What Can We Learn About the Racial Gap in the Presence of Sample Selection?

Esfandiar Maasoumi†
Emory University

Le Wang
University of Oklahoma, & IZA

Abstract

We examine the distance and relations between the distributions of wages for two exogenously identified groups (black and white women here). The literature commonly employs decomposition methods for the conditional means, to propose explanations for observed wage differentials, as "structural" components, attributable to difference in market structures, and the "composition" components, attributable to difference in characteristics and skills. Estimation of these components is often hampered by restrictive wage structure assumptions, and sample selection issues (wages are only observed for those working). We address these issues by first utilizing modern strategies in the treatment effects literature to identify the entire distributions of wages and counterfactual wages among working women, which afford a separation of composition and market effects. We avoid restrictive wage structure modeling by nonparametric inverse probability weighting methods. This approach allows for decomposition beyond the gap at the mean, and can deliver distributional statistics of interest, such as inequalities and target quantiles. Accounting for selection, we extend the basic framework to provide a computationally convenient way to identify bounds on the decomposed components for the whole population. We employ these methods to understand the sources and dynamics of the racial gap in the U.S.. Our analysis reveals that what may be learned about racial gap is impacted by labor force participation, and is also sensitive to the choice of population of interest. Our results question what may be gleaned from the commonly reported point estimates when sample selection is neglected.

JEL Classifications: C13; C14; C21; J31

Keywords: Decomposition, Distributional Analysis, Racial Wage Gap, Sample Selection, Partial Identification, Treatment Effects, Women

---

† Corresponding address: Esfandiar Maasoumi, Department of Economics, Emory University. E-mail: esfandiar.maasoumi@emory.edu. This paper is in celebration of Robert (Bob) Basmann, great friend and mentor, and renowned econometrician/economist. We hope our work demonstrates the continuing central importance of Robert Basmann tradition of rigorous examination of underlying statistical distributions, and of his care for formal decision/welfare theoretic analysis of observed data. His enduring standards are exemplified by his pioneering work on finite sample distribution theory and identification (Basmann (JASA (1960), Econometrica 1962a-b), and on innovative utility functions and inequality measurement (e.g, Econometrica 1956). The authors would also like to thank Dan Slottje and two anonymous referees for their helpful comments and suggestions.

1. Introduction

Wages are an important source of individual incomes, which in turn determine one's well-being. The wage gap between groups is an important measure of between-group inequality. Policymakers and labor economists have long been interested in measuring and understanding the distance between the distributions of wages for advantaged and disadvantaged groups, for example, men and women (the gender wages gap), or whites and blacks (the racial gap). It is essential to explicitly define and understand any measure of disparity, and what can account for it when outcomes are distributed and vary over individuals and groups?

There are generally two sources or types for the wage gap: the differences in market value of skills and characteristics, and the differences in wage structures (different returns to individual characteristics). The former is referenced as "composition effects", and the latter as "structural effects". In a perfectly competitive labor market, individual characteristics should be rewarded the same, and if they are not, as captured by the structural effects, there may exist some form of discrimination. As such, identification and estimation of these two effects not only can help us understand better the sources of the gap, but it may also help inform policy making to reduce the gap and to improve workers' well-being. For example, the existence of composition effects suggests that policies aimed at improving individuals' productivity-enhancing characteristics may be beneficial, while the existence of structural point to the need for other policies.

Wage gap decomposition methods date back to the seminal works of Oaxaca (1973) and Blinder (1973). The conventional Oaxaca-Blinder decomposition method, however, relies on linear specifications of the wage functions and focuses on the decomposition values for the average worker, composition at the conditional mean. Misspecification of the wage function leads to inconsistent estimates of the structural and compositional components. Further, there has been evidence that the gender gap or the racial gap can differ across the wage distribution (e.g., Albrecht et al. (2003)), and focusing on the mean could mask potentially substantial heterogeneity in the relative importance of different components across the distribution. A number of decomposition methods have been developed to relax both the functional form assumption and to estimate the decomposition components at the distributional level; for instance, the "plug-in" procedure of Juhn et al. (1993), the reweighing procedure of DiNardo et al. (1996), the quantile approach of Machado and Mata (2005a), and the influence function approach of Firpo et al. (2007).

Sample selection or missing outcomes are frequently a challenge to inference about the whole population. Sample selection arises because wages are only observed for those working. This issue can be particularly severe in empirical applications when the disadvantaged group is women. As shown in Blau and Kahn (2016), while there was a sharp increase in female participation rates before the 1990s, the participation rates have plateaued since then and remain relatively low. Failure to take into account the selection issue has been shown to significantly bias the estimation of wage functions (Heckman, 1974), and hence the estimation of the counterfactual distributions and corresponding decomposition components.

While there has been an increasing awareness of the selection issue in the decomposition methods literature, there have been few developments to accommodate it. One of the main challenges is lack of a widely accepted approach to estimate the "entire" counterfactual distribution in the presence of sample selection. Reweighting and influence function methods have not been extended to accommodate sample selection. One alternative is to appropriately estimate conditional quantiles, and invert to find conditional and marginal distributions. Albrecht et al. (2009) extend Machado and Mata (2005a)'s quantile approach by incorporating Buchinsky (1998)'s quantile selection model in the first-stage estimations, while Maasoumi and Wang (2016) adopt a similar approach but by using Arellano and Bonhomme (2017)'s quantile-copula approach in the first stage. Both approaches *point* identify the decomposition terms based on assumptions on the wage structure or on the dependence between unobservable determinants of wages and employment decisions (in addition to assuming availability of an exclusion restriction).[1]

In this paper, we first extend the reweighting approach to identify the counterfactual distribution for working women. This group of individuals is of great interest, and is indeed what the literature has been focusing on. As in the related treatment effects literature, we show that variants of conditional independence and common support assumptions, along with an assumption on the selection/participation mechanism, are required to identify some counterfactual distributions. Provided that these assumptions are met, we can identify both structural and composition effects across the distribution for the working group. Empirical implementation of this method requires an additional step of estimation of participation/selection propensity scores.

As noted in Newey (2007) and Huber (2014), without further assumptions, we cannot *point* identify the distributions for other groups than the working group since we do not observe the wages for nonworkers. We extend the basic framework to consider partial identification of the counterfactual distribution and the corresponding decomposition components. The conditional quantile bounds preserve the ranking of the corresponding (but unknown) quantiles. We exploit this simple result to provide a computationally convenient way to identify and estimate the bounds on *marginal* distributions and their counterfactuals for the *whole* population in the presence of sample selection.[2] Our approach can be considered as a partial identification version of Machado and Mata (2005a)'s approach. Our results can be employed to examine a wide range of characteristics of the wage distributions, such as inequality measures (e.g., inter-quartile range values), in the presence of sample selection.

This paper is focused on the difference in the wage distributions of black and white women (i.e., the racial gap) in the U.S.. The racial gap in standards of living "remains a fact of life in

[1] Buchinsky (1998)'s method assumes that wages are a separable and additive function of observable and unobservable characteristics. More important, it also implicitly assumes independence between the error term and the regressors conditional on the selection probability." (Melly and Huber, 2008). This assumption leads to parallel quantile curves, in other words, all quantile functions are identical and equal to the conditional mean function (Melly and Huber, 2008 and Arellano and Bonhomme, 2017).

[2] For an account of estimation of average treatment effects in the presence of sample selection, see Lee (2009).

the U.S." (Neal, 2006). Notwithstanding progress in politics, medicine, and social sciences, black women have continued to see significant disparity in their observed wage outcomes, not only compared to white women, but also to minority men (Altonji and Blank, 1999). Neal (2004)'s work points out that labor force participation rates are relatively low among women, and that such disparity could be even larger if we consider "potential" wages for the "whole" population. The author uses several methods to address sample selection and imputes wages for women who do not work and thus have no observed wages. He finds that the *median* racial gap for the whole population (including those who do not work) may be more than 60% larger than the gap among working women. Using more recent data to revisit this issue, Albrecht et al. (2014) find that there exists a substantial median racial gap, and both the median log wage gap and the selection-corrected gap have increased over time. More importantly, they confirm that selection matters, and that the observed racial gap among working women underestimates the racial gap for the larger population. Existence of such a substantial gap has fueled "much public debate about the government's role in regulating the labor market" (Neal, 2006). The presumption is in favor of more regulation to combat racial discrimination in the labor market by employers in hiring and pay. Empirical verification hinges on how much of the gap can be explained by the differences in wage structure and the differences in typical measures of "skills". Our approach sheds light on these questions, and more importantly, reveals what we can learn about the racial gap with and without imposing certain assumptions in the presence of sample selection. Our paper therefore also contributes to the important yet understudied topic of the role of race and ethnicity among women in the labor market (Altonji and Blank, 1999; Black et al., 2008).

Using data from the Current Population Survey March supplement, we find that the view of the racial gap is impacted by selection and is sensitive to the definition of the population of interest. When focusing only on working women, we find substantial heterogeneity in the gap at different quantiles, with different dynamic patterns over time. The gap was larger in the lower tail but smaller and even non-existent in the upper tail during the 70s. This pattern is reversed during recent years, with the gap being the largest among high-earning women. Similarly, the sources of the gap vary across the distribution and over time as well. Addressing the selection leads to larger estimates of structural effects for the upper tail of the distribution. This result suggests that failure to address the participation decision may understate the extent of potential discrimination against high-earning black women in the labor market, especially in more recent years. On the other hand, when focusing on the whole population, bounds are generally wide and include both positive and negative values for both structural and composition effects. Assuming the availability of an exclusion restriction tightens the bounds, but they continue to be wide. The inability to infer clear patterns is valuable and powerful. It indicates that any empirical consensus regarding the extent and the sources of the racial gap for all women is likely reached under strong assumption(s). Our results also question strong inferences based on point estimates when ignoring sample selection.

The remainder of the paper is organized as follows. Section 2 introduces general notation and definitions. Section 3 discusses the identification of structure and composition effects for the selected sample. Section 4 discusses partial identification results. Section 5 describes the data, while Section 6 presents the results for our empirical example. Section 7 summarizes and further discusses the results and finally Section 8 concludes.

