Gender Wage Differentials in China from 1995 to 2018: Distributional Evidence Accounting for Employment Composition using Partial Identification

Rui Xu[*] Xintong Wang[†] Alfonso Flores-Lagunes[‡]

January 30, 2024

(For the latest version please click here)

Abstract

This paper examines changes in the gender gap of the wage distribution in China from 1995 to 2018. We use data from the China Household Income Survey (CHIP) 1995-2013 and the China Family Panel Studies (CFPS) 2014 and 2018. To effectively account for changes in employment composition, we employ nonparametric bounds. We account for the labor supply's intensive margin by computing workers' working hours and hourly wage using available information in CHIP and CFPS. Our methodology adopts a weak quartile dominance assumption and a stochastic dominance assumption to tighten the bounds. The results show statistically significant evidence that, over the years from 1995 to 2018, the median gender wage gap for the young workers (age 25-45) who are non-college-educated has increased by 0.17 - 0.62 log points. To estimate potential changes in the evolution of the gender wage gap suggested in the literature, we split up our analysis into two periods from 1995 - 2007 and 2007 - 2018. The results show larger changes in the gender wage gap compared to estimates in existing studies. In the eailer period, we find a significant increase by 0.15 - 0.27 log points in the median gender wage gap among the young workers who are college-educated. In the second period, the bounds estimates are less conclusive and suggest a decrease in the median gender wage gap among the college-educated young workers by 0.05 - 0.19 log points, but the 95% CI does not exclude a zero change. The estimates of the gender wage gap at the $75th$ wage percentile show a similar pattern as the changes at the median wage, while the statistical implications at the $25th$ percentile are inconclusive.

Key words and phrases: Gender Wage Gap; Wage Inequality; China; Partial Identification; Bounds

JEL classification: J3, C31, C36

[*] China Economic Research Institution, Liaoning University, Shenyang, Liaoning; Email: ruixu1993econ@hotmail.com.

[†] Department of Accounting, Economics, and Finance, Slippery Rock University of Pennsylvania, Slippery Rock, PA; Telephone: +1 724-738-2579; Email: xintong.wang@sru.edu.

[‡] Department of Economics and Center for Policy Research, Syracuse University; IZA, and GLO; 426 Eggers Hall, Syracuse NY 13244-1020; Telephone: +1 315-443-9045; Email: afloresl@maxwell.syr.edu.

1 Introduction

Reducing the gender wage gap brings multiple benefits to the economy such as promoting economic growth (Schober and Winter-Ebmer, 2011), potentially improving women's healthcare access (Fee, 1991) and mental health (Platt and Keyes, 2016), reducing domestic violence against women (Aizer, 2010), and increasing women's fertility autonomy (Qian and Jin, 2018). To reduce the gender wage gap, it is necessary to estimate its level in recent decades and its trend. Researchers have documented a substantial reduction in the gender wage gap in the United States during the 1980s and a stable gender wage gap from 1980 to 2010 (Blau and Kahn, 2017).

The story is quite different in China. In recent years, China has experienced a transition of gender pay gaps. The observed wage earnings gap between males and females has progressively widened since 1988 (Gustafsson and Li, 2000; Gustafsson and Wan, 2020). Gustafsson and Li (2000) use the Urban Household Income Survey and find that the average gender wage gap has increased from 15.6% in 1988 to 17.5% in 1995. For a later period, Chi and Li (2014) find that the average gender earnings gap has increased from 2005 to 2009; estimates from Heckman's selection-correction model, which accounts for selection into employment, suggest an overall underestimated raw observed gender earnings wage gap by 12 - 14%. In more recent years, Song et al. (2019) used China Household Income Survey (CHIP) and recorded a temporary narrowing in the gender earnings gap from 29% in 2007 to 25% in 2013.

The existing literature has mostly focused on measuring the average gender earnings gaps conditional on employment. Instead, this study aims to re-examine changes in the gender wage differentials at the median, the $25th$ and the $75th$ wage quantiles in China from 1995-2018, while effectively accounting for changes in employment composition and the intensive margin of labor supply (i.e., hours worked). We use data from the China Household Income Survey (CHIP) 1995-2013 and the China Family Panel Studies (CFPS), 2014 and 2018.

Controlling for selection into employment is particularly important in estimating the gender wage gap in China. Since 1988, the labor market structure in China has gone through dramatic structural changes (e.g., Li et al., 2012; Meng, 2012). Before 1995, China's unemployment rate was lower than other countries' unemployment rate. Since the mid-1990s, the Chinese government began privatizing small and medium-sized state-owned enterprises (SOEs), which triggered large-scale layoffs. The unemployment rate jumped to a level even higher than that of the high-income countries, peaking above 10% in 2002-2003, then slowly drifting down (Feng et al., 2017). In the same period when the unemployment rate increased, the overall urban labor participation rate dropped from over 82% to around 75%. The labor force participation rate has remained low ever since. These changes fell most heavily on the unskilled women (Feng et al., 2017), which can be potentially due to the increase of the returns to education and the high wage elasticity of women (Hare, 2019). Additionally, in late 2015, the Chinese government relaxed the one-child policy in China and replaced it with the two-child policy, which may have profound labor market impacts on women. For example, employers may be concerned that they need to pay for maternity leaves multiple times for each female employee and may be more reluctant to hire women after the two-child policy took effect. Importantly, the estimated gender wage gap may be biased due to changes in labor force participation by gender over the years. For example, some highly educated and likely high-wage women might be deterred by discrimination in the labor market as a result of child-bearing. If high-wage women are increasingly exiting the labor market, the observed gender wage gap may be inflated.

In the literature on gender wage gap estimation, methods employed to control for selection into employment include the Heckman selection-correction model (Blau and Beller, 1988; Mulligan and Rubinstein, 2008; Chi and Li, 2014), semiparametric quantile-copula (Maasoumi and Wang, 2019), the sample restriction and identification at infinity (Mulligan and Rubinstein, 2008; Machado, 2017), imputation of unobserved wage offers (Blau and Kahn, 2006; Blau and Comey, 2023), and bounding techniques (Blundell et al., 2007). Each method has its respective strengths and drawbacks. The Heckman selection-correction model yields precise estimates for gender wage gaps; however, the identification relies on strong assumptions about instrumental variables that affect employment but not wages (i.e., the exclusion restriction assumption). The nonparametric quantile-copula approach deals with selection into employment by computing the reservation wages of the non-working and allows for time-varying selection. However, it also relies on the exclusion restriction of the instrumental variables. The identification at infinity does not impose restrictions on the direction of the selection to employment; however, it restricts the sample among a population group that would "always work", which may not be representative of the population. The wage imputation method relies on the assumption that selection into employment is based on observed variables. Therefore, rich panel data with individuals' wage histories is usually needed for the imputation method, and this requirement may not be satisfied in all settings. The nonparametric bounds method we employ does not require exclusion restriction assumptions, although sometimes it may lead to wide bounds.

To account for differences in labor force participation (employment composition), we use bounds introduced by Manski (1994), Manski and Pepper (2000), and Blundell et al. (2007). We start with the worst-case bounds on the wage distribution in Manski (1994) and then employ additional assumptions substantiated by economic theory to tighten the bounds. The first assumption we use is the quartile dominance assumption. This assumption requires that, conditional on age, education, and sex, the quartiles of the wage distribution (wages at the $25th$, $50th$, $75th$ percentiles) of the non-working population not be higher than the corresponding quartiles of the wage distribution of the working population. We also employ a stronger version of this assumption – the stochastic dominance assumption, which requires the wage distribution of the working population to stochastically dominate the non-working population. These two assumptions are based on a positive selection into labor force participation, which is implied by standard models of labor supply (e.g., Gronau, 1974; Blundell et al., 2007). To assess those assumptions, we estimate the log residual wage conditional on age, education, and survey year using CHIP 1995-2013 and CFPS 2014-

2018. For males and females, respectively, the residual wage distribution of those who are continuously employed is higher than the residual wage distribution of those who have non-working spells across all percentiles, except for three incidences – the $90th$ percentile for males over 45, and the $90th$ and the $95th$ percentiles for females under 45. Aside from the above exceptions, which occur at very high wage percentiles, the evidence from the residual wage analysis supports our quartile and stochastic dominance assumptions. [1]

After controlling for labor force participation and the hours worked, our bounds estimates show stronger evidence of an increase in the gender wage gap in the 1995-2007 period. The increase in the gender wage gap is most statistically significant among the young (under age 45), the college-educated, and at the median and high percentiles of the wage distribution. Specifically, the bounds estimates suggest a statistically significant increase in the gender wage gap for the young college-educated at the median wage of at least 0.15 log points, and at the $75th$ percentile of at least 0.19 log points. The estimated bounds at the $25th$ percentile for young college graduates also suggest an increase in the gender wage gap of at least 0.07 log points, however, this 95% confidence interval (CI) does not exclude a zero change. The estimates for the 2007-2018 period do not exclude a zero change for most age and education groups. The bounds at the $75th$ wage percentile suggest an at least 0.05 log points decrease in the gender wage gap for the young college-educated, while the 95% confidence intervals (CIs) do not exclude a zero change.

The main contributions of this paper are in four aspects. First, to the best of our

[1] We also employ the income of other household members as a monotone instrumental variable (MIV) for the wage of individuals. Specifically, we assume that for individuals with higher-income family members, their wage distribution would likely first-order stochastically dominate those with relatively lower-income family members. A theoretical justification of this assumption rests on the notion of assortative mating (Becker (1973); Nie and Xing (2019)) and the inter-generational income persistence (Feng et al., 2021; Gong et al., 2010). We also improve statistical inference on the bounds using MIVs in Blundell et al. (2007). Bounds that use MIVs involve maximum and minimum operators, for which the standard inference breaks down (Hirano and Porter, 2012). We adopt a method proposed by Chernozhukov et al. (2013) to bias-correct and obtain asymptotically valid confidence intervals for these bounds. However, since the survey data structure, we could not consistently construct the average household member's income for an individual throughout the study period. In addition, employing the Chernozhukov et al. (2013) will lead to wider bounds and confidence intervals for these bounds, which makes the estimated bounds hard to interpret and compare to the estimated bounds under other assumptions. Therefore, we do not include these results in the main paper, and all the details and estimations under the MIV assumption are in the Online Appendix B.

knowledge, we are the first to use bounds as the primary method to control for selection into employment in estimating the gender wage gap in China. Second, to conduct our analysis, we harmonize two nationally representative datasets to estimate the gender wage gap from 1995 to 2018. Different from previous literature that used earnings as the measure for the gender wage gap (e.g.,Chi and Li, 2014; Song et al., 2019), we construct a measure for the hourly wage. In this way we provide statistical evidence of changes in the gender wage gap avoiding biases due to labor supply's intensive (hours worked) and extensive (employed v.s. unemployed) margins, respectively. Third, we go beyond the median gender wage gap by analyzing the gender wage gap dynamics in China at the $25th$ and $75th$ percentiles of the wage distribution, thereby providing a fuller picture that includes the lower and upper sides of the wage distribution.

2 Bounds on the Wage Distribution Accounting for Employment

Let W be the log wage and X be control variables such as gender, age, education, and the survey year. Let E indicate whether a person is employed, with $E = 1$ being employed and $E = 0$ otherwise. The probability of being employed given characteristics $X = x$ is written as $P(x)$. We write the cumulative distribution function (CDF) of W given $X = x$ by $F(w|x)$, given $X = x$ and $E = 1$ by $F(w|x, E = 1)$, and given $X = x$ and $E = 0$ by $F(w|x, E = 0)$. We have

$$F(w|x) = F(w|x, E = 1)P(x) + F(w|x, E = 0)[1 - P(x)] \quad (1)$$

In equation (1), data only identifies $F(w|x, E = 1)$ and $P(x)$. $F(w|x, E = 0)$, which is the wage distribution of the population who did not take up employment, is not observed in the data. We partially identify the wage distribution of the unemployed, $F(w|x, E = 0)$, using comparably weak assumptions.

2.1 The Worst Case Bounds

The worst case bounds following Manski (1994) and Blundell et al. (2007) substitute the inequality that follows from the definition of a CDF

$$0 \leq F(w|x, E = 0) \leq 1$$

into equation (1) to bound the log wage cumulative distribution function of the total population ($F(w|x)$) as:

$$F(w|x, E = 1)P(x) \leq F(w|x) \leq F(w|x, E = 1)P(x) + [1 - P(x)] \tag{2}$$

The bounds can then be translated to give the worst case bounds on the conditional quantiles. Denote the q-th quantile of $F(w|x)$ by $w^q(x)$, then

$$w^{q(l)}(x) \leq w^q(x) \leq w^{q(u)}(x)$$

where the log wage $w^{q(l)}(x)$ is the lower bound and the log wage $w^{q(u)}(x)$ is the upper bound that respectively solve the following two equations with respect to w,

$$q = F(w|x, E = 1)P(x) + [1 - P(x)] \tag{3}$$

and

$$q = F(w|x, E = 1)P(x) \tag{4}$$

Since $F(w|x, E = 1)P(x)$ cannot be smaller than zero, equation (3) cannot be smaller than $[1 - P(x)]$; likewise, since $F(w|x, E = 1)$ cannot be greater than 1, equation (4) cannot be larger than $P(x)$. Due to the lower limit of equation (3) and the upper limit of equation (4), using the worst case bounds, we can only identify the lower bounds to log wage quantiles $q \geq 1-P(x)$ and upper bounds for quantiles $q \leq P(x)$ (Blundell et al., 2007). The worst-case bounds are likely to be wide in practice. Therefore, we impose restrictions on the log wage distribution to obtain narrower bounds.

2.2 Stochastic Dominance and Quartile Dominance

The standard labor supply model suggests that when the substitution effect of a change in the wage dominates its income effect, individuals that command higher wages will be more likely to work, ceteris paribus (Blundell and MaCurdy, 1999). Thus, as in Blundell et al. (2007), we impose a stochastic dominance assumption between the wage distributions of the workers and non-workers. That is, we assume that conditional on $X = x$, the wages of those observed working first-order stochastically dominate those of the non-workers. This assumption is based on the notion that workers are more productive than non-workers; therefore, at each percentile of the distribution, the workers' observed wages would not be lower than non-workers' potential wages. Blundell et al. (2007) show that this positive selection into employment requires that the difference between the observed wage and the reservation wage, denoted by $w - w^R$ should be positively correlated with w. Intuitively, we can expect individuals with a higher preference to work to have a low reservation wage w^R and have invested more in human capital in the past, and the accumulated human capital yields higher wages w and greater differences from w^R (Blundell et al., 2007).

This assumption seems plausible in the case of China. In the recent decades of China's labor market, the increase in the non-working population has mostly been driven by unskilled workers (e.g., Feng et al., 2017; Gustafsson and Ding, 2011), which implies that the working population consists of workers with relatively higher human capital. In addition, Li et al. (2016) show that the college premiums from 1990-2000 in China have increased. Li et al. (2017) predict that with investment in physical capital and skill-biased technological change, the return to human capital in China will continue to increase. If individuals with more human capital are more likely to be employed and paid more, this increase in return to human capital in China continues to make the stochastic dominance assumption more convincing.

Following Blundell et al. (2007), we formulate the stochastic dominance assumption in our application as

$$F(w|x, E=1) \leq F(w|x, E=0) \quad \forall w, \quad \forall x \tag{5}$$

for each w with $0 \leq F(w|x) \leq 1$ or, equivalently,

$$Pr(E=1|W \leq w, x) \leq Pr(E=1|W > w, x).$$

Under this assumption, the wage distribution of the unemployed $F(w|x, E=0)$ in the total wage distribution in equation (1) is lower-bounded by the wage distribution of the employed $F(w|x, E=1)$. We can replace $F(w|x, E=0)$ with $F(w|x, E=1)$ in the lower bound of equation (1) and the bounds on the distribution of the wage become

$$F(w|x, E=1) \leq F(w|x) \leq F(w|x, E=1)P(x) + [1 - P(x)] \tag{6}$$

Similar to the case of the worst case bounds, the bounds for the conditional wage quantiles under the stochastic dominance assumptions are $w_s^{q(l)}(x) \leq w^q(x) \leq w_s^{q(u)}(x)$, where $w_s^{q(l)}(x)$ and $w_s^{q(u)}(x)$ respectively solve the following two equations with respect to w,

$$q = F(w|x, E=1)P(x) + [1 - P(x)] \tag{7}$$

and

$$q = F(w|x, E=1) \tag{8}$$

The stochastic dominance assumption may not be satisfied in some scenarios. For example, for individuals in households who have accumulated financial assets and human capital, a negative correlation between $w - w^R$ and w might undermine the stochastic dominance assumption (Blundell et al., 2007). In light of these scenarios in which positive selection into employment may not be satisfied, we employ a weaker restriction - a quartile dominance assumption. This assumption restricts the $25th$, $50th$, and the $75th$ wage quantiles for those not working to be not higher than the corresponding wage quantiles of the observed wage distribution. This assumption implies the following bounds for the distribution of log wages of the unemployed, where $w^{q(E=1)}$ denotes the $q-th$ quantile wage of the employed.

$$\begin{aligned} 0 \leq & F(w|x, E=0) \leq 1, \quad \text{if} \quad w < w^{25(E=1)}(x), \\ 0.25 \leq & F(w|x, E=0) \leq 1, \quad \text{if} \quad w^{25(E=1)}(x) \leq w < w^{50(E=1)}(x), \\ 0.5 \leq & F(w|x, E=0) \leq 1, \quad \text{if} \quad w^{50(E=1)}(x) \leq w < w^{75(E=1)}(x), \\ 0.75 \leq & F(w|x, E=0) \leq 1, \quad \text{if} \quad w \geq w^{75(E=1)}(x), \end{aligned} \tag{9}$$

