Causal Effects of Time-constant Variables with Counterfactual Mediation Modeling:

Evidence from the Gender Wage Gap in Japan*

Shoki OKUBO
(Osaka University)

Abstract

The decomposition of inequality among time-constant variables (e.g., gender) into explained and unexplained components has been a major research agenda in social inequality studies. Explained components of the gender wage gap in the decomposition method correspond to the indirect wage effect of gender that operates through these mediators. Furthermore, unexplained components correspond to the direct wage effect of gender on wage that operates through unobserved mediators (e.g., discrimination).This study links together conventional decomposition methods, such as Blinder-Oaxaca (BO) and DiNardo-Fortin-Lemieux (DFL), based on propensity-score weighting and counterfactual mediation modeling together by introducing the sequential ignorability assumption discussed in Imai et al. (2010). The reason for linking these two methods is that conventional decomposition methods typically control for post-birth variables that lie on the causal pathway from gender or race (which are basically randomly assigned at birth) to wage but neglect the potential endogeneity that may arise from this approach. Moreover, we never directly test the assumptions of conventional decomposition methods and mediation modeling. Therefore, based on the more recent literature on counterfactual mediation modeling, this study presents more reasonable identifying assumptions and the sensitivity of the results to different sets of assumptions.In addition, this study aims to re-focus on time-constant variables in statistical causal analysis of social inequality studies. The analysis focuses on the decomposition of the gender wage gap in Japan. Empirical results indicate that explained components with four mediators (education, occupation, employment status, and post) account for nearly 2%-20% of the gender wage gap in hourly wages.

* This research was supported by JSPS KAKENHI Grant Number JP25000001. Updated version of this paper is available via `https://osf.io/sx47e/?view_only=3caefba079ec49d798609ab2c8bd0b5b`

Keywords: Gender Wage Gap, Counterfactual Modeling, Sequential Ignorability Assumption, Sensitivity Analysis

1. Introduction

This paper links conventional decomposition methods such as DiNardo-Fortin-Lemieux (DFL) based on propensity-score weighting and counterfactual mediation modeling together by introducing sequential ignorability assumption discussed in Imai et al. (2010). The reason for linking these two methods is that conventional decomposition methods typically control for post-birth variables that lie on the causal pathway from gender or race which are basically randomly assigned at birth to wage but neglect the potential endogeneity that may arise from this approach. Moreover, we never directly test the assumptions that lie in conventional decomposition methods and mediation modeling. Based on the newer literature on counterfactual mediation modeling, this paper therefore shows more attractive identifying assumptions and the sensitivity of the results to different sets of assumptions.

The decomposition of inequality among time-constant variables (e.g., gender) into explained and unexplained components has been a major research agenda in social inequality studies. The classical linear decomposition for the difference in mean outcomes for two groups A and B assumes following equations (Blinder 1973 and Oaxaca 1973).

$$Y_A = \boldsymbol{\beta}'_A \boldsymbol{X}_A + \epsilon \tag{1}$$

$$Y_B = \boldsymbol{\beta}'_B \boldsymbol{X}_B + \epsilon \tag{2}$$

This decomposition assumes that the outcome variable Y is linearly related to the covariates, $\boldsymbol{X}$, and that the error term is conditionally independent of $\boldsymbol{X}$. Subtracting equation (1) from (2), we get

$$\bar{Y}_B - \bar{Y}_A = [\ \underbrace{\boldsymbol{\beta}'_B(\bar{\boldsymbol{X}}_B - \bar{\boldsymbol{X}}_A)}_{explained\ component}\ ] + [\ \underbrace{(\boldsymbol{\beta}'_B - \boldsymbol{\beta}'_A)\bar{\boldsymbol{X}}_A}_{unexplained\ component}\ ] \tag{3}$$

Equation (3) shows that explained component is the group difference between A and B in the mean of Y, which would be eliminated if B had A's covariate distribution. Unexplained component is the group difference between A and