2. Notation and Definitions

We consider two exclusive groups indexed by $D = d \in \{0,1\}$, with N_d individuals in each group ($N_1 + N_0 = N$). Let $Y(d)$ be the outcome of interest for any individual in group d (for example, wage *offers*), given by

$$Y(d) = m^d(X,U) \quad \text{for } d \in \{0,1\}, \quad U|X \sim \text{Uniform(0,1)} \tag{2.1}$$

where $\tau \mapsto m^d(X,\tau)$ is strictly increasing and continuous in τ.[3] X is a vector of observable characteristics distributed from F_X^d defined over the support $\mathcal{X}_d \subset \mathcal{X}$ for each group d. U is an unobserved error term, normalized and typically interpreted as ability (Doksum, 1974). This nonseparable model is by construction a quantile function.[4]. We focus on this for two reasons. First, this allows us to go beyond decomposition at the mean level and consider distributional levels. Second, as noted in Melly and Huber (2011) "quantiles are naturally bounded while bounding mean functions requires bounding the support of the outcome". This feature will become useful when we consider partial identification of the decomposition terms. Note that this model nests the parametric quantile regression model $m(X,U) = X'\beta(U)$ that has been widely used in the literature (see, e.g., Machado and Mata (2005a)).

Let $F_{Y(d)|D=d}$ and $F_{Y(d)|D=d,X}$ be the unconditional and conditional distributions of the outcome for group d, respectively, defined over the support $(y_{\min}, y_{\max})$; $y_{\min}$ is the smallest possible value that $Y(d)$ can take on and can be $-\infty$, while $y_{\max}$ is the largest possible value that $Y(d)$ can take on and can be ∞. To facilitate the connection to the treatment effects literature, we can consider $Y(d)$ as the potential outcome for those in group d, given characteristics X, U. One cannot be in both groups at the same time, and we observe an outcome when she belongs to a particular group $Y = Y(1) \cdot D + Y(0) \cdot (1 - D)$.

Let S be a binary indicator for sample selection, equal to one if an individual's outcome is observed in the sample and zero otherwise. The observed outcome is equal to Y for individuals who are selected ($S = 1$), but missing for those who are not ($S = 0$). We can similarly define the unconditional and conditional distributions of the outcome for group d in the selected sample, $F_{Y(d)|D=d,S=1}$ and $F_{Y(d)|X,D=d,S=1}$, and the distribution of observable

[3] Note that this follows from the Skorohod representation of random variables, see, e.g. Chernozhukov and Hansen (2005) and Galichon (2016).

[4] As noted in Sasaki (2014), if X is exogenous and the underlying structural function $Y = g(X,\epsilon)$ is monotone with respect to the scalar error term,ϵ, the quantile function can be used to represent the underlying structural function. Even in the general case, in which monotonicity fails or ϵ is multidimensional, the quantile partial derivatives still identify "a weighted average of heterogeneous structural partial effects among the subpopulation of individuals at the conditional quantile of interest."

characteristics, $F^d_{X|S=1}$.

We typically characterize the differences in the outcome between two groups by examining some functionals of the unconditional distributions. Let $v : \mathcal{F} \to \mathbb{R}$ be a functional from a class of distribution functions to the real line. In this paper, we focus on τ^{th} quantile function, with $0 < \tau < 1$, $v : F \mapsto F^{-1}(\tau)$. Our goal is to measure and understand the *overall* difference in the functionals of the distributions for two groups for the target population $T = t$ (which can be all the members in each group or a selected population).

$$\Delta^v_{O|T=t} = v(F_{Y(1)|D=1,T=t}) - v(F_{Y(0)|D=0,T=t}) \tag{2.2}$$

The counterfactual distribution of $Y(0)|D = 1, T = t \sim F_{C|T=t}$ is defined as follows[5]

$$F_{C|T=t} = \int F_{Y(0)|X,D=1,T=t}(y|x, 1)dF^1_{X|T=t}(x) \tag{2.3}$$

In the case of the racial gap, D is an indicator for whether an individual is white or black, and Y her log wage offer is observed only if she participates in the labor market (i.e., $S = 1$). A "treatment" here is receiving a wage structure under which her characteristics would be valued in the labor market. One may consider a hypothetical counterfactual state where a black woman can be treated as a white woman in the labor market. The counterfactual distribution of wages, $F_{C|T=t}$, can be thought of as the distribution of wages for black women that would prevail, were the distribution of observable characteristics among blacks remain the same, but paid under the wage structure for whites.[6] We decompose the overall difference

[5] Note that this is different from the counterfactual distribution function of the outcome when the values of the covariates for the entire population are exogenously switched to a fixed value in Hsu et al. (2015)

[6] Note that following the convention in this literature, our definition of the counterfactual distribution is a simple counterfactual treatment. There could be alternative counterfacutals based on alternative states of the world, for instance, that involve general equilibrium effects ((Fortin et al., 2011)).

in the functionals of the distributions as follows

$$\Delta^v_{O|T=t} = v(F_{Y(1)|D=1,T=t}) - v(F_{Y(0)|D=0,T=t}) \quad (2.4)$$

$$= v(F_{Y(1)|D=1,T=t}) - v(F_{C|T=t}) + v(F_{C|T=t}) - v(F_{Y(0)|D=0,T=t}) \quad (2.5)$$

$$= \Delta^v_{S|T=t} + \Delta^v_{X|T=t} \quad (2.6)$$

The index $T = t$ is dropped when referring to the whole population as the target population. Equation (2.6) suggest that the difference between two distributions may be made of two terms: the first is typically called *structural effects*, reflecting differences in wage structures, $m(\cdot, \cdot)$, between groups, while the second term is includes *composition effects*, reflecting the effect of individual characteristics, X. As has been noted, (Fortin et al., 2011), our definition of counterfactuals allows a link to the decompositions in the treatment effects literature, and a discussion of identification issues, with the structural effects equivalent to the treatment effects on the treated.[7]

3. Identification of Structure and Composition Effects for the observed sample of workers

This section discusses how we can *point* identify the structure and composition effects for the selected sample. Without correction for sample selection less restrictive parametric models can be entertained. Our results allow for nonseparable models. Not surprisingly, our *point* identified results below align with the results reported in the treatment literature that addresses sample selection, such as Newey (2007) and Huber (2014) (While Huber (2014) focuses on other parameters of interest, the results here can be considered as an application of his results).

[7] As noted in (Fortin et al., 2011), treatment effects and decomposition methods do not "seek to recover behavioral relationships or "deep" parameters". In this sense, the decomposition of the overall difference in the outcome distributions can be considered an "accounting method". However, it is useful to delineate the relative importance of various factors in explaining a difference in outcomes. The results are useful indication of what explanations may be more important and worth further explorations.

Assumption 1 (Ignorability or Conditional Independence Assumption)**.** *Let* (D, X, U) *have a joint distribution. For all* x *in the support of* $\mathcal{X}$*:* $U \perp D | X = x$*.*

Assumption 2 (Overlapping Support)**.** *For all* x *in the support of* $\mathcal{X} : 0 < \Pr[D = 1|X = x] < 1$ *and* $\Pr[D = 1] > 0$*.*

Assumptions (1) and (2) are commonly invoked in the decomposition literature. The first assumption requires that, conditioned on observable characteristics, the distribution of unobservable characteristics such as ability should be uncorrelated with race (group membership). The assumption permits a causal interpretation of the decomposition terms. The second assumption ensures that there be an overlap in observable characteristics between the two groups, or $\mathcal{X}_1 \subset \mathcal{X}_0$ for the integral in Equation (2.3) to be well-defined. Note that as shown in Fortin et al. (2011), these assumptions are sufficient to identify the distributions of the outcomes of interest for each group and the counterfactual distributions. Consequently, distributional statistics are identified, and so are the structure and composition effects, absent sample selection. Formally,

Proposition 1 (Theorem 1 in Firpo et al. (2007): Inverse Probability Weighting)**.** *Under Assumptions (1) and (2)*

1. $F_{Y(d)}(y) = \mathbb{E}[\omega_d(D) \cdot \mathbf{1}[Y \leq y]]$, $d = (0, 1)$
2. $F_C(y) = \mathbb{E}[\omega_C(D, X) \cdot \mathbf{1}[Y \leq y]]$

The corresponding weighting functions are given by

$$\omega_1 = \frac{D}{p}, \omega_0 = \frac{1 - D}{1 - p}, \omega_C = \frac{1 - D}{p} \cdot \frac{\pi(X)}{1 - \pi(X)}$$

where $p = \Pr[D = 1]$ *and* $\pi(X) = \Pr[D = 1|X]$ *.*

Proof. A similar proof can be found in Firpo (2007). □

With sample selection, "independence" in Assumption (1) may fail to hold in the selected sample even though it may hold in the population (Angrist, 1997). As a result, we cannot identify the underlying wage functions even for the selected sample, as well as the counterfactual distributions with meaningful interpretations. We require further (but standard) assumptions to point identify the structure and composition effects even for the selected sample. The first set of assumptions is concerned with the selection mechanism, while the second set of assumptions is concerned with the exclusion restriction. These assumptions together allow us to control for the selection bias by simply conditioning on an additional variable in the conventional approach, the selection-propensity score, as defined below.

Assumption 3 (Selection Mechanism)**.** *The selection mechanism is given by*

$$S = \mathbf{1}[V \leq \Pi(X, Z)] \tag{3.1}$$

where $\Pi(\cdot)$ *is an unknown function,* V *is an unobservable error term that could be correlated with* U*, and its distribution,* $F_V(v)$ *is strictly monotonic.*[8]

Let $W \equiv (X, Z)$. Equation (3.1) implies that $p(W) \equiv \Pr[S = 1|X, Z] = F_V(\Pi(X, Z))$; this selection probability is often called the selection propensity score. Z is an exclusion restriction that satisfies the following conditions

Assumption 4 (Exclusion Restriction)**.**

1. *(Existence of Correlation)* $\mathbb{E}[Z \cdot S|X] \neq 0$
2. *(Conditional Independence)* $(U, V) \perp (D, Z)|X$

Assumption (4) requires that conditional on X, an excluded IV be correlated with the selection decision but independent of the unobservable error terms. While the first part of the assumption can be readily verified, the second is not. The validity of the exclusion restriction needs to be examined and verified. In our empirical analysis below, we will use a commonly used exclusion restriction from the literature (see data section below for detailed discussion of the variable and the rationale behind our choice). The second part of the assumption

[8] Note that X is assumed to be independent of both error terms throughout the paper.

further requires that conditional on X, the group membership, D, should be independent of the unobservable error terms, U and V. In the case of the racial gap, the unobservable determinants of the wages, U, and the unobservable determinants of employment, V, should be highly correlated and overlap to a large extent. This assumption can thus be considered as a natural extension of Assumption (1) that is typically imposed in the literature without sample selection.

