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‡

December 14, 2022

(For the latest version please click here)

Abstract

This paper aims to examine 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, we employ the nonparametric bounds. To also account for the labor supply's intensive margin, we compute workers' working hours and hourly wage using available information in CHIP and CFPS. Our methodology adopts a weak quartile dominance assumption, a monotone instrumental variable, 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 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 survey period between 1995-2007, we find a significant increase by 0.19 - 0.63 log points in the **median** gender wage gap among the young workers who are college-educated. In 2007 - 2018, the bounds estimates are less conclusive and imply a decrease in the **median** gender wage gap among the college-educated young workers by 0.12 - 0.59 log points, while the 95% CI does not exclude zero change. The estimates of the gender wage gap at the $\mathbf{75^{th}}$ wage percentile show a similar pattern as the changes at the median wage, with the statistical implications at the $\mathbf{25^{th}}$ percentile inconclusive.

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

JEL classification: J3, C31, C36

* Department of Economics, Syracuse University, Syracuse, NY; Telephone: +1 315-443-3843; Email: rxu101@syr.edu.

† 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 the gender wage gap changes 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). Chi and Li (2014) find that the average gender earnings gap has increased from 1988 to 2009; estimates from Heckman's selection-correction model, which accounts for selection into employment, suggest an overall smaller gap than the raw observed gender earnings wage gap. In more recent years, Song et al. (2019) record a temporary narrowing in the gender earnings gap from 2007 to 2013.

While the existing literature has mostly focused on measuring the average gender earnings gaps conditional on employment, this study aims to re-examine changes in the gender wage differentials at the median, the 25^{th} and the 75^{th} wage quantiles in China from 1995-2018, while effectively accounting for changes in employment 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 to date, 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' average 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 drifted 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, and 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 taking effect. In addition, the estimated gender wage gap may be biased due to changes in labor participation over the years. For example, some highly-educated and likely high-wage women might be deterred by discrimination in the labor market based on their child-bearing demand. If high-wage women are increasingly exiting the labor market, the observed gender wage gap may be inflated.

In the literature of 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 et al., 2021), 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" that is not representative of the total 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 does not require exclusion restriction assumptions, although sometimes it may lead to imprecise and non-informative implications.

To account for differences in labor force participation, we use bounds introduced by Manski (1994), Manski and Pepper (2000), and Blundell et al. (2007). We start with the worst-case bounds of 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 quartile wages (wages at the $25th$, $50th$, $75th$ percentile) of the non-working population not be higher than the quartile wages of the working population. We also employ a stronger version of the dominance assumption – the stochastic dominance assumption, which requires the wage distribution of the working population stochastically dominates the non-working population's. 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 of those who are continuously employed is higher than the residual wage of those who have non-working spells across all percentiles, except for three incidences – the 90th percentile for males over 45, the 90th and the 95th percentiles for females under 45. Besides the above exceptions at very high wage percentiles, evidence from the residual wage is in line with our quartile and stochastic dominance assumptions. The third assumption employs the income of other household members as a monotonic instrumental variable (MIV) for the wage of individuals which is different from the traditional instrumental variable. Specifically, we assume that a higher value of other household members' income leads to a distribution of wage for individuals that first-order stochastically dominates the distribution of wages of individuals with lower values of other household members' income. A theoretical justification of this assumption rests on the notion of assortative mating (Becker (1973); Nie and Xing (2019)) and inter-generational income persistence (Feng et al., 2021; Gong et al., 2010).

After controlling for labor force participation and the hours worked, our bounds estimates show strong evidence of an increase in the gender wage gap in 1995-2007. 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 significant increase of the gender wage gap for the young college-educated at the median wage by at least 0.19 log points, and at the 75th percentile by at least 0.21 log points; the bounds at the 25th percentile for the young college graduates also suggest an increase in the gender wage gap by at least 0.11 log points, however the 95% confidence interval (CI) does not exclude a zero change. The estimates in 2007-2018 do not exclude a zero change for all age and education groups; two exceptions are that the bounds at the median wage suggest an at least 0.12 log points decrease in the gender wage gap of the young college graduates, and at the 75th wage percentile a 0.05 log points decrease for the same group, while the 95% confidence intervals (CIs) does not exclude a zero change.

The main contributions of this paper are in four aspects. First, to the best of our 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, we harmonize two different 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 use the measure of the hourly wage. In this way, by using hourly wages and bounds, we provide statistical evidence of changes in the gender wage gap avoiding biases due to labor supply's intensive (hours worked) and extensive margins (employed v.s. unemployed), respectively. Third, in addition to the median gender wage gap, we provide additional information on the gender wage gap dynamics in China at the 25^{th} and 75^{th} percentiles of the wage distribution, which provides a fuller picture of both the lower side and the upper sides of the wage distribution. Fourth, we 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-corrected and obtain asymptotically valid confidence intervals for these bounds.

2 Bounds on the Wage Distribution

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

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

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 data. To partially identify the wage distribution of the unemployed, $F(w|x, E = 0)$, we construct informative bounds for $F(w|x, E = 0)$ 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

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

into equation (1) to bound the wage cumulative distribution function of the total population 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 following Blundell et al. (2007). 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). In addition, the worst-case bounds are likely to be too wide to be informative. Therefore, we impose restrictions on the wage distribution to obtain narrower bounds for the wage distribution of all quantiles.

2.2 Stochastic Dominance and Quartile Dominance

The standard labor supply model suggests that when the substitution effect of wage dominates its income effect, individuals that command higher wages will be more likely to work, ceteris paribus (Blundell and MaCurdy, 1999). Thus, following Blundell et al. (2007), we impose the stochastic dominance assumption on the wage distribution of the non-workers. That is, we assume that conditional on $X = x$, the wages of those observed working first-order stochastically dominates those of the non-workers. This assumption is based on the notion that workers are more productive than non-workers; therefore, at each percentile, 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 reservation wage, denoted by $w - w^R$ should be positively correlated with w. One justification of the positive correction between $w - w^R$ and w is that 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).

In the recent decades of China's labor market, the increase in the non-working population is most heavily fallen on the unskilled workers (e.g., Feng et al., 2017; Gustafsson and Ding, 2011), which in turn relates employment to workers with relatively higher human capital. In addition, Li et al. (2016) show that the college premiums from 1990-2000 in China have been increased. Li et al. (2017) predict that with investment in physical capital and skill-biased technological changes, the return to human capital in China will continue to increase. If populations with more human capital are more likely to be employed and paid more, this increase in return to human capital in China makes 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 becomes

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

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). Considering the possibilities that positive labor selection may not be satisfied, we employ a weaker restriction - a quartile dominance assumption, which is similar to the median dominance assumption in Blundell et al. (2007). This assumption restricts that the 25th, 50th, and the 75th wage quantiles offered for those not working is not higher than the corresponding wage quantiles of the observed wage. This assumption implies the following bounds for the distribution of wage of the unemployed.

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

Under the 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. The bounds for the wage distribution are:

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

In the set of bounds of equation (10), the bounds for $w^{25(E=1)}(x) \leq w < w^{50(E=1)}(x)$ 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 $w^{50(E=1)}(x) \leq w < w^{75(E=1)}(x)$ and $w \geq w^{75(E=1)}(x)$ 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 $w_m^{q(l)}(x) \leq w^q(x) \leq w_m^{q(u)}(x)$, where $w_m^{q(l)}(x)$ and $w_m^{q(u)}(x)$ 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

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

We find empirical evidence in our data that supports the stochastic and quartile dominance assumptions. In Figure 1, we compare 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 China Family Panel Studies (CFPS), 2014 and 2018. The residual wages are obtained in a regression controlling for the age and the quadratic of age, whether obtained a college degree while controlling for province and survey year dummies. If the wage percentiles of workers without unemployment spell are higher than workers with unemployment spells, we consider it is in line with the positive selection into employment. 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, the 90th and the 95th percentiles for females under 45. The above exceptions at very high wage percentiles indicate 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 support the weaker quartile dominance assumption in Figure 1 in all samples.

2.3 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. 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 depend on an exclusion restriction condition- to tighten the bounds:

$$F(w|x, z') \leq F(w|x, z), \quad \forall w, x, z, z' \quad with \quad z < z'. \tag{13}$$

Equation (13) assumes that a higher value of the instrument Z will lead to a distribution of wages that first-order stochastically dominates the distribution of wages with lower values of Z. This MIV Z in our application is the average income from the other household members. The rationale of the MIV assumption is based on the human capital assortative mating behavior in China (Han, 2010; Nie and Xing, 2019) and the inter-generational income persistence (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 should first-order stochastically dominate those whose spouses have lower income. Second, inter-generational income persistence may also contribute to the monotone relationship in equation (13). 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.