Under the quartile dominant assumption, in equation (9), since the three wage quartiles (i.e., the $25th$, $50th$, and $75th$ wage quantiles) of the employed should not be lower than the respective counterpart wage quartiles of the unemployed, when wage w is higher than the $25th$ quantile wage of the employed ($w^{25(E=1)}$), the wage distribution of the unemployed $F(w|x, E=0)$ is lower-bounded by 0.25, and similarly when w is higher than the $50th$ or the $75th$ quartile wages of the employed. Therefore, the bounds for the wage distribution are:

$$
\begin{aligned}
&F(w|x,E=1)P(x) \\
&\qquad \le F(w|x) \\
&\qquad \le F(w|x,E=1)P(x)+(1-P(x)), \quad \text{if} \quad w<w^{25(E=1)}(x), \\
&F(w|x,E=1)P(x)+0.25(1-P(x)) \\
&\qquad \le F(w|x) \\
&\qquad \le F(w|x,E=1)P(x)+(1-P(x)), \quad \text{if} \quad w^{25(E=1)}(x)\le w<w^{50(E=1)}(x), \\
&F(w|x,E=1)P(x)+0.5(1-P(x)) \\
&\qquad \le F(w|x) \\
&\qquad \le F(w|x,E=1)P(x)+(1-P(x)), \quad \text{if} \quad w^{50(E=1)}(x)\le w<w^{75(E=1)}(x), \\
&F(w|x,E=1)P(x)+0.75(1-P(x)) \\
&\qquad \le F(w|x) \\
&\qquad \le F(w|x,E=1)P(x)+(1-P(x)), \quad \text{if} \quad w\ge w^{75(E=1)}(x)
\end{aligned}
\tag{10}
$$

In the set of bounds in equation (10), the bounds for $w^{25(E=1)}(x) \le w < w^{50(E=1)}(x)$ is obtained by replacing $F(w|x,E=0)$ with 0.25 in the lower bound of the total wage distribution in equation (1). Similarly, the bounds when $w^{50(E=1)}(x) \le w < w^{75(E=1)}(x)$ and $w \ge w^{75(E=1)}(x)$ are obtained by replacing $F(w|x,E=0)$ with 0.5 and 0.75 respectively. The corresponding bounds for the conditional wage quantiles under the quartile dominance assumptions are $w_q^{q(l)}(x) \le w_q^q(x) \le w_q^{q(u)}(x)$, where $w_q^{q(l)}(x)$ and $w_q^{q(u)}(x)$ respectively solve the following two equations (11) and (12) with respect to w,

$$q = F(w|x,E=1)P(x) + [1-P(x)] \tag{11}$$

and

$$
\begin{aligned}
q &= F(w|x, E=1)P(x), && \text{if} \quad w < w^{25(E=1)}(x), \\
q &= F(w|x, E=1)P(x) + 0.25(1 - P(x)), && \text{if} \quad w^{25(E=1)}(x) \leq w < w^{50(E=1)}(x), \\
q &= F(w|x, E=1)P(x) + 0.5(1 - P(x)), && \text{if} \quad w^{50(E=1)}(x) \leq w < w^{75(E=1)}(x), \\
q &= F(w|x, E=1)P(x) + 0.75(1 - P(x)), && \text{if} \quad w \geq w^{75(E=1)}(x).
\end{aligned} \tag{12}
$$

For our quartile dominance bounds, we assume each log wage quartile of the employed individuals should be no lower than the respective quartile of the unemployed. The difficulty in justifying this assumption is that we do not observe the log wage distribution for those not employed. To find the closest substitute, we use the China Family Panel Studies (CFPS) panel data to identify individuals who have experienced an unemployment spell before. We then compare the log wage distribution for individuals who have continuously worked during the observed periods and the log wage distribution for those who had an unemployment spell. The rationale is that the observed wage of an individual after an unemployment spell should be no lower than the reservation wage during the unemployment spell, which in turn should be no lower than the wage offers available to the individual during the unemployment spell. Therefore, the difference in wage quantiles between those continuously employed and those with unemployment spells will be no greater than the unobserved difference in wage quantiles between the employed and the unemployed. If the wage quantiles of workers without unemployment spells are higher than those of workers with unemployment spells, it should imply the wage quantiles of the employed are also higher than those of the unemployed.

We find empirical evidence in our data that supports the stochastic and quartile dominance assumptions. In Figure 1, we present the distribution of residual wages by gender, age, and work history of workers who have been continuously employed and of workers with spells of unemployment using the CFPS, 2014 and 2018 [2]. The residual wages are obtained in a regression controlling for age, age squared, college degree attainment, province of residence,

[2]See the Data Section for details of the sample

and survey year dummies. The darker lines indicate the residual wages across percentiles for workers who do not have spells of unemployment in their work history. The lighter lines are for the workers with spells of unemployment. The results show that the residual wages of males and females who do not have unemployment spells are consistently higher than the wages of males and females who do have unemployment spells from the $5th$ quantile to the $95th$ quantile, except for three incidences – the 90th percentile for males over 45, and the $90th$ and the $95th$ percentiles for females under 45. The above exceptions occur at very high wage percentiles, suggesting that the stochastic dominance assumption, which implies that any wage quantiles of the unemployed should not be higher than the employed, may fail at very high wage quantiles for young women and older men. In Figure 1, we use boxes to indicate the $25th$, $50th$ and the $75th$ wage quantiles. The residual wage quantile estimates offer support for the weaker quartile dominance assumption in all samples.

2.3 Bounds on the Gender Wage Gap and its Change over Time

Our goal is to conduct inference on the gender wage gap dynamics from 1995-2018 in China. We use the bounds of males and females' wage quantiles to estimate the gender wage gap over the wage distribution and its changes over different points in time. For example, let the lower bound and the upper bound for males' wage quantile q with education and age characteristics x in year t be $w^{q(l)}(male, x, t)$ and $w^{q(u)}(male, x, t)$, and the female's equivalent bounds be $w^{q(l)}(female, x, t)$ and $w^{q(u)}(female, x, t)$. The bounds for the gender wage gap at the quantile q, $D_t^q(x) = w^q(male, x, t) - w^q(female, x, t)$ are:[3]

$$w^{q(l)}(male, x, t) - w^{q(u)}(female, x, t) \leq D_t^q(x) \leq w^{q(u)}(male, x, t) - w^{q(l)}(female, x, t). \quad (13)$$

Similarly, the lower bound of the change in the gender wage gap from year t to year s,

[3]These bounds can be computed under different combinations of the assumptions presented in Section 2.2 and 2.3.

$\Delta D_{st}^{q(l)}$, where $s > t$, is given by,

$$\{w^{q(l)}(male, x, s) - w^{q(u)}(female, x, s)\} - \{w^{q(u)}(male, x, t) - w^{q(l)}(female, x, t)\}, \quad (14)$$

and the upper bound, $\Delta D_{st}^{q(u)}$, where $s > t$, is given by,

$$\{w^{q(u)}(male, x, s) - w^{q(l)}(female, x, s)\} - \{w^{q(l)}(male, x, t) - w^{q(u)}(female, x, t)\}. \quad (15)$$

Our main focus will be the bounds on the quantiles of the wage distribution. To estimate these, we first estimate the bounds on the distribution of wages. We now describe the nonparametric estimation procedure we have used. The conditioning vector X includes gender, education, age, and time. Estimating the worst case bounds and the bounds with monotonicity requires estimating the employment probability and the distribution of wages observed amongst the workers for each possible set of characteristics X. We define two education groups: those who with at most a high school degree (Non-College Group) and those who with at least a college degree or a Dazhuan (equivalent to a vocational or associate degree in the U.S.) degree (College Group). We also limit the number of age groups to two: those below 45 (young) and those above 45 (old). We construct confidence intervals for the changes in the differentials over time using the bootstrap and applying the results of Imbens and Manski (2004).

3 Data and Variable Definitions

This study uses both household-level and individual-level data from two surveys. We use the Chinese Household Income Project (CHIP) for the years of 1995, 2002, 2007, 2013, and the China Family Panel Study (CFPS) for the years of 2014 and 2018. Using CHIP and CFPS together enables us to analyze the dynamics of the gender wage gap in China from the mid-1990s to the late 2010s. This section describes CHIP and CFPS, discusses the challenges we encounter while using data from those two surveys together, explains how we construct our key variables, and introduces our analytic sample.

3.1 CHIP and CFPS

CHIP was carried out as part of a collaborative research project on income and inequality in China organized by Chinese and international researchers and institutions, including the Chinese Academy of Social Sciences and the School of Economics and Business Administration at Beijing Normal University. CHIP is a nationally representative household-level survey aimed at estimating income, wealth, consumption, and related economic measures in rural and urban areas in China. CHIP uses a stratified random sampling process to collect data for three different samples – rural, urban, and migrant groups in 22 provinces, all at household and individual levels. CHIP samples are cross-sectional and are subsamples taken from the National Bureau of Statistics (NBS) samples used to obtain the official household statistics published in the annual Statistical Yearbook of China. CFPS is a nationally representative, bi-annual longitudinal survey of the Chinese communities, families, and individuals, conducted by the Institution of Social Science Survey of Peking University since 2010. Both CHIP and CFPS include individual-level demographics and detailed information on wage income and wealth, making it possible to analyze the national trend of wage inequality.

3.2 CHIP and CFPS Data Harmonization

Although both CHIP and CFPS are nationally representative surveys, their samples are drawn from different provinces in China.[4] Therefore, we need to make sure we use the correct sampling weights to make those two samples comparable. In the CFPS samples, we use "the individual-level national sampling weights" provided in the data set. In CHIP, we use the sample weights based on regional and provincial total population for CHIP samples, following Li et al. (2017) for CHIP 2007 and 2013. Since Li et al. (2017) only provide the

[4] Table A.10 in the Appendix lists the covered provinces for each survey by year.

sampling weight information for the years 2007 and 2013 but not for the earlier years, we do not apply weights for the CHIP 1995 and 2002.[5]

To construct the hourly wage variable given yearly earnings, information about each individual's working hours is necessary. Since CHIP 1988 does not have information about hours worked, we are forced to exclude it from our analysis. Additionally, we exclude CFPS 2010, 2012, and 2016 from our analysis due to missing values in key variables. Specifically, in CFPS 2010 and 2012, we found abnormal employment rates, especially for non-college-educated females in the raw sample. As a reference, the employment-to-population ratio was 67.75% in 2010 for individuals aged 15+ according to the World Bank; however, in CFPS 2010, after applying sampling weights, the employment-to-population ratio is only 55.41% for the same age group, and 63.25% for individuals aged 25 – 55. We also noticed that, compared to the CHIP sample, the CFPS sample generally has a lower employment rate. However, compared to CHIP 2007, CHIP 2013, and CFPS 2014, non-college-educated females in CFPS 2012 experienced an extremely low employment rate. The employment ratio for non-college-educated females is between 60 - 75% for CHIP 2007, CHIP 2013, and CFPS 2014; however, the employment ratio is below 60% in CFPS 2012, which we have not found any reference in explaining. Therefore, we exclude CFPS 2010 and CFPS 2012 from our analysis. In CFPS 2016, an improper operation failed to collect main-job-related information for individuals who did not experience work changes between CFPS 2014 and CFPS 2016 (see CFPS Database Clean Report), which makes these data not usable to us as we would not be able to measure earnings and hours worked accurately for everyone in the sample. Therefore, we use data from CHIP 1995, 2001, 2007, 2013 together with CFPS 2014 and 2018 to construct our analytic sample. This sample includes Chinese urban residents aged 25 to 55 with an urban hukou who do not work in the agriculture sector.

[5]Not applying these sampling weights is also consistent with the previous studies that used CHIP 1995 and 2002 (for example, Xing and Li, 2012; Zhu, 2016; Yang and Gao, 2018), which also makes our results more comparable to the literature.

3.3 Key Variables Construction

There are some differences between CHIP and CFPS in the income and employment variables. Following Kanbur et al. (2021) and Li and Wan (2015) both of whom use CFPS and CHIP data to analyze the evolution of household income inequality, we break down different income sources in CHIP (for both individual's income and household income) and reconstruct them into the same income definition as in CFPS. Below we discuss how we construct each key variable.

3.3.1 Hourly Wage

In our analysis, earnings are measured in an accounting period of one year. They include regular wages, overtime compensation, allowances, and bonuses. This is the same definition employed in Gustafsson and Wan (2020) and Zhu (2016). We use an individual's earnings from the major/primary job as the earnings measure in our analysis. For cases where the survey does not specify a major/primary job for an individual, we used the earnings from the job where an individual spent the most time and which had the highest-earning. Earnings are adjusted to the 2018 prices level using the national urban consumer price index provided by the National Bureau of Statistics of China.

To construct the hourly wage, information about hours worked is needed. Among all the surveys, only CHIP 2002 has yearly earnings with working hours per day, working days per month, and months worked to accurately construct hourly wage. In other surveys, where the annual working hours are not directly provided, we compute annual working hours by using either worked hours per week or worked hours per month, whichever is available, assuming workers work four weeks per month and 52 weeks per year. We then construct the hourly wage for our primary analysis by dividing the annual primary income by the annual total working hours, following Hering and Poncet (2010), Kamal et al. (2012), and Lovely et al. (2019). Constructing hourly wages helps us account for the intensive margin of labor supply.

The left panel of Figure 2 presents the observed log wage gender gap at the median, and the right panel presents the observed log hourly wage estimates by gender. From the graph, we can tell that there is a progressive increase in the gap before 2007, and after 2007 the direction changes and shows a decreasing trend.

3.4 Sample and Summary Statistics

Our sample includes Chinese urban residents aged 25 to 55 with an urban hukou and not working in the agriculture sector. We focus on urban households to mitigate the differences in social benefits between households with urban and rural hukou (Xing and Li, 2012). We exclude individuals with no household registrations or foreign residents for similar reasons. An individual is classified as employed ($E_i = 1$) if he/she is reported to have been employed during the past year. Since we use the hourly wage in our analysis, we treat self-employed individuals as employed ($E_i = 1$) but exclude them from calculating the observed wage distribution. The observed wage distribution is conditional on the employed individuals ($E = 1$) after controlling for the observed individual characteristics x, $F(w|x, E = 1)$. We control for age and education in the analysis. We divide our sample into two age groups and two education groups. We define individuals older than 45-years-old as in the old age group and individuals aged 45 or younger as in the young age group. For those with at most a high school degree, we define them as non-college degree holders, and for those with either a Dazhuan degree or at least a college degree as college degree holders.[6]

Figure 3 shows the changes in employment (including self-employed) against age by gender. Compared to 1995, the probability of employment for males under age 45 and females under age 40 increased in 2018. However, there is a dramatic drop in the employment probability for males around 50 and females around 45. This is correlated with the statutory retirement age in China – 60 for males and 55 for females in China.

Figure 4 illustrates that the changes in employment have been heavily skill-and-gender-biased. The employment gap between college-educated and non-college-educated females

[6] We do not use finer age and education groups because constructing bounds on the wage distribution requires a large number of observations.

is larger than their male counterparts'. Moreover, the non-college females' employment dropped greatly in 2013. If low-skilled women are exiting employment, we anticipate the gender wage gap would be larger after considering the employment composition in the 2010s.

4 Results

4.1 Changes in the Median Gender Wage Gap

This section presents the results of estimated bounds on the changes in the median gender wage gap in China under different assumptions. Importantly, the estimated bounds account for employment composition. Figure 5 shows the results for changes from 1995 to 2018.[7] In each figure, the space between the two dots represents the bounds of the change in the gender wage gap between 1995 and 2018. The thin outer lines denote the 95% confidence interval for the change in the gender wage gap. Panel A presents the estimated results for young people without a college degree. The worst-case bounds to the change in the gender wage gap for this group include zero change. To narrow the worst-case bounds, we separately impose the quartile and stochastic dominance restrictions. With the quartile dominance restriction alone, the estimated bounds for the young non-college educated group indicate an increase in the gender wage gap differetials of at least 0.10 log points and by at most 0.65 log points. However, the 95% confidence intervals (CIs) do not exclude a zero change. Using the stronger stochastic dominance assumption, the bounds are tighter. The bounds of the young non-college indicate an increase of the gender wage gap of at least 0.17 log points to at most 0.62 log points, with the 95% CI excluding zero. Panel B shows the estimated results for young college graduates. Similar to the estimates for their non-college graduated peers, the worst-case bounds include a zero change. With the imposed quartile dominance restriction, we find an increase in the gender wage gap by at least 0.03 log points

[7]Table A.1 in the appendix reports the values for the upper and lower bounds and the corresponding 95% confidence intervals (CIs) of the bounds in Figure 5.

to at most 0.21 log points. Using a stronger stochastic dominance assumption to further tighten the bounds, we find an increase in the gender wage gap of at least 0.05 log points to at most 0.20 log points. However, the 95% CIs cannot exclude a zero change under either of the restrictions. Panel C shows the estimated changes in the gender wage gap for old non-college graduates. Similar to the estimates for the young groups, the worst-case bounds include a zero change. Even though under the quartile and stochastic dominance assumption, the estimated bounds are tighter, none of them exclude a zero change for the changes in the gender wage gap for old non-college graduates. Panel D shows the estimated changes for the old college-educated group. Neither of the worst-case bounds nor the bounds under the quartile dominance assumption exclude zero. With a stronger assumption, we find an increase in the gender wage gap among college graduates over age 45 for at least 0.12 log points to at most 0.47 log points. However, the 95 CIs could not exclude a zero change in the gender wage gap.

Overall, the worst-case bounds to the change in the gender wage gap all include zero change. In addition, we find that the worst-case bounds are with a large width, especially for the non-college-educated groups. These larger widths are partially due to the low employment rates, as shown in Figure 4; there is also a wide difference in the employment rate between females with and without a college degree. We impose quartile dominance restriction and stochastic dominant assumption to tighten the bounds. With the quartile dominance restriction alone, we do not find statistically significant evidence of changes in the gender wage gap. Under the stochastic dominance assumption, the bounds of the young non-college indicate an increase of the gender wage gap of at least 0.17 log points to 0.62 log points, with the 95% CI excluding zero. We find that some bounds are wide, especially for older workers; those wide bounds are mainly due to the low employment of females and older workers (Figure 4).