Under Assumptions (3) and (4), the dependence between the unobservable error term (and hence the outcome of interest) and the group membership within the working sample is controlled for by conditioning on both the covariates, X and the selection propensity score, $p(W)$ in the selected sample. Formally, we can show that for any bounded function $\alpha(U)$,

$$\mathbb{E}[\alpha(U)|D, X, p(W), S] = \mathbb{E}[\alpha(U)|X, p(W), S] \tag{3.2}$$

See a proof in Newey (2007) and Huber (2014).

Since the dependence between the outcome of interest and group membership is controlled by conditioning on both X and $p(W)$, we also require that group membership not be perfectly predicted by these variables . Formally, we impose a common support requirement similar to Assumption (2):

Assumption 5 (Overlapping Support for the Selected Sample)**.** *For all $x, p(w)$ in the support of $\mathcal{X} \times P$, $0 < \Pr[D = 1|X = x, p(W) = p(w), S = 1] < 1$, and $\Pr[S = 1|D = d] > 0$, where $d = \{0, 1\}$.*

Note that the overlapping support assumptions in (2) and (5) can be weakened, as in the treatment effects literature. From the observed sample of (Y, D, X, Z, S), we can nonparametrically identify the marginal distributions of the outcomes for the selected sample, $F_{Y(d)|D=d,S}$. Under the assumptions above, we can also identify the counterfactual distribution for group 1 in the selected sample, $F_{C|S} = F_{Y(0)|D=1,S}$. This result is formally stated in the following proposition.

Proposition 2 (Inverse Probability Weighting). *Under Assumptions (3)-(5),*

1. $F_{Y(d)|D=d,S}(y|d,s) = \mathbb{E}[\omega_d(D) \cdot \mathbf{1}[Y \leq y]|S=1]$, $d = (0,1)$
2. $F_{C|S}(y|s) = \mathbb{E}[\omega_C(D,W) \cdot \mathbf{1}[Y \leq y]|S=1]$

The corresponding weighting functions are given by

$$\omega_1 = \frac{D}{p}, \omega_0 = \frac{1-D}{1-p}, \omega_C = \frac{1-D}{p} \cdot \frac{\pi(X,p(W))}{1-\pi(X,p(W))}$$

where $p = \Pr[D=1|S=1]$ *and* $\pi(X,p(W)) = \Pr[D=1|X,p(W),S=1]$.

Proof. Our proof is analogous to Firpo (2007) and thus omitted but available from the authors upon request. □

Comparing the results in Proposition (2) to the results in Proposition (1), it is seen that sample selection does not impact the identification and estimation of the marginal distributions of the outcome for observed working groups, but selection impacts the identification of the counterfactual distribution and, in turn, the identification of structural and composition effects. The difference between our approach and the usual reweighting method for identification of the counterfactual distribution lies in whether or not the propensity score for group membership controls for selection propensity score.

We need to estimate the selection propensity scores first. Following Huber (2014), we can augment the standard reweighting approach with an additional step. Specifically, we employ a four-step estimation procedure to estimate the distributional statistics, $v(F_{Y(1)|D=1,S=1}), v(F_{C|S=1}), v(F_{Y(0)|D=0,S=1})$. The first and second steps involve estimating various propensity scores, $p(W)$ and $\pi(X,p(W))$, and the related weighting functions, $\omega_1, \omega_0, \omega_C$, while the third step computes the distributional statistics directly from the reweighted samples. Specifically,

1. Step 1: Estimate either a parametric or nonparametric binary choice model of S on X, Z and obtained the predicted values as the estimates for $p(W)$. Among popular choices are a logistic or probit model, estimated via maximum likelihood.

2. Step 2: Again, estimate either a parametric or nonparametric binary choice model of D on $X, \widehat{p(W)}$ and obtain the predicted values as the estimates for $\pi(X, p(W))$.

3. Step 3: The first two weighting functions are estimated by their normalized sample analogs:[9]

$$\widehat{\omega_1(D_i)} = \frac{D_i}{\widehat{p}} / \sum_{i=1}^{N} \left[\frac{D_i}{\widehat{p}} \right] \quad \text{and} \quad \widehat{\omega_0(D_i)} = \frac{1-D_i}{1-\widehat{p}} / \sum_{i=1}^{N} \left[\frac{1-D_i}{1-\widehat{p}} \right]$$

where $\widehat{p} = \frac{1}{N} \sum_{i=1}^{N_s} D_i$. The weighting function for the counterfactual distribution can be similarly estimated as

$$\widehat{\omega_C(D_i)} = \frac{1-D_i}{\widehat{p}} \cdot \frac{\pi(X_i, \widehat{p(W_i)})}{1-\pi(X_i, \widehat{p(W_i)})} / \sum_{i=1}^{N} \left[\frac{1-D_i}{\widehat{p}} \cdot \frac{\pi(X_i, \widehat{p(W_i)})}{1-\pi(X_i, \widehat{p(W_i)})} \right]$$

4. Step 4: Obtain the quantiles of relevant distributions from the reweighted samples using the weights obtained in Step 3. Standard errors can be obtained via bootstrapping.[10]

4. Bounds on Structure and Composition Effects for the Whole Population

The previous section shows identification of the structure and composition effects for the sample of workers. We may be interested in the effects for the whole population. Without further restrictions we cannot identify the unconditional distributions, let alone the counterfactual distributions. It is, however, possible to construct bounds on these effects.

Let $\underline{b}_d(y)$ and $\bar{b}_d(y)$ be sharp lower and upper bounds for $F_{Y(d)|D=d,X}(y|d,x)$. For the τ-th quantile of the conditional distribution, its corresponding sharp lower and upper bounds are given by $\underline{y}^d \equiv \bar{b}_d^{-1}(\tau|x) = \underline{m}^d(x,\tau)$ and $\bar{y}^d \equiv \underline{b}_d^{-1}(\tau|x) = \overline{m}^d(x,\tau)$, respectively.

[9] Note that $\widehat{p}$ in the formula is not necessary and cancels out. We keep it here to facilitate the comparison to our results in Proposition 2, as well as the connection to the treatment effects literature (in which $\widehat{p}(\cdot)$ is often observation specific and cannot be cancelled out).

[10] Note that our estimand is directly linked to the treatment effects literature, in which the asymptotic normality is shown for nonparametric estimators in e.g., Firpo (2007), and for parametric estimators like ours in e.g., Huber (2014). In practice, we use bootstrap to obtain the standard errors, which also takes into account the first step estimation.

Proposition 3 (Bounds on Unconditional Quantile for Each Group)**.** *Let* $\overline{Y}(d) \equiv \overline{m}^d(X, U)$ *and* $\underline{Y}(d) \equiv \underline{m}^d(X, U)$ *be sharp upper and lower bounds for* $m^d(X, U)$*.* $F_{\overline{Y}(d)|D=d}$ *and* $F_{\underline{Y}(d)|D=d}$ *denote the distributions of the upper and lower bounds for group* $d = \{0, 1\}$*, respectively. Let* $0 < \tau < 1$ *, and the unconditional quantile for group* d *is defined as* $y^d_\tau = F^{-1}_{Y(d)|D=d}(\tau)$*. The corresponding sharp upper and lower bounds,* $\overline{y}^d_\tau$ *and* $\underline{y}^d_\tau$ *are given by:*

$$\overline{y}^d_\tau = \inf_{\overline{Y}} \left\{ F_{\overline{Y}(d)|D=d}(\overline{y}) \geq \tau \right\} \tag{4.1}$$

$$\underline{y}^d_\tau = \inf_{\underline{Y}} \left\{ F_{\underline{Y}(d)|D=d}(\underline{y}) \geq \tau \right\} \tag{4.2}$$

Proof. Note the following useful property: Let $\overline{y}_i$ and $\underline{y}_i$ be sharp upper and lower bounds for y_i. $y_1 \geq y_2 \implies \overline{y}_1 \geq \overline{y}_2$ and $\underline{y}_1 \geq \underline{y}_2$. In other words, the tight upper (lower) bound on y_1 cannot be lower than the tight upper (lower) bound on y_2. A proof of a similar property can be found in Lemma 2 in Lechner and Melly (2010). This property means that quantiles of the upper (lower) bound distribution correspond to the respective quantiles of the bounded distribution.

First, note that $F_{Y(d)|X}(y|x) = \int_0^1 \mathbf{1}[Y(d) \leq y]du = \int_0^1 \mathbf{1}[m^d(x, u) \leq y]du$, and $F_{Y(d)|D=d}(y) = \int \left[\int_0^1 \mathbf{1}[m^d(x, u) \leq y]du \right] dF^d_X$.

Second, let y^d_τ be the τ-th quantile such that $F_{Y(d)|D=d}(y^d_\tau) = \int \left[\int_0^1 \mathbf{1}[m^d(x, u) \leq y_\tau]du \right] dF^d_X = \tau$. This implies that $F_{\overline{Y}(d)|D=d}(\overline{y}^d_\tau) = \int \left[\int_0^1 \mathbf{1}[\overline{m}^d(x, u) \leq \overline{y}^d_\tau]du \right] dF^d_X \geq \tau$. That is, $\overline{y}^d_\tau \in \{\overline{y} : F_{\overline{Y}(d)|D=d}(\overline{y}) \geq \tau\}$.

Next, we show the result is true by contradiction. Suppose $\overline{y}^d_\tau \neq \inf\{\overline{y} : F_{\overline{Y}(d)D=d}(\overline{y}) \geq \tau\}$. There exists $\overline{y}' < \overline{y}^d_\tau$ such that $\overline{y}' = \inf_{\overline{Y}}\{F_{\overline{Y}(d)|D=d}(\overline{y}) \geq \tau\}$. Let y' be a value whose upper bound is $\overline{y}'$. It follows that $y' < y^d_\tau$, and that $\alpha \equiv F_{Y(d)|D=d}(y') = \int \left[\int_0^1 \mathbf{1}[m^d(x, u) \leq y']du \right] dF^d_X < \tau$.

For $F_{\overline{Y}(d)|D=d}(\overline{y}') \geq \tau$ and $F_{Y(d)|D=d}(y') < \tau$ to hold simultaneously, there must exist a set of values, $A \equiv \{Y : y' < Y < y^d_\tau$ whose upper bounds are also $\overline{y}'\}$. And $\int\limits_{Y \in A} \left[\int_0^1 [y' < m^d(x, u) < y^d_\tau]du \right] dF^d_X = \tau - \alpha$. Take $y'' = \max(A)$. Then, $F_{Y(d)|D=d}(y'') = \tau$, which is contradicting to that y^d_τ is the τ-th quantile.