B in the mean of Y, which would be eliminated if A had B's covariate effects. Many scholars has proposed the weaker-assumption based nonparametric decomposition methods (DiNardo et al. 1996; Frölich 2007). Let ϕ denotes an unspecified function and θ_A and θ_B are covariate effects parameters. We express

$$Y_A = \phi(\boldsymbol{X}_A, \boldsymbol{\theta}_A) + \epsilon \tag{4}$$

$$Y_B = \phi(\boldsymbol{X}_B, \boldsymbol{\theta}_B) + \epsilon \tag{5}$$

Subtracting equation (5) from (4), we get

$$\bar{Y}_B - \bar{Y}_A = \underbrace{[\bar{\phi}(\boldsymbol{X}_B, \boldsymbol{\theta}_B) - \bar{\phi}(\boldsymbol{X}_A, \boldsymbol{\theta}_B)]}_{\textit{explained component}} + \underbrace{[\bar{\phi}(\boldsymbol{X}_A, \boldsymbol{\theta}_B) - \bar{\phi}(\boldsymbol{X}_A, \boldsymbol{\theta}_A)]}_{\textit{unexplained component}} \tag{6}$$

We obtain $\bar{\phi}(\boldsymbol{X}_B, \boldsymbol{\theta}_B)$ and $\bar{\phi}(\boldsymbol{X}_A, \boldsymbol{\theta}_A)$ simply by calculating the sample means. $\bar{\phi}(\boldsymbol{X}_A, \boldsymbol{\theta}_B)$ is obtainable from following integration (Fortin et al. 2010).

$$\begin{aligned}
\bar{\phi}(\boldsymbol{X}_A, \boldsymbol{\theta}_B) &\equiv \int_{\boldsymbol{V}} \phi(\boldsymbol{V}, \boldsymbol{\theta}_B) f(\boldsymbol{x}|T = 0) d\boldsymbol{x} \\
&= \int_{\boldsymbol{X}} E(Y_B|\boldsymbol{x}) f(\boldsymbol{x}|T = 0) d\boldsymbol{x} \\
&= \int_{\boldsymbol{x}} \omega(\boldsymbol{x}) E(Y_B|\boldsymbol{x}) f(\boldsymbol{x}|T = 1) d\boldsymbol{x} \\
&= E_\omega(Y_B)
\end{aligned} \tag{7}$$

Suppose the treatment $T = 1$ is Group A and $T = 0$ is Group B. E_ω is the weighted mean with weights $\omega(\boldsymbol{x})$

$$\begin{aligned}
\omega(\boldsymbol{x}) &\equiv \frac{f(\boldsymbol{x}|T = 0)}{f(\boldsymbol{x}|T = 1)} \\
&= \frac{p(T = 0|\boldsymbol{x}) f(\boldsymbol{x})/p(T = 0)}{p(T = 1|\boldsymbol{x}) f(\boldsymbol{x})/p(T = 1)} \\
&= \frac{p(T = 1) p(T = 0|\boldsymbol{x})}{p(T = 0) p(T = 1|\boldsymbol{x})}
\end{aligned} \tag{8}$$

Scholars have also dealt with decompositions at quantiles in the outcome distribution (Firpo et al. 2009). As Fortin et al. (2011) pointed out, however, most papers in the decomposition literature jump directly to the estimation issues without addressing the identification strategy. In terms of Rubin Causal Model (RCM) using potential outcome framework, Yamaguchi (2015) also indicate that the conventional decomposition methods fails to handle endogenous covariates. Figure 1 shows why the conventional methods do not handle potential endogeneity.

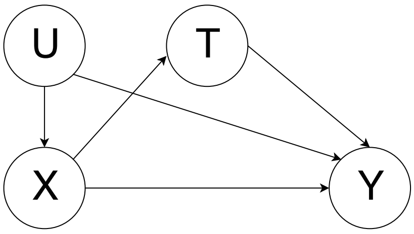

Diagram (a)

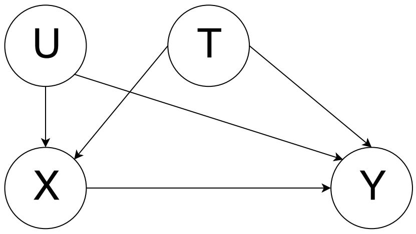

Diagram (b)