To explore any potential changes in the trend of the gender wage gap through the 23 years between 1995 - 2018, we split our study period into 1995 – 2007 and 2007 – 2018. The break in 2007 is motivated by the finding in Song et al. (2019) of a temporary narrowing in the gender wage gap from 2007 to 2013. Figure 6 presents the estimated bounds from 1995 to 2007.[8] During 1995 - 2007, we cannot conclude any change in the gender wage gap among young people without a college education using either the worst-case bound or under the quartile dominant restriction (Panel A). Using the stronger stochastic dominance assumption, we find an increase in the gender wage gap of at least 0.04 log points to at most 0.48 log points. However, the 95% CIs cannot exclude a zero change. At the same time, it is striking to see that the worst-case bounds for the young college graduates indicate a 0.07 - 0.32 log points increase in the gender wage gap, and the 95% CIs exclude zero (Panel B). Since worst-case bounds do not utilize any restrictions on the wage distribution, we consider this a strong indication of an increase in the gender wage gap for this group. Under the quartile restriction, the bounds show similar results as the worst-case bounds, with tighter bounds for the young college graduates showing an increase of the gender wage gap of 0.13 - 0.28 log points, and the 95% CI excluding zero. The bounds under stochastic dominance are the narrowest, showing a statistically significant increase in the gender wage gap of 0.15 - 0.27 log points. The estimated bounds for the old non-college graduates do not exclude a zero change in the gender wage gap (Panel C), the tightest bounds under the stochastic dominance assumption showing a potential decrease of at most 0.22 log points and a potential increase of at most 0.41 log points. The worst-case bounds for older college graduates indicate an increase of 0.10 - 0.25 log points in the gender wage gap, although the CI does not exclude a zero change(Panel D). After imposing the restrictions, neither the 95% CIs under the quartile or the stochastic dominance exclude zero changes in the gender wage gap among college graduates aged above 45, even though the tightest bounds under the stochastic dominance indicate an increase in the gender wage differential for at least 0.12 log points to 0.19 log points. Overall, we find a statistically significant increase in the gender wage gap among young college graduates but not other groups from 1995 to 2007.

[8] Table A.2 in the appendix reports the corresponding estimated values for the upper and lower bounds and the corresponding 95% CIs.

Figure 7 presents the bounds of the change in the median gender wage differential from 2007 to 2018[9]. For every group under consideration, the estimated worst-case bounds, the estimated bounds under the quartile dominance restriction, and the bounds under the stochastic dominance restriction all include zero.

In summary, at the median of the wage distribution from 1995 to 2018, the estimated bounds indicate a statistically significant increase in the gender wage gap for the young workers who are non-college-educated, and this gap has increased by 0.17 - 0.62 log points. After splitting the analysis into two time periods from 1995 - 2007 and 2007 - 2018, the estimated bounds indicate a significant increase in the median gender gap among young college graduates in the earlier period. We do not find any statistically significant change in the median gender wage gap in the later period for either group under consideration.

4.2 Changes in the $25th$ Gender Wage Gap

Figure 8 to Figure 10 present the estimated bound on the gender wage gap changes over time at the $25th$ quantile of the wage distribution.[10] Except for some bounds of the old college graduates in 1995-2018 and 1995-2007 and young individuals, the estimations indicate inconclusive changes in the gender wage gap for all the age and education groups in the two different time periods.

Figure 8 shows the change from 1995 to 2018. From the figure, none of the estimated bounds excludes a zero change based on the 95% CIs. The narrowest bounds are those under the stochastic dominance assumption. From panel A, the estimated bounds for the young non-college graduates indicate an increase in the gender wage gap of 0.04 - 1.23 log points under the stochastic dominance assumption. The estimated bounds for the young college graduates (Panel B) suggest an increase in the gender wage gap of 0.04 - 0.36 log points. The bounds for the older non-college graduates (Panel C) rule out a decrease in the gap

[9]Table A.3 in the appendix report the corresponding estimated values for the upper and lower bounds and the corresponding 95% CIs.

[10]Appendix Tables A4 - A6 present the corresponding values in these figures.

by more than 0.47 log points and an increase by more than 0.89 log points. The estimated bounds for the older college-graduates (Panel D) suggest an increase in the gender wage gap of 0.06 - 0.90 log points.

Figure 9 presents the estimated bounds on the gender wage gap change between 1995 - 2007. The estimated bounds for the non-college groups (Panel A and Panel C) include zero. For the young college-graduates group(Panel B), the estimated bounds suggest similar implications as with the gender wage gap at the median wage. From the worst-case bounds to bounds under different restrictions, the estimated bounds suggest there is a statistically significant increase in the $25th$ gender wage gap for this group. Based on the estimated bounds under the stochastic dominance, the increase is at least 0.07 log points and at most 0.32 log points. Additionally, the estimated bounds under the stochastic dominance indicate an increase in the gender wage gap of 0.01 - 0.20 log points for the old college graduates (Panel D); however, the 95% CIs do not exclude zero.

Figure 10 presents the estimated bounds for the change in the gender wage gap between 2007 - 2018. The estimated bounds for all groups are inconclusive for the sign of the gender wage gap changes. The tightest bounds are under stochastic dominance. The estimated lower bounds indicate a decrease in the gender wage gap of 0.08 - 0.67 log points, and the estimated upper bounds indicate an increase in the gender wage gap by 0.25 - 1.19 log points.

In a nutshell, compared to the estimates of the changes in the median gender wage gap, the results are less conclusive for the gender wage gap over time at the $25th$ quantile of the wage distribution. Some evidence suggests an increase in the gender wage gap at the $25th$ quantile of the wage distribution for the old college-educated group and young groups, especially for young college graduates.

4.3 Changes in the $75th$ Percentile Gender Wage Gap

Figure 11 to Figure 13 present the estimated bounds on the gender wage gap change over time at the $75th$ quantile of the wage distribution.[11] Figure 11 shows the change from 1995 to 2018. The estimated bounds for the old groups and non-college-educated groups are not inclusive (Panel A, C, and D). The estimated bounds under the quartile restriction and the bounds under the stochastic dominance show an increase in the gender wage gap for young college graduates (Panel B) of 0.04 - 0.18 log points and 0.07 - 0.18 log points, respectively. However, none of the 95% CIs excludes a zero change. After we split up the study period, the estimated bounds show a consistent increase in the gender wage gap for college graduates (Figure 12 Panel B and D). The estimated worst-case bounds suggest a 0.03 to 0.38 log points increase in the gender wage gap for young college graduates(Panel B). The estimated bounds under quartile dominance restriction are tighter and suggest a statistically significant increase in the gender wage gap for young college graduates of 0.17 - 0.30 log points, and this estimated increase is 0.20 - 0.28 log points after imposing the stochastic dominance restriction. For the old college graduates (Panel D), without any further restriction, the estimated worst-case bounds indicate an increase in the gender wage gap of 0.12 to 0.35 log points, with the 95% CI including a zero change. The estimated bounds under the quartile and the stochastic dominance restrictions suggest a significant increase in the $75th$ gender wage gap for college graduates above age 45 by 0.16 - 0.24 log points and 0.17 - 0.22 log points, respectively.

Figure 13 presents the estimated results during 2007 - 2018. The estimated bounds under the quartile and the stochastic dominance restrictions suggest a decrease in the gender wage gap of 0.02 to 0.22 log points and 0.05 to 0.19 log points for young college graduates (Panel B), respectively. However, the 95% CIs do not exclude a zero change. The estimated bounds for the other education and age groups all include zero change and are inconclusive.

In summary, at the $75th$ quantile of the wage distribution, the estimated bounds indicate

[11] Appendix Tables A7 - A9 present the values in these figures

a statistically significant increase in the gender wage gap for workers who are college-educated in 1995 - 2007. After 2007 to 2018, the estimated bounds indicate a decrease in the gender gap among young college graduates at the $75th$ quantile of the wage distribution. Over the whole period of interest, the estimated bounds show a statistically significant increase in the gender wage gap for young college graduates. We do not find statistically significant changes in the gender wage gap for all the other groups.

5 Discussion

Our estimated bounds show a pattern of an increasing gender wage gap among the young workers (age 25-45) in survey years of 1995-2007 at the median, the $25th$ and the $75th$ quantile of the wage distribution, after accounting for the employment composition. The increase in the gender wage gap from 1995 to 2007 is between 0.15 - 0.28 log points. This result is in line with previous findings without fully accounting for employment composition by Gustafsson and Wan (2020), which show an increase in the gender earnings gap from 1988 - 2007 by 0.14 log points, and findings by Song et al. (2019), who estimates a 0.15 log points increase in the gender earnings gap from 1995 – 2007. By separating the estimates by different age and education groups, our results suggest that the gender wage gap increase may be larger among the young college-educated workers than the other groups.

Specifically, our estimated lower bound estimates show an increase of 0.15 - 0.27 log points at the median, of 0.07 - 0.32 log points at the $25th$ quantile and of 0.20 - 0.28 log points and at the $75th$ quantile of the wage distribution. These magnitudes are greater than the estimated gender wage gap increase in Gustafsson and Wan (2020) and Song et al. (2019), which were based on the population of age 16 - 70 and 16 - 60, respectively, and which do not account for employment composition.

Our bounds for young college graduates during the period 2007 - 2018 suggest a decrease of the gender wage gap at the 75th percentile of 0.05 - 0.19 log points, while the 95% CI does not exclude zero. This result suggests that the narrowing of the gender wage gap might be potentially larger in 2007-2018 than what Song et al. (2019) has previously documented, where they find the gender wage earnings gap narrowed between 2007 - 2013 by 0.04 log points without accounting for the employment composition. One potential explanation could be the self-selection of employment for females. Suppose more young high-skilled women choose to be self-employed or work for fewer hours in recent years. Without controlling for selection to employment and labor supply, estimates may overstate the gender wage gap and understate the decrease in the gender wage gap in more recent years. This could potentially explain a larger decrease in the gender wage gap after 2007, suggested by our bounds estimates compared to Song et al. (2019).

Our results suggestively show different trends in the evolution of the gender wage gap in two time periods. Economic factors that contribute to the gender wage gap may explain the potentially different trends. In the time period of 1995 - 2007, we find results consistent with an increase in the gender wage gap among young workers both at the median wage and at the $75th$ wage quantile. The widened gender wage gap can be explained by the privatization and marketization in the 1990s' China (Liu et al., 2000; Maurer-Fazio and Hughes, 2002). Shu et al. (2007) also show that globalization perpetuates the gender wage differential by absorbing women in exporting-orientated manufacturing jobs that offer lower wages.

Different from 1995 - 2007, in the later period 2007 - 2018, we do not find evidence of any increase in the gender wage gap, and some weak evidence of a decrease in the gender wage gap among the young workers who are college-educated both at the median wage and at the $75th$ wage quantile. One potential explanation for this slow-down of the gender wage gap growth can be higher returns to the schooling of women relative to men and an increase in the return to schooling in China (Ma and Iwasaki, 2021). Using panel data of the China population from 2011 - 2015, McGarry and Sun (2018) show that the gender schooling gap in China has been diminishing from birth cohorts born in the 1950s to those born in the late 1980s. Suppose women are gaining more years of schooling over birth cohorts while the return to schooling is increasing and higher for women than for men. In that case, the schooling factor may significantly contribute to the closing of the gender wage gap among college-educated young workers. However, other offsetting factors, such as gender discrimination, may also exist to slow down the closing of the gender wage gap. These factors include the intra-sector gender wage differential Ma (2018), as well as the increase in men's labor market return to work experience relative to females'(Hare, 2019 and Zhao et al., 2019). Future research can look into the mechanisms that contribute to those changes in the gender wage gap at different quantiles of the wage distribution while accounting for the employment composition.

6 Conclusion

This paper estimates China's distributional gender wage gap dynamics from 1995 to 2018. To control for selection into employment, we employ nonparametric bounds in the spirit of Manski (1994), Manski and Pepper (2000), and Blundell et al. (2007) under different assumptions. To tighten the bounds, we use a weak quartile dominance assumption and a stochastic dominance assumption.

We have found statistically significant evidence that over the years from 1995-2018, the median gender wage gap for young workers (age 25-45) who are non-college-educated has increased by 0.17 - 0.62 log points. By splitting the study period, in the survey period between 1995-2007, we show a significant increase in the median gender wage differentials from 1995 to 2007 among young workers who are college-educated (an increase of at least 0.15 log points).

Additionally, this paper also estimates the gender wage gap change at the $25th$ and the $75th$ percentiles of the wage distribution. At the $25th$ percentile, all bounds estimates do not statistically significantly exclude zero change in the gender wage gaps between 1995 - 2007 or 2007 - 2018. At the higher $75th$ percentile of the wage distribution, in the earlier years of 1995-2007, we find significant increases in the gender wage gap in 1995-2007 for both the young and older college-educated workers. However, we do not find evidence that the increase in the gender wage gap has persisted into the 2010s.

Although we do not find that the gender wage gap in China has continued to increase after 2007, we also do not find strong evidence that the gender wage gap is closing in more recent years in any education and age groups we considered. In addition, studies such as Song et al. (2014) and Ma (2018) show majority portion of the gender wage gap is not explained by social and labor market characteristics. To sustain economic growth and reduce gender inequality, the Chinese labor market needs more protective legislation for women, such as reinforcing equal pay for work guidelines, non-discriminatory policies in hiring, and pay data collection. Future research can look into the mediating factors of the apparent slowdown of the gender wage gap in recent years and evaluate the impacts of recent policy changes, such as the two-child policy, on the gender wage gap and women's labor market outcomes.

References

Aizer, A. (2010). The gender wage gap and domestic violence. *American Economic Review*, 100(4):1847–59.

Angrist, J. D., Krueger, A. B., Ashenfelter, O., and Card, D. (1999). Handbook of labor economics. *Causality: Statistical Perspectives and Applications (O. Ashenfelter and D. Card, eds.), edition*, 1:1277–1366.

Becker, G. (1973). A Theory of Marriage: Part I. *Journal of Political Economy*, 81(4):813–846.

Blau, F. and Beller, A. (1988). Trends in Earnings Differentials By Gender, 1971–1981. *Industrial and Labor Relations Review*, 41(4):513–529.

Blau, F. and Kahn, L. (2006). Swimming Upstream: Trends in the Gender Wage Differential in the 1980s. *Journal of Labor Economics*, 15(1):1–42.

Blau, F. and Kahn, L. (2017). The gender wage gap: Extent, trends, and explanations. *Journal of Economic Literature*, 55(3):789–865.

Blau, F., K. L. B. N. and Comey, M. (2023). The impact of selection into the labor force on the gender wage gap. *Journal of Labor Economics, forthcoming*.

Blundell, R., Gosling, A., Ichimura, H., and Meghir, C. (2007). Changes in the Distribution of Male and Female Wages Accounting for Employment Using Bounds. *Econometrica*, 75(2):323–363.

Blundell, R. and MaCurdy, T. (1999). Labor Supply: A Review of Alternative Approaches. *Handbook of Labor Economics*, 3:1559–1695.

Bratti, M. (2015). Fertility Postponement and Labor Market Outcomes: Postponed Childbearing Increases Women's Labor Market Attachment But May Reduce Overall Fertility. *IZA Working Papers*, 2015(117).

Chernozhukov, V., Lee, S., and Rosen, A. (2013). Intersection Bounds: Estimation and Inference. *Econometrica*, 81(2):667–737.

Chi, W. and Li, B. (2014). Trends in China's Gender Employment and Pay Gap: Estimating Gender Pay Gaps with Employment Selection . *Journal of Comparative Economics*, 45(3):708–725.

Fee, E. (1991). The Gender Gap in Wages and Health. *Health Affairs*, 10(4):304–305.

Feng, S., Hu, Y., and Moffitt, R. (2017). Long Run Trends in Unemployment and Labor Force Participation in Urban China. *Journal of Comparative Economics*, 45:304–324.

Feng, Y., Yi, J., and Zhang, J. (2021). Rising Inter-generational Income Persistence in China. *American Economic Journal: Economic Policy*, 13(1):202–230.

Flores, C. A. and Flores-Lagunes, A. (2013). Partial Identification of Local Average Treatment Effects With an Invalid Instrument. *Journal of Business and Economic Statistics*, 31(4):534–545.

Gong, H., Leigh, A., and Meng, X. (2010). Intergenerational Income Mobility in Urban China. *the Review of Income and Wealth*, 58(3):481–503.

Gronau, R. (1974). Wage Comparisons – A Selectivity Bias. *Journal of Political Economy*, 82(6):1119–1143.

Gustafsson, B. and Ding, S. (2011). Unemployment and the Rising Number of Non-workers in Urban China: Causes and Distributional Consequences. *CIBC Working Paper*, (2011-17).

Gustafsson, B. and Li, S. (2000). Economic Transformation and the Gender Earnings Gap in Urban China. *Journal of Population Economics*, 13(2):305–329.

Gustafsson, B. and Wan, H. (2020). Wage Growth and Inequality in Urban China 1988-2013. *China Economic Review*, 62:101462.

Han, H. (2010). Trends in Educational Assortative Marriage in China from 1970 to 2000. *Demographic Research*, 22(24):733–770.

Hare, D. (2019). Decomposing Growth in the Gender Wage Gap in Urban China: 1989-2011. *Economics of Transition and Institutional Change*, 27(4):915–941.

Hering, L. and Poncet, S. (2010). Market Access and Individual Wages: Evidence From China. *The Review of Economics and Statistics*, 92(1):145–159.

Hirano, K. and Porter, J. (2012). Impossibility Results for Non-difference Functionals. *Econometrica*, 80(4):1769–1790.

Imbens, G. W. and Manski, C. F. (2004). Confidence intervals for partially identified parameters. *Econometrica*, 72(6):1845–1857.

Kamal, F., Lovely, M., and Puman, O. (2012). Does Deeper Integration Enhance Spatial Advantages? Market Access and Wage Growth in China. *International Review of Economics and Finance*, 23:59–74.

Kanbur, R., Wang, Y., and Zhang, X. (2021). The Great Chinese Inequality Turnaround. *Journal of Comparative Economics*, 49(2):467–482.

Li, H., Li, L., Wu, B., and Xiong, Y. (2012). The End of Cheap Chinese Labor. *Journal of Economic Perspectives*, 26(4):57–74.

Li, H., Liang, J., and Wu, B. (2016). Labor Market Experience and Returns to Education in Fast Growing Economies. *Unpublished Paper, Tsinghua University*, Accessed Online At:https://www.aeaweb.org/conference/2016/retrieve.php?pdfid=21842tk=SAnGNeEy.

Li, H., Loyalka, P., Rozelle, S., and Wu, B. (2017). Human Capital and China's Future Growth. *Journal of Economic Perspectives*, 31(1):25–48.

Li, S. and Wan, H. (2015). Evolution of Wealth Inequality in China. *China Economic Journal*, 8(3):264–287.

Liu, P.-W., Meng, X., and Zhang, J. (2000). Sectoral Gender Wage Differentials and Discrimination in the Transitional Chinese Economy. *Journal of Population Economics*, 13(2):331–352.