Under the MIV assumption, for a value of $Z = z_1$, we can find the best lower bound to be the largest lower bound over $z \geq z_1$ in the support of Z for the distribution of the wage:[1]

$$F(w|x, z_1) \geq F^1(w|x, z_1) \equiv \max_{z \geq z_1}\{F(w|x, z, E = 1)P(x, z)\}. \tag{14}$$

and the best upper bound is the smallest upper bound over $z \leq z_1$ in the support of Z:

$$F(w|x, z_1) \leq F^u(w|x, z_1) \equiv \min_{z \leq z_1}\{F(w|x, z, E = 1)P(x, z) + 1 - P(x, z)\}. \tag{15}$$

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,

$$q = F^u(w|x, z_1) \equiv \min_{z \leq z_1}\{F(w|x, z, E = 1)P(x, z) + 1 - P(x, z)\}, \tag{16}$$

and

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

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

$$E_Z[w^{q(l)}_{miv}|x] \leq w^q(x) \leq E_Z[w^{q(u)}_{miv}|x]. \tag{18}$$

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

2.4 Gender Wage Gap Bounds and Changes over Time

Our goal is to estimate 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 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:[2]

$$w^{q(l)}(male, x, t) - w^{q(u)}(female, x, t) \leq D_t^q(x) \leq w^{q(u)}(male, x, t) - w^{q(l)}(female, x, t). \quad (19)$$

Similarly, the lower bound of the change in the gender wage gap from year t to year s, $\Delta D_{st}^{q(l)}$, where $s > t$, is given by,

$$\{w^{q(l)}(male, x, s) - w^{q(u)}(female, x, s)\} - \{w^{q(u)}(male, x, t) - w^{q(l)}(female, x, t)\}, \quad (20)$$

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

$$\{w^{q(u)}(male, x, s) - w^{q(l)}(female, x, s)\} - \{w^{q(l)}(male, x, t) - w^{q(u)}(female, x, t)\}. \quad (21)$$

3 Estimation and Inference

Our bounds under the MIV assumption contains maximum or minimum operators (see equations (14)-(17)). 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

[2] These bounds can be under different combinations of assumptions in Section 2.2 and 2.3.

wage percentile parameters of interest, we estimate these confidence intervals using a 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 wage distribution 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.[3] 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. Cher-

[3]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).

nozhukov et al. (2013) compute the critical value $\kappa^{l}_{n,\hat{V}^{l}_{n}}(p)$ based on simulation methods and a preliminary estimator $\hat{V}^{l}_{n} = arg\max_{v \in \mathcal{V}^{l}} \theta^{l}(v)$, 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^{u}_{0} 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 Appendix B.

4 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) 1995, 2002, 2007, 2013 and the China Family Panel Study (CFPS) 2014, 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 provides an introduction to 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.

4.1 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.

4.2 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 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 have 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 ratio is only 63.25% for individuals aged 25 −

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

[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.

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 even 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 its data not usable to us as we would not measure earnings and hours worked accurately. Therefore, we use data from CHIP 1995, 2001,2007,2013 together with CFPS 2012 and 2016 to construct our sample. Our sample includes Chinese urban residents aged 25 to 55 with an urban hukou who do not work in the agriculture sector. A more detailed summary of our sample is in Section 4.4.

4.3 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) that use both CFPS and CHIP 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.

4.3.1 Hourly Wage

Earnings in our analysis measure an accounting period of one year, including regular wages, overtime compensation, allowances, and bonuses, which is the same definition as 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 with 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 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. Figure 2 presents the observed log hourly wage estimates and the observed log wage gender gap at the median (labeled by triangles). Even though the overall trend for the observed median hourly wage gender gap seems flat, there is a slight increase in the gap before the year 2007, and after 2007 the gap appears to reach a plateau and shows a decreasing trend.

4.3.2 Other Household Members' Income

For bounds using the monotone instrumental variable (MIV) assumption, the MIV for employment in our analysis is the income from other household members. Specifically, we use the average family income minus the person's total income as the income from other family members for those living with a family.

CHIP does not report the total household income in the survey; therefore, we use all individuals' 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 in the household survey, i.e., the sum of the household total wage income, operating income, transfer income, property income, and other income. We also construct another measure for total household income by adding up a household's members' total income. In our analysis, we take the larger amount among these two income measures as the household total income measure./footnoteTheoretically, 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 added-up 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 added-up measures like the total household income. 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. The other family members' income is likely to correlate positively with yearly wage income due to assortative mating and inter-generational income transmissions, as explained in Section 2.

4.4 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 the 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 the observed individual domestic characteristics x, $F(w|x, E = 1)$. 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 (equivalent to an associate degree in the U.S.) 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 employment probability for males around 50 and females around 45. Note that the Statutory Retirement Age is 60 for males and 55 for females in China.

Figure 4 illustrates that the change in employment has been heavily skill-and-gender-biased. The employment gap between college-educated and non-college-educated females 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 selection in the 2010s.

5 Results

5.1 Median Wage Differential Change

This section presents the results of estimated bounds on the changes in the median gender wage gap in China after imposing different restrictions. Figure 5 shows the results

[6] We do not use finer age and education groups because constructing bounds of the wage distribution requires the number of the observations ideally no smaller than 100.

for changes from 1995 to 2018.[7] In each figure, the space between two dots represent the bounds to the change in the differential between 1995 to 2018. The thin outer lines denote the 95% confidence intervals for the change in the gender wage gap. The worst-case bounds to the change in the differentials (Figure 5 Panel A) all include a zero change. The wide worst-case bounds are partially due to the lower employment rates for females, as shown in Figure 4, particularly for those without a college education. To narrow the worst-case bounds, we impose the quartile and stochastic dominance restrictions and utilize the MIV assumption. With the quartile dominance restriction alone (Figure 5 Panel B), except for the non-college above 45-years-old sample, all samples' bounds indicate an increase in the gender wage differential by at least 0.03 - 0.10 log points and by at most 0.21 - 0.65 log points, although none of the 95% confidence intervals (CIs) excludes zero. Using the stronger stochastic dominance assumption (Figure 5 Panel C), the bounds are tighter across the board than those with only the quartile restrictions. Under the stochastic dominance assumption, the bounds of the young non-college indicate a potential increase of the gender wage gap by 0.17 log points to 0.62 log points, with the 95% CI excluding zero; for the young college graduates, the bounds indicate an increase in the gap of 0.05 - 0.20 log points, however, the 95% CI includes zero. For the older workers, the bounds for those without a college degree include zero change, suggesting a potential 0.07 log points decrease and a 1.06 log points increase; the bounds for older workers with a college degree suggest an increase in the gender wage gap by 0.12 - 0.47 log points, while the 95% CI does not exclude zero. In Panel D, the MIV bounds tend to be wide and not informative, where the lower bounds indicate 0.10 - 1.42 log points of decreases in the gender wage gap, the upper bounds indicate 0.23 - 1.38 log points of increase in the gender wage gap. [8].

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

[8] In the current set of the results, bounds under the MIV assumption tend to be wider than those under the median 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 to first compute bounds under the MIV assumption in each sub-sample conditional on the ten quantiles of the MIV (the 5^{th}, the 15^{th}, ..., and the 95^{th} 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

To identify any changes in the trend of the gender wage gap, we split our study period into 1995 – 2007 and 2007 – 2018. The break in 2007 is from the consideration that Song et al. (2019) find a temporary narrowing in the gender wage from 2007 to 2013. Figure 6 presents the changes from 1995 to 2007.[9] 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, with the 95% CI excluding zero (Figure 6 Panel A). Since worst-case bounds do not utilize any restrictions on wage distribution, we consider this a strong indication of an increase in the gender wage gap. The 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 zero change. The worst-case bounds for the young and older non-college graduates do not exclude zero. In Panel B, the bounds under the quartile restriction follow the pattern of the worst-case bounds, with tighter bounds for the young college graduates showing an increase of the gender wage gap by 0.13 - 0.28 log points, with the 95% CI excluding zero; an increase by 0.11 - 0.20 log points for the old college graduates, with the 95% CI not excluding zero. The stochastic bounds in Panel C are the narrowest, showing an increase in the gender wage gap for the young non-college graduates by 0.04 - 0.48 log points, for the young college graduates by 0.15 - 0.27 log points, and the old college graduates by 0.12 - 0.20 log points, while only the CI for the young college graduates excludes zero; the bounds for the old non-college graduates do not exclude zero change, showing a potential decrease by 0.22 log points and a potential increase by 0.41 log points. In Panel D, all MIV bounds include zero change except for the young college graduates, where the bounds show an increase in the gender wage gap of 0.19 - 0.63 log point and the 95% CI exclude zero change.

Figure 7 presents the bounds of the change in the median gender wage differential from 2007 to 2018. The worst-case bounds in Panel A, the bounds under the quartile dominance restriction in Panel B, and the bounds under the stochastic dominance restriction in Panel C

computation of these bounds.

[9] Table A.2 and Table A.3 in the appendix report the corresponding values for the upper and lower bounds and the 95% CIs.

include zero for every sample. Under the MIV restriction, the bounds for the young college indicate a decrease in the gender wage gap between 0.12 - 0.59 log points (Figure 7 Panel D), but the 95% CI does not exclude zero; the estimated bounds for the other age and education groups using the MIV all include zero changes of the gender wage gap from 2007 - 2018.

5.2 25^{th} Percentile Wage Differential Change

Figure 8 to Figure 10 present the gender wage gap changes over time at the 25^{th} wage quantiles.[10] Except for some bounds of the college-graduates in 1995-2018 and 1995-2007, the bounds indicate inconclusive changes in the gender wage gap among all age and education groups and two different periods.

Figure 8 shows the change from 1995 to 2018. From the figure, none of the estimated bounds excludes zero change based on the 95% CI. The narrowest bounds are those under the stochastic dominance assumption(Panel C). From the left to the right, the bounds for the young non-college graduates indicate an increase in the gender wage gap by 0.04 - 1.23 log points; the bounds for the older non-college graduates rule out a decrease in the gap greater than 0.47 log points, and an increase greater than 0.89 log points; the bounds for the young college-graduates suggest an increase in the gap by 0.04 - 0.36 log point; the bounds for the older college-graduates suggest an increase in the gap by 0.06 - 0.90 log points.

Splitting up the study period into 1995-2007 and 2007-2018, Figure 9 presents the estimated bounds of the gender wage gap change between 1995-2007. For the young college-graduates group, the bounds suggest similar implications as with the gender wage gap at the median wage. From the worst-case bounds to bounds under different restrictions, all bounds suggest an increase in the gender wage gap by 0.03 - 0.60 log points. The bounds for the old college graduates under the stochastic dominance indicate an increase in the gap by 0.01 - 0.20 log points. The bounds for the other education and age groups include zero.

Figure 10 presents the bounds for the change in the gender wage gap in 2007-2018. All

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

bounds estimations are inconclusive for the sign on the change. The tightest bounds are under the stochastic dominance restriction, where the lower bounds indicate a decrease in the gender wage gap by 0.08 - 0.67 log points, and the upper bounds indicate an increase in the gender wage gap by 0.24 - 1.19 log points.