Figure1: Causal diagrams with endogenous covariates

In Figure 1, $\boldsymbol{X}$ are endogenous covariates because there is an unobserved confounder U that affects both $\boldsymbol{X}$ and Y. In diagram (a), covariates $\boldsymbol{X}$ are confounders of treatment variable T and outcome variable Y, and $T \perp U | \boldsymbol{X}$ holds. This conditional independence is the important key to estimate causal effects of treatment variables. On the other hand, in diagram (b), covariates $\boldsymbol{X}$ are mediator variables, and even though T is assumed to be independent of U, $T \perp U | \boldsymbol{X}$ does not hold by controlling for covariates $\boldsymbol{X}$, which are common causal descendants of T and U (Pearl 2009; Morgan and Winship 2015; Yamaguchi 2015). As Yamaguchi (2015) point out, diagram (a) is the situation that the RCM assumes for causal analysis with cross-sectional data, and the condition $T \perp U | \boldsymbol{X}$ is equivalent to the ignorability assumption. Since the ignorability assumption doesn't hold in diagram (b), we need to reformulate decomposition analysis for handling endogenous covariates. This paper tackles this problem by introducing counterfactual mediation modeling based on RCM.

2. Identification Strategy

2.1 Set up

We first refer to the relationship between decomposition analysis and mediation analysis. Let's take an example of the gender wage gap. Because the sex is basically randomly assigned at birth and one's characteristics including human capital and social capital will be determined after birth, explained components of the gender wage gap in the decomposition method correspond to the indirect wage effect of gender that operates through these mediators. Meanwhile, unexplained components correspond to the direct wage effect of gender on wage that operates through unobserved mediators (e.g., discrimination). Hereafter, we reformulate the conventional decomposition methods from the point of view of counterfactual mediation modeling.

The popular mediation analysis by Baron and Kenny (1986) shows indirect effects (IE) as follows.

$$M_i = \alpha_2 + \beta_2 T_i + \xi_2 \boldsymbol{X}_i + \epsilon_{i2} \tag{9}$$

$$Y_i = \alpha_3 + \beta_3 T_i + \gamma M_i + \xi_3 \boldsymbol{X}_i + \epsilon_{i3} \tag{10}$$

$$IE = \beta_2 \gamma \tag{11}$$

However, recent literatures have pointed out that IE $\beta_2\gamma$ in equation (11) would be average causal mediation effects (ACME) only if sequential ignorability assumption holds. From the point of view of causal inference, the identification of a causal mechanism requires the specification of an intermediate variable or a mediator that lies on the causal pathway between the treatment and outcome variables (Imai et al. 2011). In RCM, average causal effects (ACE) can be expressed as following.

$$E[Y_i(T_i = 1) - Y_i(T_i = 0)] \tag{12}$$

where treatment $T_i = 1$ or the control $T_i = 0$ condition. Many of causal inference literatures mainly focus on whether one treatment variable causally affects outcome variable and fail to explain how such a causal effect arises.

Counterfactual mediation analysis instead aims to quantify the effect of a treatment variable that operates through a causal mechanism. Let $M_i(t)$ denote the potential value of a mediator of interest for unit i with the treatment condition $T_i = t$. In the same way, let $Yi_(t, m)$ denote the potential outcome if the treatment and mediating variables equal t and m. In reality, we observe only one of the potential outcomes, and thus the observed outcome Y_i is $Y_i(T_i, M_i(T_i))$. Observed outcome is dependent upon both the treatment and mediator status.

2.2 Quantities of interest

Our goal is to clarify how much of the treatment variable is transmitted by the mediator. Following Robins and Greenland (1992), Pearl (2001) and Imai et al. (2010) let ACME define as

$$\delta_i(t) \equiv E[Y_i(t, M_i(1)) - Y_i(t, M_i(0))] \tag{13}$$

for each unit i. ACME is the difference between the potential outcome that would result under treatment status t, and the potential outcome that would occur if the treatment status is the same and yet the mediator takes a value that would result under the other treatment status $M_i(1)$ and $M_i(0)$. Similarly, we define the average causal direct effect (ACDE) as

$$\zeta_i(t) \equiv E[Y_i(1, M_i(t)) - Y_i(0, M_i(t))] \tag{14}$$

for each unit i and each treatment status $t = 0, 1$. Since we observe only one mediator conditions in reality, we require additional assumption for identifying equation (13) and (14).