We can similarly show the lower bound result holds as well. $\square$

Proposition 4 (Bounds on Unconditional Quantile of Counterfactual Distribution)**.** *Let $\overline{Y}(0) \equiv \overline{m}^0(X,U)$ and $\underline{Y}(0) \equiv \underline{m}^0(X,U)$ be sharp upper and lower bounds for $m^0(X,U)$. Let $0 < \tau < 1$, and the unconditional quantile of the counterfactual distribution is defined as $y^c_\tau = F_C^{-1}(\tau)$. The corresponding sharp upper and lower bounds, $\overline{y}^c_\tau$ and $\underline{y}^c_\tau$ are given by*

$$\overline{y}^c_\tau = \inf_{\overline{Y}}\{\overline{F}_C(\overline{y}) \geq \tau\} \tag{4.3}$$

$$\underline{y}^c_\tau = \inf_{\underline{Y}}\{\underline{F}_C(\underline{y}) \geq \tau\} \tag{4.4}$$

where $\overline{F}_C(y) = \int \left[\int_0^1 \mathbf{1}[\overline{m}^0(x,u) \leq y]du\right] dF^1_X$ *and* $\underline{F}_C(y) = \int \left[\int_0^1 \mathbf{1}[\underline{m}^0(x,u) \leq y]du\right] dF^1_X$.

Proof. The proof is similar to that of Proposition 3. $\square$

Note that these results are analogous to the results in the seminal work of Machado and Mata (2005b). It provides a convenient way to obtain bounds on unconditional quantiles based on bounds on conditional quantiles, and more important, to perform similar counterfactual exercises on bounds. As we will see below, we can combine some of the results on estimation of bound functions in Chandrasekhar et al. (2012) to easily incorporate economic theories to tighten bounds on various effects. One can proceed with a two-step procedure similar to Melly (2005):

1. Suppose that we have appropriate estimates of $\widehat{\overline{m}}^d(x,\tau)$ and $\widehat{\underline{m}}^d(x,\tau)$, which we will discuss below, for the whole quantile process with a grid of τ values. As suggested in Melly (2005) , when the estimation is based upon a large dataset, a smaller number of estimations should be estimated.
2. Estimations of the τ-the quantile of the sample $\{\{\widehat{\overline{m}}^d(x_i,\tau_j)\}^J_{j=1}\}^{N_d}_{i=1}$ and $\{\{\widehat{\underline{m}}^d(x_i,\tau_j)\}^J_{j=1}\}^{N_d}_{i=1}$ by weighting each "observation" by $(\tau_j - \tau_{j-1})$ will give us $\overline{y}^d_\tau$ and $\underline{y}^d_\tau$. The weights are not necessary if we use a equidistant grid of τ values. Similar estimations are

performed on the sample $\{\{\widehat{\overline{m}}^0(x_i, \tau_j)\}_{j=1}^J\}_{i=1}^{N_1}$ and $\{\{\underline{\widehat{m}}^0(x_i, \tau_j)\}_{j=1}^J\}_{i=1}^{N_1}$ will give us $\overline{y}_\tau^c$ and $\underline{y}_\tau^c$.

With these bounds, we can then construct the bounds on both structural effects and composition effects at τ-th quantile.

$$\begin{aligned} \underline{y}_\tau^1 - \overline{y}_\tau^c &\leq \Delta_S^v \leq \overline{y}_\tau^1 - \underline{y}_\tau^c \\ \overline{y}_\tau^c - \overline{y}_\tau^0 &\leq \Delta_X^v \leq \overline{y}_\tau^c - \underline{y}_\tau^0 \end{aligned} \tag{4.5}$$

5. Data

Data are from March Current Population Survey (CPS) (available at `http://cps.ipums.org`, a large nationally representative household data that contain detailed information on labor market outcomes such as wages and other characteristics needed for decomposition analysis. The data have been widely used in the literature to study inequality including the gender gap and the racial gap (e.g., Blau and Beller 1988; Mulligan and Rubinstein 2008; and Waldfogel and Mayer 2000). In our analysis we examine four distinct periods: 1976-1979, 1995-1999, 2000-2006, and 2007-2012. Year 1976 was the first year that information on weeks worked and hours worked is available in the March CPS. Based on these variables, we can obtain log hourly wages, the outcome of interest, measured by an individual's wage and salary income for the previous year divided by the number of weeks worked and hours worked per week. The wages are adjusted for inflation based on the 1999 CPI adjustment factors (available at `https://cps.ipums.org/cps/cpi99.shtml`). Also following the literature (e.g., Lemieux 2006), we exclude extremely low values of wages (less than one unit of the log wages). We also exclude those imputed wages for those respondents with missing information on their wages. The literature has shown that *inclusion* of these imputed wages in wage studies is "problematic" and recommended that excluding them can "largely eliminate the first-order distortions resulting from imperfect matching" (Bollinger and Hirsch 2013); (Hirsch and Schumacher 2004; Bollinger and Hirsch 2006).

Our sample consists of individuals aged between 18 and 64, excluding those living in group quarters. Following the literature, our definition of full-time workers is those 1) who work only for wages and salary, 2) who worked for more than 20 weeks (inclusive) in the previous year, and 3) who worked more than 35 hours per week in the previous year.

In our illustrative analysis, the vector, x, is a typical set of wage determinants, including age (capturing potential working experience), years of schooling, dummy variables for marital status, and regional dummies. Such specification of wage functions is the standard "human capital specification" in the literature (Blau and Kahn, 2016). Our choice of exclusion restrictions for the selection equation is whether there is a child under age 5 in the household. This is a popular choice in the literature. For example, Mulligan and Rubinstein (2008) use the number of children younger than six, interacted with marital status as variables determining employment, but excluded from the wage equation. Also noted in Machado (2012), the number of children is used as an explanatory variable in the shadow price function in Heckman (1974), "one of the seminal works on female selection", and an IV in the participation equation in Heckman (1980). The number of young children may affect women's reservation wages and their labor supply decisions because it could affect "the value of leisure" for women (Keane et al. 2011) and child-rearing is time consuming and costly. Huber and Mellace (2014) has also provided empirical tests supporting the validity of the exclusion restriction used in the sample selection models similar to our setting for the inverse probability weighting estimators.

6. Results

6.1. Decomposition for Employed Women only

We first consider the decomposition for employed women using the inverse probability weighting method in Firpo et al. (2007) without addressing the sample selection. As shown in Millimet and Tchernis (2009), flexible propensity score specifications tend to perform better in finite samples. In estimations of propensity scores, $\pi(\cdot)$, we include polynomial terms of age up to 4th order, polynomial terms of years of schooling up to 2nd order, dummy variables for marital status, dummy variables for regions, an interaction between age and schooling, interactions between age and all dummy variables, and interactions between schooling and all dummy variables.

The results are presented in Table (1), and their corresponding standard errors based on 299 bootstraps in Table (2). Columns (1) and (2) estimate the wages at various quantiles (τ = .10,.25,.50, .75 and .90) quantiles for white and black female workers, respectively. Column (5) presents the overall racial gap – the difference in the wages between black and white workers at the corresponding quantiles. We first notice that the racial gap is generally negative across the *whole* distribution of log wages among workers, and the differences are statistically significant. This result indicates that black women generally fare worse compare to white women in the labor market, consistent with the literature. However, the gap is not uniform, and there exists substantial heterogeneity across the distribution. For example, in the 70s, the racial gap ranges from -0.0858 (at the 10^{th} quantile) to .0068 (at the 90^{th} quantile). In other words, among women in the lower tail of the distribution, blacks earn at least 8.5 percent less than whites, while among women in the upper tail, blacks earn even slightly more than their white counterparts (the difference is not statistically significant). This result is similar to what was reported in early studies; for example, a report entitled *The Economic Status of Black Women* by the U.S. Commission on Civil Rights in 1990 indeed finds that, among working women with at least a high school education, black women earned more than their white counterparts. Existence of such substantial heterogeneity in the racial gap across the distribution highlights the importance to examine the issue beyond the mean level.

The pattern of heterogeneity across the distribution varies over time. In the 70s, we find that the gap decreases monotonically with respect to quantiles. This pattern is less clear in the 90s, with the gap fluctuating around 7.5 percent. In more recent periods (both 2000-2006 and 2007-2012), the pattern is completely reversed. We now find that the gap *increases* with respect to quantiles. In other words, the gap becomes larger among women in the upper tail of the wage distribution than women in the lower tail. During the most recession, the racial gap at the 90^{th} percentile is nearly twice as large as the racial gap at the 10^{th} percentile.

The changes in the pattern of heterogeneity across the distribution over time are due to the changes in the gap at each quantile. Over time, the racial gap among women in the lower tail ($\tau = .10$ or $.25$) has gradually narrowed or remain relatively stable, while the gap among women in the rest of the distribution has widened. The most drastic changes are concentrated in the upper tail of the distribution ($\tau = .90$). For example, for women at the 90^{th} quantile, the racial gap increases from no difference in the 70s, to -.0710 in the 90s, to -.1061 in the period of 2007-2012. Only over the past 20 years had there been about 50% increase in the gap.

Turning to decomposition, quantiles of the counterfactual distribution (without addressing selection) are presented in Column (3) of Table (1), and their implied structural and composition effects in Columns (6) and (9), respectively. Both structural and composition effects are generally negative, and statistically significant. In other words, black workers experience different returns to their individual characteristics than their white counterparts, and they also have different level of market-valued characteristics such as education and work experience. A negative structural effect indicates that had black women been paid the same return for the their characteristics, they could have higher wages. This suggests *potential* discrimination. For example, in the 70s, had black women at the 10^{th} percentile been paid the same for their characteristics as whites, their wages could have increased by at least 7 percent, holding everything else constant. A negative composition effect indicates that had black women possessed the same level of characteristics, they could have higher wages. This suggests black women may posses lower levels of market-valued characteristics, which, for instance, could be due to less investment in education and frequent interruption in accumulation of other forms of human capital.

While structural and composition effects can help to explain the existence of the racial gap, their importance varies across the distribution. For example, in the 70s, there is a monotonically declining importance of structural effects over the distribution. Specifically, structural effects could account for the majority of the gap among women in the lower tail of the distribution, but play a much smaller role in the upper tail. At the 10^{th} percentile, structural effects account for roughly 82 percent of the gap, while composition effects account for only 18 percent. In contrast, at the 75^{th} percentile, composition effects account for nearly all the gap. An interesting result is that during the 70s, the reason why we fail to observe a difference between blacks and whites at the 90^{th} percentile is not because women are paid the same for their characteristics or are similar in their characteristics. In fact, they are paid differently *and* possess different level of characteristics, but the structural and composition effects are in opposite direction!

