A – Model 2A ②⑤⑥	Model 1B – Model 2B ②	Model 1A – Model 2A + Model 1B – Model 2A ⑤⑥

Table 5. Significance of firm strata and firm-level covariates

	Model 2			Model 3			Model 4		
Variables	Wald	d.f.	P-value	Wald	d.f.	P-value	Wald	d.f.	P-value
Employee Size (ES)	11.24	3	.010	-------	--------	---------	0.34	3	.952
Female ES	60.27	5	.000	--------	--------	---------	9.61	5	.087
Industry	2.81	2	.245	--------	---------	---------	2.00	2	.368
Firm Strata	--------	--------	--------	674.56	34	.000	615.34	34	.000

Table 6. Explanatory power of firm strata and
firm-level covariates for variance in propensity scores

	Mean	Variance	Ratio of variance to that of Model 4
Model 1	.3676	.0363	0.580
Model 2	.3676	.0391	0.625
Model 3	.3676	.0621	0.992
Model 4	.3676	.0626	1.000

Table 7. The test of independence between gender and covariates in the IPT-weighted (Model 2B) and unweighted data

Variables	LL statistics	Unweighted	IPT-weighted
Education (df=1)	L^2	476.75	1.29
	P	0.000	0.257
Age (df=6)	L^2	396.11	6.74
	P	0.000	0.346
Employment duration (df=7)	L^2	157.11	11.12
	P	0.000	0.133
Firm strata (df=34)	L^2	896.57	28.64
	P	0.000	0.728

Note: For the estimation of the propensity score $\theta(\mathbf{x}, \mathbf{s})$, category-by-category interaction effects of education and age are included because of their significance.

Table 8. Decomposition of gender income gap:

$\overline{Y}(Z=1)=526.59; \overline{Y}(Z=0)=343.60$

	Gender inequality	Gender inequality within firms	Firm selection effect by gender
(1) Human capital not equalized	182.99*** (5.06) ①②③④⑤⑥	182.02*** (4.90) ①②	0.97 ③④⑤⑥
(2))Human capital equalized within firms	119.02*** (5.38) ①③④	128.28*** (5.14) ①	–9.26* ③④
ratio: (2) /(1)	0.653	0.705	
(3) Inequality due to gender difference in human capital	63.97*** ②⑤⑥	53.74*** ②	10.23** ⑤⑥
ratio: (3)/(1)	0.347	0.295	

***p<0.001;**p<0.01;*p<0.05.

Notes. Annual income is in 10,000 yen. Standard errors in parentheses. The paths of direct and indirect effects represent the following: ① G → Y, ② G → **X** → Y, ③ G → FS → Y, ④ G → FS → **X** → Y, ⑤ G → **X** → FS → Y, and ⑥ G → **X** → FS → **X** → Y.

Table 9. Decomposition of gender gap in the proportions of managers or higher-level employees:

$\overline{Y}(Z=1)=0.339; \overline{Y}(Z=0)=0.039$

	Gender inequality	Gender inequality within firms	Firm selection effect by gender
(2) Human capital not equalized	0.300*** (0.010) ①②③④⑤⑥	0.302*** (0.011) ①②	–0.002 ③④⑤⑥
(2))Human capital equalized within firms	0.239*** (0.011) ①③④	0.246*** (0.011) ①	–0.007 ③④
ratio: (2) /(1)	0.797	0.815	
(3) Inequality due to gender difference in human capital	0.061*** ②⑤⑥	0.056*** ②	0.005 ⑤⑥
ratio: (3)/(1)	0.203	0.185	

***p<0.001;**p<0.01;*p<0.05.

Notes. The paths of direct and indirect effects represent the following: ① $G \rightarrow Y$, ② $G \rightarrow \mathbf{X} \rightarrow Y$, ③ $G \rightarrow FS \rightarrow Y$, ④ $G \rightarrow FS \rightarrow \mathbf{X} \rightarrow Y$, ⑤ $G \rightarrow \mathbf{X} \rightarrow FS \rightarrow Y$, and ⑥ $G \rightarrow \mathbf{X} \rightarrow FS \rightarrow \mathbf{X} \rightarrow Y$.

Table 10. Decomposition of gender gap in the proportions of supervisors or higher-level employees:

$\overline{Y}(Z=1)=0.676; \overline{Y}(Z=0)=0.230$

	Gender inequality	Gender inequality within firms	Firm selection effect by gender
(3) Human capital not equalized	0.446*** (0.012) ①②③④⑤⑥	0.459*** (0.012) ①②	–0.013 ③④⑤⑥
(2))Human capital equalized within firms	0.314*** (0.012) ①③④	0.341*** (0.012) ①	–0.027* ③④
ratio: (2) /(1)	0.704	0.745	
(3) Inequality due to gender difference in human capital	0.132*** ②⑤⑥	0.118*** ②	0.014 ⑤⑥
ratio: (3)/(1)	0.296	0.255	