2.3 Sequential ignorability

Sequential ignorability (SI) assumption discussed in Imai et al. (2010) identifies ACME and ACDE in equation (13) and (14). SI consist of two parts

$$\{Y_i(t', m), M_i'(t)\} \perp T_i | X_i = x \tag{15}$$

$$Y_i(t', m) \perp M_i(t) | T_i = t, X_i = x \tag{16}$$

Equation (15) means the treatment variable T_i is assumed to be ignorable given the pre-treatment covariates X_i. This assumption, held in randomized experiments, is known as unconfoundedness or no omitted variable bias. Equation (16) assumes the observed mediator is ignorable given the actual treatment status T_i and pretreatment covariates X_i. This assumption is not standard ignorability assumptions because randomizing both the treatment and mediator does not identify the ACME (Imai et al. 2011).

2.4 Sensitivity Analysis

SI can never be tested directly so sensitivity analysis is used. The idea of violation of SI leads to a correlation between ε_{i2} and ε_{i3}. Let this correlation denote ρ. Because ACME can be expressed as a function of ρ, the sensitivity analysis investigates how robust the estimation results are to the violation of the SI. Mathematical details are discussed in Imai et al. (2010).

2.5 Estimation procedure

We link the method in Yamaguchi (2015) to counterfactual mediation modeling. Let $\boldsymbol{W}$ denote potential counfounder of the treatment variable T, mediator variable M and covariates X. We follow Imai et al. (2010) in counterfactual mediation analysis from procedure 2 to 6.

1. Obtain a consistent estimate of $p(T = 1|\boldsymbol{W})$ and conduct a diagnosis for an appropriate construction of propensity scores to create a weighted sample for which statistical independence between T and $\boldsymbol{W}$ hold.
2. Fit models for the observed outcome and mediator variables with propensity score weighting[1] created in procedure 1.
3. Simulate model parameters from their sampling distribution.
4. Simulate the potential values of the mediator.
5. Simulate the PO given the simulated values of the mediator.
6. Compute quantities of interest (ACME, ACDE and TE).

3. Empirical Application

The next step is to conduct simulation studies for investigating the finite-sample performance of the estimators, however, we skip the simulation studies in this paper[2]. In this section, we apply these methods using Social Stratification and Mobility (SSM) 2015 survey data.

[1] Huber (2015) clarifies using propensity score weighting induces less biased estimates.

[2] See Okubo (2018) that simulates and compares five estimators: Blinder-Oaxaca decomposition, DiNardo-Fortin-Lemieux decomposition, Yamaguchi (2015) decomposition, ACME (in other words, average natural mediation effects) and average controlled mediation effects.

3.1 Data

SSM survey have been implemented every 10 years since 1955. The SSM 2015 survey is its seventh time, especially focusing on verification of the change of population structure represented by rapid declining birthrate and aging population in terms of social stratification. Respondents of this survey is male and female with Japanese nationality of 20 to 79 years old living in Japan at the end of December 2014. Sampling design followed the basic policy since the SSM survey in 1995 and stratified two-stage cluster sample design was used for sampling respondents. The survey was conducted from January to July in 2015. The sample size and response rates were 7817 and 50.1% for each.

3.2 Variables

The outcome variable Y is the log hourly wage of the 20-64 years old survey participants. Treatment variables T is the sex which denotes 1 for female and 0 for male. Binary mediators $\boldsymbol{M}$ are education (non-colledge graduate for 1), occupation (non-manager for 1), employment status (non-regular work for 1), post (non-post for 1). We also consider controlling for pre-treatment variables $\boldsymbol{W}$ that reflect family background and could potentially confound the treatment variable and the mediators. Specifically, $\boldsymbol{W}$ contains mother's and father's levels of education which denote 1 for colledge graduate. Similarly to the standard literature, covariates $\boldsymbol{X}$ characterizing the explained component include age, employment duration in labor market, employment duration in non-regular work ever. We restrict the sample to observed hourly wage and exclude self-employed and family workers. Table 1 provides summary statistics on these variables, namely the mean and standard deviation values for male and female.

3.3 Results

We analysed two estimation model: (1) without propensity score weighting using the potential confounder $\boldsymbol{W}$, (2) with propensity score weighting using the potential confounder $\boldsymbol{W}$. The calculation of statistical uncertainty estimates is based on the quasi-Bayesian Monte Carlo approximation (King et al. 2000) and we run 1000 times simulations for the quasi-Bayesian approximation of

Table1: Summary statistics by treatment status T.