During the 1990s, structural effects continue to be larger in the lower tail than in the upper tail of distribution, while we observe the opposite holds for the composition effects. This pattern becomes less clear in the period 2000-2006 when importance of structural effects exhibits an inverted U shape across the distribution. Structural effects explain only roughly 16 percent of the gap at the 10^{th} and 90^{th} percentiles, but more than 22 percent in the rest of the distribution.

During the period 2007-2012, the pattern observed for the 70s is reversed. Structural effects increase monotonically with quantiles, while the opposite is true for composition effects. Specifically, structural effects are least important for explaining the gap at the 10^{th} percentile (about 10 percent of the gap), but account for about 40 percent of the gap at the 90^{th} percentile. Over time, the importance of structural effects in explaining the gap has declined in the lower tail but increased for higher wage earners.

Turning to actual estimates of structural and composition effects, we find that over time, composition effects have become larger in magnitudes at every part of the distribution. In contrast, the evolution of the structural effect differs for different wage levels. Structural effects became smaller in terms of magnitude in the lower tail and became larger in the very upper tail. The increasing composition effects are offset by the decreasing structural effects to some extent for women in the lower tail. Together, the differing patterns for these two effects over time explain why we find that the racial gap continues to widen for those in the upper tail, while the gap narrowed in the lower tail.

To examine the selection affect on the counterfactual distribution and its decomposition terms among workers, we now turn to the inverse probability weighted estimators controlling for selection propensity scores. The estimates of the counterfactual distribution are presented in Column (4) of Table (1), and the corresponding structural and composition effects in Columns (7) and (10). While addressing selection does impact the estimates, the general trend observed above continues to hold. In terms of actual estimates, addressing selection has different impacts on the counterfactual distribution at different quantiles and over time. For example, in the 70s, the estimates of the quantiles of the counterfactual distribution are generally larger than the (biased) estimates without addressing selection, except in the lower tail $\tau = .10$. In contrast, addressing selection has little impact on the counterfactual distribution in the 90s. Addressing the selection leads to smaller estimates of the median wages in the counterfactual distribution, but the difference is not statistically significant at conventional levels.

Except for the 90s, accounting for selection affects the estimates of counterfactual structural and composition effects, and in different ways. To facilitate, we calculate the percentage changes in the actual estimates of structural and composition effects before and after accounting for selection in Columns (8) and (11). The impact of selection is generally larger for structural effects than the composition effects. Specifically, out of 11 cases where the estimates are impacted, percentage changes in structural effects are larger than the percentage changes in composition effects. For example, during the period 2000-2006, accounting for selection leads to an increase of 17.5 percentage change in the estimates of the structural effects at $\tau = .90$, but only 3.49 percent change in the estimates of composition effects. Selection changes the composition of the workers, and also affects the estimation of the underlying wage structure.

Selection (participation decision) has had a larger impact on the higher wages than for the lower tail, especially in more recent years (2000-2006 and 2007-2012). Failure to address selection underestimates the structural effects, but overestimates the composition effects in the upper tail. In other words, pay differences for the same level of market-valued characteristics are even larger for women in the upper tail, and high-earning black women may potentially experience even worse discrimination in the labor market than suggested by the results that ignore selectivity.

6.2. Bounds on Decomposition for the Distribution of Potential Wages for the Whole Population

The above findings are of interest, but do not consider women who are not full-time workers or do not even participate in the labor market. In general, we cannot obtain "point" estimates of the gap and the corresponding structure and composition effects for these larger populations. We now turn to the estimates of the "bounds" on these parameters of interest. We consider two types of bounds under different assumptions. The first set of bounds impose least restrictive assumptions and are conservative, while the second set of bounds assumes availability of an exclusion restriction for sample selection.

6.2.1. Worse Case Bounds

Manski (1998) shows that without any further information, the conditional distribution of the wage offers can be bounded as follows

$$F(y|x, S=1)\,\Pr[S=1|X=x] \leq F(y|x) \leq F(y|x, S=1) + Pr[S=0|x] \qquad (6.1)$$

Such bounds are often called worst bounds. The unknown distribution of interest is dominated (dominates) estimable distributions for the observed workers, adjusted by estimable probabilities. Following Horowitz and Manski (1995) we can obtain the inverse image of the upper and lower bound functions on the conditional distribution to obtain the bounds on the conditional quantile function as follows:

$$\underline{y}_\tau = \underline{m}(x,\tau) = \begin{cases} F^{-1}_{Y|X,S=1}(\frac{\tau - \Pr[S=0|x]}{\Pr[S=1|x]}|x, S=1) & \text{if } \tau \geq \Pr[S=0|x] \\ y_{\min} & \text{otherwise} \end{cases} \tag{6.2}$$

$$\overline{y}_\tau = \overline{m}(x,\tau) = \begin{cases} F^{-1}_{Y|X,S=1}(\frac{\tau - \Pr[S=0|x]}{\Pr[S=1|x]}|x, S=1) & \text{if } \tau \leq \Pr[S=1|x] \\ y_{\max} & \text{otherwise} \end{cases} \tag{6.3}$$

As shown in Chandrasekhar et al. (2012), instead of inverting the bounds on the conditional distributions, one may obtain the bounds on the corresponding quantiles, estimating the quantile functions for $\underset{\sim}{Y}^d = Y(d) \cdot S + y_{\min} \cdot (1-S)$ and $\tilde{Y}^d = Y(d) \cdot S + y_{\max} \cdot (1-S)$. The τ-th conditional quantile function of $\underset{\sim}{Y}^d$ is equivalent to $\underline{m}^d(X,\tau)$, and that of $\tilde{Y}^d$ to $\overline{m}^d(X,\tau)$. Following Chandrasekhar et al. (2012), we employ flexible specifications of both $\overline{m}^d(X,\tau)$ and $\underline{m}^d(X,\tau)$ by including polynomial terms of continuous variables, and interaction terms between each of them and other dummy variables.[11] Once these bound functions are obtained, we can use the results in Propositions (3) and (4) to obtain the bounds on both the quantiles of the marginal distributions of $Y(d)$, and those of the counterfactual distribution of Y_C. Using these bounds, we can then calculate the bounds on both structural and decomposition effects. The actual estimates of the bounds on unconditional quantiles are reported in Table (4). To facilitate the discussions, we also plot the results in Figures (1) and (2).

Examining the upper and lower bounds on the quantiles of the marginal distributions, the lower bounds are generally informative only in the upper tail of the distribution ($\tau >= .5$), while the upper bounds are informative in the lower tail of distribution ($\tau < .5$). This is consistent with the results above: the lower bounds for a specific value, x, are informative

[11] One may also employ alternative nonparametric quantile approach in Li and Racine (2008).

when $\tau \geq \Pr[S = 0|x]$, and the upper bounds are informative when $\tau \leq \Pr[S = 1|x]$. As a result, we can obtain informative upper *and* lower bounds only when $1 - \Pr[S = 1|x] \geq \tau \Pr[S = 1|x]$. And if we have a relatively low full-time employment rate among women (i.e.,$1 - \Pr[S = 1|x] > \Pr[S = 1|x]$), we can never obtain any informative bounds on both. Given the low full-time employment rates in the 70s for both black and white women (39.67 and 36.06 percent, respectively), it is then not surprising that we do not observe informative upper and lower bounds at the same time for either black or white women. As a result, we also fail to find informative bounds on the overall racial gap. The bounds include both large negative and positive values for the racial gap.

As the full-time employment increased over time, we do find more informative upper and lower bounds on central quantiles ($\tau = .5$). However, the bounds on the overall gap continue to be wide and uninformative. Without imposing further assumptions, we cannot provide a definite answer to whether black women perform better or worse than white women in the labor market, or to whether the racial gap has continued to widen or narrow over time.

Uninformative bounds on the overall gap, however, do not necessarily imply uninformative bounds on either structural or decomposition effects since the bounds also depend on the bounds on the counterfactual distribution. And we indeed find that the bounds on the counterfactual distributions are more narrow than the bounds on the distributions for whites. For example, in the 90s, the bounds on the 50^{th} quantile of the counterfactual distribution is $[1.4767, 6.2767]$, while that for the wage distribution among white women is $[1.3710, 7.1364]$. Despite this improvement, we again find the bounds on both structural and compositions effects contain both large negative and positive values. Whether or not structural effects or discrimination actually exists against black women is inconclusive.

6.2.2. Bounds in the Presence of an Exclusion Restriction

As we can see, the worst bounds are usually wide and not necessarily informative of the dynamics and the sources of the racial gap. Here we assume availability of an exclusion restriction for sample selection, as the literature often does, to tighten the bounds. In the presence of an exclusion restriction satisfying the independence assumption, we can show that

$$F(y|x,z,S=1)\Pr[S=1|x,z] \leq F(y|x) \leq F(y|x,z,S=1) + Pr[S=0|x,z] \tag{6.4}$$

We can similarly obtain the bounds on the conditional quantile function by replacing x in Equations (6.2) with x, z:

$$\underline{y}_\tau = \underline{m}(x,z,\tau) = \begin{cases} F^{-1}_{Y|X,Z,S=1}(\frac{\tau - \Pr[S=0|x,z]}{\Pr[S=1|x,z]}|x, zS=1) & \text{if } \tau \geq \Pr[S=0|x,z] \\ y_{\min} & \text{otherwise} \end{cases} \tag{6.5}$$

$$\overline{y}_\tau = \overline{m}(x,z,\tau) = \begin{cases} F^{-1}_{Y|X,Z,S=1}(\frac{\tau - \Pr[S=0|x,z]}{\Pr[S=1|x,z]}|x, S=1) & \text{if } \tau \leq \Pr[S=1|x,z] \\ y_{\max} & \text{otherwise} \end{cases} \tag{6.6}$$

As noted in Blundell et al. (2007) and Chandrasekhar et al. (2012), Equations (6.5) imply that for all values of z in its support, the bounds on the quantile functions are given by

$$\sup_z \quad \underline{m}(x,z,\tau) \leq m(x,\tau) \leq \inf_z \quad \overline{m}(x,z,\tau) \tag{6.7}$$

In practice, we again obtain estimate $\underline{m}(x,z,\tau)$ and $\overline{m}(x,z,\tau)$ by estimating the equivalent quantile functions for $\underset{\sim}{Y}^d$ and $\tilde{Y}^d$. Here, in addition to the previously included covariates, we also include the exclusion restriction and its interaction terms with other covariates. The bounds for a specific value, x, are the maximum and minimum of the estimated $\underline{m}(x,z,\tau)$ and $\overline{m}(x,z,\tau)$ over the support of z. The actual estimates of the bounds on unconditional quantiles are reported in Table (5).