Lovely, M., Liang, Y., and Zhang, H. (2019). Economic Geography and Inequality in China: Did Improved Market Access Widen Spatial Wage Differences? *China Economic Review*, 54(2):306–323.

Ma, X. (2018). Ownership sector segmentation and the gender wage gap in urban China during the 2000s. *Post-Communist Economies*, 30(6):775–804.

Ma, X. and Iwasaki, I. (2021). Return to Schooling in China: a Large Meta Analysis. *Education Economics*, 29(4):379–410.

Maasoumi, E. and Wang, L. (2019). The Gender Gap between Earnings Distributions. *Journal of Political Economy*, 127(5):2438–2504.

Machado, C. (2017). Unobserved Selection Heterogeneity and the Gender Wage Gap. *Journal of Applied Econometrics*, 32(7):1348–1366.

Manski, C. (1994). The Selection Problem. In *Advances in Econometrics, Sixth World Congress*, volume 1, pages 143–170, U.K., Cambridge University Press.

Manski, C. and Pepper, J. (2000). Monotone Instrumental Variables: With Application to the Return to Schooling. *Econometrica*, 68(4):997–1010.

Maurer-Fazio, M. and Hughes, J. (2002). The Effects of Market Liberalization on the Relative Earnings of Chinese Women . *Journal of Comparative Economics*, 30(4):709–731.

McGarry, K. and Sun, X. (2018). Three Generations of Changing Gender Patterns of Schooling in China . *Journal of the Asia Pacific Economy*, 23(4):584–605.

Meng, X. (2012). Labor Market Outcomes and Reforms in China. *Journal of Economic Perspectives*, 26(4):75–102.

Mulligan, C. and Rubinstein, Y. (2008). Selection, Investment, and Women's Relative Wages over Time. *Quarterly Journal of Economics*, 123(3):1061–1110.

Nie, H. and Xing, C. (2019). Education Expansion, Assortative Marriage, and Income Inequality in China. *China Economic Review*, 55(3):37–51.

Platt, J., P. S. B. L. and Keyes, K. (2016). Unequal Depression for Equal Work? How the Wage Gap Explains Gendered Disparities in Mood Disorders. *Social Science and Medicine*, 149(1).

Qian, Y. and Jin, Y. (2018). Women's Fertility Autonomy in Urban China: The Role of Couple Dynamics Under the Universal Two-Child Policy. *Chinese Sociological Review*, 50(3).

Schober, T. and Winter-Ebmer, R. (2011). Gender Wage Inequality and Economic Growth: Is There Really a Puzzle? - A Comment. *World Development*, 39(8):1476–1484.

Shu, X., Zhu, Y., and Zhang, Z. (2007). Global Economy and Gender Inequalities: The Case of the Urban Chinese Labor Market. *Social Science Quarterly*, 88(5):1307–1332.

Song, J., Sicular, T., and Gustafsson, B. (2019). A Review of China's Urban Gender Wage Gap from 1995 to 2013. *Japan Labor Issues*, 3(17):4–12.

Song, S., Zhu, E., and Chen, Z. (2014). Equal Work Opportunity but Unequal Income: Gender Disparities Among Low-Income Households in Urban China. *The Chinese Economy*, 44(1):39–45.

Xing, C. and Li, S. (2012). Residual wage inequality in urban china, 1995-2007. *China Economic Review*, 23(1):205–222.

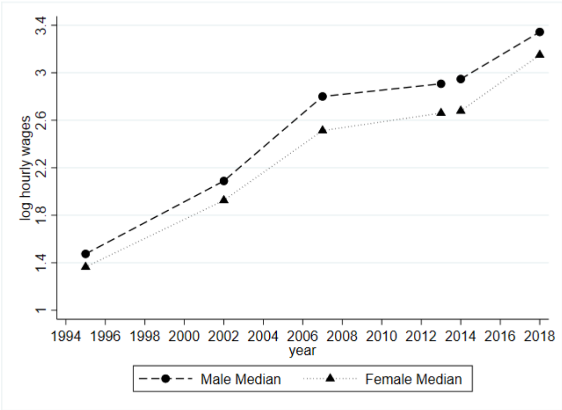
Yang, J. and Gao, M. (2018). The impact of education expansion on wage inequality. *Applied Economics*, 50(12):1309 – 1323.

Zhao, X., Zhao, Y., Chou, L., and Leivang, B. (2019). Changes in gender wage differentials in china: a regression and decomposition based on the data of chips 1995–2013. *Economic Research-Ekonomska Istraživanja*, 32(1):3162–3182.

Zhu, R. (2016). Wage differentials between urban residents and rural migrants in urban china during 2002-2007: A distributional analysis. *China Economic Review*, 37(1):2–14.

Figure 1: Distribution of Residual Wage by Gender, Age and Work History

Figure 2: Unconditional Gender Wage Gap at the Median and the Median Log Hourly Wage by Gender



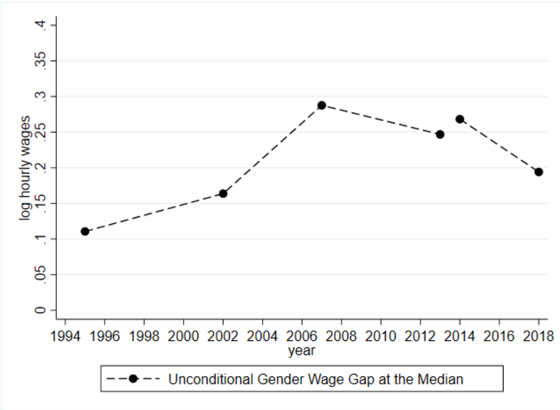


Figure 3: Age Profile for Employment for 1995 and 2018

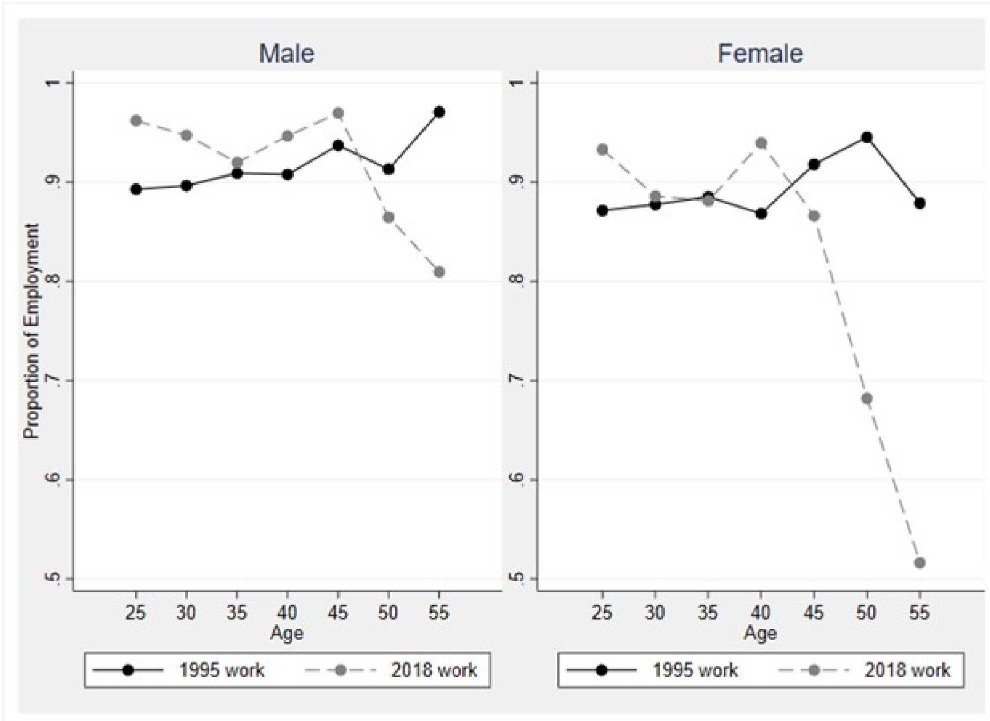


Figure 4: Employment by Education for Males and Females from 1995 to 2018

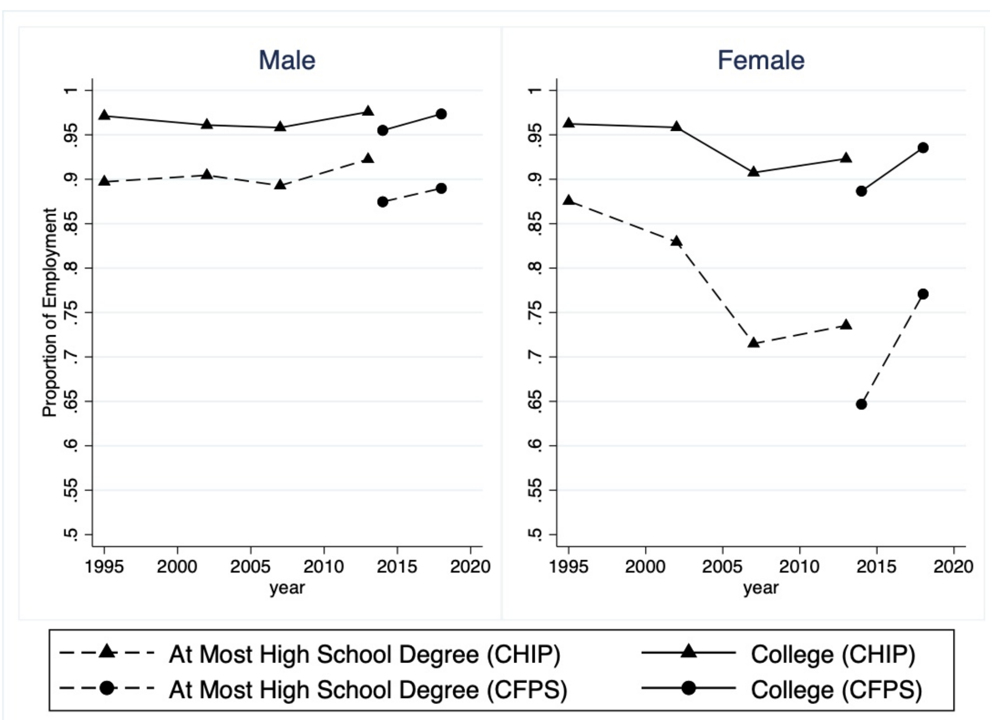


Figure 5: Changes in Median Gender Wage Gap under Various Assumptions for Different Groups (1995 - 2018)

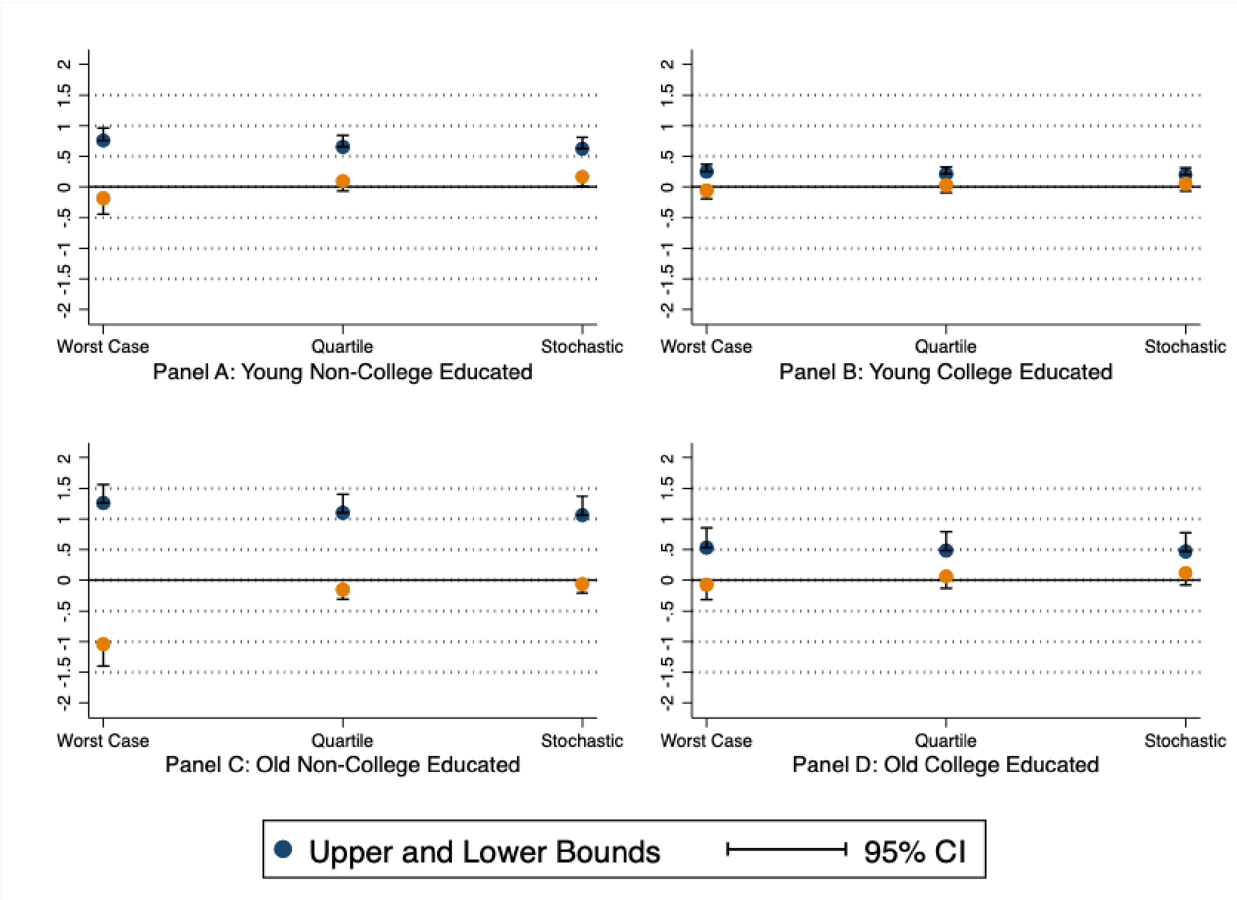


Figure 6: Changes in Median Gender Wage Gap under Various Assumptions for Different Groups (1995 - 2007)

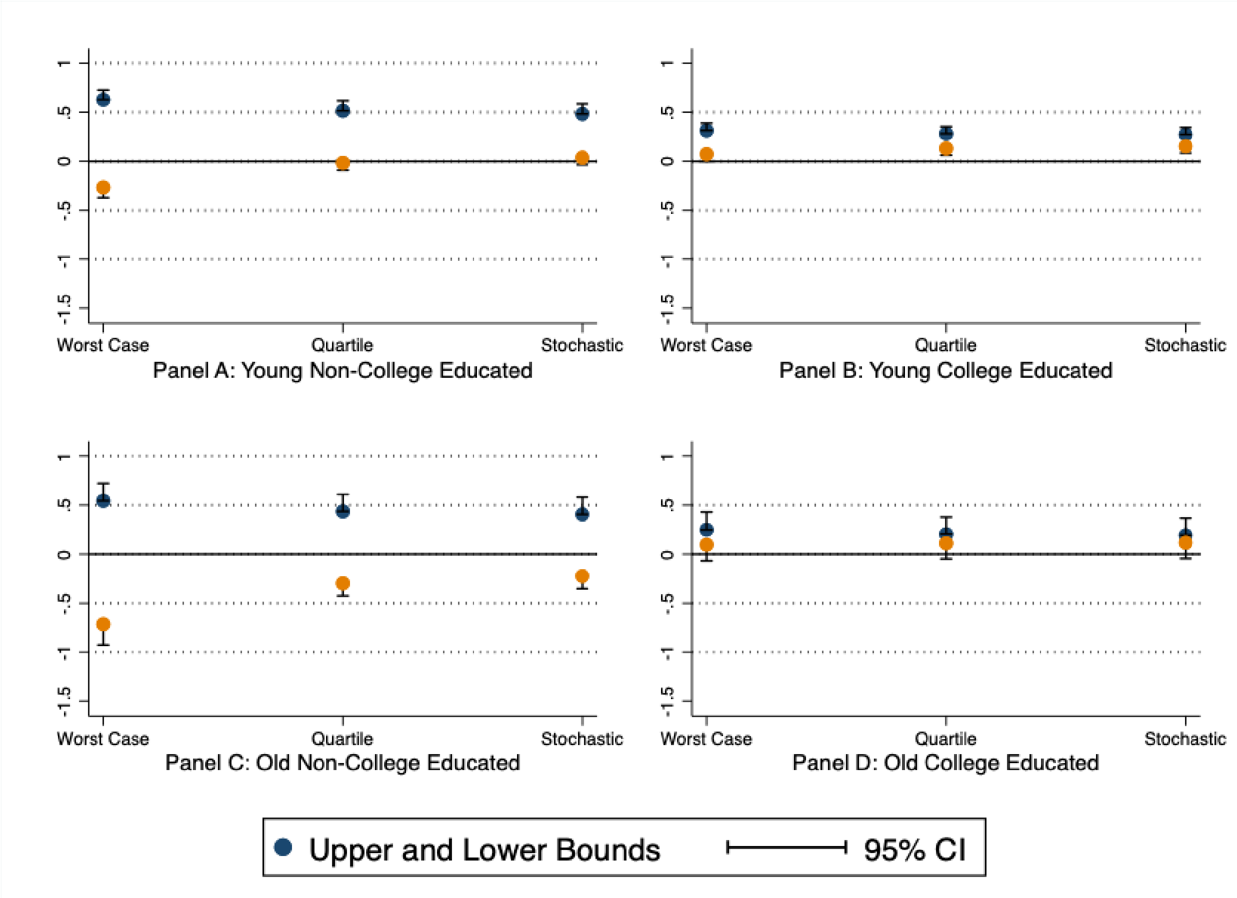


Figure 7: Changes in Median Gender Wage Gap under Various Assumptions for Different Groups (2007 - 2018)