Variables	Female ($T=1$) Mean	S.D.	Male ($T=0$) Mean	S. D.	Difference
Y					
Log Hourly Wage（1,000 YEN）	7.022	0.628	7.582	0.600	0.560
$\boldsymbol{M}$					
Non-colledge graduate	0.775	0.418	0.558	0.498	0.217
Non-manager	0.996	0.065	0.939	0.240	0.057
Non-regular work	0.521	0.500	0.144	0.351	0.377
Non-post	0.866	0.341	0.548	0.498	0.318
$\boldsymbol{X}$					
Age	43.069	11.110	44.081	11.348	1.012
Regular employment duration	11.010	9.941	20.599	12.230	9.589
Non-regular employment duration	7.197	7.887	1.746	4.576	5.451
$\boldsymbol{W}$					
Father's education	12.022	3.512	11.799	3.547	0.223
Mother's education	11.626	2.795	11.514	2.960	0.112
n	1421		1408		

NOTE: On the right hand side column *Difference* calculates the absolute mean difference between female's and male's mean.

parameter uncertainty. Results in table 2 show that estimates vary across two models especially in Panel A. Model 1 without PSW overestimates ACME. This is because mediator *Non-colledge graduate* is confounded by $\boldsymbol{W}$, parent's education. Though we are able to use only two variables for $\boldsymbol{W}$ in SSM 2015 data, obtaining more $\boldsymbol{W}$, pre-treatment variable, would induce different estimates. Focusing on the percentages of TE mediated, equivalent to the explained component in conventional decomposition analysis, mediator *Non-post* best explains the gender wage gap followed by *Non-colledge*. Employment status in non-regular work mediate TE around 12 percent and only 2 percent for non-manager.

The estimates are identified if the SI holds. We now move on the sensitivity analysis for investigating how robust these estimation results are to the violation of the SI because the assumption SI can never be directly tested. Figure 3 shows the results of the sensitivity analysis for each counterfactual mediation

