Gender pay gap in Vietnam: a propensity score matching analysis

Gabriel Obermann
Institute for Public Sector Economics,
Vienna University of Economics and Business, Vienna, Austria
Nguyen Hoang Oanh
Faculty of Economics, National Economics University, Hanoi, Vietnam, and
Nguyen Hong Ngoc
The University of Queensland, Saint Lucia, Australia

Abstract

Purpose – This paper investigates the extent, the determinants and the change in the gender pay gap in Vietnam in the period 2010–2016 in order to provide suggestions for policy adjustment to narrow gender pay inequality more effectively.

Design/methodology/approach – This study employs the propensity score matching (PSM) method to examine inequality in pay between female and male earners sharing identical characteristics. The analysis is conducted for both the full sample and various characteristic-based subsamples. This procedure is conducted for 2010 and 2016 separately to discover the change in gap and inequality during this period.

Findings – The matching results based on the data sets taken from the Vietnam Household Living Standards Survey (VHLSS) 2010 and 2016 affirm that gender income inequality in Vietnam, though persisted, decreased significantly in 2016 compared to 2010, and was insignificant in many subsamples in 2016. In addition to the observable determinants including educational level, occupation, economic sector and industry, unobservable factors are proved to also play an important role in creating the gender pay gap in Vietnam.

Practical implications – The research findings suggest that policies aimed at mitigating gender pay inequality should take into account both observable characteristics and unobservable factors such as unobservable gender differences that affect wages and gender discrimination in pay.

Originality/value – This is the first study using a matching technique to investigate gender wage gap in Vietnam. With up-to-date data, longer research period and the superiority of the method used in dealing with sample selection bias, the results obtained are more robust, more detailed and reliable.

Keywords Gender pay gap, Propensity score matching, Vietnam

Paper type Research paper

1. Introduction

Labor income not only plays a particularly important role in ensuring the life of laborers and reproducing labor power but also serves as a management tool to improve labor quality and productivity. If labor income is distributed equally among individuals in society, it will be a driving force for sustainable growth.

In fact, gender inequality in earnings is very common. Across the world, women only make 77 cents for every dollar that men earn. Even for jobs that require the same or higher efforts and

skills, women are still undervalued and underpaid. Especially, for women of color, immigrant women and mothers, the pay gap is even larger. The so-called "motherhood penalty" pushes women into informal economy and casual and part-time jobs, and this trend is more common and to a greater extent in developing countries than in developed ones (UN Women, 2020). There are many causes of this situation, in which the root cause comes from the traditional view and stereotypical prejudices against women in feudal societies that limit opportunities for women to access to education and training, career selection and professional qualification improvement. In addition, women also have fewer opportunities than men to access basic services and resources to assert their economic position such as transportation, markets, funds and so on.

In Vietnam, the current system of laws and policies has created a relatively comprehensive legal framework to ensure gender equality in economic activities, labor supply and employment. Although Vietnam has gained remarkable achievements in narrowing the gender inequality, the gender pay gap still exists. Statistics indicate that the labor force participation rate of females is increasing and remains high compared to other countries in the region and in the world. If the labor force participation rate of women in 2010 was 76.92%, this rate increased to 78.63% in 2016 (ILOSTAT, 2020). However, the average monthly salary of female workers in 2004 was only 80 percent of that of male workers, but this rate increased gradually and reached 88.3 percent in 2016 (ILSSA and ILO, 2018).

The issue of gender pay inequality has attracted great attention of researchers all around the world. A common method used to investigate the gender pay gap is the Blinder–Oaxaca (B–O) decomposition method. This method has been improved by Juhn *et al.* (1993) and Machado and Mata (2001) to investigate the change in pay gap over time and to account for differences in labor market features. However, these methods have certain limitations, including the commonly seen limitations of the parametric methods and also the problem of comparability in the supports, thus leading to overestimates of the component of the gap attributable to differences in individuals' characteristics (Nopo, 2004). Thus, Nopo (2004) proposed a new non-parametric matching technique to explain gender pay gap as an alternative to the traditional B–O method. The most important advantage of this method is that it produces more robust and reliable results since it overcomes the heterogeneity of the samples under investigation and avoids diseases commonly faced by the parametric methods. Until now, the number of international researches on income inequality employing the matching approach is still small.

In Vietnam, there have been some empirical studies on this topic. These studies, regardless of being conducted for the whole economy, by economic sector or by income group, all confirm that gender inequality in pay persists. However, the results of these studies may not be very convincing given the limitations of the methods applied. This paper examines the gender pay inequality of wage earners in Vietnam for the period 2010–2016 using the propensity score matching (PSM) method. This is the first study on gender pay gap in Vietnam that follows this approach. The results from the matching analysis using the Vietnam Household Living Standards Survey (VHLSS) data sets of 2010 and 2016 allow us to affirm with certainty that gender pay inequality persists, though the extent has decreased significantly over time. In addition to the observable determinants including educational qualification level, occupation, economic sector and industry, unobservable factors are proved to also play an important role in creating the gender pay gap.

The following section of the paper explores the related literature. Section 3 presents the matching comparison technique and the data used in the study. Section 4 describes the estimation results. Section 5 discusses the results and concludes.

2. Literature review
From an economic perspective, wages are determined by the two important forces – labor supply and labor demand – in the labor market. Labor supply is affected by individuals' demographic characteristics and the factors related to human capital, which are likely

influenced by discrimination, while labor demand is determined by factors related to firm attributes and discrimination. Based on such idea of a market, Becker's work (1971) laid the foundation for the economics of discrimination. According to him, discrimination occurs on both the supply and demand sides, and discrimination creates costs to society – the discriminated worker is paid less, and the discriminator incurs greater expense to hire a worker with the same productivity with the discriminated one. Thus, according to market rules, competitiveness will reduce gender discrimination over time. In contrast to taste-based theories of discrimination represented by Becker, Phelps (1972) and Arrow (1973) introduced the theory of statistical discrimination. According to them, due to the problem of imperfect information in the labor market, employers are forced to rely on statistical information on different groups of applicants to infer their productivity, and thus being able to make recruitment decisions. In the absence of assumptions on racial or gender animosity against members of a targeted group, recruitment and employment decisions based on the group average, by accident, create inequality for minority groups.