5.3 75^{th} Percentile Wage Differential Change

Figure 11 to Figure 13 present the gender wage gap change at the 75^{th} wage percentile.[11] Figure 11 shows the change from 1995 to 2018. The bounds under the quartile restriction (Panel B) and the bounds under the stochastic dominance (Panel C) show an increase in the gender wage gap for the young college graduates by 0.04 - 0.18 log points and 0.07 - 0.17 log points respectively, while none of the 95% CIs excludes a zero change. After we split up the total study period, in Figure 12 for 1995 - 2007, the bounds estimates show a consistent increase in the gender wage gap for those who have attended college. The worst-case bounds of the young-college suggest a 0.03 to 0.38 log points increase, and for the old-college a 0.12 to 0.35 log points increase, while their 95% CIs do not exclude zero (Figure 12 Panel A). After the quartile dominance restriction is imposed, the bounds estimates become more precise, which suggest an increase in the gender wage gap for the young college graduates by 0.17 - 0.30 log points and for the older college graduates by 0.16 - 0.24 log points (Figure 12 Panel B), and in both cases the 95% CI exclude zero (Figure 12 Panel B). Under the stochastic dominance restriction, the bounds suggest an increase of the gender wage gap by 0.20 - 0.28 log points for the young college graduates and an increase of the gap by 0.17 - 0.22 log points for the old college graduates, with the 95% CI excluding zero (Figure 12 Panel C). The MIV bounds in Figure 12 Panel D also suggest an increase in the gender wage gap for the young college graduates by 0.21 - 0.62 log points, with the 95% CI excluding zero.

In Figure 13 for 2007 - 2018, the bounds under the quartile and the stochastic dominance restrictions for the young college graduates suggest a decrease in the gender wage gap by

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

0.02 to 0.22 log points (Figure 13 Panel B and C). However, the 95% CI does not exclude a zero change for these bounds. Bounds estimates for the other education and age groups all include a zero change.

6 Discussion

Our estimated bounds show a pattern of increase in the gender wage gap among the young workers (age 25-45) in survey years of 1995-2007 at the median wage and the 75^{th} wage quantile. This result is in line with previous findings 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), whom 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 total workers average. Specifically, under the MIV assumption, our lower bound estimates show an increase by 0.19 - 0.63 log points at the median wage and by 0.21 - 0.62 log points and at the 75^{th} quantile wage, which are both greater than the estimated gender wage gap increase in Gustafsson and Wan (2020) and Song et al. (2019) based on the population of age 16 - 70 and 16 - 60, respectively.

Our bounds for young college graduates during 2007 - 2018 suggest a decrease of the gender wage gap at the median wage by 0.12 - 0.59 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 has narrowed between 2007 - 2013 by 0.04 log points. Suppose more young high-skilled women choose to be self-employed or work for less 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).

The results show different trends in the change of the gender wage gap in two time periods. Economic factors that contribute to the gender wage gap may explain the different trends. In the time period of 1995 - 2007, we find results consistent with an increase in the gender wage gap among the young workers both at the median wage and at the 75^{th} wage quantile. The widened gender wage gap can be explained by 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 75^{th} wage quantile. One potential explanation for this slow-down of the gender wage gap growth can be higher returns to the schooling of women than 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. If women are gaining more years of schooling over birth cohorts while the return to schooling is increasing and higher for women than for men, 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. Song et al. (2014) look at the urban low-income workers in China in 2007 and show that the gender wage gap unexplained by marital status, age, and education accounts for 60% of the total gender wage gap. Ma (2018) uses the China Household Income Panel (CHIP) 2002-2013 and shows that intra-sector gender wage differential contributed more to the observed wage differential and up to 80% of the gender wage differential is unexplained by education, occupation, and working experience.

Hare (2019) finds that the increase in men's labor market return to work experience has offset the closing of the gender wage gap from the increase in females' gain in observed labor market characters, and Zhao et al. (2019) show similar findings.

7 Conclusion

This paper estimates China's distributional gender wage gap dynamics from 1995 to 2018. To control for selection into employment, we employ the nonparametric bounds in Manski (1994), Manski and Pepper (2000), and Blundell et al. (2007). To tighten the bounds, we use a weak quartile dominance assumption, a stochastic dominance assumption, and a monotone instrumental variable (MIV). We have found statistically significant evidence that over the years from 1995-2018, 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. 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 by at least 0.19 log points).

Additionally, this paper also estimates the gender wage gap change at the 25^{th} and the 75^{th} percentiles of the wage distribution. At the 25^{th} 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 75^{th} 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 group. 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 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.

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., Kahn, L., Boboshko, N., and Comey, M. (2021). The Impact of Selection into the Labor Force on the Gender Wage Gap. *NBER Working Paper*, 28855.

Blundell, R., Gosling, A., Ichimura, H., and Meghir, C. (2007). Changes in the Distribution of Male and Female Wages Accounting for Employment Using Bounds. *Journal*, 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.

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 Log Hourly Wage

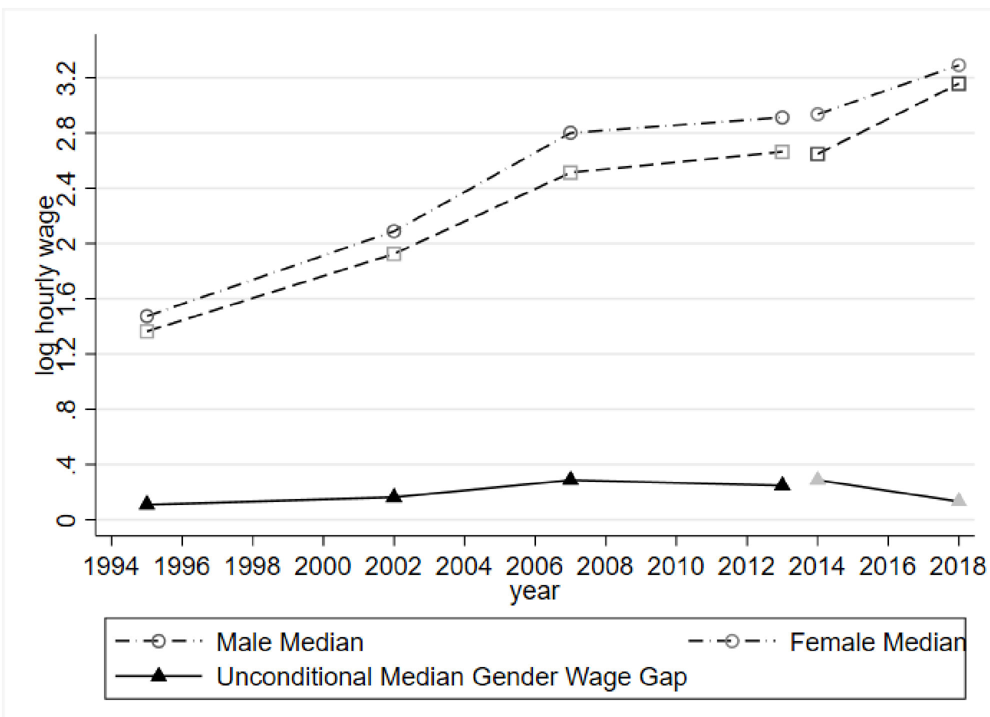


Figure 3: Age Profile for Employment for 1995 and 2018

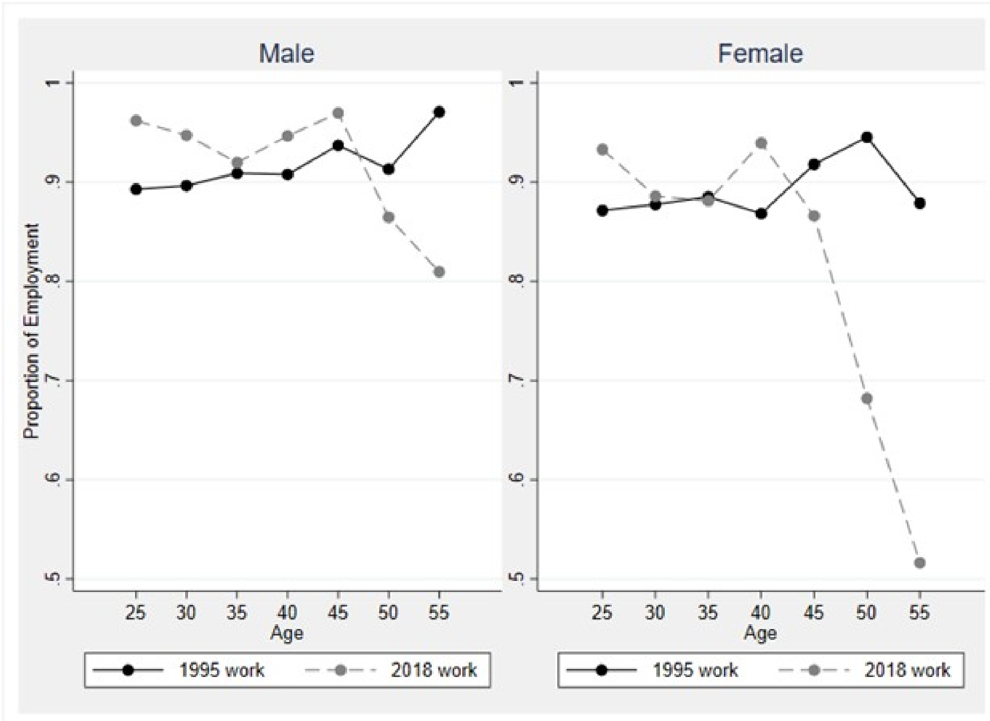


Figure 4: Employment by College Attendance for Males and Females in 1995 and 2018