The availability of the exclusion restriction indeed tightens the bounds on the quantiles of the marginal distributions and the counterfactual distributions. Some of the lower and upper bounds become more informative at the quantiles where the worst bounds take on only the minimum and maximum values of the support (for instance, the median wages for white women in the 70s). Such improvement also lead to tightened bounds on structural and composition effects. To facilitate the comparison to the worst bounds, we also report the percentage difference between two sets of bounds in the appendix. The extent of the improvement varies across the distribution. We find that the larger improvements generally occur in the lower half of the distribution ($\tau \leq 0.5$). For central quantiles $\tau = .5$, the improvements can be substantial. For example, during the period 2000-2006, both the upper and lower bounds on structural effects are tightened by at least 25%. However, despite the significantly tightened bounds, we continue to find that the bounds are wide and include both positive and negative values for both structural and composition effects. In other words, whether the gap actually exists in favor of white women, or if it does, whether the racial gap is a result of differences in wage structures or of differences in market-valued characteristics again remains inclusive.

7. Discussion

For working women, we find a clear pattern that the racial gap has been widening in the upper tail of the distribution, and narrowing in the lower tail. Since the 1990s, the differences in market-valued characteristics are a major part of the racial gap across the distribution. The widening gap between high-earning black and white women is attributable to the widening differences in *both* their wage structures *and* their market-valued characteristics, while the differences in wages structure component among low-earning women have gradually shrunk. These results seem to suggest that high-earning black women have experienced more discrimination in the labor market over time. The differing impact of structural effects in explaining the racial gap at different parts of the wage distribution suggests that policies aimed at improving pay structures may be most effective among high earners. Also, the increased importance of composition effects suggests that policies aimed at improving women's market-valued characteristics, such as education and work experience, could lead to potentially smaller racial inequality.

When we examine the distribution of potential wages for *all* women, the bounds are wide and uninformative. One cannot provide a definitive answer concerning the evolution and the sources of the racial gap. The inability to infer clear patterns is valuable and informative: even assuming the availability of an exclusion restriction, difficult to find and controversial in practice, one is still unable to obtain tight bounds on the sources of the racial gap. Our results suggest that any empirical consensus regarding the extent and the sources of the racial gap for all women is probably reached based on highly subjective assumptions. For example, another often-invoked assumption, positive selection into employment among women as proposed in Blundell et al. (2007) to potentially help to tighten the bounds. While plausible and potentially useful, its validity has been questioned and refuted, for example in Mulligan and Rubinstein, 2008 and Maasoumi and Wang, 2016).

8. Conclusions

We have offered an examination of the racial wage gap among women in the U.S, and its dynamics. This is complicated by sample selection, as workers self-select into the labor force, and we do not observe wages for women "out of work". To address this issue, we first provide nonparametric identification of the counterfactual distribution for working women by extending the commonly used inverse probability weighting method to accommodate sample selection. We provide decompositions of the observed racial gap into structural and composition effects at the distributional level. We then extend the basic framework to provide a computationally convenient approach to identify the bounds on the decomposition components. We illustrate how to incorporate different assumptions in such analysis. Our methodology is general and can be applied broadly to similar situations. Using the March CPS, we find that both structural and composition effects contribute to observed racial gap among working women, but we have very loose bounds on the importance of these effects among "all" women. Altogether, our results suggest that the empirical analysis of the racial gap is sensitive to the definition of the population of interest, and so are their policy implications.

References

Albrecht, James, Anders Björklund, and Susan Vroman. 2003. "Is There a Glass Ceiling in Sweden?" *Journal of Labor Economics* 21:145–177.

Albrecht, James, Aico van Vuuren, and Susan Vroman. 2009. "Counterfactual distributions with sample selection adjustments: Econometric theory and an application to the Netherlands." *Labour Economics* 16:383 – 396.

Albrecht, James, Aico Van Vuuren, and Susan Vroman. 2014. "Selection and the Measured Black-White Wage Gap Among Young Women Revisited." *IZA DP Series* 8005.

Altonji, J.G. and R.M. Blank. 1999. *Handbook of Labor Economics*, volume 3, chapter Race and Gender in the Labor Market, pp. 3143–3259. New York: Elsevier.

Angrist, Joshua D. 1997. "Conditional independence in sample selection models." *Economics Letters* 54:103–112.

Arellano, M. and S. Bonhomme. 2017. "Quantile Selection Models: with an Application to Understanding Changes in Wage Inequality." *Econometrica* 85:1–28.

Black, Dan A., Amelia M. Haviland, Seth G. Sanders, and Lowell J. Taylor. 2008. "Gender Wage Disparities among the Highly Educated." *Journal of Human Resources* 43:630–659.

Blau, F.D. and A.H. Beller. 1988. "Trends in Earnings Differentials By Gender, 1971-1981." *Industrial and Labor Relations Revew* 41:513–529.

Blau, F.D. and L.M. Kahn. 2016. "The Gender Wage Gap: Extent, Trends, and Explanations." *NBER Working Papers* .

Blinder, Alan S. 1973. "Wage Discrimination: Reduced Form and Structural Estimates." *The Journal of Human Resources* 8:436.

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

Bollinger, Christopher R. and Barry T. Hirsch. 2006. "Match Bias from Earnings Imputation in the Current Population Survey: The Case of Imperfect Matching." *Journal of Labor Economics* 24:483–519.

Bollinger, Christopher R. and Barry T. Hirsch. 2013. "Is Earnings Nonresponse Ignorable?" *The Review of Economics and Statistics* 95:407–416.

Buchinsky, M. 1998. "The Dynamics of Changes in the Female Wage Distribution in the U.S.A.: A Quantile Regression Approach." *Journal of Applied Econometrics* 13:1–30.

Chandrasekhar, Arun, Victor Chernozhukov, Francesca Molinari, and Paul Schrimpf. 2012. "Inference for best linear approximations to set identified functions." .

Chernozhukov, V. and C. Hansen. 2005. "An IV Model of Quantile Treatment Effects." *Econometrica* 73:245–261.

DiNardo, J., N. Fortin, and T. Lemieux. 1996. "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica* 64:1001–1044.

Doksum, Kjell. 1974. "Empirical Probability Plots and Statistical Inference for Nonlinear Models in the Two-Sample Case." *The Annals of Statistics* 2:267–277.

Firpo, S. 2007. "Efficient Semiparametric Estimation of Quantile Treatment Effects." *E-conometrica* 75:259–276.

Firpo, S., N. Fortin, and T. Lemieux. 2007. "Decomposing Wage Distributions using Recentered Influence Function Regressions." *Unpublished Manscript* .

Fortin, Nicole, Thomas Lemieux, and Sergio Firpo. 2011. "Chapter 1 C Decomposition Methods in Economics." In *Handbook of Labor Economics*, volume 4, pp. 1–102.

Galichon, A. 2016. *Optimal Transport Methods in Economics*. Princeton University Press.

Heckman, J.J. 1974. "Shadow Price, Market Wages, and Labor Supply." *Econometrica* 42:679–694.

Heckman, J.J. 1980. *Female Labor Supply: Theory and Estimation*, chapter Sample Selection Bias as a Specification Error with an Application to the Estimation of Labor Supply Functions. NJ: Princeton University Press.

Hirsch, Barry T. and Edward J. Schumacher. 2004. "Match Bias in Wage Gap Estimates Due to Earnings Imputation." *Journal of Labor Economics* 22:689–722.

Horowitz, Joel L. and Charles F. Manski. 1995. "Identification and Robustness with Contaminated and Corrupted Data." *Econometrica* 63:281.

Hsu, Y., K. Kan, and T. Lai. 2015. "Distribution and Quantile Structural Functions in Treatment Effect Models: Application to Smoking Effects on Wages." *IEAS Working Paper, Institute of Economics, Academia Sinica* Academic Research 15-A001.

Huber, M. 2014. "Treatment Evaluation in the Presence of Sample Selection." *Econometric Reviews* 33:869–905.

Huber, M. and G. Mellace. 2014. "Testing Exclusion Restrictions and Additive Separability in Sample Selection Models." *Empirical Economics* 2011-45.

Juhn, C., K.M. Murphy, and B. Pierce. 1993. "Wage Inequality and the Rise in Returns to Skill." *Journal of Political Economy* 101:410–442.

Keane, M.P., K.I. Wolpin, and P. Todd. 2011. *Handbook of Labor Economics*, volume 4, chapter The Structural Estimation of Behavioral Models: Discrete Choice Dynamic Programming Methods and Applications, pp. 331–461. Elsevier.

Lechner, Michael and Blaise Melly. 2010. "Partial Idendification of Wage Effects of Training Programs." *Working Papers* .

Lee, David S. 2009. "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects." *The Review of Economic Studies* 76:1071.

Lemieux, Thomas. 2006. "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?" *The American Economic Review* 96:461–498.

Li, Q. and J.S. Racine. 2008. "Nonparametric Estimation of Conditional CDF and Quantile Functions With Mixed Categorical and Continuous Data." *Journal of Business & Economic Statistics* 26:423–434.

Maasoumi, E. and L. Wang. 2016. "The Gender Gap Between Earnings Distributions." *Working Paper* .

Machado, C. 2012. "Selection, Heterogeneity and the Gender Wage Gap." *IZA Discussion Papers* 7005.

Machado, J. and J. Mata. 2005a. "Counterfactual Decompositions of Changes in Wage Distributions Using Quantile Regression." *Journal of Applied Econometrics* 20:445–65.

Machado, Jos A. F. and Jos Mata. 2005b. "Counterfactual decomposition of changes in wage distributions using quantile regression." *Journal of Applied Econometrics* 20:445–465.

Manski, C.F. 1998. "Anatomy of the Selection Problem." *Journal of Human Resources* 24:343–360.

Melly, Blaise. 2005. "Decomposition of differences in distribution using quantile regression." *Labour Economics* 12:577–590.

Melly, B. and M. Huber. 2008. "Sample Selection, Heteroskedasticity, and Quantile Regression." *Unpublished Manuscript* .