Various empirical studies all around the world have been done in an attempt to not only measure the extent but also identify the causes and discover the trends in pay gap and inequality. Blau and Kahn (2017) provide a relatively comprehensive literature review on the explanations for gender wage gap. In addition to traditional market-based factors, Blau and Kahn (2017) investigate the impact of psychological attributes or non-cognitive skills on gender wage inequality.

The literature on gender wage gap has recorded various decomposition methods, ranging from parametric (Blinder, 1973; Oaxaca, 1973; Juhn et al., 1993; Machado and Mata, 2001), semi-parametric (DiNardo et al., 1996; Donald et al., 2000; Bourguignon et al., 2002), to non-parametric (Rosenbaum and Rubin, 1983; Nopo, 2004) methods. From another angle, Fortin et al. (2011) classify decomposition methods into mean decomposition and beyond the mean decomposition. A representative of the former group is the traditional B–O method, while the latter includes the extended versions of the B–O method.

Some economists argue that as gender inequality in earnings, in essence, means the difference between earnings of male and female workers despite the same characteristics of labor capacity and productivity, it is very important to strictly control the differences in characteristics of the two comparative groups. Thus, the comparison must be restricted to a common support, where there is sufficient overlap in the characteristics of treated and untreated individuals to find adequate matches (Rosenbaum and Rubin, 1983). Nopo (2004) pointed out the limitations of the traditional B–O method in controlling the differences between the two comparative groups and introduced an alternative method, a new matching technique, to explore the determinants of wage disparity between male and female groups in Peru during the period 1986–2000. Applying the PSM technique, Frolich (2007) conducted an analysis on gender wage gap of UK graduates. Meara et al. (2017) investigated the gender wage gap in the USA, applying the extended PSM method in combination with the inverse probability weighted regression adjustment to address the problems of sample selection bias.

Gender pay gap in Vietnam has been studied using different methods, but, so far, there have been no studies using matching techniques. It is noteworthy that regardless of approach, scope, point and span of time, all studies assert that gender pay gap persists in Vietnam, but the determinants are diverse. Amy (2004) investigates the sectoral gender wage gap in Vietnam during the period 1997–1998 using the decomposition model proposed by Juhn et al. (1993). Pham and Reilly (2007) conducted a mean and quantile regression analysis to explore the gender pay gap for the wage earners in Vietnam over the period 1993–2002. Gian (2014) conducted a comparison analysis of the gender wage gap between Korea and Vietnam using the decomposition method developed by Juhn et al. (1993). Nguyen and Hoang (2018) applied the B–O method to identify the determinants of the gender wage gap in Vietnam for the period 2012–2014. Vu and Yamada (2018)

decompose the gender equality in terms of wage distribution in Vietnam during the period 2002–2014 using two methods: one is suggested by Chernozhukov *et al.* (2013) and the other is proposed by Firpo *et al.* (2009).

3. Methodology and data

3.1 The PSM technique

This study employs the PSM technique to investigate the gender pay gap in Vietnam in the period 2010–2016.

Based on the concept "propensity scores" introduced for the first time in experimental designs by Rosenbaum and Rubin (1983), Nopo (2004) develops a simple matching procedure, a fully non-parametric method, to decompose gender wage gap. This procedure helps to select two subsamples of males and females having the same characteristics, and thus constructing a counterfactual wage for the common support. By seeking for matched samples with "identical" observable characteristics, this technique helps to solve the problems of sample selection bias created by the traditional B–O method and its later versions.

In this study, female earners constitute the treatment group and male earners form the untreated one. The control group is established as a subset of the untreated group by selecting only males that are identical in all other key characteristics to the treated group. As usual, a probit model is developed to identify the key characteristics. By defining an outcome variable (log of hourly pay) and a binary treatment variable, the matching technique seeks to establish whether a statistically significant difference exists in the log of hourly pay between the treated group and the control one, to investigate the inequality in pay between females and males sharing the same characteristics. This procedure is conducted for 2010 and 2016 separately to discover the change in gap and inequality during this period.

The main parameters in our study are, therefore, the average treatment effect (ATE) for the population, which represents the pay gap between the treated and the untreated groups, and the average treatment effect for "treated" individuals (ATET), which represents the pay gap between the treated group and the control one. The latter is obtained with the use of matching method, and the former is for comparison purpose. Let Y be the log of the hourly pay of an individual, D indicates whether or not "treatment" is received, "1" denotes individuals who are "treated" (females) and "0" denotes those who are untreated (males); thus, these two parameters can be given by:

$$\text{ATE} = E(Y_{1i} - Y_{0i}) \equiv E(\beta_i),$$

$$\text{ATET} = E(Y_{1i} - Y_{0i}|D_i = 1) \equiv E(\beta|D_i = 1).$$

According to Nopo (2004), ATET can be considered as a component of ATE. By performing further algebraic transformation, the pay gap (ATE) can be decomposed into four elements and can be expressed in a simple way as follows

$$\Delta = \Delta_M + \Delta_F + \Delta_X + \Delta_0,$$

where Δ represents the pay gap (ATE), Δ_M the part that can be explained by differences between two groups of males matched and unmatched to female characteristics, Δ_F the gap explained by the differences in characteristics between two groups of females whose characteristics can be matched and unmatched to male characteristics, Δ_X the gap explained by differences in the distribution of (observable) characteristics of males and females over the common support, which can be eliminated using the matching approach, and Δ_0 is the gap that cannot be explained, presumably due to unobservable characteristics or gender discrimination. Δ_0 represents ATET and is of utmost interest to our study. Also it should be noted that in this study, all females are matched to the corresponding control groups, and therefore Δ_F does not exist ($\Delta_F = 0$).