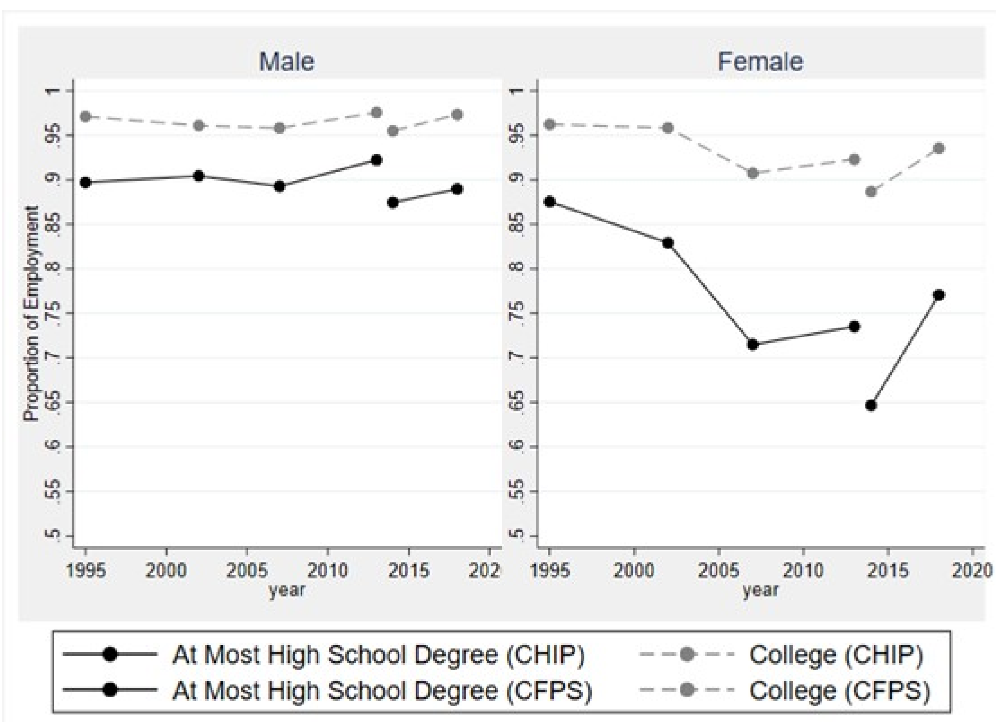


Figure 5: Changes in Median Gender Wage Gap with Various restrictions (1995 - 2018)

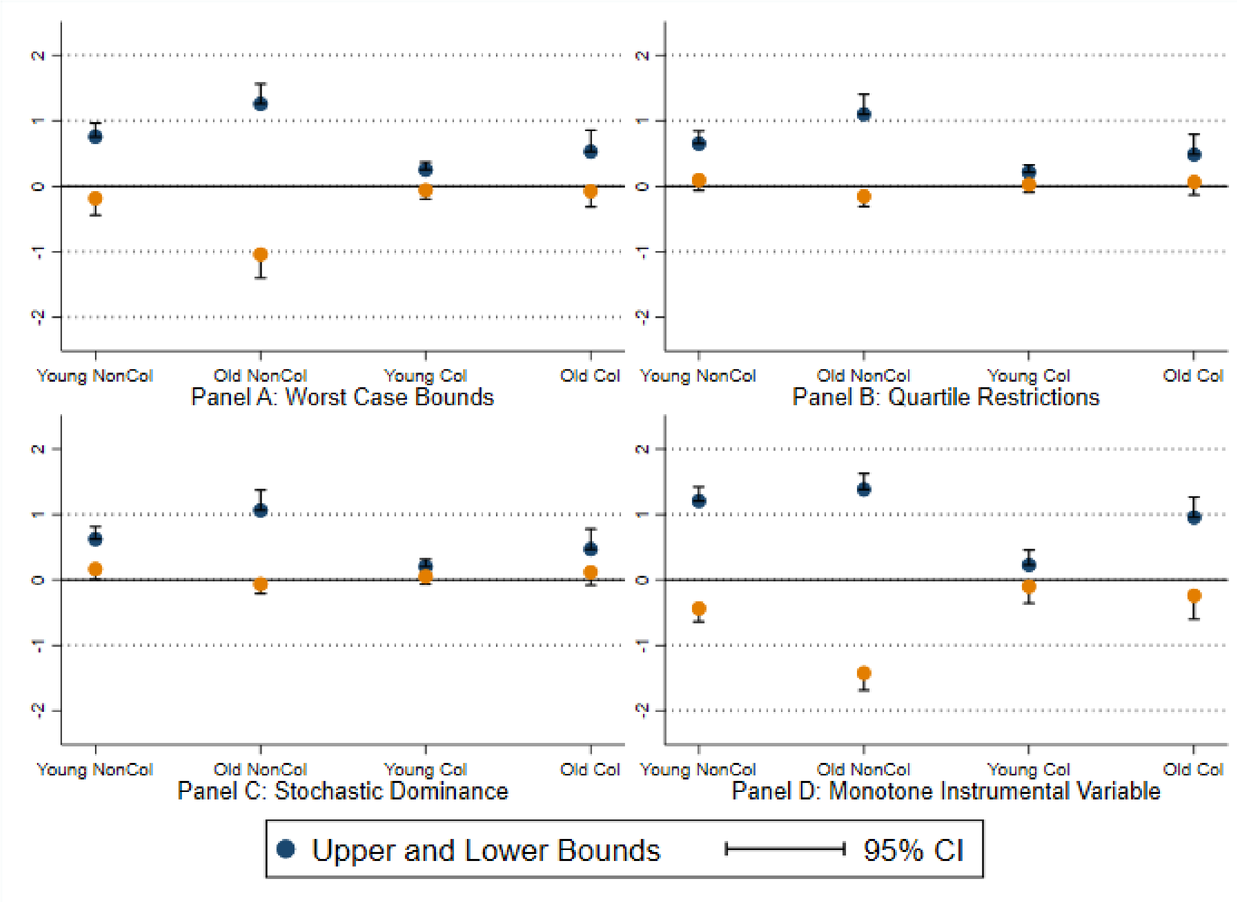


Figure 6: Changes in Median Gender Wage Gap with Various restrictions (1995 - 2007)

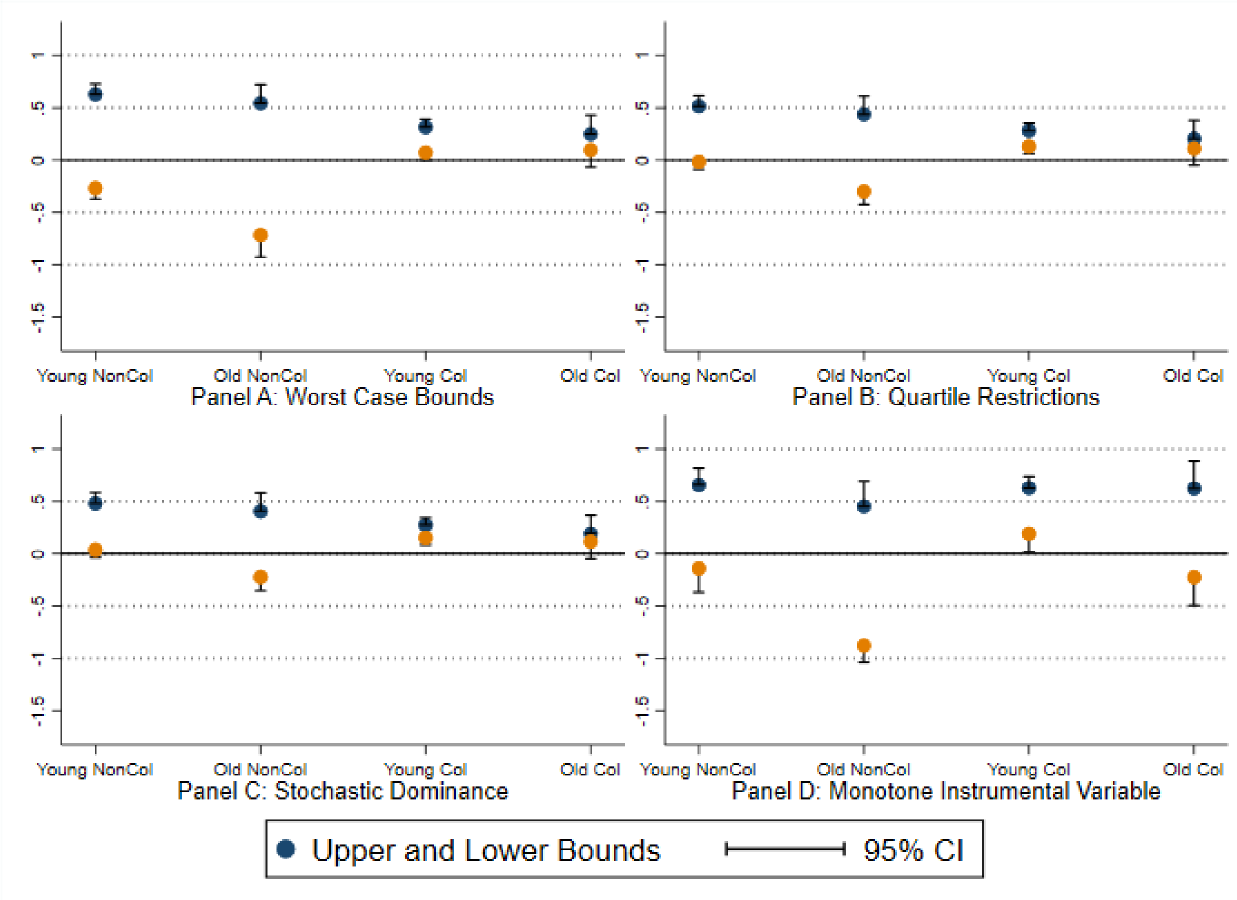


Figure 7: Changes in Median Gender Wage Gap with Various restrictions (2007 - 2018)