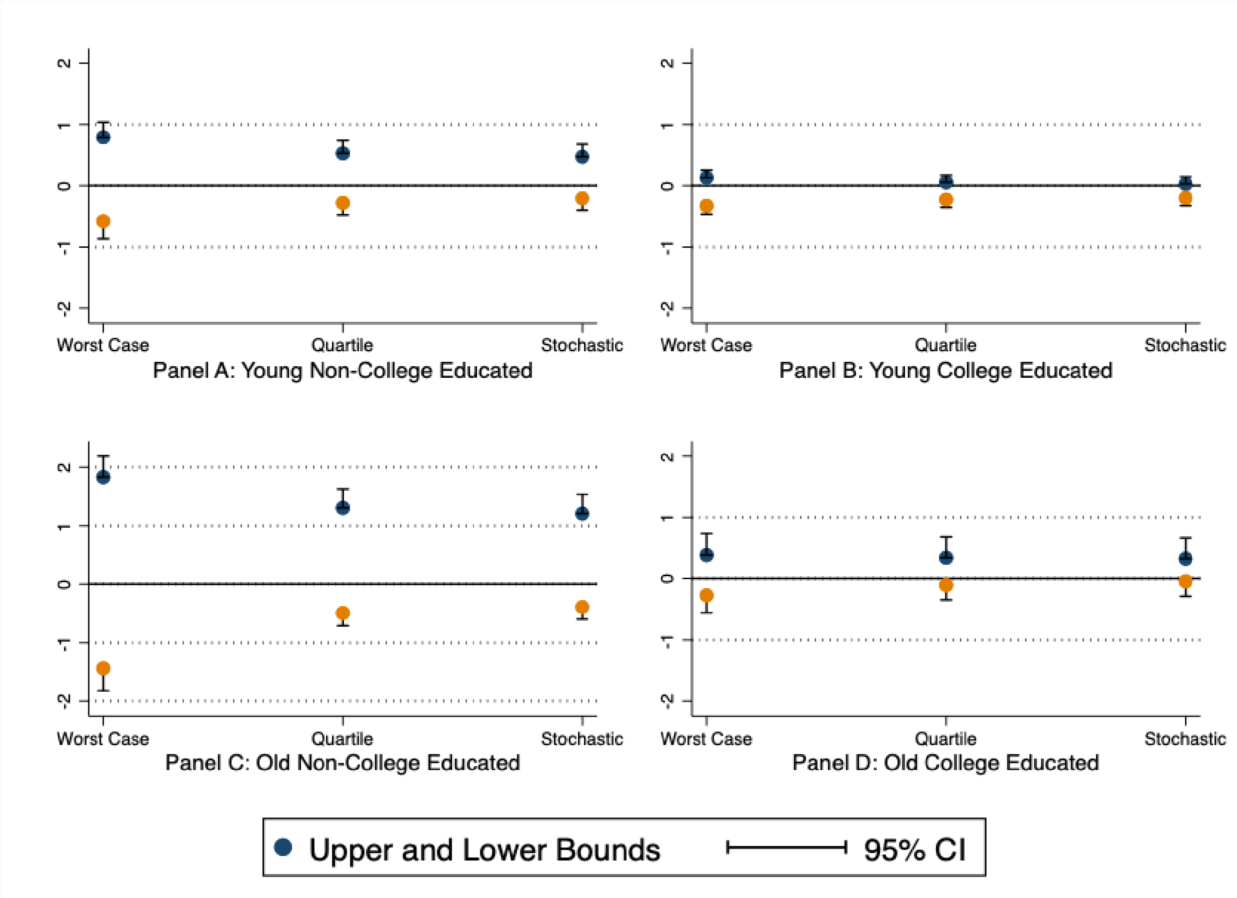


Figure 8: Changes in Gender Wage Gap under Various Assumptions for Different Groups at $25th$ Percentile (1995 - 2018)

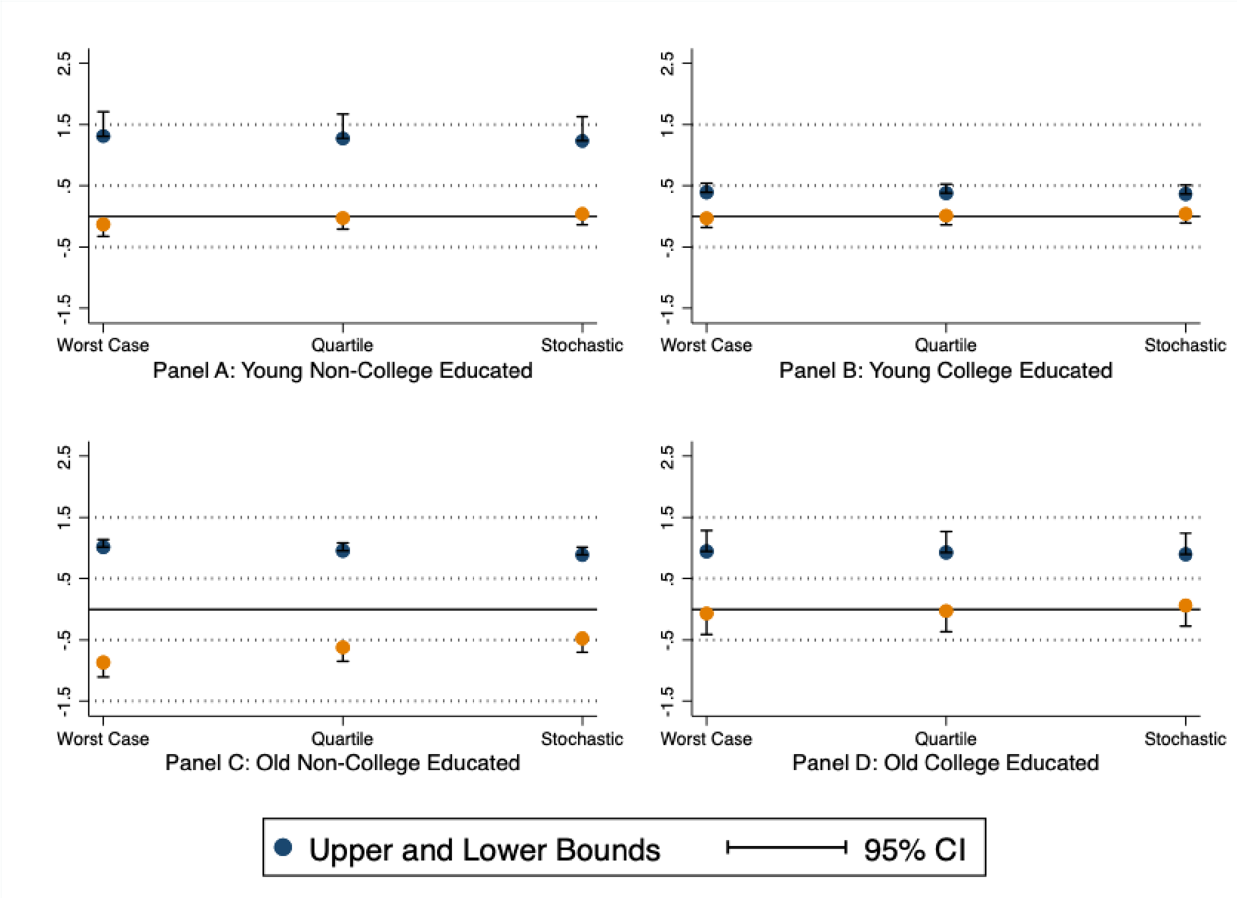


Figure 9: Changes in Gender Wage Gap under Various Assumptions for Different Groups at $25th$ Percentile (1995 - 2007)

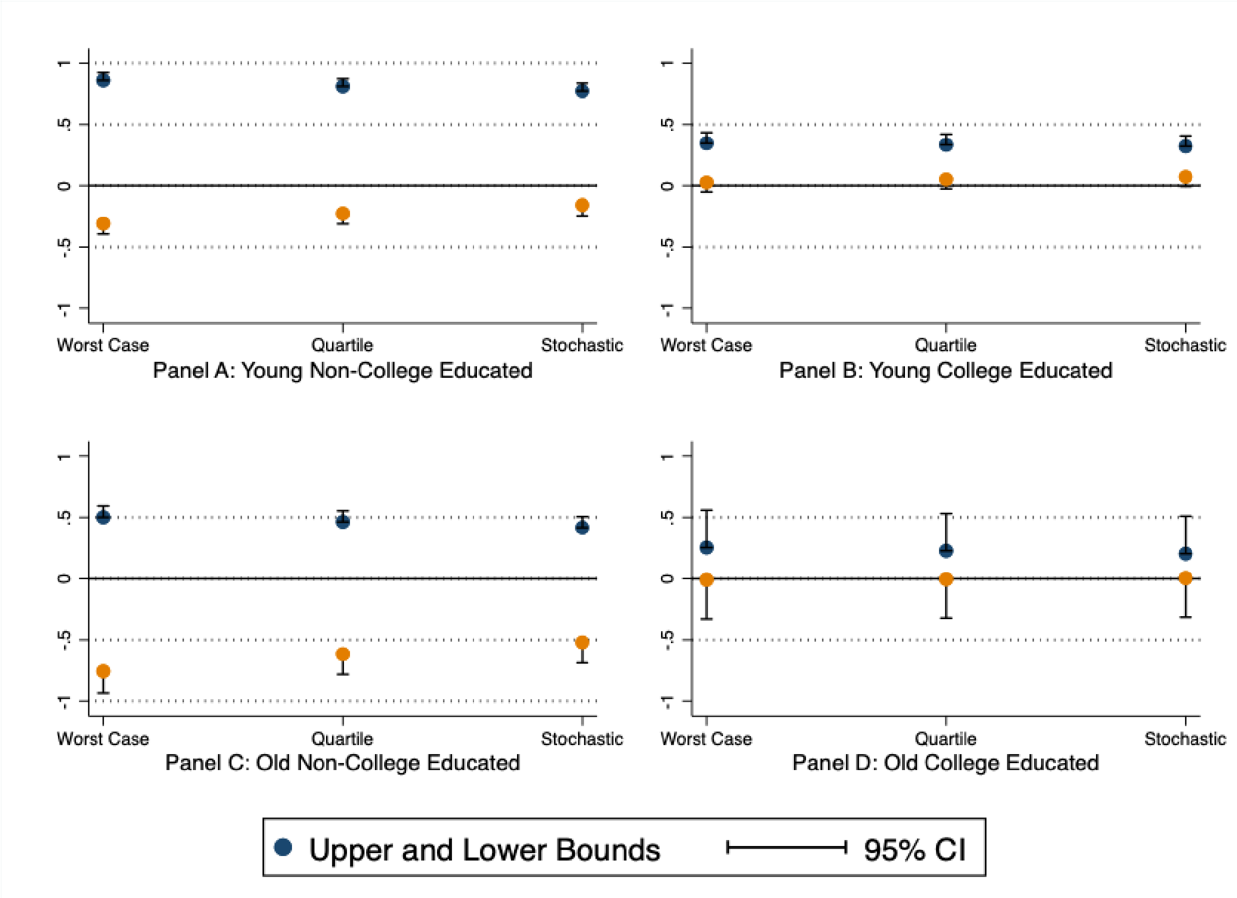


Figure 10: Changes in Gender Wage Gap under Various Assumptions for Different Groups at 25*th* Percentile (2007 - 2018)

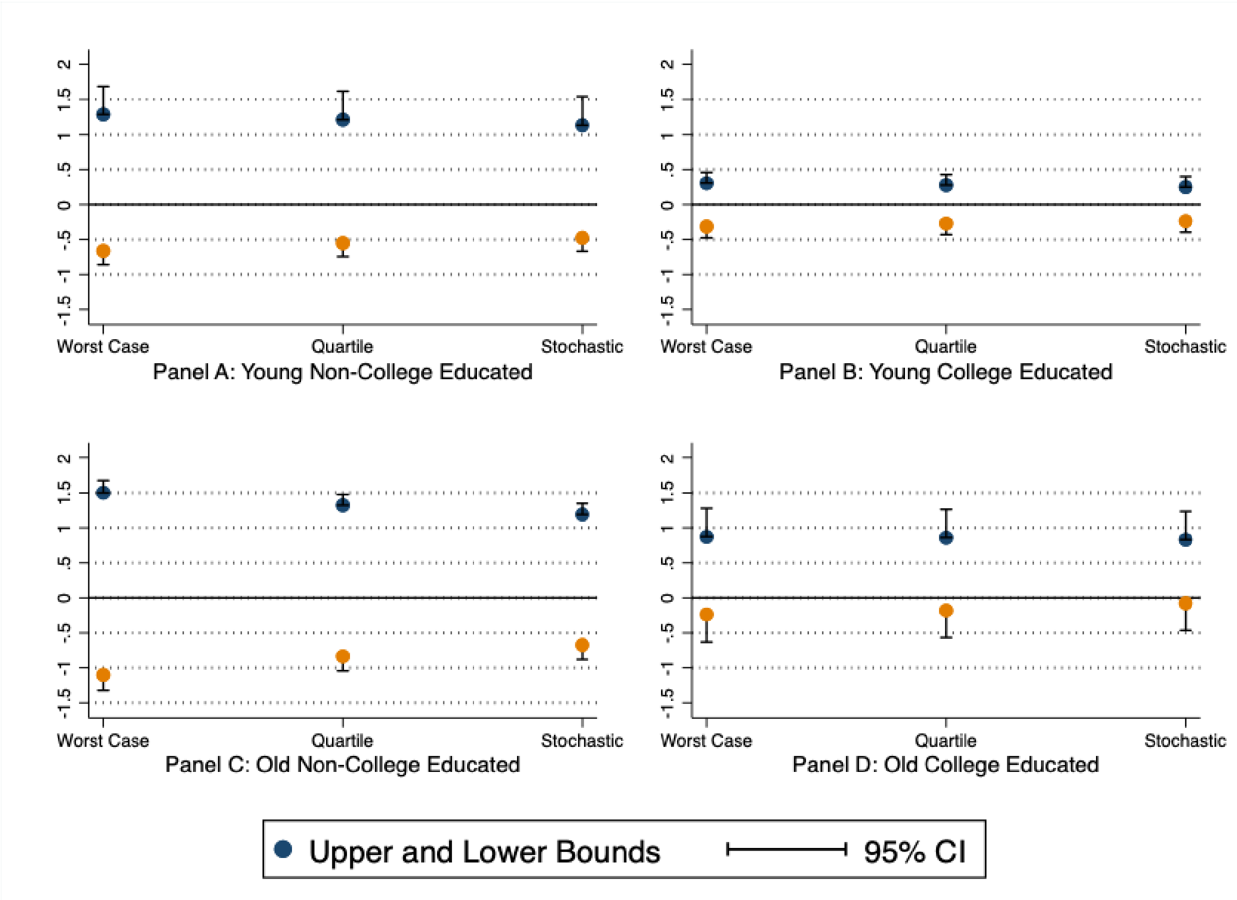


Figure 11: Changes in Gender Wage Gap under Various Assumptions for Different Groups at 75*th* Percentile (1995 - 2018)

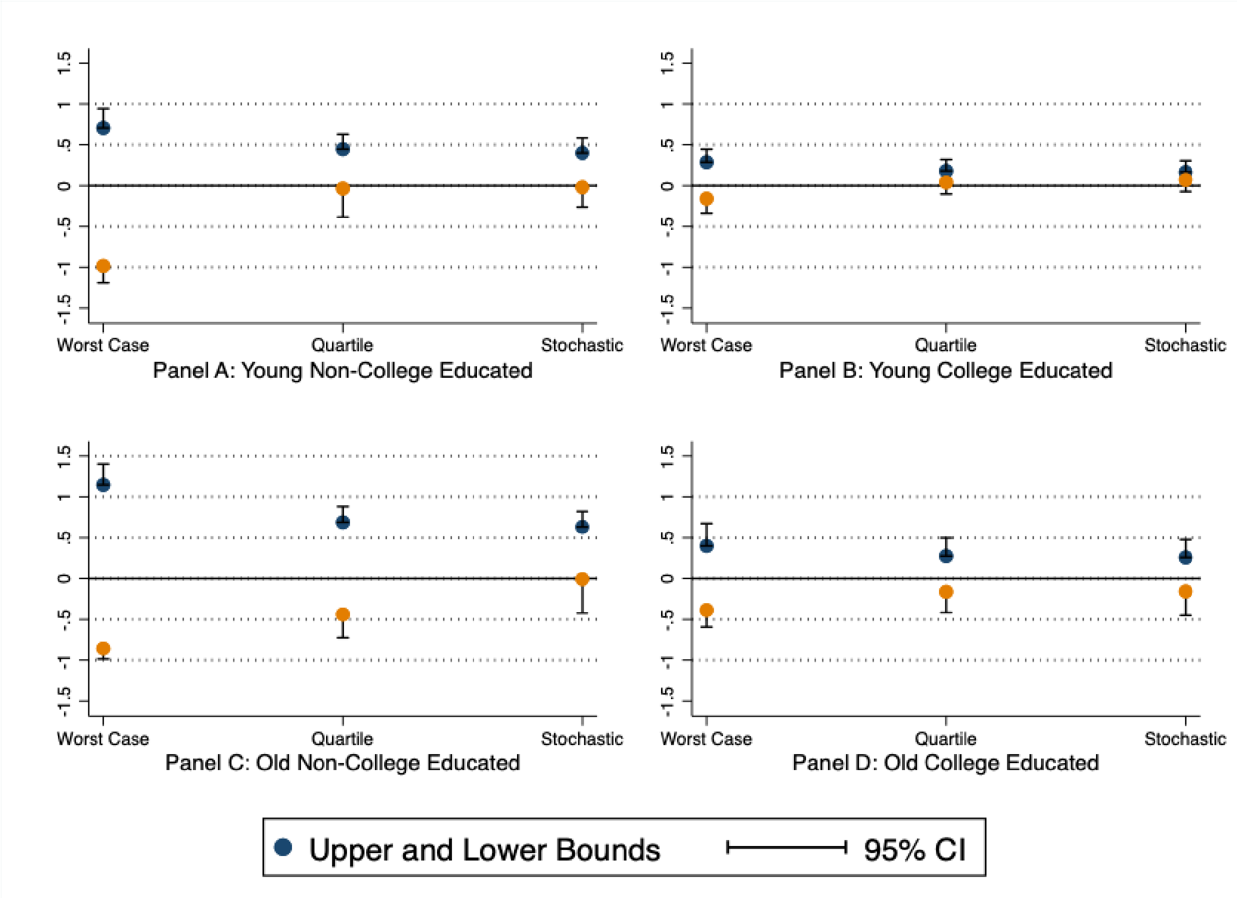


Figure 12: Changes in Gender Wage Gap under Various Assumptions for Different Groups at $75th$ Percentile (1995 - 2007)