Melly, B. and M. Huber. 2011. "Nonseparable Sample Selection Models." *Unpublished Manuscript* .

Millimet, Daniel L. and Rusty Tchernis. 2009. "On the Specification of Propensity Scores, With Applications to the Analysis of Trade Policies." *Journal of Business & Economic Statistics* 27:397–415.

Mulligan, C.B. and Y. Rubinstein. 2008. "Selection, Investment, and Women's Relative Wages Over Time." *Quarterly Journal of Economics* 123:1061–1110.

Neal, D.A. 2004. "The Measured Black-White Wage Gap among Women Is Too Small." *Journal of Political Economy* S1:S1–S28.

Neal, Derek. 2006. "Chapter 9 Why Has BlackCWhite Skill Convergence Stopped?" *Handbook of the Economics of Education* 1:511–576.

Newey, W.K. 2007. "Nonparametric Countinuous/Discrete Choice Models." *International Economic Review* 48:1429–1439.

Oaxaca, R.L. 1973. "Male/Female Wage Differentials in Urban Labor Markets." *International Economoic Review* 14:693–709.

Sasaki, Yuya. 2014. "What Do Quantile Regressions Identify for General Structural Functions?" *Econometric Theory* 31:1–15.

Waldfogel, J. and S.E. Mayer. 2000. *Finding Jobs: Work and Welfare Reform*, chapter Gender Differences in the Low-Wage Labor Market, pp. 193–232. New York, New York: Russell Sage Foundation.

Table 1: Decomposition Using Inverse Probability Weighting Estimators (With and Without Selection Correction)

Periods	Quantile	Whites	Blacks	Counterfactual		Overall Gap	Structural Effects		Difference (Percentage)	Composition Effects		Difference (Percentage)
				Without	With		Without	With		Without	With	
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)=100* [(7)-(6)]/(6)	(9)	(10)	(11)= 100* [(10)-(9)]/(9)
1976-1979	10	1.7997	1.7140	1.7844	1.7844	-0.0858	-0.0704	-0.0704	0.0000	-0.0154	-0.0154	0.0000
	25	2.0467	1.9763	2.0266	2.0273	-0.0704	-0.0503	-0.0510	1.3670	-0.0201	-0.0194	-3.4199
	50	2.3485	2.2975	2.3191	2.3211	-0.0509	-0.0216	-0.0236	**9.3773**	-0.0294	-0.0273	**-6.8878**
	75	2.6460	2.6126	2.6037	2.6063	-0.0334	0.0089	0.0063	-29.4536	-0.0423	-0.0397	-6.2105
	90	2.9069	2.9137	2.8605	2.8647	0.0068	0.0532	0.0491	**-7.8626**	-0.0464	-0.0422	**-9.0203**
1995-1999	10	1.7351	1.6591	1.7028	1.7028	-0.0759	-0.0437	-0.0437	0.0000	-0.0323	-0.0323	0.0000
	25	2.0620	1.9927	2.0236	2.0236	-0.0693	-0.0309	-0.0309	0.0000	-0.0384	-0.0384	0.0000
	50	2.4404	2.3513	2.3877	2.3868	-0.0892	-0.0364	-0.0356	-2.2898	-0.0528	-0.0536	1.5801
	75	2.8031	2.7290	2.7518	2.7518	-0.0741	-0.0228	-0.0228	0.0000	-0.0513	-0.0513	0.0000
	90	3.1344	3.0635	3.0734	3.0734	-0.0710	-0.0099	-0.0099	0.0000	-0.0610	-0.0610	0.0000
2000-2006	10	1.8010	1.7516	1.7596	1.7583	-0.0494	-0.0080	-0.0067	-15.5621	-0.0414	-0.0427	3.0017
	25	2.1436	2.0811	2.1008	2.1008	-0.0625	-0.0197	-0.0197	0.0000	-0.0427	-0.0427	0.0000
	50	2.5132	2.4408	2.4569	2.4567	-0.0723	-0.0160	-0.0159	-0.7958	-0.0563	-0.0564	0.2267
	75	2.8814	2.8056	2.8259	2.8276	-0.0758	-0.0203	-0.0220	**8.4470**	-0.0555	-0.0537	**-3.0921**
	90	3.2358	3.1494	3.1638	3.1663	-0.0864	-0.0144	-0.0169	**17.5122**	-0.0720	-0.0695	**-3.4916**
2007-2012	10	1.8221	1.7649	1.7710	1.7710	-0.0572	-0.0061	-0.0061	0.0000	-0.0510	-0.0510	0.0000
	25	2.1697	2.0844	2.1064	2.1064	-0.0853	-0.0219	-0.0219	0.0000	-0.0633	-0.0633	0.0000
	50	2.5462	2.4610	2.4811	2.4811	-0.0852	-0.0201	-0.0201	0.0000	-0.0651	-0.0651	0.0000
	75	2.9273	2.8391	2.8719	2.8760	-0.0883	-0.0328	-0.0370	**12.6796**	-0.0555	-0.0513	**-7.5016**
	90	3.2920	3.1858	3.2263	3.2310	-0.1061	-0.0405	-0.0451	**11.4507**	-0.0657	-0.0610	**-7.0600**

[1] Data Source: Current Population Survey.
[2] The column "without" ("with") refers to the results without (with) selection correction. The cells shaded in gray are statistically *insignificant*, while the percentage differences in bold are statistically *significant* at 10 percent levels. Actual standard errors obtained based on 299 bootstraps are presented in the appendix.

Table 2: Bootstrapped Standard Errors for Decomposition using Inverse Probability Weighting Estimators (With and Without Selection Correction)

Periods	Quantile	Whites	Blacks	Counterfactual		Overall Gap	Structural Effects		Difference	Composition Effects		Difference
				Without	With		Without	With		Without	With	
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)=[(7)-(6)]	(9)	(10)	(11)=[(10)-(9)]
1976-1979	10	0.0041	0.0099	0.0044	0.0045	0.0107	0.0102	0.0102	0.0039	0.0040	0.0013	0.0013
	25	0.0026	0.0090	0.0040	0.0041	0.0096	0.0093	0.0092	0.0035	0.0036	0.0010	0.0010
	50	0.0033	0.0086	0.0046	0.0047	0.0089	0.0083	0.0084	0.0040	0.0041	0.0012	0.0012
	75	0.0019	0.0108	0.0036	0.0046	0.0110	0.0107	0.0108	0.0031	0.0040	0.0020	0.0020
	90	0.0030	0.0063	0.0063	0.0064	0.0070	0.0082	0.0082	0.0057	0.0058	0.0015	0.0015
1995-1999	10	0.0045	0.0076	0.0049	0.0048	0.0091	0.0090	0.0090	0.0048	0.0047	0.0011	0.0011
	25	0.0012	0.0071	0.0046	0.0044	0.0073	0.0078	0.0078	0.0042	0.0041	0.0013	0.0013
	50	0.0029	0.0067	0.0057	0.0056	0.0076	0.0075	0.0075	0.0054	0.0053	0.0014	0.0014
	75	0.0039	0.0063	0.0056	0.0052	0.0075	0.0073	0.0071	0.0056	0.0054	0.0012	0.0012
	90	0.0030	0.0096	0.0037	0.0032	0.0102	0.0096	0.0095	0.0037	0.0035	0.0016	0.0016
2000-2006	10	0.0032	0.0051	0.0051	0.0049	0.0059	0.0069	0.0068	0.0042	0.0041	0.0010	0.0010
	25	0.0027	0.0058	0.0027	0.0027	0.0062	0.0059	0.0059	0.0027	0.0027	0.0008	0.0008
	50	0.0031	0.0047	0.0041	0.0041	0.0056	0.0055	0.0054	0.0040	0.0040	0.0009	0.0009
	75	0.0029	0.0056	0.0032	0.0032	0.0061	0.0054	0.0054	0.0035	0.0035	0.0008	0.0008
	90	0.0019	0.0052	0.0047	0.0046	0.0055	0.0061	0.0060	0.0045	0.0043	0.0014	0.0014
2007-2012	10	0.0025	0.0070	0.0055	0.0054	0.0075	0.0086	0.0085	0.0053	0.0052	0.0014	0.0014
	25	0.0035	0.0075	0.0039	0.0038	0.0082	0.0075	0.0075	0.0040	0.0040	0.0012	0.0012
	50	0.0024	0.0065	0.0034	0.0034	0.0069	0.0064	0.0066	0.0033	0.0033	0.0010	0.0010
	75	0.0011	0.0063	0.0044	0.0044	0.0063	0.0068	0.0069	0.0042	0.0042	0.0018	0.0018
	90	0.0029	0.0065	0.0053	0.0054	0.0069	0.0072	0.0074	0.0049	0.0051	0.0020	0.0020

[1] The column “without” (“with”) refers to the results without (with) selection correction.
[2] Standard errors are obtained based on 299 bootstraps.

Table 3: IMPORTANCE OF STRUCTURAL VS COMPOSITION EFFECTS

			Actual Estimates				Share of the Gap			
Periods	Quantile	Overall Gap	Structural Effects		Composition Effects		Structural Effects		Composition Effects	
			Without	With	Without	With	Without	With	Without	With
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1976-1979	10	-0.0858	-0.0704	-0.0704	-0.0154	-0.0154	0.8207	0.8207	0.1793	0.1793
	25	-0.0704	-0.0503	-0.0510	-0.0201	-0.0194	0.7144	0.7242	0.2856	0.2758
	50	-0.0509	-0.0216	-0.0236	-0.0294	-0.0273	0.4235	0.4632	0.5765	0.5368
	75	-0.0334	0.0089	0.0063	-0.0423	-0.0397	-0.2672	-0.1885	1.2672	1.1885
	90	0.0068	0.0532	0.0491	-0.0464	-0.0422	7.7916	7.1790	-6.7916	-6.1790
1995-1999	10	-0.0759	-0.0437	-0.0437	-0.0323	-0.0323	0.5752	0.5752	0.4248	0.4248
	25	-0.0693	-0.0309	-0.0309	-0.0384	-0.0384	0.4456	0.4456	0.5544	0.5544
	50	-0.0892	-0.0364	-0.0356	-0.0528	-0.0536	0.4083	0.3990	0.5917	0.6010
	75	-0.0741	-0.0228	-0.0228	-0.0513	-0.0513	0.3081	0.3081	0.6919	0.6919
	90	-0.0710	-0.0099	-0.0099	-0.0610	-0.0610	0.1401	0.1401	0.8599	0.8599
2000-2006	10	-0.0494	-0.0080	-0.0067	-0.0414	-0.0427	0.1617	0.1365	0.8383	0.8635
	25	-0.0625	-0.0197	-0.0197	-0.0427	-0.0427	0.3157	0.3157	0.6843	0.6843
	50	-0.0723	-0.0160	-0.0159	-0.0563	-0.0564	0.2215	0.2198	0.7785	0.7802
	75	-0.0758	-0.0203	-0.0220	-0.0555	-0.0537	0.2680	0.2906	0.7320	0.7094
	90	-0.0864	-0.0144	-0.0169	-0.0720	-0.0695	0.1662	0.1954	0.8338	0.8046
2007-2012	10	-0.0572	-0.0061	-0.0061	-0.0510	-0.0510	0.1072	0.1072	0.8928	0.8928
	25	-0.0853	-0.0219	-0.0219	-0.0633	-0.0633	0.2572	0.2572	0.7428	0.7428
	50	-0.0852	-0.0201	-0.0201	-0.0651	-0.0651	0.2357	0.2357	0.7643	0.7643
	75	-0.0883	-0.0328	-0.0370	-0.0555	-0.0513	0.3717	0.4189	0.6283	0.5812
	90	-0.1061	-0.0405	-0.0451	-0.0657	-0.0610	0.3814	0.4251	0.6186	0.5749

[1] See notes in 1.