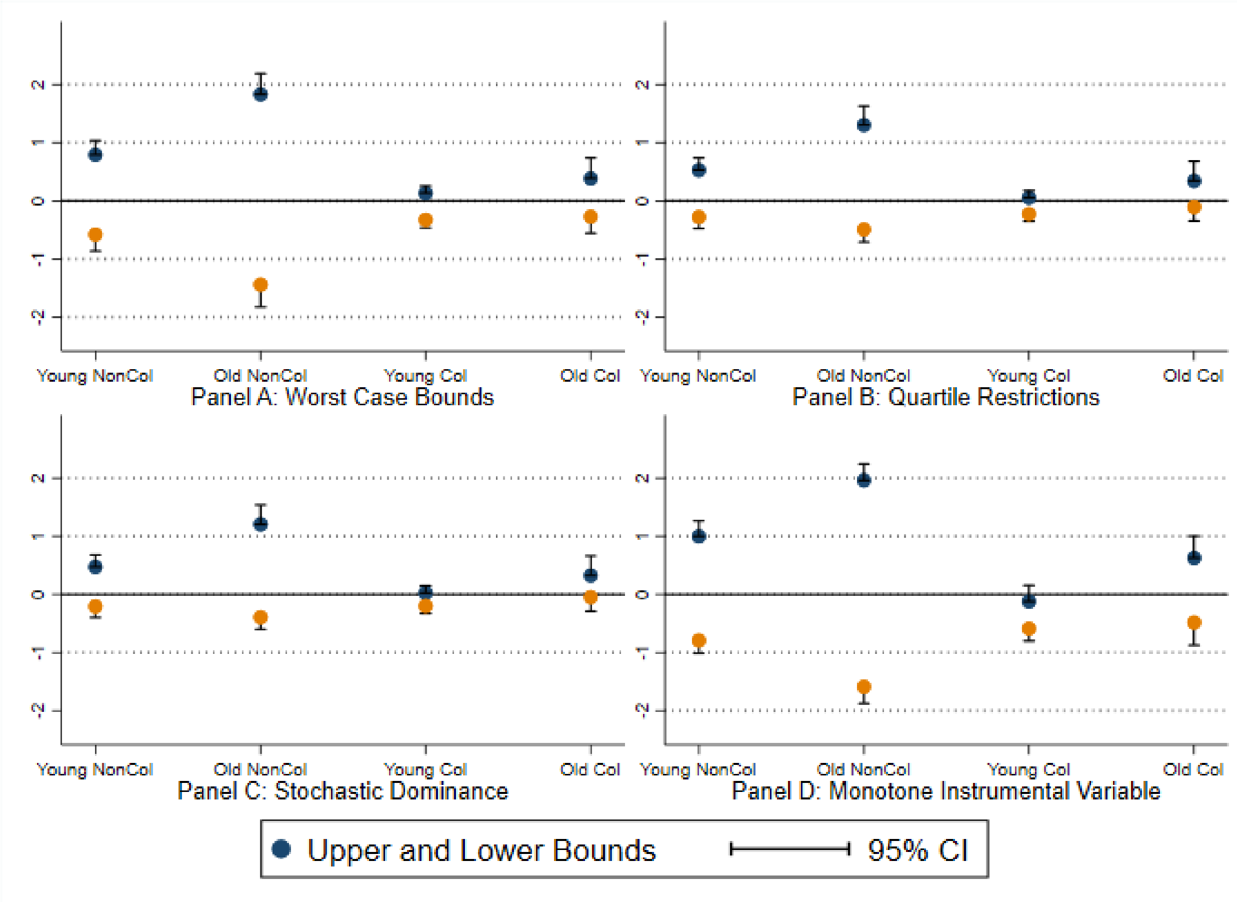


Figure 8: Changes in Gender Wage Gap with Various restrictions at 25th percentile (1995 - 2018)

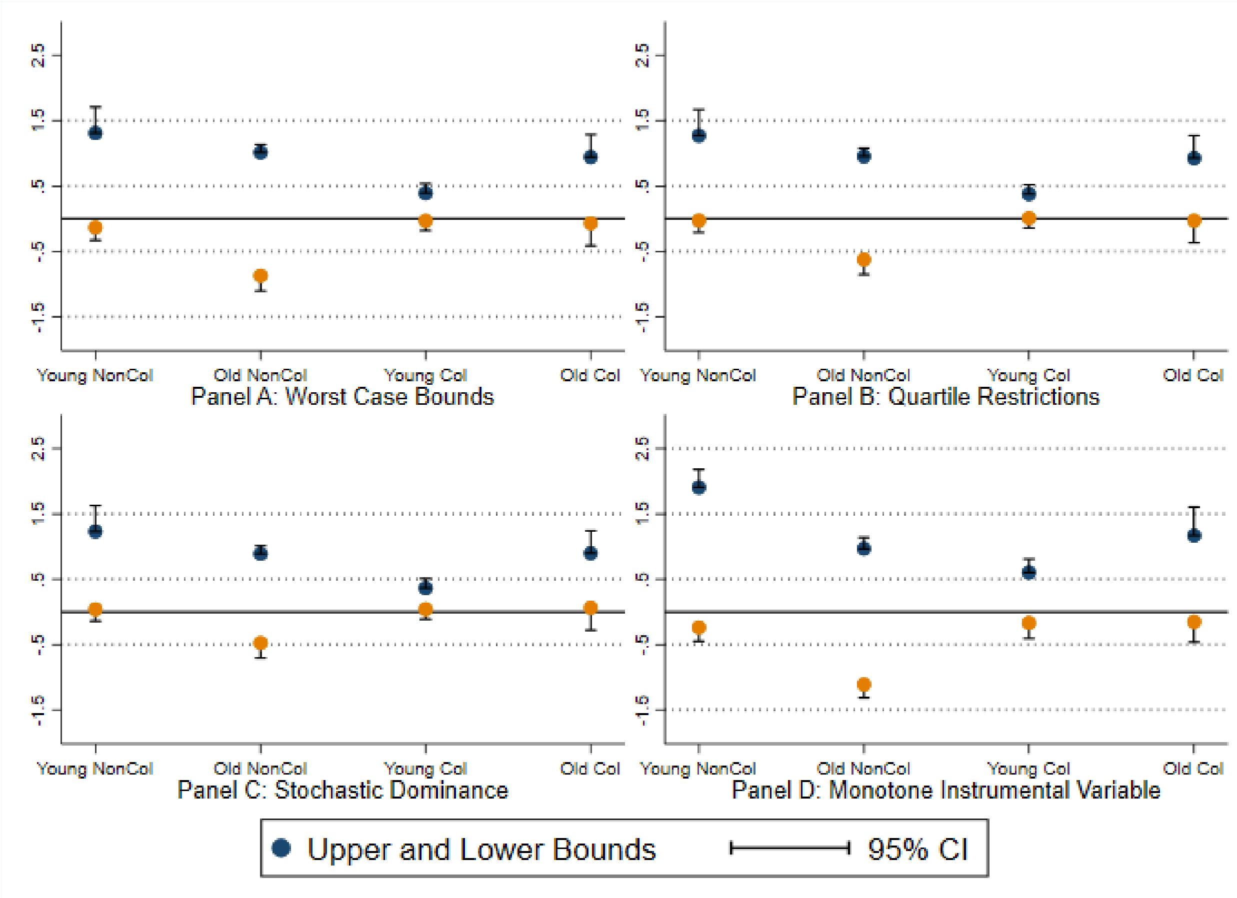


Figure 9: Changes in Gender Wage Gap with Various restrictions at 25^{th} percentile (1995 - 2007)