The PSM procedure is applied to the full sample as well as various subsamples based on a wide range of key characteristics selected in order to investigate comprehensively the equality in pay between the two genders. Using the propensity scores, a comparable control group comprising males who are similar in terms of all the chosen key observable characteristics is selected for each treatment group of females under investigation. The goal is to identify whether there exists a statistically significant pay gap and the magnitude of the gap between these two essentially similar groups, which is tested using bootstrapped standard errors. A variety of subsamples are, in fact, employed in order to obtain even finer matching of control and treatment groups (and therefore more accurate comparison) and also to evaluate the influences of each selected characteristics on the gender pay gap. Note that, regarding occupation, since there are ten occupational categories, division of the sample by these ten categories may lead to subsamples of insufficient size. Consequently, only three main groups of occupation (namely, professionals, skilled and unskilled workers) are taken into account.

To begin with, the propensity scores are obtained for the full sample by running a probit model with the dependent variable being gender, which takes the value of 1 for females and 0 otherwise, and independent variables consisting of the following key common characteristics:

(1) Marital status (1 if married and 0 otherwise).

(2) Family role (1 for household heads and 0 otherwise).

(3) Household size (number of people).

(4) Dependency ratio (ratio of the number of dependent people (children and the elderly) to household size).

(5) Migrant (1 for migrants and 0 otherwise).

(6) Urban (1 for urban workers and 0 for rural ones).

(7) Ethnicity (1 if from Kinh ethnic group and 0 otherwise).

(8) Educational levels (no qualification; primary school; secondary school; high school; and college and higher qualification).

(9) Occupation categories (leaders in all fields and levels; top-level professionals; mid-level professionals; staff (elementary professionals, white-collar technical personnel); skilled workers in personal services, security protection and sales; skilled workers in agriculture, silviculture and aquaculture; skilled handicraftsmen and other manual workers; assemblers and machine operators; unskilled workers; and army force).

(10) Industry (agriculture, silviculture and aquaculture; industry and construction; and service).

(11) Economic sector (non-state, state, Foreign Direct Investment (FDI)).

(12) Region [1] (Red River Delta, Northern Midland and Mountainous, North Central and Central Coast, Central Highlands, Southeast, Mekong River Delta).

For each subsample based on certain selected characteristics, the probit model used is the same as for the full sample except for the fact that the corresponding selected characteristic is no longer included as an independent variable in the regression.

3.2 Data
This study examines and compares the gender pay gap in Vietnam between two separate years, 2010 and 2016, using the VHLSS data sets of the corresponding years. Consequently, only employed individuals (wage earners) with available information on incomes are included. We employ the hourly wage of females and males to measure gender pay gap, which accounts for the differences in both the total earnings and the number of hours worked of the two genders. The hourly wage of each individual is obtained by dividing his/her annual earnings from his/her primary job by total working hours spent on that job in the corresponding year. Data on annual income are calculated as the sum of his/her salary and all types of accompanying bonuses, allowances and other benefits from the job during 12 months prior to the survey; the total number of working hours is computed by multiplying the number of his/her working days in 12 months preceding the survey by the average number of hours per working day he/she spends on that job.

The resulting sample in 2010 comprises a total of 7,561 individuals with 3,018 females and 4,543 males (accounting for 39.92% and 60.08% of the total, respectively), while the 2016 sample is considerably larger, with 8,508 individuals and a slightly higher proportion of females (41.58%). This might imply a relative improvement in working opportunities for women between 2010 and 2016.

Regarding labor income, which is the utmost focus of this paper, the average hourly earnings for employed females was 13.87 compared to 15.51 thousand VND for males in 2010. The figures nearly doubled for 2016, with female and male earnings of 27.52 and 30.52 thousand VND, respectively. Accordingly, gender pay gaps, though still significant, reduces from 11.89% to 10.91% between the two years, proving an equality improvement.

Considering potential determinants of gender pay gap, while the characteristics regarding marital status, size of household and rate of dependency are rather comparable for both groups, the percentage of men assigned to be household head is exceedingly higher than that of women in both samples. This somewhat reflects social perception of gender roles in families and may suggest a higher sense of responsibility for males to work and earn more. In fact, the average hourly income of household heads was higher than that of other members of the family, but the gap decreased by almost half, from 16.03% in 2010 to 8.29% in 2016 (see Table 2).

On the contrary, Table 1 shows relatively higher proportions of women who are migrants and from the Kinh group than those for men in both samples. More remarkably, urban rates,

| | 2010 | | 2016 | |
	Females	Males	Females	Males
Observations	3,018	4,543	3,538	4,970
Average hourly income (VND 1,000)	13.87	15.51	27.52	30.52
% married[a]	74.29	72.38	99.27	99.38
% household heads	19.09	48.40	14.73	48.29
Average household size (number of people)	4.31	4.36	4.38	4.32
Average dependency ratio[b]	27.71	26.03	29.96	27.89
% migrants[c]	2.68	2.05	1.16	1.11
% urban	42.01	36.19	41.92	38.03
% Kinh ethnic	90.76	89.76	89.88	87.34
Average years of schooling (years)	8.98	8.84	9.33	8.99
Average educational level	2.34	2.17	2.50	2.22

Note(s): [a]Married and currently living with their partners (not widowed, divorced or separated)
[b]Calculated as the ratio of the number of dependent people (children and the elderly) to household size
[c]In this study, migrants are defined as those who have been living in their current city/province for less than five years
Source(s): VHLSS 2010 and 2016

Table 1.
Selected characteristics
of employed women
versus employed men
in the two samples