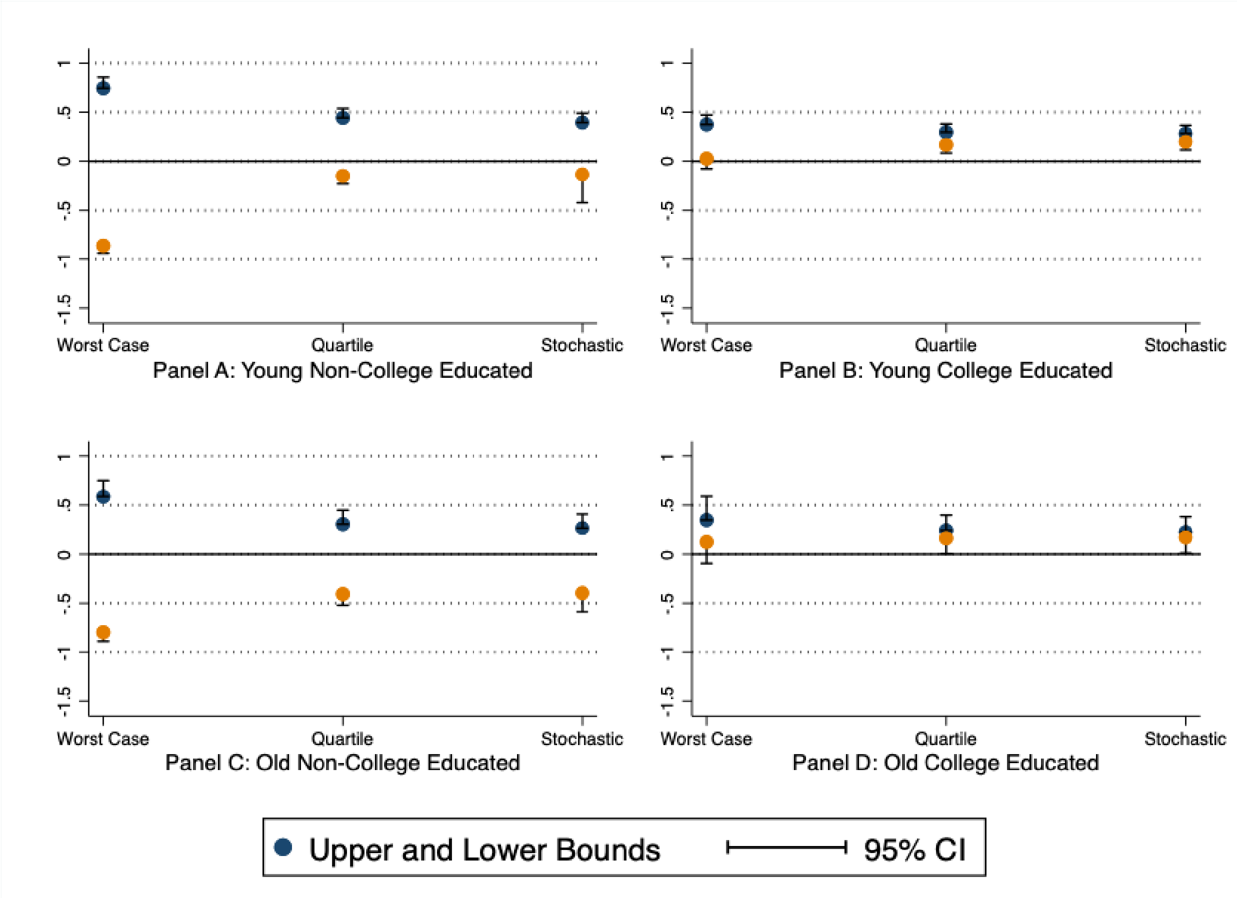
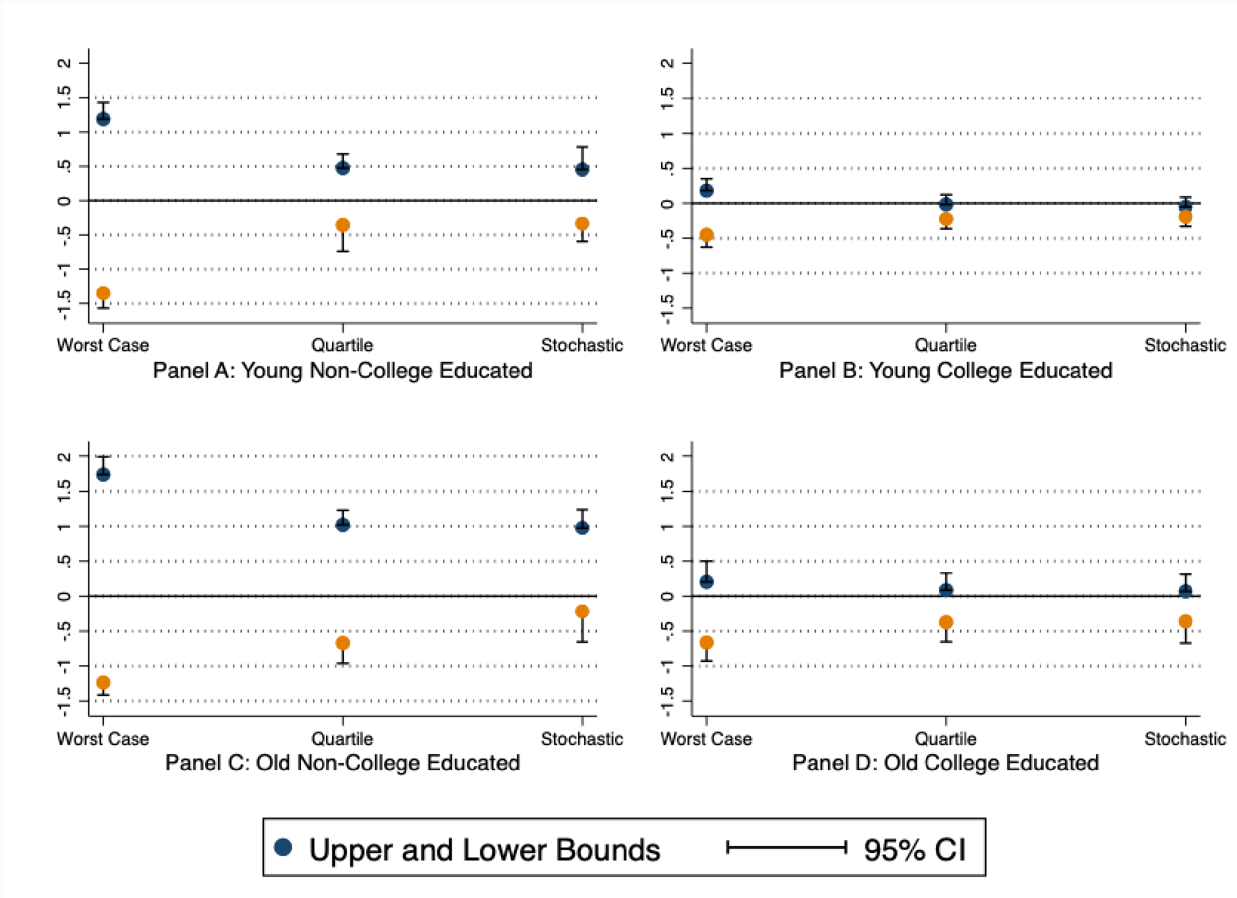


Figure 13: Changes in Gender Wage Gap under Various Assumptions for Different Groups at 75th Percentile (2007 - 2018)



Online Appendix

Appendix A. Tables

Table A.1 : Bounds on Changes in Gender Wage Differential (1995 - 2018)

	Worst Case	**Quartile Restrictions**	**Stochastic Dominance**
Young Non-College	(-0.1839, 0.7582)	(0.0966, 0.6525)	(0.1661, 0.6218)
	[-0.4410, 0.9636]	[-0.0665, 0.8449]	[0.0072, 0.8138]
Young-College	(-0.0583, 0.2532)	(0.0291,0.2138)	(0.0535, 0.2029)
	[-0.1934, 0.3697]	[-0.0943,0.3263]	[-0.0683, 0.3154]
Old Non-College	(-1.0392, 1.2585)	(-0.1525, 1.1007)	(-0.0652, 1.0626)
	[-1.4006 ,1.5633]	[-0.3081, 1.4067]	[-0.2080, 1.3711]
Old-College	(-0.0731,0.5305)	(0.0655, 0.4845)	(0.1200, 0.4692)
	[-0.3145, 0.8582]	[-0.1322, 0.7956]	[-0.0791, 0.7792]

Table A.2 : Bounds on Changes in Gender Wage Differential (1995 - 2007)

	Worst Case	**Quartile Restrictions**	**Stochastic Dominance**
Young Non-College	(-0.2684, 0.6283)	(-0.0159, 0.5145)	(0.0374, 0.4822)
	[-0.3720, 0.7277]	[-0.0887, 0.6150]	[-0.0345, 0.5840]
Young-College	(0.0727, 0.3150)	(0.1309,0.2821)	(0.1525, 0.2740)
	[0.0004, 0.3891]	[0.0617,0.3536]	[0.0837, 0.3445]
Old Non-College	(-0.7159, 0.5440)	(-0.2982,0.4364)	(-0.2237, 0.4062)
	[-0.9265 ,0.7223]	[-0.4255,0.6110]	[-0.3529, 0.5805]
Old-College	(0.0961, 0.2484)	(0.1104,0.2037)	(0.1164, 0.1919)
	[-0.0680, 0.4299]	[-0.0478,0.3783]	[-0.0440, 0.3681]

Table A.3 : Bounds on Changes in Gender Wage Differential (2007 - 2018)

	Worst Case	**Quartile Restrictions**	**Stochastic Dominance**
Young Non-College	(-0.5779, 0.7923)	(-0.2787, 0.5293)	(-0.2050, 0.4734)
	[-0.8652, 1.0367]	[-0.4741, 0.7383]	[-0.3974, 0.6780]
Young-College	(-0.3249, 0.1322)	(-0.2244, 0.0543)	(-0.1978, 0.0277)
	[-0.4630, 0.2567]	[-0.3513, 0.1727]	[-0.3234, 0.1459]
Old Non-College	(-1.4397, 1.8309)	(-0.4932,1.3032)	(-0.3910, 1.2060)
	[-1.8237, 2.1884]	[-0.7108,1.6285]	[-0.5981, 1.5356]
Old-College	(-0.2722, 0.3851)	(-0.1062,0.3422)	(-0.0458, 0.3267)
	[-0.5574, 0.7369]	[-0.3468,0.6784]	[-0.2886, 0.6622]

Table A.4 : Bounds on Changes in Gender Wage Differential at $25th$ Percentile (1995 - 2018)

	Worst Case	**Quartile Restrictions**	**Stochastic Dominance**
Young Non-College	(-0.1327, 1.3107)	(-0.0268, 1.2722)	(0.0406, 1.2321)
	[-0.3272, 1.7072]	[-0.2100, 1.6693]	[-0.1400, 1.6268]
Young-College	(-0.0302, 0.3920)	(0.0113, 0.3788)	(0.0419, 0.3640)
	[-0.1838, 0.5378]	[-0.1415, 0.5228]	[-0.1113, 0.5078]
Old Non-College	(-0.8702, 1.0143)	(-0.6224, 0.9549)	(-0.4750, 0.8891)
	[-1.1050, 1.1352]	[-0.8521, 1.0807]	[-0.7052, 1.0122]
Old-College	(-0.0677, 0.9448)	(-0.0278, 0.9266)	(0.0635, 0.8977)
	[-0.4119, 1.2891]	[-0.3674, 1.2723]	[-0.2770, 1.2424]

Table A.5 : Bounds on Changes in Gender Wage Differential at $25th$ Percentile (1995 - 2007)

	Worst Case	**Quartile Restriction**	**Stochastic Dominance**
Young Non-College	(-0.3083, 0.8612)	(-0.2266, 0.8123)	(-0.1577, 0.7747)
	[-0.3941, 0.9234]	[-0.3115, 0.8738]	[-0.2459, 0.8379]
Young-College	(0.0275, 0.3475)	(0.0527, 0.3349)	(0.0733, 0.3226)
	[-0.0514, 0.4338]	[-0.0255, 0.4202]	[-0.0063, 0.4074]
Old Non-College	(-0.7554, 0.5009)	(-0.6158, 0.4637)	(-0.5202, 0.4166)
	[-0.9360, 0.5927]	[-0.7808, 0.5530]	[-0.6872, 0.5056]
Old-College	(-0.0094, 0.2532)	(-0.0042, 0.2268)	(0.0051, 0.2036)
	[-0.3311, 0.5586]	[-0.3237, 0.5312]	[-0.3165, 0.5098]

Table A.6 : Bounds on Changes in Gender Wage Differential at $25th$ Percentile (2007 - 2018)

	Worst Case	**Quartile Restrictions**	**Stochastic Dominance**
Young Non-College	(-0.6600, 1.2851)	(-0.5504, 1.2101)	(-0.4767, 1.1324)
	[-0.8613, 1.6839]	[-0.7410, 1.6172]	[-0.6660, 1.5416]
Young-College	(-0.3169, 0.3037)	(-0.2719, 0.2745)	(-0.2367, 0.2467)
	[-0.4765, 0.4591]	[-0.4309, 0.4286]	[-0.3970, 0.4011]
Old Non-College	(-1.1004, 1.4991)	(-0.8352, 1.3199)	(-0.6717, 1.1894)
	[-1.3180, 1.6742]	[-1.0434, 1.4771]	[-0.8818, 1.3463]
Old-College	(-0.2373, 0.8705)	(-0.1810, 0.8571)	(-0.0759, 0.8284)
	[-0.6306, 1.2769]	[-0.5661, 1.2619]	[-0.4651, 1.2328]

Table A.7 : Bounds on Changes in Gender Wage Differential at $75th$ Percentile (1995 - 2018)

	Worst Case	**Quartile Restrictions**	**Stochastic Dominance**
Young Non-College	(-0.9838, 0.7089)	(-0.0345, 0.4471)	(-0.0195, 0.4006)
	[-1.1875, 0.9431]	[-0.3839, 0.6326]	[-0.2612, 0.5864]
Young-College	(-0.1580, 0.2887)	(0.0418, 0.1824)	(0.0717, 0.1671)
	[-0.3350, 0.4461]	[-0.1021, 0.3206]	[-0.0722, 0.3064]
Old Non-College	(-0.8550, 1.1503)	(-0.4380, 0.6905)	(-0.0062, 0.6338)
	[-0.9853, 1.4019]	[-0.7229, 0.8818]	[-0.4213, 0.8222]
Old-College	(-0.3852, 0.4004)	(-0.1610, 0.2774)	(-0.1562, 0.2602)
	[-0.5926, 0.6754]	[-0.4157, 0.4984]	[-0.4506, 0.4781]

Table A.8 : Bounds on Changes in Gender Wage Differential at $75th$ Percentile (1995 - 2007)

	Worst Case	**Quartile Restrictions**	**Stochastic Dominance**
Young Non-College	(-0.8616, 0.7468)	(-0.1496, 0.4436)	(-0.1346, 0.3954)
	[-0.9371, 0.8578]	[-0.2299, 0.5375]	[-0.4215, 0.4897]
Young-College	(0.0261, 0.3759)	(0.1672, 0.2975)	(0.1983, 0.2836)
	[-0.0771, 0.4717]	[0.0836, 0.3812]	[0.1150, 0.3668]
Old Non-College	(-0.7978, 0.5870)	(-0.4069, 0.3063)	(-0.3966, 0.2663)
	[-0.8887, 0.7512]	[-0.5204, 0.4487]	[-0.5883, 0.4087]
Old-College	(0.1239, 0.3484)	(0.1621, 0.2428)	(0.1707, 0.2245)
	[-0.0931, 0.5886]	[0.0052, 0.3986]	[0.0108, 0.3828]

Table A.9 : Bounds on Changes in Gender Wage Differential at $75th$ Percentile (2007 - 2018)

	Worst Case	**Quartile Dominance**	**Stochasrtic Dominance**
Young Non-College	(-1.3482, 1.1881)	(-0.3575, 0.4761)	(-0.3343, 0.4547)
	[-1.5674, 1.4269]	[-0.7370, 0.6772]	[-0.5933, 0.7788]
Young-College	(-0.4504, 0.1791)	(-0.2242, -0.0163)	(-0.1923, -0.0508)
	[-0.6259, 0.3459]	[-0.3656, 0.1215]	[-0.3322, 0.0863]
Old Non-College	(-1.2313, 1.7375)	(-0.6675, 1.0205)	(-0.2173, 0.9753)
	[-1.4153, 1.9888]	[-0.9620, 1.2267]	[-0.6526, 1.2343]
Old-College	(-0.6596, 0.2026)	(-0.3736, 0.0851)	(-0.3603, 0.0691)
	[-0.9286, 0.5011]	[-0.6505, 0.3303]	[-0.6693, 0.3116]

Table A.10: Provinces Covered by Each Survey

Survey	Covered Provinces
CHIP 1995	Beijing, Shanxi, Liaoning, Jiangsu,Anhui, Henan, Hubei, Guangdong, Sichuan, Yunan, Gansu
CHIP 2002	Beijing, Shanxi, Liaoning, Jiangsu,Anhui, Henan, Hubei, Guangdong, Chongqing, Yunan, Gansu
CHIP 2007	Shanghai,Jiangsu, Zhejiang, Anhui, Henan, Hubei, Guangdong, Chongqing, Sichuan
CHIP 2013	Beijing, Shanxi, Liaoning, Jiangsu, Anhui, Henan, Hubei, Hunan, Guangdong, Chongqing, Sichuan, Yunan, Gansu
CFPS 2014 **CFPS 2018**	Beijing, Tianjin, Hebei, Shanxi, inner Mongolia, Liaoning, Jilin, Heilongjiang, Shanghai, Jiangsu, Zhejiang, Anhui, Fujian, Jiangxi, Shandong, Henan, Hubei, Hunan, Guangdong, Guangxin, Hainan, Chongqing, Sichuan, Guizhou, Yunan, Shaanxi, Gansu, Ningxia, Xinjiang

Appendix B. Bounds Estimation using Monotone IV (MIV) Assumption

B.1 Monotone Instrumental Variables

Under the exclusion restriction (ER), traditional instrumental variables can help to tighten the bounds in equation (2) (Manski, 1994; Blundell et al., 2007). The literature has used instrumental variables (IVs) to tackle the employment selection, such as an indicator of a young child aged less than six years (Chi and Li, 2014), and the number of young children in the household (Mulligan and Rubinstein, 2008). However, these instrumental variables may not satisfy the ER, which requires that the IV can only affect wages through employment (Angrist et al., 1999). For example, in cases of using the number of young children as the IV, fertility decisions may affect wage and earnings independently of employment status. For example, Bratti (2015) shows that postponing fertility raises women's wages, in which case the number of children may affect earnings independently of employment, violating the ER.