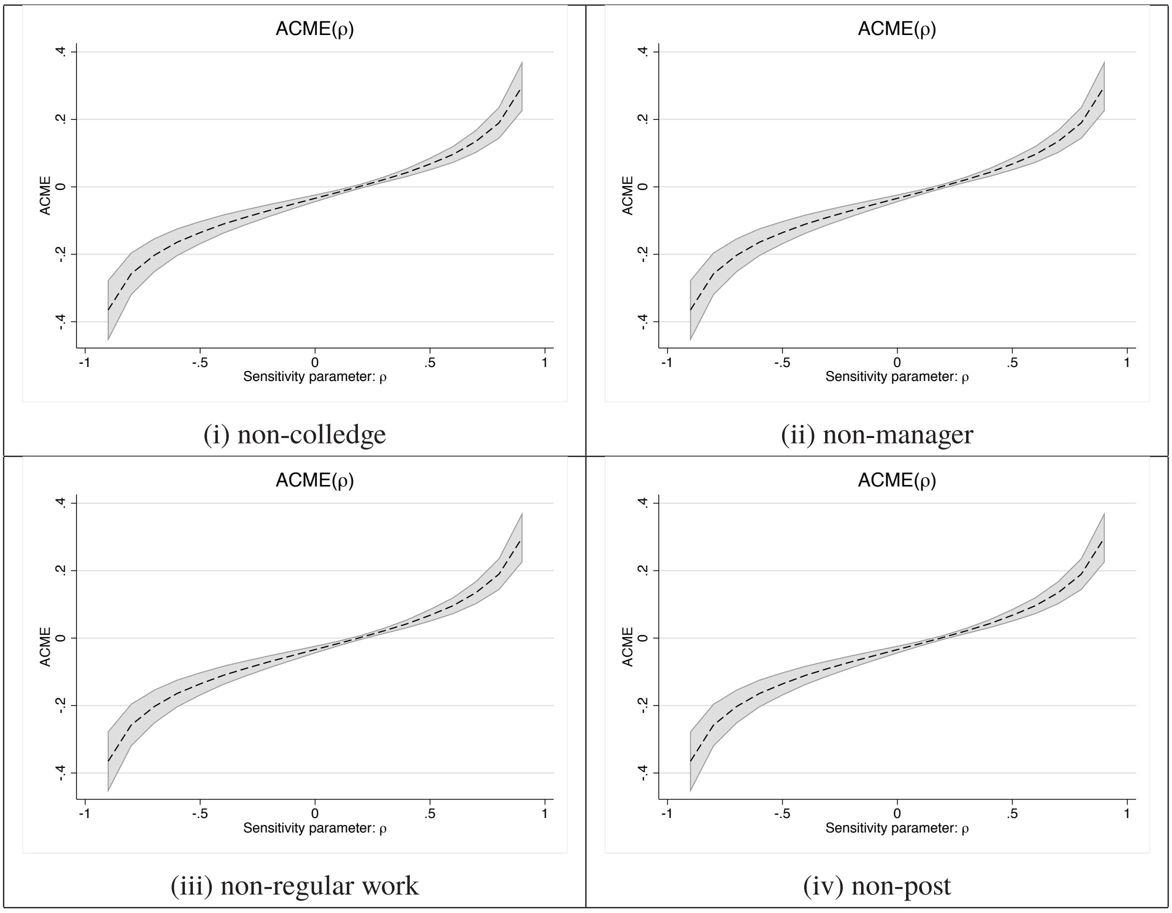

Figure2: Sensitivity analysis for ACME(ρ): The dashed line represents the estimated ACME of the sensitivity parameter ρ. The gray area represents the 95% confidence interval.

estimation. The analysis indicates that the results about the direction of the ACME under SI would be maintained unless ρ is less than -0.150 for (i) non-colledge, -0.098 for (ii) non-manager, -0.207 for (iii) non-regular work and -0.220 for (iv) non-post. This means that the estimation results are plausible except for (ii) non-manager[3], given even large departures from the ignorability assumption of the mediator $\boldsymbol{M}$.

4. Conclusion

This paper aimed to clarify the identification strategy using counterfactual mediation modeling instead of conventional decompositions methods such as Blinder (1973), Oaxaca (1973) and DiNardo et al. (1996). Literatures have

[3] As shown in Table 1, the percentage of manager is less than 7% both for female and male. This small percents indicate there is a limitation for controlling potential confounder of $\boldsymbol{T}$ and $\boldsymbol{M}$, which induces the small absolute value of ρ.

Table2: Counterfactual Mediation Modeling Estimates

	(1) Without PSW		(2) With PSW	
Panel A				
Mediator: Non-colledge				
ACME	-.0409	[-.0514, -.0302]	-.0435	[-.0561, -.0318]
ADME	-.1978	[-.2424, -.1547]	-.1987	[-.2414, -.1527]
TE	-.2387	[-.2820, -.1948]	-.2422	[-.2873, -.1967]
% of TE mediated	.1717	[.1736, .2466]	.1791	[.1513, .2210]
Panel B				
Mediator: Non-manager				
ACME	-.0048	[-.0097, -.0012]	-.0044	[-.0082, .-0015]
ADME	-.1978	[-.2424, -.1547]	-.1976	[-.2419, -.1549]
TE	-.2026	[-.2455, -.1592]	-.2020	[-.2446, -.1589]
% of TE mediated	.0238	[.0196, .0303]	.0217	[.0179, .0276]
Panel C				
Mediator: Non-regular work				
ACME	-.0265	[-.0369, -.0175]	-.0285	[-.0400, .-0184]
ADME	-.1978	[-.2424, -.1547]	-.1976	[-.2419, -.1549]
TE	-.2243	[-.2668, -.1806]	-.2262	[-.2682, -.1825]
% of TE mediated	.1186	[.0995, .1470]	.1266	[.1065, .1564]
Panel D				
Mediator: Non-post				
ACME	-.0503	[-.0626, -.0386]	-.0494	[-.0615, -.0375]
ADME	-.1978	[-.2424, -.1547]	-.1976	[-.2419, -.1549]
TE	-.2481	[-.2900, -.2041]	-.2471	[-.2877, -.2041]
% of TE mediated	.2039	[.1736, .2466]	.2008	[.1718, .2422]

NOTE: Sample size $n = 2829$. Parenthesis [ ] shows 95% confidence interval. TE = ACME + ADME.

clarified conventional decompositions do not control for confounders of the time-constant treatment variable and/or mediator variables. This is the standard case in estimating the parameters of time-constant variables, as gender or ethnicity are determined at or prior to birth and therefore precede mediators like education or profession (Yamaguchi 2015; Huber 2015). We proposed the linking of counterfactual mediation modeling that accounts for causal mechanisms, discussed in Imai et al. (2010), and propensity score weighting decomposition in Yamaguchi (2015).