		2010 Hourly wage	% female		2016 Hourly wage	% female
	Obs	wage	female	Obs	wage	female
Household role						
Head	2,775	16.28	20.76	2,921	30.82	17.84
Member	4,786	14.03	51.02	5,587	28.46	54.00
Living area						
Urban	2,912	19.45	43.54	3,373	35.05	43.97
Rural	4,649	11.98	37.64	5,135	25.47	40.02
Education level						
None	746	10.02	40.62	768	20.67	37.89
Primary school	1,543	10.74	37.20	1,646	22.53	36.39
Middle school	1,922	11.62	33.04	2,098	24.22	37.23
High school	1,846	15.53	43.07	1,959	29.55	40.53
College and higher	1,504	24.78	47.27	2,037	42.89	52.68
Economic sector						
Non-state	4,932	12.14	34.94	5,636	25.58	36.00
State	2,173	20.30	45.84	2,036	38.51	46.56
FDI	456	18.22	65.57	836	31.66	67.11
Industry						
Agriculture	183	10.67	15.30	272	22.62	20.22
Industry and construction	3,468	13.14	34.83	3,998	26.80	38.27
Service	3,910	16.57	45.58	4,238	32.03	46.08
Occupation						
Unskilled workers	2,262	9.80	42.09	2,214	20.22	40.74
Leaders	235	23.81	23.40	231	52.53	20.78
Top-level professionals	847	28.96	47.58	1,004	47.82	55.58
Mid-level professionals	733	20.66	61.53	604	37.86	59.11
Staff	310	14.42	51.94	384	30.34	53.39
Skilled workers in service	583	11.57	44.94	790	23.41	45.06
Skilled workers in agriculture	123	16.17	26.83	145	24.75	30.34
Skilled workers in industry and construction	1,730	11.61	28.67	2,001	25.64	29.64
Assemblers and machine operators	684	15.23	29.68	1,063	29.26	44.03
Army	54	22.13	3.70	72	71.01	9.72
Region						
Red River Delta	–	–	–	2,054	33.32	43.18
Northern Midland and Mountainous	–	–	–	1,208	29.84	39.82
North Central and Central Coast	–	–	–	2,181	28.93	39.89
Central Highlands	–	–	–	286	24.83	44.41
Southeast	–	–	–	1,005	31.36	43.78
Mekong River Delta	–	–	–	1,774	24.14	41.32

Note(s): Data on regions where individuals were from in the VHLSS data set in 2010 are not sufficient for calculation

Source(s): Authors' calculation using data from VHLSS 2010 and 2016

Table 2.
Average hourly wage and proportion of females by selected criteria

average years of schooling and educational level of female earners are significantly larger compared to males in both years. However, they are, on average, paid lower than males, although those factors have been widely proved to result in higher income in literature. This can also be seen in our sample. Table 2 shows that, on average, urban workers earned 1.62

and 1.38 times higher than rural ones in 2010 and 2016, respectively. Earnings are also shown to increase sharply with education level, with salary at the highest educational level (college or higher degree) being 2.47 and 2.08 times higher than that of the lowest level (no degree) in 2010 and 2016, respectively.

Also, while improvement in education takes place for both groups over time, it is more significant for female workers. In terms of education, economic sector, industry or occupation, the shares of women in the categories that yield higher income tend to be larger. Regarding occupation, the two categories that offer the highest income, leaders and the army force, however, have strikingly low proportions of females, but these also have the smallest numbers of observations in the sample.

Overall, Table 2 indicates that the mean hourly earnings vary greatly across different categories in each criterion, proving the significant impacts of these characteristics on the level of earnings earned by individuals. Nonetheless, according to the facts discussed above, women, though having characteristics that are likely to lead to higher wages, are still paid less than men. This implies that these determinants of earnings are insufficient to explain the gender pay gap in Vietnam in the two periods under study and suggests the contribution of other factors, presumably unobserved ones. This judgment is further strengthened in the following section.

4. Empirical results

Matching results for the full sample and most subsamples in 2010, indeed, provide a strong evidence of a statistically significant pay gap between women and men that share more or less the same characteristics. In other words, women are firmly proved to be paid significantly less than men in this year, and the differences in earnings arise from other (perhaps unobservable) factors besides the wide range of observable ones that are taken into account.

Nevertheless, inequality is shown to be improved dramatically in 2016. Though still significant, the overall wage gap reduces by half from 28.20% in 2010 to 14.29% in 2016 between females and their matched male group. Regarding the subsamples, the gap is insignificant in not only the ones with small sizes but also many others, most notable of which are the state sector and service industry. It, also, has low magnitudes associated with low significance levels in many subsamples, particularly the one of professionals.

When having a closer look at the full sample of 2010, we find that females were paid 28.20% less than males sharing similar characteristics (at the significance level of 0.01). This gap is considerable and much larger than the gap for unmatched groups (16.88%) (See Table 3). This means that the males in the control group, who are identical to the treated group of females in terms of characteristics, on average, earn more than the bigger untreated (male) group, which, in turn, implies that, overall, women possess characteristics that would lead to considerably higher earnings but were treated very unequally.

Despite large variation across different subsamples, the estimated gaps between matched groups, when significant, are also higher compared to unmatched ones for all subsamples,

Sample	Treated	Controls	Difference	S.E.	T-stat	Observations Untreated (male)	Treated (female)
Full sample							
Unmatched	9.2628	9.4316	−0.1688	0.0156	−10.84	4,543	3,018
ATET	9.2628	9.5448	−0.2820***	0.0289	−9.77		
Note(s): ***indicates statistical significance at 0.01 level							

except for two groups relying on educational level (the two lowest levels) (See Tables 4 and 5). This further supports the existence of a profound gender pay inequality.

Considering the differences in earnings between females and their matched groups, most subsamples see lower pay gap compared to the full sample, which shows the importance of controlling the distribution of characteristics of the samples to the accuracy of the results (see Tables 4 and 5). Three out of the four exceptions are industry and construction, non-state and skilled workers, all of which are all male-dominated, with the number of males nearly doubling females. They witness the highest pay gaps of 35.21%, 32.96% and 31.90% (all at significance level of 0.01), respectively. The remaining exception is the Kinh group, which accounts for more than 90% of the whole sample, with pay gap only slightly higher than the overall one.

The subsamples with the lowest significant gaps and, therefore, highest income equality, on the other hand, are high school level workers, service workers, professionals and workers in the state sector (with pay gaps of 17.46%, 19.46%, 20.03% and 20.50%, respectively, all at significance level of 0.01). These also suggest that educational level, industry, occupation and economic sector offer good explanations for the gender pay gap in 2010. It is also noteworthy that the wage gap is insignificant in the three smallest subsamples, namely, ethnic minorities, agriculture and FDI workers, with very small and even positive gaps for the first two groups.