Given that it is hard to find a valid traditional IV for employment that is independent of $F(w|x)$, we instead follow Manski and Pepper (2000) and adopt the following weaker monotone IV (MIV) assumption, which does not require an exclusion restriction condition-to tighten the bounds:

$$F(w|x,z') \leq F(w|x,z), \quad \forall w,x,z,z' \quad with \quad z < z'. \tag{16}$$

Equation (16) assumes that a higher value of the MIV Z will lead to a distribution of wages that first-order stochastically dominates the distribution of wages with lower values of Z. In our application, Z is the average income of the other household members in an individual's household. The rationale of the MIV assumption is predicated on the human capital assortative mating behavior in China (Han, 2010; Nie and Xing, 2019) and the documented inter-generational income persistence in China (Feng et al., 2021; Gong et al., 2010). First, people tend to marry spouses with similar human capital and earning potential (assortative mating). For people with higher-income spouses, their wage distribution would likely first-order stochastically dominate those whose spouses have lower income. Second, inter-generational income persistence may also contribute to the monotone relationship in equation (16). Specifically, if children with higher-income parents are likely to earn more, the wage distribution of workers who live with their high-income parents will stochastically dominate the workers who live with their lower-income parents.

To exploit the MIV restriction, we can find the tightest bounds over the spourt of Z and ehn integreate out Z. Therefore, under the MIV assumption, for a value of $Z = z_1$, we can find the highest lower bound ($F^1(w|x, z_1)$) for the distribution of the wage[12] over $z \geq z_1$ in the support of Z:

$$F(w|x, z_1) \geq F^1(w|x, z_1) \equiv \max_{z \geq z_1}\{F(w|x, z, E = 1)P(x, z)\}. \tag{17}$$

and the lowest upper bound($F^u(w|x, z_1)$) over $z \leq z_1$ in the support of Z:

$$F(w|x, z_1) \leq F^u(w|x, z_1) \equiv \min_{z \leq z_1}\{F(w|x, z, E = 1)P(x, z) + 1 - P(x, z)\}. \tag{18}$$

Regarding the bounds on the wage quantiles, for a value of $Z = z_1$, we have $w^{q(l)}_{miv}(x, z_1) \leq w^q(x, z_1) \leq w^{q(u)}_{miv}(x, z_1)$, where $w^{q(l)}_{miv}(x, z_1)$ and $w^{q(u)}_{miv}(x, z_1)$ respectively solve the following two equations with respect to w,

$$q = F^u(w|x, z_1) \equiv \min_{z \leq z_1}\{F(w|x, z, E = 1)P(x, z) + 1 - P(x, z)\}, \tag{19}$$

and

$$q = F^l(w|x, z_1) \equiv \max_{z \geq z_1}\{F(w|x, z, E = 1)P(x, z)\}. \tag{20}$$

The bounds on $w^q(x)$ can then be constructed by integrating over the distribution of Z given $X = x$, that is,

$$E_Z[w^{q(l)}_{miv}|x] \leq w^q(x) \leq E_Z[w^{q(u)}_{miv}|x]. \tag{21}$$

Our approaches to estimating the gender wage differentials are motivated by the fact that the assumptions needed for point identification are not easy to justify and satisfy in practice. The worst-case bounds do not rely on any assumptions; therefore, bounds derived under other weak assumptions are theoretically narrower than the worst-case bounds. The stochastic dominance and quartile dominance assumption express the notion that workers are likely to be more productive than nonworkers, and we show evidence of this positive selection. Since the quartile dominance assumption is a weaker version of the stochastic dominance assumption, the estimated bounds should be narrower under the stochastic dominance assumption. We also relax the exclusion restriction and use a weaker monotonicity

[12]Please see Appendix B for computation and inference details.

assumption that allows for the positive relationship between wages and the instrument, which is the average of other members' income in the worker's household. Theoretically, the tightest bounds should be under the combination of stochastic assumption and MIV.

Our bounds under the MIV assumption contains maximum or minimum operators (see equations (17)-(20)). Hirano and Porter (2012) show that for bounds that contain maximum or minimum operators, standard inference breaks down, which prevent us from using the confidence intervals in Blundell et al. (2007). To obtain valid confidence regions for the true wage percentile parameters of interest, we estimate these confidence intervals using the method proposed by Chernozhukov et al. (2013). In this section we briefly describe Chernozhukov et al. (2013) as applied to our bounds.

Let the bounds for a parameter θ_0 (e.g., the median wage) be given by $[\theta_0^l, \theta_0^u]$, where $\theta_0^l = \max_{\upsilon \in \mathcal{V}^l = \{1,\ldots,m^l\}} \theta^l(\upsilon)$ and $\theta_0^u = \min_{\upsilon \in \mathcal{V}^u = \{1,\ldots,m^l\}} \theta^u(\upsilon)$. Chernozhukov et al. (2013) calls $\theta^l(\upsilon)$ and $\theta^u(\upsilon)$ bounding functions. We follow Flores and Flores-Lagunes (2013) and let υ index the bounding functions and m^l and m^u be, respectively, the number of terms inside the max and min operators. For example, suppose the wage distribution $F(w_1|x, z_1)$ has two lower bound candidates $\max_{z \geq z_1}\{F(w_1|x, z_1, E = 1)P(x, z_1), F(w_1|x, z_2, E = 1)P(x, z_2)\}$, and we can write $\theta_0^l = \max_{\upsilon \in \mathcal{V}^l = \{1,2\}} \theta^l(\upsilon) = \max\{\theta^l(1), \theta^l(2)\}$, with $\theta^l(1) = F(w_1|x, z_1, E = 1)P(x, z_1)$ and $\theta^l(2) = F(w_1|x, z_2, E = 1)P(x, z_2)$. The sample analog estimators of the bounding functions $\theta^l(\upsilon)$ and $\theta^u(\upsilon)$ are consistent and asymptotically normally distributed, because they are simple functions of proportions.

Chernozhukov et al. (2013) employ precision-corrected estimates of the bounding functions to construct the confidence regions for the bounds $[\theta_0^l, \theta_0^u]$. Specifically, the precision adjustment is done by adding to each estimated bounding function (i.e., each bound candidates) the product of its pointwise standard error and an appropriate critical value, $\kappa(p)$. With different choices of $\kappa(p)$, we may obtain the confidence regions for either the true parameter value or the identified set, and half-median unbiased estimators for the lower and the upper bounds.[13] The bounding function estimates that have higher standard errors receive larger adjustments. For example, the precision-corrected estimator of the lower bound θ_0^l is given by

$$\hat{\theta}^l(p) = \max_{\upsilon \in \mathcal{V}^l} [\hat{\theta}^l(\upsilon) - \kappa^l_{n,\hat{V}^l_n}(p) s^l(\upsilon)], \tag{22}$$

where $\hat{\theta}^l(\upsilon)$ is the sample analog estimator of $\theta^l(\upsilon)$ and $s^l(\upsilon)$ is its standard error. Chernozhukov et al. (2013) compute the critical value $\kappa^l_{n,\hat{V}^l_n}(p)$ based on simulation methods and a preliminary estimator $\hat{V}^l_n = arg\max_{\upsilon \in \mathcal{V}^l} \theta^l(\upsilon)$, and p is determined by the confidence level of choice. Intuitively, $\hat{V}^l_n$ selects those bounding functions that are close enough to binding to affect the asymptotic distribution of the estimator of the lower bound. We obtain the precision-corrected estimator of the upper bound θ_0^u in a similar way. Since the critical value and the standard error in equation (22) are both non-negative, the bias-corrected bounds tend to be wider than the uncorrected ones. Further details on our specific implementation steps are provided in Online Appendix B.2.

[13]The property half-median-unbiasedness means that the lower bound estimator is less than the true value of the lower bound with probability at least one half asymptotically, while the reverse holds for the upper bound (Chernozhukov et al., 2013).

B.2 Other Household Members' Income

For bounds using the monotone instrumental variable (MIV) assumption, the MIV for employment in our analysis is the income of other household members. Specifically, we use the family income minus the person's total income and average other the size of the household minus one as the income from other family members in the household. For individuals without a family, this other member's income would be zero.

CHIP does not report the total household income; therefore, we use the sum of every household member's individual total income as the total household income. In CHIP samples, an individual's total income includes the yearly income, the subsidy from minimum living standard, living hardship subsidies from the work unit, second job, sideline income, and the monetary value of income in kind.

In CFPS, we are able to calculate the total household income directly, i.e., the sum of the household total wage income, operating income, transfer income, property income, and other income. We also construct another measure of total household income by adding up the total income of all household members. In our analysis, we take the larger amount among these two income measures as the household total income measure.[14] Similarly, we also use the larger amount between an individual's total income provided by the survey and the individual's income added up from different sources as the individual's total income in the analysis. In CFPS, the added-up individual income is the sum of wage income from all sources, operating income, subsidies, and bonuses. We assign zero to the other family members' income for individuals who live alone.

B.3 Inference for Bounds under the MIV assumption

Previously, we have briefly described the method in Chernozhukov et al. (2013) to compute confidence regions for bounds with maximum and minimum operators. In Section B.1, we explain the computation of bounds under the MIV assumption, and in this Section, we explain the detailed steps we use to compute the half-median unbiased bounds and the confidence intervals, following the implementation in Flores and Flores-Lagunes (2013).

The Chernozhukov et al. (2013) method requires us to apply the maximum and the minimum operators over all the bound candidates inside the lower bound $\theta^l(v)$ and the upper bound $\theta^u(v)$ bounding functions. This requirement cause a computational challenge for bounds under the monotone instrumental variable (MIV) assumption.

Specifically, under the MIV assumption, the bounds of the wage distribution and the wage quantiles are first constructed conditional on each quantile of the MIV Z. In our application, we used 10 MIV quantiles (i.e., the $5th$, the $15th$, ..., the $95th$ quantile of income from other household members). we would need to integrate these lower bounds and the upper bounds that are conditional on the MIV quantiles over the ten quantiles of the MIV to obtain the lower bounds and the upper bounds in Equation 18. In this scenario, the total number of

[14]Theoretically, the added-up total household income from the household survey should be the same as the added-up total income from all household members from the individual survey. However, when we use the CFPS sample, those two numbers are not always consistent, and there are cases where we have missing values in one of the two. Therefore, we use the larger amount among those two measures as the total household income.

lower and upper bounds candidates for Equation 18 may respectively surpass 3.5 million, which cause a computational challenge for us when implementing the Chernozhukov et al. (2013)

To see this issue in an example, when we compute the half-median unbiased upper bound for $w^q(x)$ in Equation 18, the bounding function of $\theta^u(v)$ contains the upper bound candidates at each of the 10 quantiles of MIV Z. **(1)** Conditional on the first MIV quantile $z = z_{5th}$, there will be 10 bound candidates, i.e., $w^q(x, z = z_{5th})$ that is solved from $q = F(w|x, z_{5th}, E = 1)P(x, z_{5th})$; $w^q(x, z = z_{15th})$ that is solved from $q = F(w|x, z_{15th}, E = 1)P(x, z_{15th})$; $w^q(x, z = z_{25th})$ that is solved from $q = F(w|x, z_{25th}, E = 1)P(x, z_{25th})$; $w^q(x, z = z_{35th})$ that is solved from $q = F(w|x, z_{35th}, E = 1)P(x, z_{35th})$, ..., and $w^q(x, z = z_{95th})$ that is solved from $q = F(w|x, z_{95th}, E = 1)P(x, z_{95th})$. **(2)** Conditional on the second MIV quantile, $z = z_{15th}$, there will be 9 bound candidates, i.e., $w^q(x, z = z_{15th})$ that is solved from $q = F(w|x, z_{15th}, E = 1)P(x, z_{15th})$; $w^q(x, z = z_{25th})$ that is solved from $q = F(w|x, z_{25th}, E = 1)P(x, z_{25th})$; $w^q(x, z = z_{35th})$ that is solved from $q = F(w|x, z_{35th}, E = 1)P(x, z_{35th})$, ..., and $w^q(x, z = z_{95th})$ that is solved from $q = F(w|x, z_{95th}, E = 1)P(x, z_{95th})$. Similarly, conditional on $25th$ quantile of the MIV, $z = z_{25th}$, there will be 8 bound candidates, and so forth for the bounds conditional on the higher MIV quantiles.

Continuing with our example, after obtaining the upper bounds for each $w^q(x, z)$, where $z = z_{5th}, z = z_{15th}, ..., z = z_{95th}$, the bounding function of the upper bound in Equation 19, $E_Z[w^q(u)_{miv}|x]$, includes bound candidates that are made of all possible combinations of the bounds conditional on the 10 MIV quantiles, which are totally $10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 3,628,800$ bound candidates. The large sizes of the matrices that contain the bounds candidates and the variance-covariance matrices of the bounds candidates make the computation time-consuming and not practical for our estimation purpose.

In practice, we first estimate the half-median unbiased MIV bounds and confidence intervals conditional on each of the ten MIV quantiles, with the total number of the bounds candidates not exceeding 10. We then average out the half-median unbiased MIV bounds and confidence interval estimates over the ten MIV quantiles.

B.4 Computation Steps of the Confidence Interval

In this section, we follow Flores and Flores-Lagunes (2013) and describe the detailed steps followed to implement the methodology used by Chernozhukov et al. (2013) to obtain the confidence interval for the true parameter and the half-median unbiased estimators for our lower and upper bounds.

As discussed in the paper, the precision adjustment in Chernozhukov et al. (2013) is done by subtracting or adding to each estimated bounding function (i.e., each bound candidates) the product of its pointwise standard error and an appropriate critical value, $\kappa(p)$. $\kappa(p)$ is selected based on a standardized Gaussian process $Z_n^*(v)$. For any compact set $V \in \mathcal{V}$, Chernozhukov et al. (2013) approximate using simulation the p-th quantile of $sup_{v \in V} Z_n^*(v)$, denoted by $\kappa_{n,V}(p)$, and use it in place of $\kappa(p)$. Since setting $V = \mathcal{V}^l$ for the lower bound leads to asymptotically valid but conservative inference, Chernozhukov et al. (2013) propose a preliminary set estimator $\hat{V}_n^l$ of $V_0^l = \arg max_{v \in \mathcal{V}^l} \theta^l(v)$ that they refer to an adaptive inequality selector. This preliminary set estimator $\hat{V}_n^l$ selects those bounding functions that are close enough to binding to affect the asymptotic distribution of the estimator of the lower bound. For the same reason, a preliminary set estimator $\hat{V}_n^u$ of $V_0^u = \arg min_{v \in \mathcal{V}^u} \theta^u(v)$ is used for the upper bound. The precision-corrected estimator of the lower bound θ_0^l is

$$\hat{\theta}^l(p) = \max_{v \in \mathcal{V}^l} [\hat{\theta}^l(v) - \kappa^l_{n,\hat{V}_n^l}(p) s^l(v)], \tag{23}$$

where $\hat{\theta}^l(v)$ is the sample analog estimator of $\theta^l(v)$ and $s^l(v)$ is its standard error.

Let $\gamma_n = [\theta_n^l(1), ..., \theta_n^l(m^l)]'$ be the vector of bounding functions and let $\hat{\gamma}_n$ be its sample analog estimator. The steps we follow to compute the set estimator $\hat{V}_n^l$ and the critical value $\kappa^l_{n,\hat{V}_n^l}(p)$ in Equation 1 are as follows.

(1) We obtain by bootstrapping a consistent estimate $\hat{\Omega}_n$ of the asymptotic variance of $\sqrt{n}(\hat{\gamma}_n - \gamma_n)$. Let $\hat{g}_n(v)'$ denote the v^{th} row $\hat{\Omega}_n^{1/2}$ and let $s_n^l(v) = \|\hat{g}_n(v)\| / \sqrt{n}$.

(2) We estimate R draws from $\mathcal{N}(0, I_{m^l})$, denoted $Z_1, ..., Z_R$, where I_{m^l} is the $m^l \times m^l$ identity matrix, and we calculate $Z_r^*(v) = \hat{g}_n(v)' Z_r / \|\hat{g}_n(v)\|$ for $r = 1, ..., R$.

(3) Let $Q_p(X)$ denote the p-th quantile of a random variable X and, following CLR, let $c_n = 1 - (.1/\log n)$. We compute $\kappa^l_{n,\mathcal{V}^l}(c_n) = Q_{c_n}(\max_{v \in \mathcal{V}^l} Z_r^*(v), r = 1, ..., R)$; that is, for each replication r we calculate the maximum of $Z_r^*(1), ..., Z_r^*(m^l)$ and take the c-th quantile of those R values. We then use $\kappa^l_{n,\mathcal{V}^l}(c_n)$ to compute $\hat{V}_n^l = \{v \in \mathcal{V}^l : \hat{\theta}^l(v) \geq max_{\tilde{v} \in \mathcal{V}^l} \{[\hat{\theta}^l(\tilde{v}) - \kappa^l_{n,\mathcal{V}^l(c_n)} s_n^l(\tilde{v})] - 2\kappa^l_{n,\mathcal{V}^l(c_n)} s_n^l(\tilde{v})\}\}$.

(4) We compute $\kappa^l_{n,\hat{V}_n^l}(p) = Q_p(max_{v \in \hat{V}_n^l} Z_r^*(v), r = 1, ..., R)$, so the critical value is based on $\hat{V}_n^l$ instead of $\mathcal{V}^l$.

The precision-corrected estimator of the upper bound θ_0^u is given by

$$\hat{\theta}^u(p) = \min_{v \in \mathcal{V}^l} [\hat{\theta}^u(v) + \kappa^u_{n,\hat{V}_n^l}(p) s^u(v)], \tag{24}$$

where $\hat{\theta}^u(v)$ is the sample analog estimator of $\theta^u(v)$ and $s^u(v)$ is its standard error. To compute $\kappa^u_{n,\hat{V}_n^l}(p)$ in (2), we follow the same steps above but in step (3) we replace $\hat{V}_n^l$ by $\hat{V}_n^u = \{v \in \mathcal{V}^u : \hat{\theta}^u(v) \geq \min_{\tilde{v} \in \mathcal{V}^u} [\hat{\theta}^u(\tilde{v}) + \kappa^u_{n,\mathcal{V}^u}(c_n) s_n^u(\tilde{v})] + 2\kappa^u_{n,\mathcal{V}^u}(c_n) s_n^u(v)\}$. Since the normal distribution is symmetric, we don't have to make any changes when computing the quantiles in step 3 and 4.