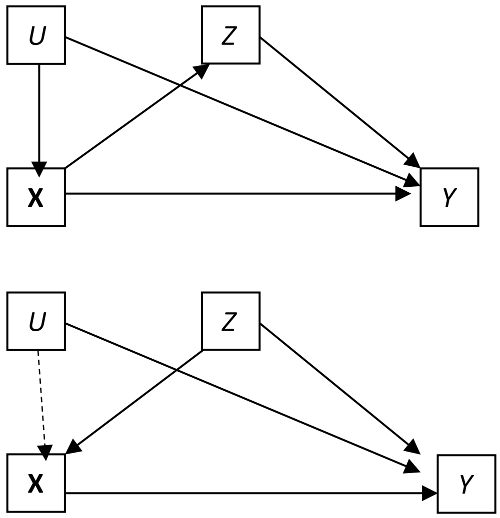
$^{***}p<0.001; ^{**}p<0.01; ^{*}p<0.05$.

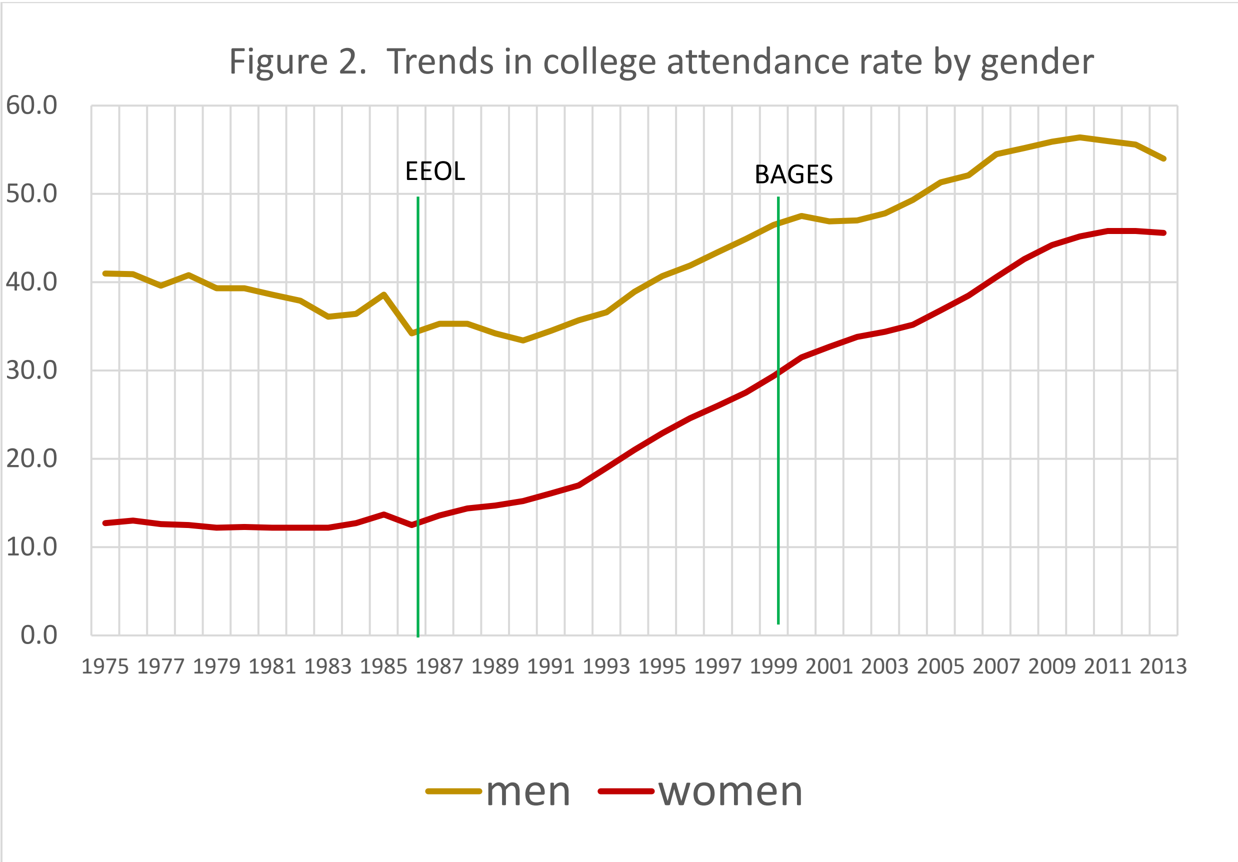
Notes. The paths of direct and indirect effects represent the following: ① G → Y, ② G → **X** → Y, ③ G → FS → Y, ④ G → FS → **X** → Y, ⑤ G → **X** → FS → Y、 and ⑥ G → **X** → FS → **X** → Y.

Figure 1. Two causal diagrams with covariates

Causal Analysis

Decomposition Analysis





Notes:
EEOL: Equality of Employment Opportunity Law
BAGES: Basic Act of Gender Equal Society

Figure 3. Assumed causal structure