By the selected key characteristics, the (significant) estimated pay differences are rather comparable between married and unmarried workers but much lower for household heads than other members of the household, for service than for industry and construction, and for state workers than for their non-state counterparts. Rural areas also witness lower pay gap than urban ones. Among five educational categories, the gap estimated is the lowest for the high school level and the highest for those at middle school level. Pay equality also differs greatly among the three major occupations, of which professionals have the highest equality, followed by unskilled workers, while skilled female workers are treated highly unequally.

Compared to 2010, gender pay differences decreased greatly for both full sample and all subsamples in 2016 (see Tables 6–8). From 2010 to 2016, overall, the estimated gender pay gap between the treatment and control groups for the full sample falls dramatically to 14.29% (at the significance level of 0.01) in 2016, nearly half of that in 2010. This fact proves a sharp progress in enhancing gender pay equality. However, the gap between matched groups still exceeds that of unmatched ones, at 9.15%, indicating that unequal treatment at work between males and females, though has considerably lessened, persisted in 2016.

Similar to the pattern observed in the results for the 2010 sample, in 2016, the use of matching method produces higher gaps for many of subsamples considered, which shows inequality in earnings persists between females and their male counterparts (see Tables 7 and 8). Exceptions are all the subsamples relying on educational level (except for the high school level), professionals and workers in non-state sector (and also workers in agriculture and FDI sector and from the Northern Midland and Mountainous, but the estimated gaps for these are not significant, largely due to insufficient observations). This indicates that, in these subsamples, gender pay differences can indeed be partly explained by the observed characteristics chosen.

It can be seen from Tables 7 and 8 that gender pay inequality varies enormously across different subsamples. The biggest (significant) estimated gender pay gaps were for workers from the Central Highlands (20.18%), the Red River Delta (19.95%) and the Mekong River Delta (18.41%), and those without any educational qualification (19.05%). Many other subsamples in 2016 also have larger gender gaps than the full sample, though most of them made considerable progress compared to the previous six years. These facts suggest the persistence of a profound gender inequality in pay in 2016, when differences are even greater among the same subgroups compared to the entire sample.

On the contrary, gender pay gap is insignificant in many subsamples, not only those with few observations but also those of large size. They include the subsamples of workers from

Sample	Treated	Controls	Difference	S.E.	T-stat	Observations Untreated (male)	Treated (female)
Sample divided by marital status							
(a) Married							
Unmatched	9.3248	9.4944	−0.1696	0.0185	−9.16	3,288	2,242
ATET	9.3248	9.5967	−0.2719***	0.0284	−9.56		
(b) Unmarried							
Unmatched	9.0836	9.2694	−0.1858	0.0272	−6.83	1,250	776
ATET	9.0836	9.3510	−0.2674***	0.0532	−5.03		
Sample divided by family role							
(a) Head							
Unmatched	9.4071	9.4619	−0.0548	0.0311	−1.76	2,187	576
ATET	9.4071	9.6326	−0.2255***	0.0550	−4.1		
(b) Not head							
Unmatched	9.2288	9.4014	−0.1726	0.0191	−9.04	2,344	2,442
ATET	9.2288	9.5043	−0.2755***	0.0342	−8.05		
Sample divided by living area							
(a) Urban							
Unmatched	9.5068	9.6922	−0.1854	0.0261	−7.12	1,644	1,268
ATET	9.5068	9.7828	−0.2760***	0.0312	−8.85		
(b) Rural							
Unmatched	9.0860	9.2817	−0.1956	0.0178	−11.01	2,872	1,750
ATET	9.0860	9.3337	−0.2477***	0.0340	−7.29		
Sample divided by ethnicity							
(a) Kinh people							
Unmatched	9.2728	9.4610	−0.1882	0.0164	−11.45	4,078	2,739
ATET	9.2728	9.5691	−0.2964***	0.0307	−9.66		
(b) Ethnic minorities							
Unmatched	9.1652	9.1683	−0.0031	0.0456	−0.07	461	279
ATET	9.1652	9.1582	0.0070	0.0868	0.08		
Sample divided by educational level							
(a) None							
Unmatched	8.8287	9.1515	−0.3228	0.0434	−7.44	441	303
ATET	8.8287	9.0688	−0.2401***	0.0903	−2.66		
(b) Primary school							
Unmatched	8.9704	9.2546	−0.2842	0.0265	−10.74	963	573
ATET	8.9704	9.1829	−0.2125***	0.0449	−4.73		
(c) Middle school							
Unmatched	9.0245	9.3004	−0.2759	0.0264	−10.46	1,283	635
ATET	9.0245	9.3026	−0.2781***	0.0521	−5.34		
(d) High school							
Unmatched	9.3866	9.4970	−0.1104	0.0287	−3.84	1,051	795
ATET	9.3866	9.5613	−0.1746***	0.0531	−3.29		
(e) College and higher							
Unmatched	9.7570	9.9346	−0.1777	0.0378	−4.7	793	711
ATET	9.7570	9.9744	−0.2174***	0.0674	−3.22		
Sample divided by occupation							
(a) Professionals							
Unmatched	9.8332	9.9812	−0.1480	0.0330	−4.48	726	854
ATET	9.8332	10.0335	−0.2003***	0.0600	−3.34		
(b) Skilled workers							
Unmatched	9.0998	9.3972	−0.2975	0.0201	−14.83	2,275	1,155
ATET	9.0998	9.4188	−0.3190***	0.0314	−10.15		
(c) Laborers							
Unmatched	8.9169	9.1430	−0.2261	0.0218	−10.37	1,310	952
ATET	8.9169	9.1541	−0.2372***	0.0369	−6.43		