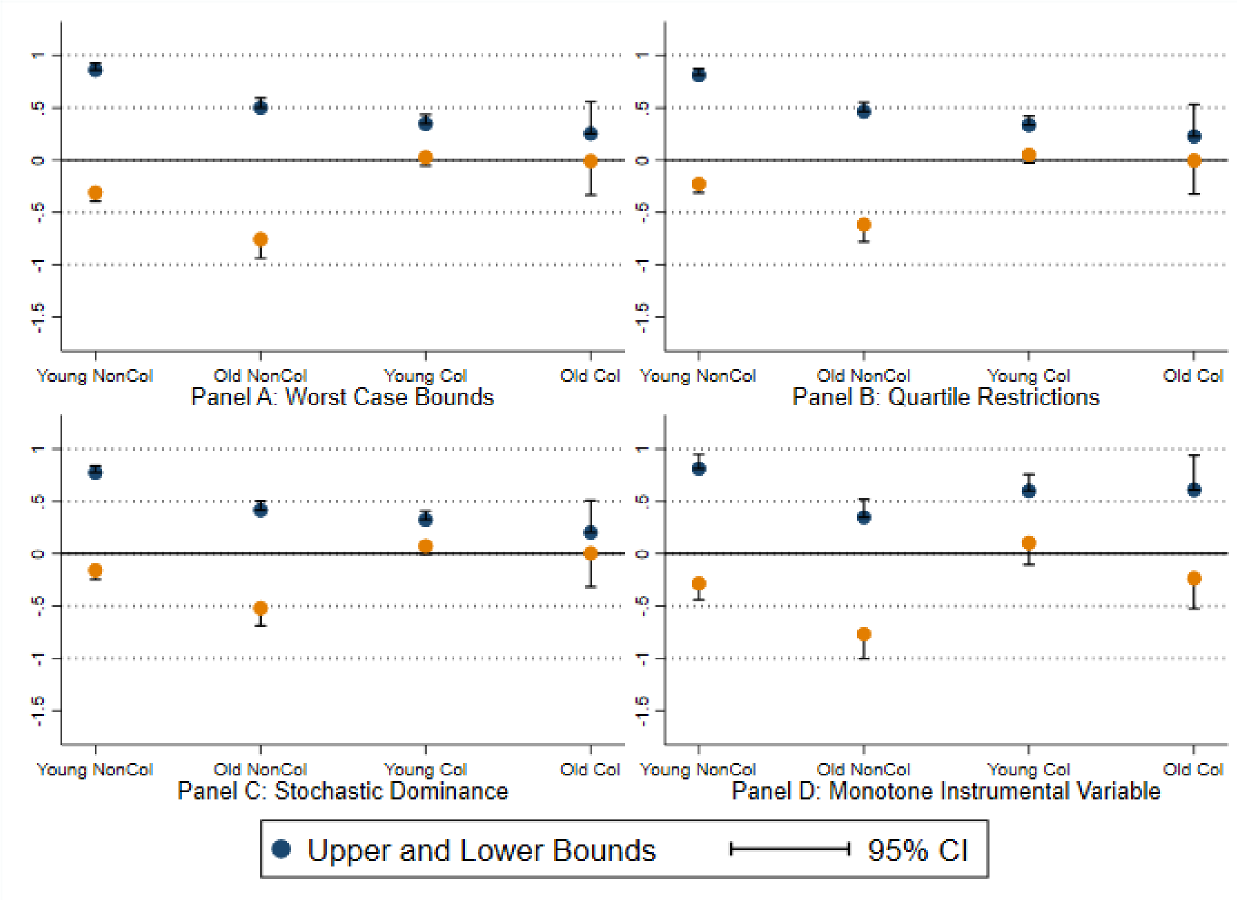


Figure 10: Changes in Gender Wage Gap with Various restrictions at 25^{th} percentile (2007 - 2018)

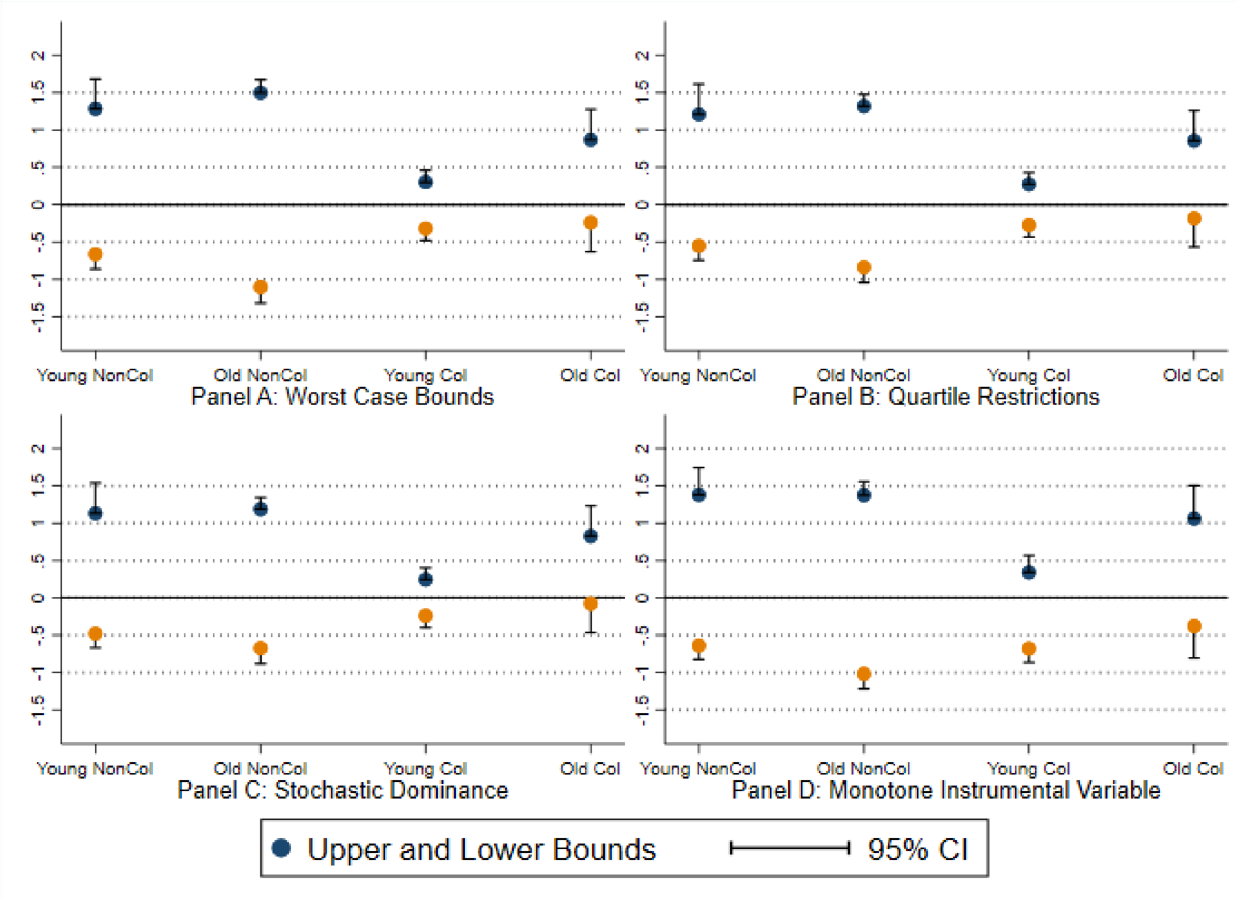


Figure 11: Changes in Gender Wage Gap with Various restrictions at 75^{th} percentile (1995 - 2018)

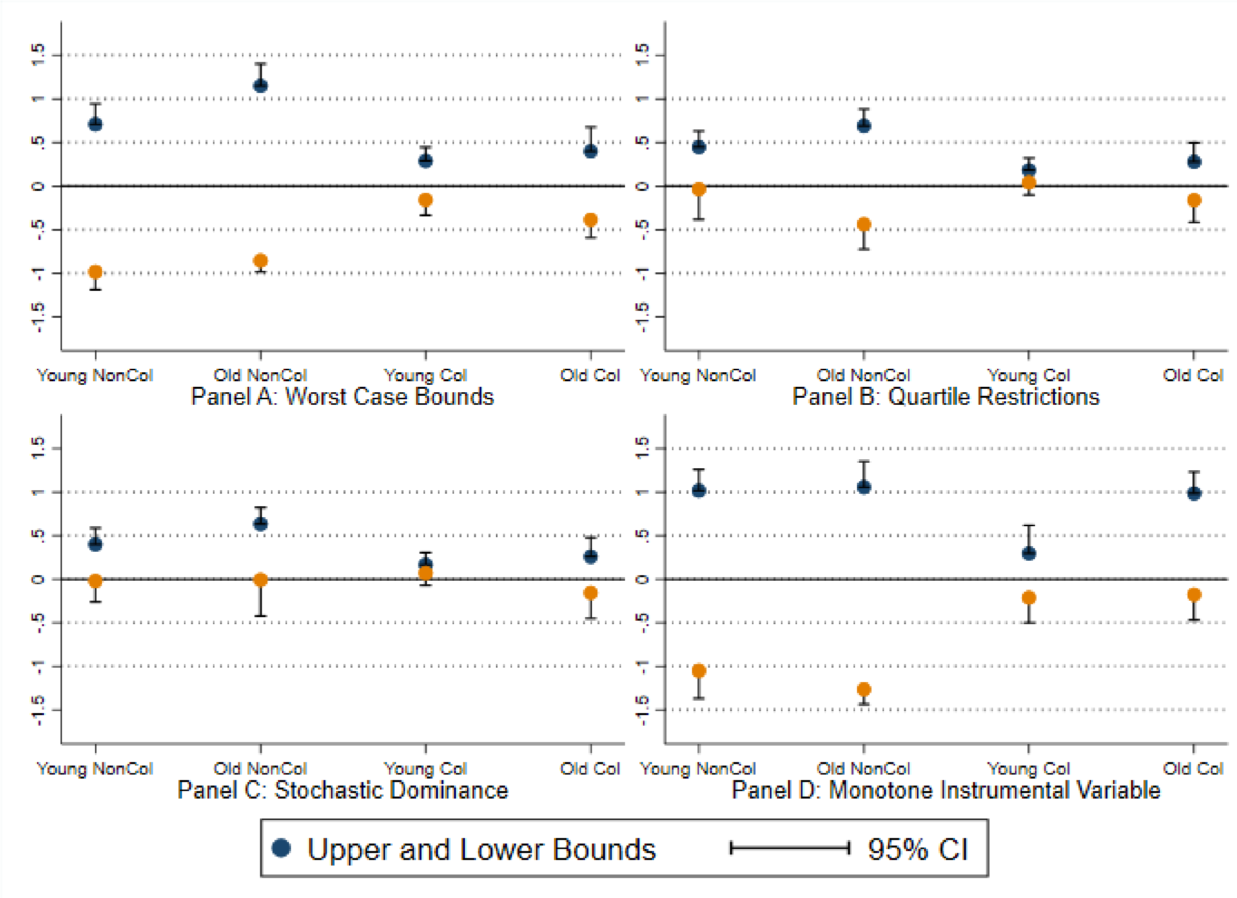


Figure 12: Changes in Gender Wage Gap with Various restrictions at 75^{th} percentile (1995 - 2007)