We conducted the sensitivity analysis for inspecting how robust the estimation results are to the violation of the SI. The sensitivity analysis implies that the conclusion about the direction of the ACME in gender wage gap under SI would be maintained. It is the fate of causal inference to examine the uncertainty and robustness because the assumption in causal inference mostly can not be directly verified. Considering the situation sociologists tries to answer the question asking how robust empirical results are to sensible changes in model specification (Western 1996; Young 2009; Young and Holsteen 2017), it is essential procedure to conduct some diagnosis for investigating the uncertainty and robustness.

This paper also aimed to bring time-constant variables back to the center of statistical causal analysis in social inequality studies. We agree with the statement in Yamaguchi (2015: 426).

> ... the use of panel survey data for causal analysis seems to have diminished the importance of gender or race in statistical causal analyses because of the lack of a methodological framework to handle such time-constant exogenous variables as the treatment variable in causal analysis. I believe that the discussion and the method presented in this paper will lead to a reconsideration of such trends, and it will be complementary to the experimental audit method, because gender and racial inequality are a major substantive research topic in sociology.

Though the counterfactual mediation analysis without randomizing the treatment variables and mediators might not sharply identifies the quantities of interest, comparing to the natural experiment such as sharp regression discontinuity, it's much better than not doing it.

References

Baron, Rueben M., and David A. Kenny. 1986. "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations." *Journal of Personality and Social Psychology* 51(6): 1173-1182.

Blinder, Alan S. 1973. "Wage Discrimination: Reduced Form and Structural Estimates." *Journal of Human Resources* 8(4): 436-455. DiNardo, John, Nicole M. Fortin, and Thomas Lemieux. 1996. "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica* 64(5): 1001-1044.

Fortin, Nicole, Thomas Lemieux, and Sergio Firpo. 2011. "Decomposition Methods in Econometrics." Orley Ashenfelter and David Card eds., *Handbook of Labor Economics vol. 4a*, North Holland, 1-102.

Frölich, Markus. 2007. "Propensity Score Matching Without Conditional Independence Assumption with an Application to the Gender Wage Gap in the United Kingdom." *Econometrics Journal* 10(2): 359-407.

Huber, Martin. 2015. "Causal Pitfalls in the Decomposition of Wage Gaps." *Journal of Business and Economic Statistics* 33(2): 179-191.

Imai, Kosuke, Luke Keele, and Teppei Yamamoto. 2010. "Identification, Inference and Sensitivity Analysis for Causal Mediation Effects." *Statistical Science* 25(1): 51-71.

Imai, Kosuke, Luke Keele, Dustin Tingley, and Teppei Yamamoto. 2011. "Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies." *American Political Science Review* 105(4): 765-789.

Morgan, Stephen L., and Christopher Winship. 2015. *Counterfactuals and Causal Inference: Methods and Principles for Social Research.* 2nd ed. Cambridge, UK: Cambridge University Press.

Oaxaca, Ronald. 1973. "Male-Female Wage Differences in Urban Labour Markets." *International Economic Review* 14(3): 693-709.

Okubo, Shoki. 2018. "Simulating Average Causal Mediation Effects: A Comparison of BO, DFL, ANME and ACME." *mimemo*.

Pearl, Judea. 2001. "Direct and Indirect Effects." in *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence*, San Francisco: Morgan Kaufman, pp.411-420.

Pearl, Judea. 2009. *Causality.* 2nd ed. Cambridge, UK: Cambridge University Press.

Robins, James. M., and Sander Greenland. 1992. "Identifiability and Exchangeability for Direct and Indirect Effects." *Epidemiology* 3(2): 143-155.

Western, Bruce. 1996. "Vague Theory and Model Uncertainty in Macrosociology." *Sociological Methodology* 26(1): 165-92.

Yamaguchi, Kazuo. 2015. "Decomposition of Gender or Racial Inequality with Endogenous Intervening Covariates: An Extension of the DiNardo-Fortin-Lemieux Method." *Sociological Methodology* 45(1): 388-428.

Young, Cristobal. 2009. "Model Uncertainty in Sociological Research: An Application to Religion and Economic Growth." *American Sociological Review* 74(3): 380-97.

Young, Cristobal, and Katherine Holsteen. 2017. "Model Uncertainty and Robustness: A Computational Framework for Multimodel Analysis." *Sociological Methods and Research* 46(1): 3-40.