Table 4.
The PSM results across
different subsamples
in 2010

Sample	Treated	Controls	Difference	S.E.	T-stat	Observations Untreated (male)	Treated (female)
Sample divided by industry							
(a) Agriculture							
Unmatched	9.0546	9.0843	−0.0297	0.1154	−0.26	147	27
ATET	9.0546	9.0298	0.0247	0.2294	0.11		
(b) Industry and construction							
Unmatched	9.0769	9.4051	−0.3282	0.0205	−16	2,260	1,208
ATET	9.0769	9.4291	−0.3521***	0.0395	−8.91		
(c) Service							
Unmatched	9.3912	9.4839	−0.0926	0.0231	−4.01	2,128	1,782
ATET	9.3912	9.5858	−0.1946***	0.0355	−5.48		
Sample divided by economic sector							
(a) Non-state							
Unmatched	9.0062	9.3248	−0.3185	0.0172	−18.53	3,208	1,723
ATET	9.0062	9.3359	−0.3296***	0.0316	−10.43		
(b) State							
Unmatched	9.6791	9.6796	−0.0005	0.0308	−0.01	1,177	996
ATET	9.6791	9.8841	−0.2050***	0.0578	−3.54		
(c) FDI							
Unmatched	9.3547	9.7169	−0.3622	0.0590	−6.14	154	299
ATET	9.3547	9.4668	−0.1121	0.0934	−1.2		

Table 5.
The PSM results across different subsamples in 2010 (continued)

Table 6.
The PSM results for full sample in 2016 (outcome variable = log of hourly earnings, treatment = female, bootstrapped standard errors)

Sample	Treated	Controls	Difference	S.E.	T-stat	Observations Untreated (male)	Treated (female)
Full sample							
Unmatched	10.0386	10.1301	−0.0915	0.0129	−7.1	4,970	3,538
ATET	10.0386	10.1815	−0.1429***	0.0248	−5.75		

the Northern Midland and Mountainous, FDI sector, agriculture, state sector, service industry, ethnic minorities and the Southeast, the first two of which witness positive wage gaps (i.e. women are paid more than their male counterparts). The lowest significant inequality is also seen at low significant level and small magnitude, for professionals (2.29%, with the low significance level of 0.10), workers whose highest qualifications are from high school (8.70%) and middle school (9.50%) and workers from the North Central and Central Coast (10.05%) (all with significance levels of 0.05). The estimated gender pay gaps for most of the above-mentioned all have dropped sharply between 2010 and 2016, even though workers in these subgroups are also those experiencing the lowest gender gaps in 2010, except for professionals and middle school educational level. They, in fact, are those who observe the largest significant improvement among all subgroups in 2016, with wage gap falling from 20.03% to 2.29% and from 27.81% to 9.50%, respectively, between the two years.

The high level of income equality observed in these subsamples also indicates that in this sample, gender pay discrepancies can indeed be partly explained by occupation, educational background, region, industry and economic sector.

By each characteristic, pay differences are much lower for professionals than for skilled workers and laborers and even more substantially for service compared to industry and construction and for state compared to non-state, which is in line with the results in 2010. Likewise, by educational level, gender inequality is also lowest for high school-level workers,

Sample	Treated	Controls	Difference	S.E.	T-stat	Observations Untreated (male)	Treated (female)
Sample divided by marital status							
(a) Married							
Unmatched	10.0441	10.1346	−0.0905	0.0129	−7.03	4,939	3,512
ATET	10.0441	10.1666	−0.1225***	0.0215	−5.7		
(b) Unmarried							
Unmatched	9.2639	9.4174	−0.1536	0.1591	−0.97	28	23
ATET	9.2639	8.8857	0.3782	0.4610	0.82		
Sample divided by family role							
(a) Head							
Unmatched	10.0283	10.1567	−0.1284	0.0283	−4.54	2,379	521
ATET	10.0283	10.2056	−0.1774***	0.0500	−3.55		
(b) Not head							
Unmatched	10.0404	10.0993	−0.0590	0.0156	−3.78	2,570	3,017
ATET	10.0404	10.1703	−0.1299***	0.0245	−5.29		
Sample divided by living area							
(a) Urban							
Unmatched	10.1961	10.3007	−0.1047	0.0209	−5.01	1,890	1,483
ATET	10.1961	10.3218	−0.1257***	0.0362	−3.47		
(b) Rural							
Unmatched	9.9249	10.0254	−0.1005	0.0156	−6.44	3,080	2,055
ATET	9.9249	10.0638	−0.1388***	0.0281	−4.95		
Sample divided by ethnicity							
(a) Kinh people							
Unmatched	10.0512	10.1669	−0.1156	0.0134	−8.64	4,341	3,180
ATET	10.0512	10.1786	−0.1274***	0.0245	−5.2		
(b) Ethnic minorities							
Unmatched	9.9262	9.8712	0.0551	0.0418	1.32	623	358
ATET	9.9262	9.9998	−0.0736	0.0815	−0.9		
Sample divided by educational level							
(a) None							
Unmatched	9.6424	9.8944	−0.2521	0.0396	−6.36	476	291
ATET	9.6424	9.8329	−0.1905**	0.0811	−2.35		
(b) Primary school							
Unmatched	9.7700	9.9749	−0.2048	0.0258	−7.93	1,044	599
ATET	9.7700	9.9283	−0.1583***	0.0433	−3.65		
(c) Middle school							
Unmatched	9.9135	10.0169	−0.1034	0.0215	−4.82	1,312	781
ATET	9.9135	10.0085	−0.0950**	0.0411	−2.31		
(d) High school							
Unmatched	10.0810	10.1666	−0.0855	0.0246	−3.48	1,165	794
ATET	10.0810	10.1680	−0.0870**	0.0373	−2.33		
(e) College and higher							
Unmatched	10.3556	10.5295	−0.1739	0.0281	−6.2	964	1,073
ATET	10.3556	10.4901	−0.1345**	0.0601	−2.24		

Table 7.
The PSM results across
different subsamples
in 2016

(*continued*)

Sample	Treated	Controls	Difference	S.E.	T-stat	Observations Untreated (male)	Treated (female)
Sample divided by occupation							
(a) Professionals							
Unmatched	10.4951	10.5829	−0.0878	0.0271	−3.24	693	915
ATET	10.4951	10.5181	−0.0229*	0.0405	−0.57		
(b) Skilled workers							
Unmatched	9.9521	10.1069	−0.1548	0.0154	−10.04	2,717	1,666
ATET	9.9521	10.1122	−0.1601***	0.0299	−5.35		
(c) Laborers							
Unmatched	9.7013	9.8604	−0.1591	0.0218	−7.31	1,312	902
ATET	9.7013	9.8657	−0.1644***	0.0364	−4.52		

Table 7.