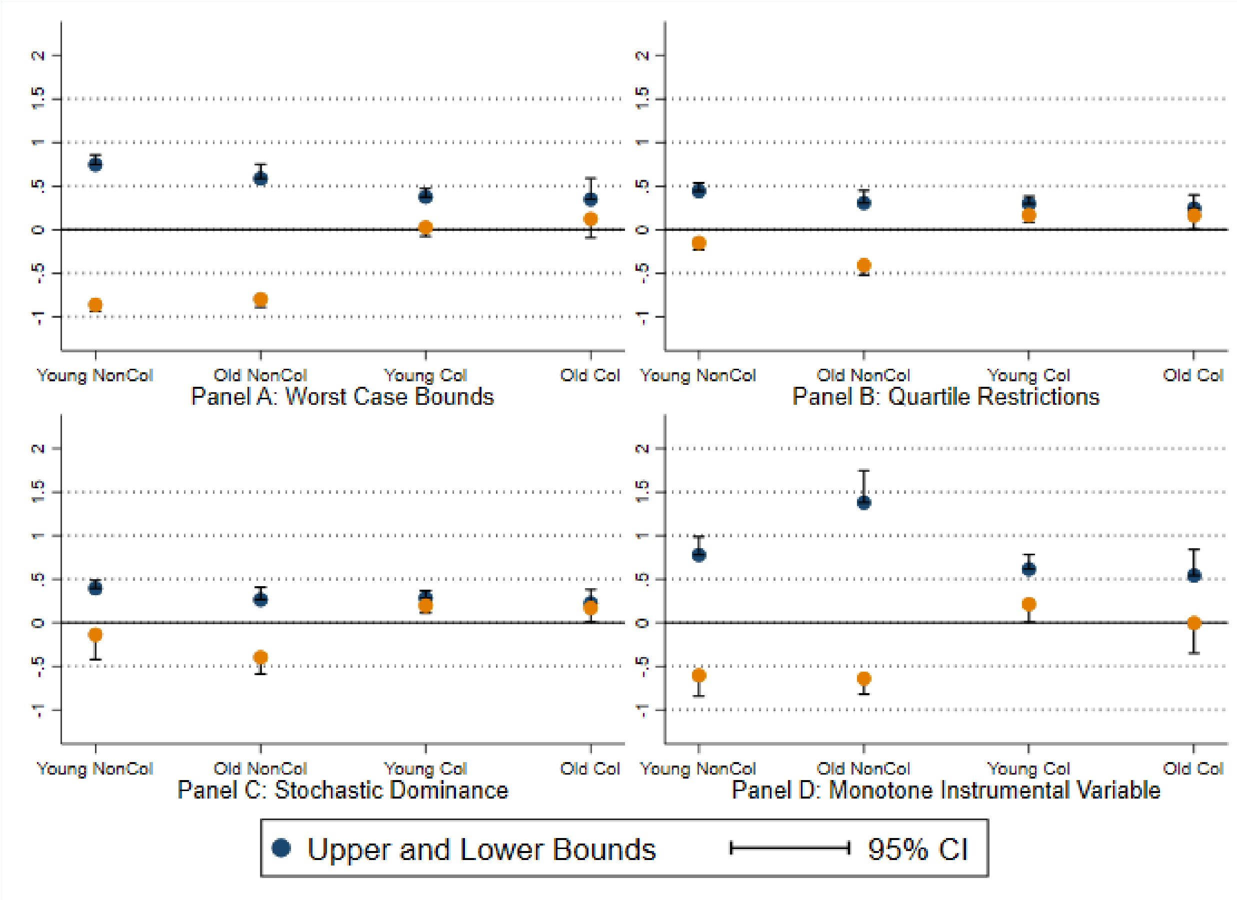
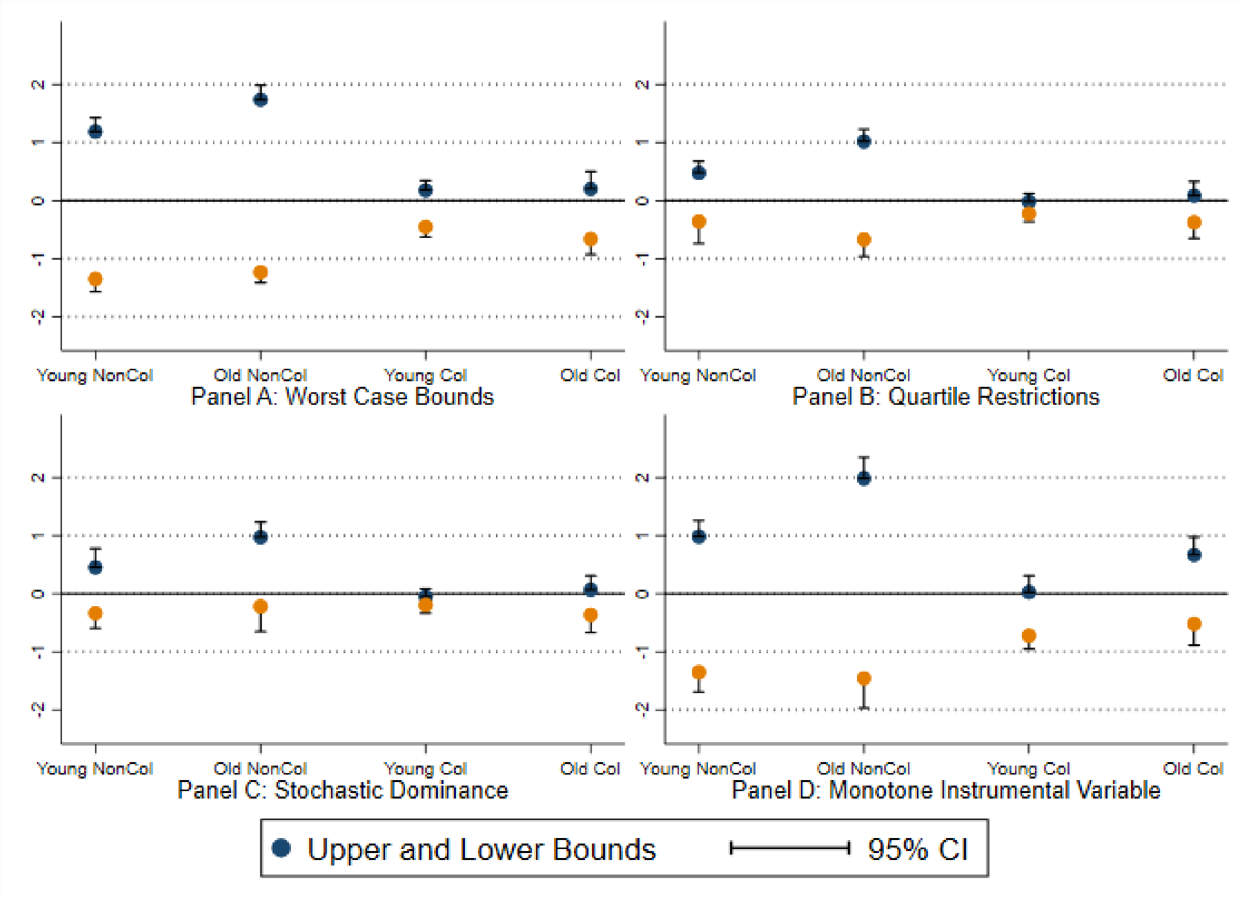


Figure 13: Changes in Gender Wage Gap with Various restrictions at 75^{th} percentile (2007 - 2018)



Appendix A

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

	Worst Case	Quartile Restrictions	Stochastic Dominance	MIV
Young Non-College	(-0.1839, 0.7582)	(0.0966, 0.6525)	(0.1661, 0.6218)	(-0.4363, 1.2050)
	[-0.4410, 0.9636]	[-0.0665, 0.8449]	[0.0072, 0.8138]	[-0.6382, 1.4219]
Old Non-College	(-1.0392, 1.2585)	(-0.1525, 1.1007)	(-0.0652, 1.0626)	(-1.4235, 1.3831)
	[-1.4006 ,1.5633]	[-0.3081, 1.4067]	[-0.2080, 1.3711]	[-1.6875, 1.6209]
Young-College	(-0.0583, 0.2532)	(0.0291,0.2138)	(0.0535, 0.2029)	(-0.0988, 0.2270)
	[-0.1934, 0.3697]	[-0.0943,0.3263]	[-0.0683, 0.3154]	[-0.3595, 0.4579]
Old-College	(-0.0731,0.5305)	(0.0655, 0.4845)	(0.1200, 0.4692)	(-0.2389, 0.9548)
	[-0.3145, 0.8582]	[-0.1322, 0.7956]	[-0.0791, 0.7792]	[-0.6021, 1.2656]

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

	Worst Case	Quartile Restrictions	Stochastic Dominance	MIV
Young Non-College	(-0.2684, 0.6283)	(-0.0159, 0.5145)	(0.0374, 0.4822)	(-0.1410, 0.6568)
	[-0.3720, 0.7277]	[-0.0887, 0.6150]	[-0.0345, 0.5840]	[-0.3695, 0.8190]
Old Non-College	(-0.7159, 0.5440)	(-0.2982,0.4364)	(-0.2237, 0.4062)	(-0.8779, 0.4525)
	[-0.9265 ,0.7223]	[-0.4255,0.6110]	[-0.3529, 0.5805]	[-1.0366, 0.6917]
Young-College	(0.0727, 0.3150)	(0.1309,0.2821)	(0.1525, 0.2740)	(0.1914, 0.6283)
	[0.0004, 0.3891]	[0.0617,0.3536]	[0.0837, 0.3445]	[0.0188, 0.7363]
Old-College	(0.0961, 0.2484)	(0.1104,0.2037)	(0.1164, 0.1919)	(-0.2255, 0.6236)
	[-0.0680, 0.4299]	[-0.0478,0.3783]	[-0.0440, 0.3681]	[-0.4937, 0.8873]

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

	Worst Case	Quartile Restrictions	Stochastic Dominance	MIV
Young Non-College	(-0.5779, 0.7923)	(-0.2787, 0.5293)	(-0.2050, 0.4734)	(-0.7906, 0.9989)
	[-0.8652, 1.0367]	[-0.4741, 0.7383]	[-0.3974, 0.6780]	[-1.0104, 1.2641]
Old Non-College	(-1.4397, 1.8309)	(-0.4932,1.3032)	(-0.3910, 1.2060)	(-1.5870, 1.9620)
	[-1.8237, 2.1884]	[-0.7108,1.6285]	[-0.5981, 1.5356]	[-1.8733, 2.2426]
Young-College	(-0.3249, 0.1322)	(-0.2244, 0.0543)	(-0.1978, 0.0277)	(-0.5864, -0.1206)
	[-0.4630, 0.2567]	[-0.3513, 0.1727]	[-0.3234, 0.1459]	[-0.7991, 0.1534]
Old-College	(-0.2722, 0.3851)	(-0.1062,0.3422)	(-0.0458, 0.3267)	(-0.4810, 0.6281)
	[-0.5574, 0.7369]	[-0.3468,0.6784]	[-0.2886, 0.6622]	[-0.8719, 1.0035]