Half-median-unbiased estimators of the upper and lower bounds are obtained by setting $p = 1/2$ in the steps above and using Equations (1) and (2) to compute, respectively, $\hat{\theta}^l(1/2)$ and $\hat{\theta}^u(1/2)$. To construct confidence intervals for the parameter θ_0, it is important to take into account the length of the identified set. Following Chernozhukov et al. (2013) and Flores and Flores-Lagunes (2013), let $\hat{\Gamma}_n = \hat{\theta}^u_n(1/2) - \hat{\theta}^l_n(1/2)$, $\hat{\Gamma}^+_n = \max(0, \hat{\Gamma}_n)$, $\rho_n = \max\{\hat{\theta}^u_n(3/4) - \hat{\theta}^u_n(1/4), \hat{\theta}^l_n(1/4) - \hat{\theta}^l_n(3/4)\}$, $\tau_n = 1/(\rho_n \log n)$ and $\hat{p}_n = 1 - \Phi(\tau_n \hat{\Gamma}^+_n)\alpha$, where $\Phi(.)$ is the standard normal CDF. Note that $\hat{p}_n \in [1-\alpha, 1-\alpha/2]$, with $\hat{p}_n$ approaching $1-\alpha$ when $\hat{\Gamma}_n$ grows large relative to sampling error and $\hat{p}_n = 1 - \alpha/2$ when $\hat{\Gamma}_n = 0$. An asymptotically valid confidence interval at the confidence level of $1-\alpha$ is given by $[\hat{\theta}^l_n(\hat{p}_n), \hat{\theta}^u_n(\hat{p}_n)]$.

B.5 Results: Changes in the Gender Wage Gap using the MIV Bounds

Figure B.1 presents the estimated results using the monotone instrumental variable for changes the median gender wage gap, the changes in the gender wage gap at the $25th$, and the $75th$ percentile.[15] Compared to the bounds under quartile and stochastic dominance restrictions(Figure 2.5 - 2.7), the MIV bounds are considerably wide for all the age-education groups in all the considered study periods. For example, the lower bounds indicate 0.10 - 1.42 log points of decrease in the gender wage gap and the upper bounds indicate 0.23 - 1.38 log points of increase in the gender wage gap from 1995 to 2018.[16] All MIV bounds estimates at the median include a zero change except for the young college graduates from the year 1995 to 2007(Panel B), where the estimated bounds show a statistically significant increase in the gender wage gap of 0.19 - 0.63 log points. Additionally, for the same group, the estimated MIV bounds also suggest a decrease in the median gender wage gap of 0.12 to 0.59 log points after 2007; however, the 95% CIs include zero change. The estimated MIV bounds do not provide any inclusive suggestions about the changes in the gender wage gap at the $25th$ wage percentile (Figure B.2). As of the changes in the gender wage gap at the $75th$ wage percentile, the estimated MIV bounds in Figure B.3 Panel B suggest a statistically significant increase in the gender wage gap for young college graduates by 0.21 - 0.62 log points during 1995 - 2007.

[15] Table B.1 - B.3 reports the values for the upper and lower bounds and the corresponding 95% confidence intervals (CIs) of the bounds shown in Figure B.1 - B.3, respectively.

[16] Bounds under the MIV assumption tend to be wider than those under the quartile dominance assumption and sometimes the worst-case bounds. It may be attributed to the computation procedure explained in Appendix B.1. In brief, due to a computational constraint, we needed first to compute bounds under the MIV assumption in each sub-sample conditional on the ten quantiles of the MIV (the $5th$, the $15th$, ..., and the $95th$ quantiles), and then obtain the average of the ten bounds to obtain the bounds for each education and age group. We are in the process of improving the efficiency in the computation of these bounds.

Figure B.1: Changes in Median Gender Wage Gap using the MIV Bounds for Different Groups

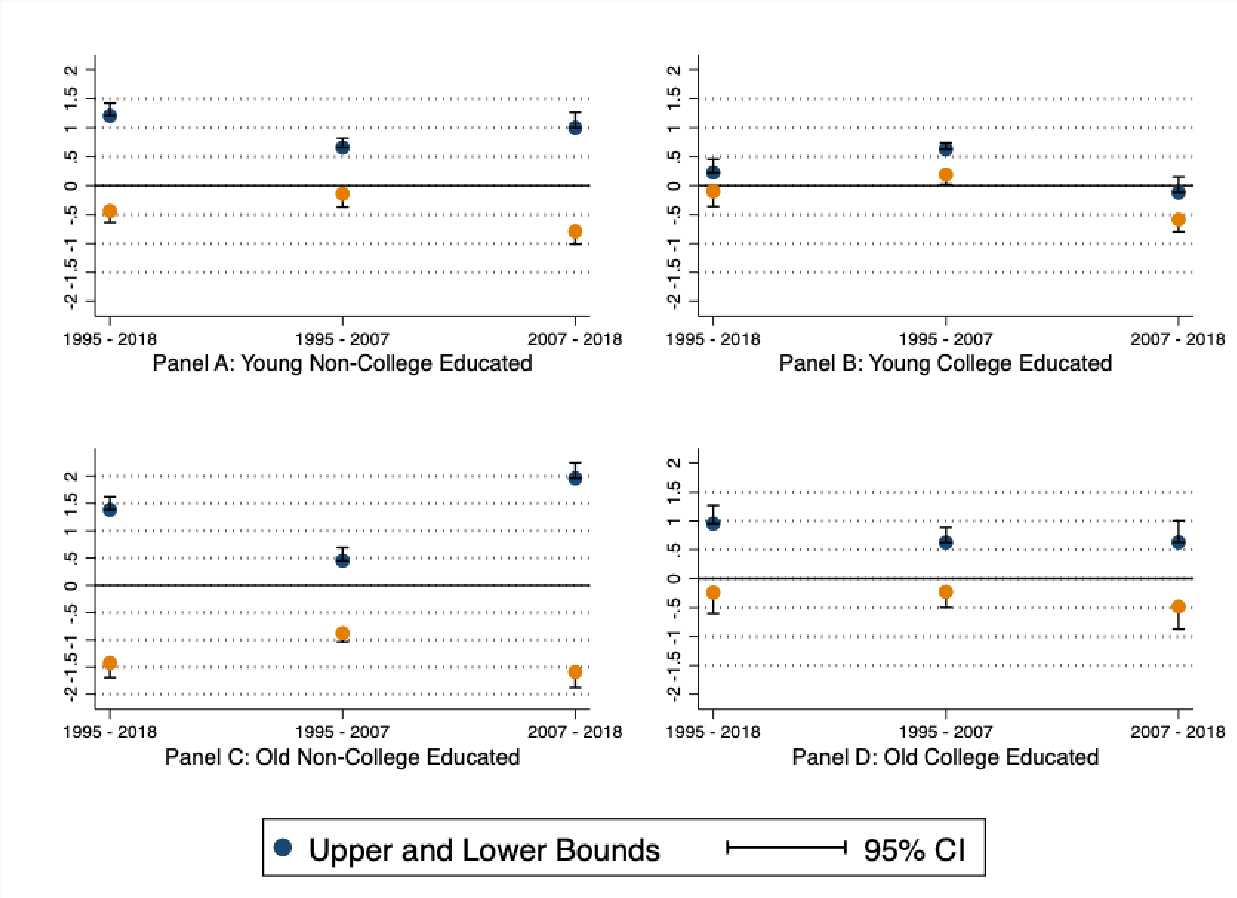


Figure B.2: Changes in Gender Wage Gap using the MIV Bounds for Different Groups at 25*th* Percentile

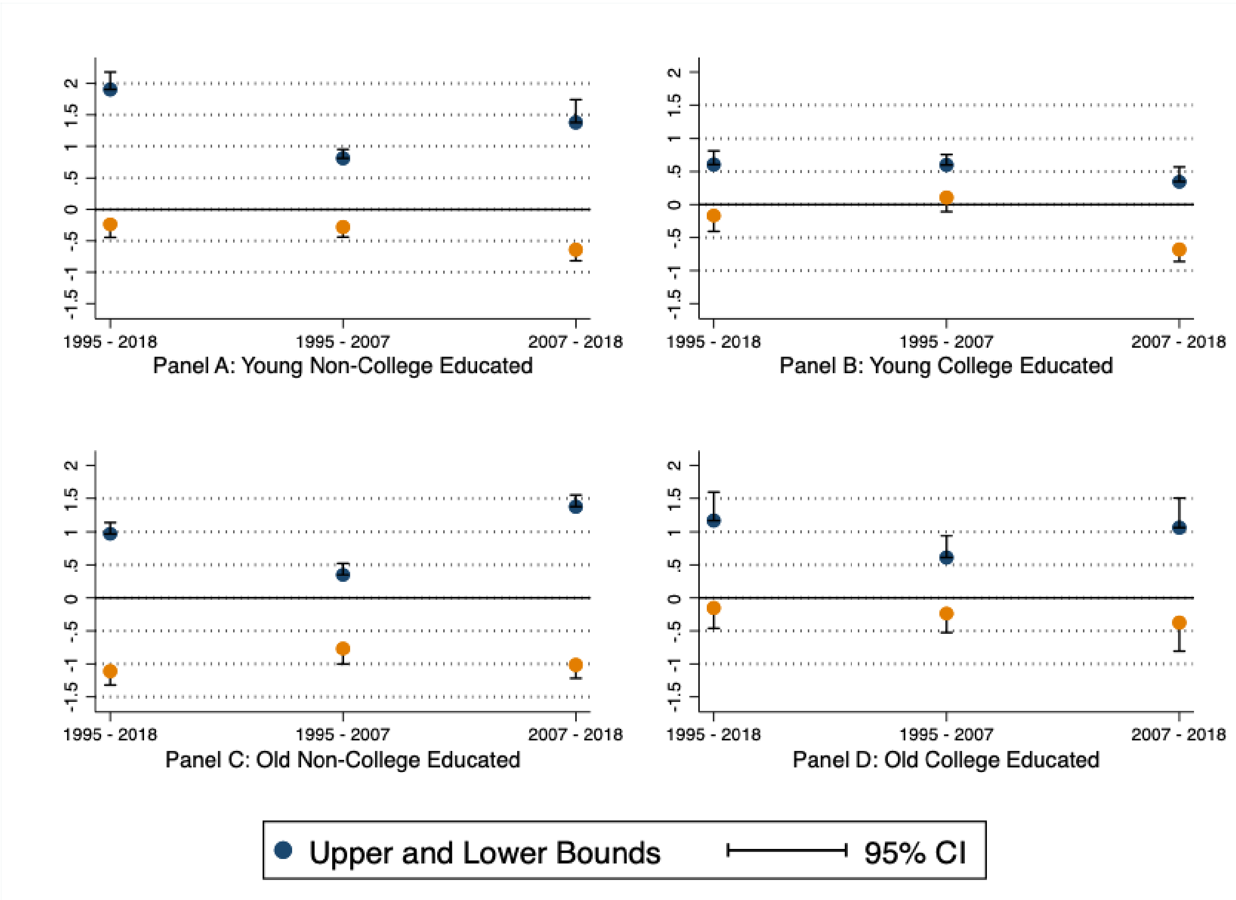
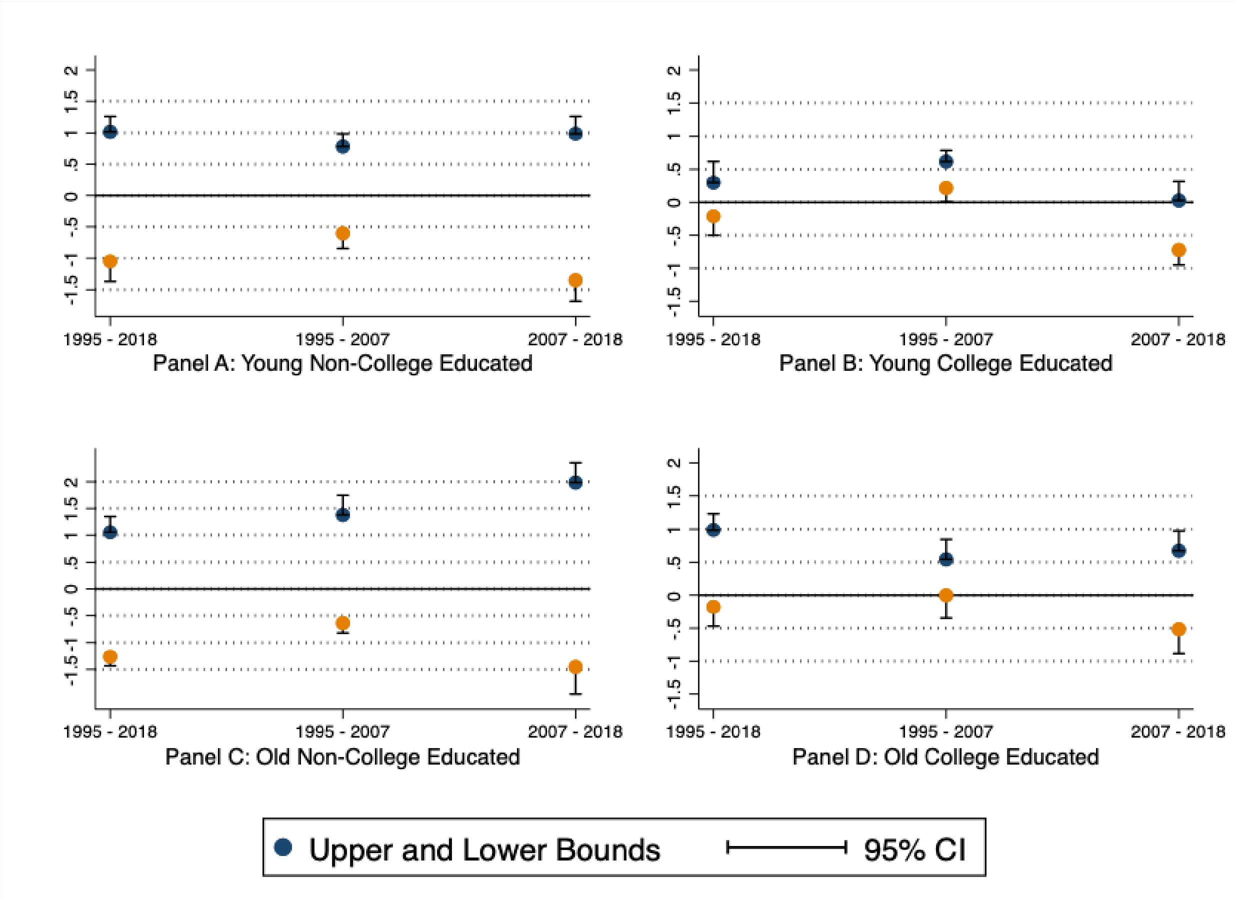


Figure B.3: Changes in Gender Wage Gap using the MIV Bounds for Different Groups at $75th$ Percentile



B.1 Changes in Median Gender Wage Gap using the MIV Bounds for Different Groups

	1995 - 2018	**1995 - 2007**	**2007 - 2018**
Young Non-College	(-0.4363, 1.2050)	(-0.1410, 0.6568)	(-0.7906, 0.9989)
	[-0.6382, 1.4219]	[-0.3695, 0.8190]	[-1.0104, 1.2641]
Young-College	(-0.0988, 0.2270)	(0.1914, 0.6283)	(-0.5864, -0.1206)
	[-0.3595, 0.4579]	[0.0188, 0.7363]	[-0.7991, 0.1534]
Old Non-College	(-1.4235, 1.3831)	(-0.8779, 0.4525)	(-1.5870, 1.9620)
	[-1.6875, 1.6209]	[-1.0366, 0.6917]	[-1.8733, 2.2426]
Old-College	(-0.2389, 0.9548)	(-0.2255, 0.6236)	(-0.4810, 0.6281)
	[-0.6021, 1.2656]	[-0.4937, 0.8873]	[-0.8719, 1.0035]

B.2 Changes in Gender Wage Gap using the MIV Bounds for Different Groups at 25th Percentile

	1995 - 2018	**1995 - 2007**	**2007 - 2018**
Young Non-College	(-0.2383, 1.9061)	(-0.2834, 0.8105)	(-0.6384, 1.3798)
	[-0.4464, 2.1788]	[-0.4415, 0.9504]	[-0.8204, 1.7461]
Young-College	(-0.1677, 0.6040)	(0.1050, 0.5982)	(-0.6773, 0.3430)
	[-0.4073, 0.8086]	[-0.1058, 0.7543]	[-0.8646, 0.5683]
Old Non-College	(-1.1090, 0.9674)	(-0.7669, 0.3469)	(-1.0152, 1.3758)
	[-1.3158, 1.1387]	[-1.0022, 0.5211]	[-1.2153, 1.5565]
Old-College	(-0.1508, 1.1682)	(-0.2364, 0.6094)	(-0.3765, 1.0626)
	[-0.4622, 1.6009]	[-0.5254, 0.9379]	[-0.8031, 1.5044]

B.3 Changes in Gender Wage Gap using the MIV Bounds for Different Groups at 75th Percentile

	1995 - 2018	**1995 - 2007**	**2007 - 2018**
Young Non-College	(-1.0494, 1.0169)	(-0.6036, 0.7791)	(-1.3497, 0.9854)
	[-1.3701, 1.2590]	[-0.8429, 0.9855]	[-1.6889, 1.2598]
Young-College	(-0.2114, 0.2965)	(0.2143, 0.6158)	(-0.7191, 0.0276)
	[-0.4995, 0.6178]	[0.0077, 0.7828]	[-0.9513, 0.3127]
Old Non-College	(-1.2634, 1.0571)	(-0.6384, 1.3798)	(-1.4526, 1.9861)
	[-1.4339, 1.3512]	[-0.8204, 1.7461]	[-1.9638, 2.3489]
Old-College	(-0.1763, 0.9858)	(-0.0020, 0.5429)	(-0.5155, 0.6714)
	[-0.4697, 1.2311]	[-0.3475, 0.8417]	[-0.8867, 0.9756]