Note(s): *, **, *** indicate statistical significance at 0.10, 0.05, 0.01 levels, respectively

but highest for those without any qualification. Nonetheless, it is higher for rural than urban workers and far higher for household heads than other members of the household, which is in stark contrast to the earlier results.

Besides, the availability of the data on regions in the sample of 2016 allows us to make further comparisons across six economic regions in Vietnam. Particularly, the region witnessing the highest pay gap was Central Highlands, followed by the Red River Delta and Mekong River Delta. In stark contrast, the estimated gaps are insignificant for both the Southeast and the Northern Midland and Mountainous.

5. Discussion and conclusion remarks

Studying gender pay gap is simply to conduct a comparison of the pay level of the two groups which are identical in terms of identified characteristics. However, establishing two comparable groups meets some difficulties. Traditional social norms create gender specialization in many areas of life, thereby limiting the role of women and giving priority to men in the labor market. In addition, the creator also contributes to gender specialization in the labor market. It is because health and some gender characteristics make women suitable for only some jobs, and the same is true for men. This fact shows that "not all males are comparable to all females" even in the support. In other words, the distributions of characteristics can be different in the support. Matching techniques are a good tool to control such differences and ensure the comparability of research groups and, therefore, can provide a reliable assessment of gender pay inequality. This study differs from the previous studies on gender pay gap in Vietnam in that it employs the PSM technique with a focus on the common supports only.

The estimation results based on the data sets of VHLSS 2010 and 2016 indicate that there exists a statistically significant pay gap and inequality between comparable groups of male and female earners, but the extent has been narrowed by half in this period for the full sample and also remarkably for all subsamples. While the pay disparity is significant for all subsamples but those of small size in 2010, it is either insignificant or significant at low significance level for many subsamples. This further reinforces the sharp progress in enhancing gender pay equality over this period.

We find that differences in pay in both years arose from not only the observable but also unobservable factors. Of which, the most significant observable determinants are shown to be educational background, occupation, industry and economic sector. Nevertheless, in both years, the estimated gap for matched groups is considerably higher than the gap for unmatched ones for both the full sample and many of the

Sample	Treated	Controls	Difference	S.E.	T-stat	Observations Untreated (male)	Treated (female)
Sample divided by industry							
(a) Agriculture							
Unmatched	9.7131	9.9111	−0.1980	0.0726	−2.73	206	55
ATET	9.7131	9.7372	−0.0240	0.0980	−0.25		
(b) Industry and construction							
Unmatched	9.9653	10.1200	−0.1547	0.0158	−9.8	2,468	1,530
ATET	9.9653	10.1373	−0.1720***	0.0257	−6.69		
(c) Service							
Unmatched	10.1052	10.1593	−0.0541	0.0205	−2.64	2,285	1,953
ATET	10.1052	10.1671	−0.0620	0.0429	−1.44		
Sample divided by economic sector							
(a) Non-state							
Unmatched	9.8423	10.0464	−0.2042	0.0151	−13.48	3,607	2,029
ATET	9.8423	10.0187	−0.1764***	0.0304	−5.8		
(b) State							
Unmatched	10.3549	10.3609	−0.0060	0.0279	−0.22	1,087	948
ATET	10.3549	10.3926	−0.0377	0.0522	−0.72		
(c) FDI							
Unmatched	10.2142	10.3078	−0.0936	0.0315	−2.97	272	560
ATET	10.2142	10.2113	0.0029	0.0589	0.05		
Sample divided by region							
(a) Red River Delta							
Unmatched	10.1124	10.2427	−0.1303	0.0256	−5.09	1,167	887
ATET	10.1124	10.3119	−0.1995***	0.0544	−3.66		
(b) Northern Midland and Mountainous							
Unmatched	10.1913	10.1017	0.0896	0.0322	2.78	710	481
ATET	10.1913	10.1674	0.0239	0.0709	0.34		
(c) North Central and Central Coast							
Unmatched	10.0235	10.1187	−0.0952	0.0256	−3.72	1,311	870
ATET	10.0235	10.1240	−0.1005**	0.0511	−1.97		
(d) Central Highlands							
Unmatched	9.8998	9.9291	−0.0294	0.0884	−0.33	157	127
ATET	9.8998	10.1016	−0.2018*	0.1215	−1.66		
(e) Southeast							
Unmatched	10.1690	10.2399	−0.0709	0.0338	−2.1	565	440
ATET	10.1690	10.2429	−0.0739	0.0573	−1.29		
(f) Mekong River Delta							
Unmatched	9.8127	9.9995	−0.1868	0.0274	−6.81	1,041	733
ATET	9.8127	9.9968	−0.1841***	0.0482	−3.82		

Table 8.
The PSM results across
different subsamples in
2016 (continued)

subsamples. This fact reveals that the women possess characteristics that should be paid higher wages but are treated highly unfairly compared to men. Wage earners in state sector and in service are among those experiencing the lowest gender gaps in both years while also witnessing sharp progress in reducing inequality in earnings between 2010 and 2016. They are followed by professionals (in all fields) and workers with highest qualifications at high school level. In stark contrast, non-state workers, industry and construction workers, along with skilled workers, all male-dominated, face highest pay inequality in both years despite considerable improvements. The decrease in pay disparity over time can be explained by a number of unobserved factors such as social norms on the role of women in society have gradually changed; women's health,

educational and employment opportunities have also improved along with the rise in living standards; and women's right protection policies have gradually come into effect. This conclusion provides an important clue for further research.

Our conclusions generally do not appear to be different from those of previous studies. However, the results are more robust and reliable than those obtained by the B–O method found in Amy (2004) and Nguyen and Hoang (2018) because the matching technique eliminates the heterogeneity of the samples which the B–O method fails to do. Also, our results are more robust than those obtained by the improved B–O models found in Pham and Reilly (2007), Gian (2014) and Vu and Yamada (2018) in the sense that the results from matching techniques are immune from the diseases commonly faced by parametric estimation techniques in terms of specification error and uncertainty of functional form. In addition, the results obtained from a matching allow for a deeper insight into the extent and causes of pay inequality. Although matching techniques have many advantages over the traditional parametric methods, they suffer from the dimensionality problem that arises when there are many explanatory variables in non-parametric setups (Nopo, 2004).