Table A.4 : Bounds on Changes in Gender Wage Differential at 25^{th} percentile (1995 - 2018)

	Worst Case	Quartile Restrictions	Stochastic Dominance	MIV
Young Non-College	(-0.1327, 1.3107)	(-0.0268, 1.2722)	(0.0406, 1.2321)	(-0.2383, 1.9061)
	[-0.3272, 1.7072]	[-0.2100, 1.6693]	[-0.1400, 1.6268]	[-0.4464, 2.1788]
Old Non-College	(-0.8702, 1.0143)	(-0.6224, 0.9549)	(-0.4750, 0.8891)	(-1.1090, 0.9674)
	[-1.1050, 1.1352]	[-0.8521, 1.0807]	[-0.7052, 1.0122]	[-1.3158, 1.1387]
Young-College	(-0.0302, 0.3920)	(0.0113, 0.3788)	(0.0419, 0.3640)	(-0.1677, 0.6040)
	[-0.1838, 0.5378]	[-0.1415, 0.5228]	[-0.1113, 0.5078]	[-0.4073, 0.8086]
Old-College	(-0.0677, 0.9448)	(-0.0278, 0.9266)	(0.0635, 0.8977)	(-0.1508, 1.1682)
	[-0.4119, 1.2891]	[-0.3674, 1.2723]	[-0.2770, 1.2424]	[-0.4622, 1.6009]

Table A.5 : Bounds on Changes in Gender Wage Differential at 25^{th} percentile (1995 - 2007)

	Worst Case	Quartile Restriction	Stochastic Dominance	MIV
Young Non-College	(-0.3083, 0.8612)	(-0.2266, 0.8123)	(-0.1577, 0.7747)	(-0.2834, 0.8105)
	[-0.3941, 0.9234]	[-0.3115, 0.8738]	[-0.2459, 0.8379]	[-0.4415, 0.9504]
Old Non-College	(-0.7554, 0.5009)	(-0.6158, 0.4637)	(-0.5202, 0.4166)	(-0.7669, 0.3469)
	[-0.9360, 0.5927]	[-0.7808, 0.5530]	[-0.6872, 0.5056]	[-1.0022, 0.5211]
Young-College	(0.0275, 0.3475)	(0.0527, 0.3349)	(0.0733, 0.3226)	(0.1050, 0.5982)
	[-0.0514, 0.4338]	[-0.0255, 0.4202]	[-0.0063, 0.4074]	[-0.1058, 0.7543]
Old-College	(-0.0094, 0.2532)	(-0.0042, 0.2268)	(0.0051, 0.2036)	(-0.2364, 0.6094)
	[-0.3311, 0.5586]	[-0.3237, 0.5312]	[-0.3165, 0.5098]	[-0.5254, 0.9379]

Table A.6 : Bounds on Changes in Gender Wage Differential at 25^{th} percentile (2007 - 2018)

	Worst Case	Quartile Restrictions	Stochastic Dominance	MIV
Young Non-College	(-0.6600, 1.2851)	(-0.5504, 1.2101)	(-0.4767, 1.1324)	(-0.6384, 1.3798)
	[-0.8613, 1.6839]	[-0.7410, 1.6172]	[-0.6660, 1.5416]	[-0.8204, 1.7461]
Old Non-College	(-1.1004, 1.4991)	(-0.8352, 1.3199)	(-0.6717, 1.1894)	(-1.0152, 1.3758)
	[-1.3180, 1.6742]	[-1.0434, 1.4771]	[-0.8818, 1.3463]	[-1.2153, 1.5565]
Young-College	(-0.3169, 0.3037)	(-0.2719, 0.2745)	(-0.2367, 0.2467)	(-0.6773, 0.3430)
	[-0.4765, 0.4591]	[-0.4309, 0.4286]	[-0.3970, 0.4011]	[-0.8646, 0.5683]
Old-College	(-0.2373, 0.8705)	(-0.1810, 0.8571)	(-0.0759, 0.8284)	(-0.3765, 1.0626)
	[-0.6306, 1.2769]	[-0.5661, 1.2619]	[-0.4651, 1.2328]	[-0.8031, 1.5044]

Table A.7 : Bounds on Changes in Gender Wage Differential at 75^{th} percentile (1995 - 2018)

	Worst Case	Quartile Restrictions	Stochastic Dominance	MIV
Young Non-College	(-0.9838, 0.7089)	(-0.0345, 0.4471)	(-0.0195, 0.4006)	(-1.0494, 1.0169)
	[-1.1875, 0.9431]	[-0.3839, 0.6326]	[-0.2612, 0.5864]	[-1.3701, 1.2590]
Old Non-College	(-0.8550, 1.1503)	(-0.4380, 0.6905)	(-0.0062, 0.6338)	(-1.2634, 1.0571)
	[-0.9853, 1.4019]	[-0.7229, 0.8818]	[-0.4213, 0.8222]	[-1.4339, 1.3512]
Young-College	(-0.1580, 0.2887)	(0.0418, 0.1824)	(0.0717, 0.1671)	(-0.2114, 0.2965)
	[-0.3350, 0.4461]	[-0.1021, 0.3206]	[-0.0722, 0.3064]	[-0.4995, 0.6178]
Old-College	(-0.3852, 0.4004)	(-0.1610, 0.2774)	(-0.1562, 0.2602)	(-0.1763, 0.9858)
	[-0.5926, 0.6754]	[-0.4157, 0.4984]	[-0.4506, 0.4781]	[-0.4697, 1.2311]

Table A.8 : Bounds on Changes in Gender Wage Differential at 75^{th} percentile (1995 - 2007)

	Worst Case	Quartile Restrictions	Stochastic Dominance	MIV
Young Non-College	(-0.8616, 0.7468)	(-0.1496, 0.4436)	(-0.1346, 0.3954)	(-0.6036, 0.7791)
	[-0.9371, 0.8578]	[-0.2299, 0.5375]	[-0.4215, 0.4897]	[-0.8429, 0.9855]
Old Non-College	(-0.7978, 0.5870)	(-0.4069, 0.3063)	(-0.3966, 0.2663)	(-0.6384, 1.3798)
	[-0.8887, 0.7512]	[-0.5204, 0.4487]	[-0.5883, 0.4087]	[-0.8204, 1.7461]
Young-College	(0.0261, 0.3759)	(0.1672, 0.2975)	(0.1983, 0.2836)	(0.2143, 0.6158)
	[-0.0771, 0.4717]	[0.0836, 0.3812]	[0.1150, 0.3668]	[0.0077, 0.7828]
Old-College	(0.1239, 0.3484)	(0.1621, 0.2428)	(0.1707, 0.2245)	(-0.0020, 0.5429)
	[-0.0931, 0.5886]	[0.0052, 0.3986]	[0.0108, 0.3828]	[-0.3475, 0.8417]

Table A.9 : Bounds on Changes in Gender Wage Differential at 75^{th} percentile (2007 - 2018)

	Worst Case	Quartile Dominance	Stochasrtic Dominance	MIV
Young Non-College	(-1.3482, 1.1881)	(-0.3575, 0.4761)	(-0.3343, 0.4547)	(-1.3497, 0.9854)
	[-1.5674, 1.4269]	[-0.7370, 0.6772]	[-0.5933, 0.7788]	[-1.6889, 1.2598]
Old Non-College	(-1.2313, 1.7375)	(-0.6675, 1.0205)	(-0.2173, 0.9753)	(-1.4526, 1.9861)
	[-1.4153, 1.9888]	[-0.9620, 1.2267]	[-0.6526, 1.2343]	[-1.9638, 2.3489]
Young-College	(-0.4504, 0.1791)	(-0.2242, -0.0163)	(-0.1923, -0.0508)	(-0.7191, 0.0276)
	[-0.6259, 0.3459]	[-0.3656, 0.1215]	[-0.3322, 0.0863]	[-0.9513, 0.3127]
Old-College	(-0.6596, 0.2026)	(-0.3736, 0.0851)	(-0.3603, 0.0691)	(-0.5155, 0.6714)
	[-0.9286, 0.5011]	[-0.6505, 0.3303]	[-0.6693, 0.3116]	[-0.8867, 0.9756]

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. Estimation and Inference Implementation

In Section 3 of the paper, we have briefly described the method in Chernozhukov et al. (2013) to compute confidence regions for bounds with maximum and minimum operators. In Section B1, we explain the computation of bounds under the MIV assumption, and in Section B2, 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).

B.1 Inference for Bounds under the MIV assumption

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 $\theta^l(\upsilon)$ 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 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.2 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

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

where $\hat{\theta}^u(v)$ is the sample analog estimator of $\theta^u(v)$ and $s^u(v)$ is its standard error. To compute $\kappa^u_{n,\hat{V}^l_n}(p)$ in (2), we follow the same steps above but in step (3) we replace $\hat{V}^l_n$ by $\hat{V}^u_n = \{v \in \mathcal{V}^u : \hat{\theta}^u(v) \geq \min_{\tilde{v} \in \mathcal{V}^u} [\hat{\theta}^u(\tilde{v}) + \kappa^u_{n,\mathcal{V}^u}(c_n) s^u_n(\tilde{v})] + 2\kappa^u_{n,\mathcal{V}^u}(c_n) s^u_n(v)\}$. 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 $[\hat{\theta}^l_n(\hat{p}_n), \hat{\theta}^u_n(\hat{p}_n)]$.