Our findings suggest an important policy implication that policies aimed at narrowing gender inequality in labor income are effective only when taking into account both observable characteristics and unobservable factors such as unobservable gender differences that determine wages and gender discrimination in pay.

Note

1. Only data in 2016 is sufficient for calculation.

References

Amy, L. (2004), "Sectoral gender wage gap in Vietnam", *Oxford Development Studies*, Vol. 32 No. 2, pp. 225-239.

Arrow, K. (1973), "The theory of discrimination", in Ashenfelter, O. and Rees, A. (Eds), *Discrimination in Labor Markets*, Princeton University Press, Princeton and Oxford, pp. 3-33.

Becker, G. (1971), *The Economics of Discrimination*, 2nd ed., University of Chicago Press, Chicago and London.

Blau, F. and Kahn, L. (2017), "The gender wage gap: extent, trends, and explanations", *Journal of Economic Literature*, Vol. 55 No. 3, pp. 789-865.

Blinder, A. (1973), "Wage discrimination: reduced form and structural estimates", *The Journal of Human Resources*, Vol. 7 No. 4, pp. 436-455.

Bourguignon, F., Francisco, F. and Phillipe, L. (2002), "Beyond Oaxaca-blinder: accounting for differences in household income distributions", *The Journal of Economic Inequality*, Vol. 6 No. 2, pp. 117-148.

Chernozhukov, V., Fernández-Val, I. and Melly, B. (2013), "Inference on counterfactual distributions", *Econometrica*, Vol. 81 No. 6, pp. 2205-2268, doi: 10.3982/ECTA10582.

DiNardo, J., Fortin, N. and Lemieux, T. (1996), "Labor market institutions and the distribution of wages, 1973-1992: a semiparametric approach", *Econometrica*, Vol. 64 No. 5, pp. 1001-1044.

Donald, S., Green, D. and Paarsch, H. (2000), "Differences in wage distributions between Canada and the United States: an application of a flexible estimator of distribution functions in the presence of covariates", *Review of Economic Studies*, Vol. 67, pp. 609-633.

Frolich, M. (2007), "Propensity score matching without conditional independence assumption – with an application to the gender wage gap in the United Kingdom", *The Econometrics Journal*, Vol. 10 No. 2, pp. 359-407.

Firpo, S., Fortin, N.M. and Lemieux, T. (2009), "Unconditional quantile regressions", *Econometrica*, Vol. 77 No. 3, pp. 953-973.

Fortin, N., Lemieux, T. and Firpo, S. (2011), "Chapter 1 – Decomposition methods in economics", in Ashenfelter, O. and Card, D. (Eds), *Handbook of Labor Economics*, North Holland, Amsterdam, pp. 1-102, doi: 10.1016/S0169-7218(11)00407-2.

Gian, T.C. (2014), "Gender wage gap: the comparison analysis between Korea and Vietnam", *E-Labor News*, Vol. 94, available at: https://www.kli.re.kr/kli_eng/selectBbsNttView.do?key=220&bbsNo=31 &nttNo=100750&searchY=&searchCtgry=&searchDplcCtgry=&searchCnd=all&searchKrwd=&pa geIndex=9&integrDeptCode=.

ILOSTAT (2020), *Database*, ILO's Department of Statistics, United Nations.

ILSSA and ILO (2018), *Labour and Social Trends in Vietnam 2012–2017*, ILO Hanoi.

Juhn, C., Murphy, K. and Pierce, B. (1993), "Wage inequality and the rise in returns to skill", *The Journal of Political Economy*, Vol. 101 No. 3, pp. 410-442.

Machado, J. and Mata, J. (2001), "Earning functions in Portugal 1982-1994: evidence from quantile regressions", *Empirical Economics*, Vol. 26 No. 1, pp. 115-134.

Meara, K., Pastore, F. and Webster, A. (2017), "Is the gender pay gap in the US just the result of gender segregation at work?", Working Paper, IZA Discussion Paper No. 10673, Institute of Labor Economics (IZA), Bonn.

Nguyen, H.O. and Hoang, T.H. (2018), "Gender gap in income of wage-earners in Vietnam in the period 2012–2014", *Journal of Economics and Development*, Vol. 254, pp. 10-20 (in Vietnamese).

Nopo, H.R. (2004), "Matching as a tool to decompose wage gaps", Working Paper, IZA Discussion Paper No. 981, Institute of Labor Economics (IZA), Bonn.

Oaxaca, R. (1973), "Male-female wage differentials in urban labor market", *International Economic Review*, Vol. 14 No. 3, pp. 693-709.

Pham, H. and Reilly, B. (2007), "The gender pay gap in Vietnam, 1993–2002: a quintile regression approach", *Journal of Asian Economics*, Vol. 185 No. 5, pp. 775-806.

Phelps, E. (1972), "The statistical theory of racism and sexism", *The American Economic Review*, Vol. 62 No. 4, pp. 659-661.

Rubin, D. and Rosenbaum, P. (1983), "The central role of the propensity score in observational studies for causal effects", *Biometrika*, Vol. 70 No. 1, pp. 41-55.

UN Women (2020), "Equal pay for work of equal value", available at: https://www.unwomen.org/en/ news/in-focus/csw61/equal-pay.

Vu, T.M. and Yamada, H. (2018), "Decomposing Vietnamese gender equality in terms of wage distribution", *Pacific Economic Review*, Vol. 23 No. 5, pp. 705-731.

Further reading

General Statistical Office of Vietnam (2011), *Vietnam Household Living Standards Survey 2010*, Hanoi.

General Statistical Office of Vietnam (2017), *Vietnam Household Living Standards Survey 2016*, Hanoi.

Corresponding author
Nguyen Hoang Oanh can be contacted at: oanhnh@neu.edu.vn