[2] The share of the gap is calculated as the ratio of the corresponding effects to the overall gap. For example, Column (6)=Column (2)/ Column (1).

Table 4: Bounds on Decomposition Terms (Worst Bounds)

Periods	Quantile	Whites		Blacks		Counterfactual		Overall Gap		Structural Effects		Composition Effects	
		Lower (1)	Upper (2)	Lower (3)	Upper (4)	Lower (5)	Upper (6)	Lower (7)	Upper (8)	Lower (9)	Upper (10)	Lower (11)	Upper (12)
1976-1979	10	1.0002	2.0819	1.0002	1.9475	1.0002	1.9636	-1.0817	0.9473	-0.9634	0.9473	-1.0817	0.9634
	25	1.0002	2.7200	1.0002	2.7216	1.0002	2.5042	-1.7198	1.7214	-1.5040	1.7214	-1.7198	1.5040
	50	1.0002	8.5717	1.0869	8.0391	1.0002	8.5717	-7.4848	7.0389	-7.4848	7.0389	-7.5716	7.5716
	75	2.0234	8.5717	1.9919	8.5717	2.0457	8.5717	-6.5799	6.5483	-6.5799	6.5260	-6.5260	6.5483
	90	2.5822	8.5717	2.5617	8.5717	2.5398	8.5717	-6.0100	5.9895	-6.0100	6.0319	-6.0319	5.9895
1995-1999	10	1.0011	1.9832	1.0011	1.8567	1.0011	1.9002	-0.9821	0.8555	-0.8990	0.8555	-0.9821	0.8990
	25	1.0011	2.4788	1.0011	2.3507	1.0011	2.3584	-1.4777	1.3495	-1.3572	1.3495	-1.4777	1.3572
	50	1.3710	7.1364	1.5147	5.3765	1.4767	6.2767	-5.6217	4.0055	-4.7620	3.8998	-5.6596	4.9056
	75	2.3598	10.4849	2.3245	10.4849	2.3328	10.4849	-8.1604	8.1251	-8.1604	8.1521	-8.1521	8.1251
	90	2.8580	10.4849	2.8049	10.4849	2.7929	10.4849	-7.6799	7.6268	-7.6799	7.6920	-7.6920	7.6268
2000-2006	10	1.0004	2.0556	1.0004	1.9385	1.0004	1.9622	-1.0552	0.9381	-0.9618	0.9381	-1.0552	0.9618
	25	1.0004	2.5373	1.0004	2.3684	1.0004	2.4002	-1.5369	1.3680	-1.3998	1.3680	-1.5369	1.3998
	50	1.4115	6.6213	1.7174	4.2540	1.5444	5.8218	-4.9039	2.8425	-4.1043	2.7096	-5.0769	4.4103
	75	2.4599	10.2548	2.4598	10.2548	2.4314	10.2548	-7.7950	7.7949	-7.7950	7.8234	-7.8234	7.7949
	90	2.9511	10.2548	2.9197	10.2548	2.8881	10.2548	-7.3351	7.3037	-7.3351	7.3667	-7.3667	7.3037
2007-2012	10	1.0001	2.0857	1.0001	1.9723	1.0001	1.9924	-1.0856	0.9722	-0.9923	0.9722	-1.0856	0.9923
	25	1.0001	2.6117	1.0001	2.4702	1.0001	2.4878	-1.6115	1.4701	-1.4876	1.4701	-1.6115	1.4876
	50	1.3608	6.4965	1.5683	5.0456	1.4219	6.1925	-4.9282	3.6849	-4.6242	3.6237	-5.0746	4.8318
	75	2.4602	8.7710	2.4272	8.7710	2.4145	8.7710	-6.3437	6.3108	-6.3437	6.3565	-6.3565	6.3108
	90	2.9904	8.7710	2.9284	8.7710	2.9097	8.7710	-5.8426	5.7806	-5.8426	5.8613	-5.8613	5.7806

[1] The column "lower" ("upper") refers to the lower (upper) bounds for the corresponding measures.

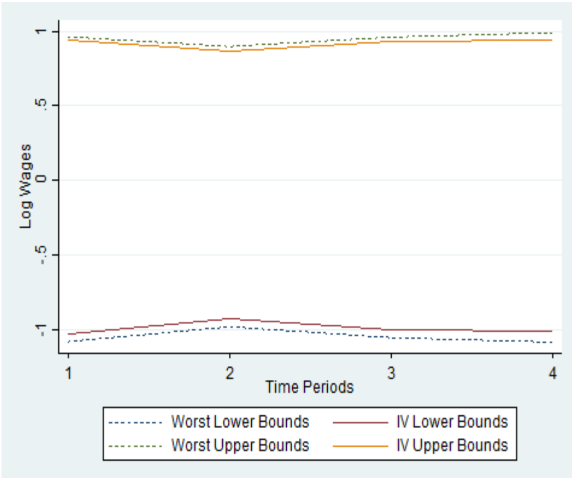
[2] The lower bound (column (7)) for the overall gap is calculated as the difference between the lower bound for blacks and the upper bound for whites. The upper bounds (column (8)) for the overall gap is calculated as the difference between the upper bound for black and the lower bound for whites. The upper and lower bounds for structural and composition effects are similarly calculated. See Equation (4.5) in the text for details.

Table 5: Bounds on Decomposition Terms (IV Bounds)

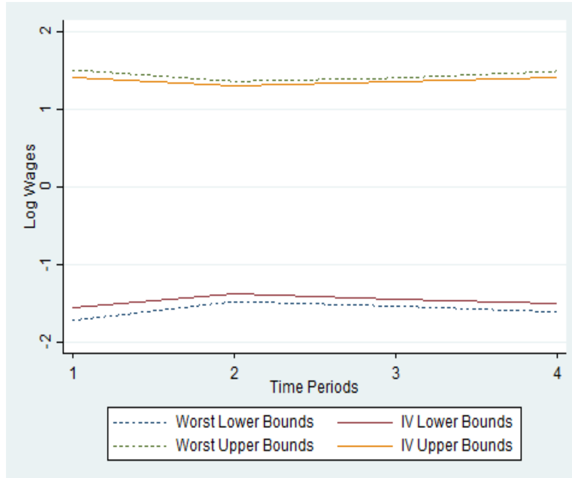
Periods	Quantile	Whites		Blacks		Counterfactual		Overall Gap		Structural Effects		Composition Effects	
		Lower (1)	Upper (2)	Lower (3)	Upper (4)	Lower (5)	Upper (6)	Lower (7)	Upper (8)	Lower (9)	Upper (10)	Lower (11)	Upper (12)
1976-1979	10	1.0002	2.0275	1.0002	1.4404	1.0002	1.9373	-1.0273	0.4402	-0.9371	0.4402	-1.0273	0.9371
	25	1.0002	2.5591	1.0002	2.1558	1.0002	2.4050	-1.5589	1.1556	-1.4048	1.1556	-1.5589	1.4048
	50	1.0349	8.4475	1.2763	7.1611	1.1138	7.9692	-7.1713	6.1261	-6.6929	6.0472	-7.3337	6.9343
	75	2.1446	8.5717	2.1130	8.5717	2.1316	8.5717	-6.4587	6.4271	-6.4587	6.4402	-6.4402	6.4271
	90	2.6351	8.5717	2.6962	8.5717	2.5756	8.5717	-5.8756	5.9366	-5.8756	5.9962	-5.9962	5.9366
1995-1999	10	1.0011	1.9292	1.0011	1.8216	1.0011	1.8678	-0.9280	0.8205	-0.8666	0.8205	-0.9280	0.8666
	25	1.0011	2.3864	1.0011	2.2806	1.0011	2.3042	-1.3853	1.2794	-1.3031	1.2794	-1.3853	1.3031
	50	1.5160	5.6737	1.5553	4.4874	1.5460	5.2529	-4.1184	2.9714	-3.6976	2.9414	-4.1277	3.7369
	75	2.4192	10.4849	2.3376	10.4849	2.3731	10.4849	-8.1473	8.0656	-8.1473	8.1117	-8.1117	8.0656
	90	2.8825	10.4849	2.8157	10.4849	2.8110	10.4849	-7.6692	7.6024	-7.6692	7.6739	-7.6739	7.6024
2000-2006	10	1.0004	2.0040	1.0004	1.7768	1.0004	1.9337	-1.0036	0.7764	-0.9333	0.7764	-1.0036	0.9333
	25	1.0004	2.4503	1.0004	2.2554	1.0004	2.3502	-1.4499	1.2550	-1.3498	1.2550	-1.4499	1.3498
	50	1.5878	5.1003	1.7936	3.6524	1.6394	4.8191	-3.3068	2.0646	-3.0256	2.0131	-3.4610	3.2313
	75	2.5162	10.2548	2.5309	10.2548	2.4652	10.2548	-7.7239	7.7386	-7.7239	7.7896	-7.7896	7.7386
	90	2.9879	10.2548	3.0016	10.2548	2.9137	10.2548	-7.2532	7.2669	-7.2532	7.3411	-7.3411	7.2669
2007-2012	10	1.0001	2.0156	1.0001	1.8669	1.0001	1.9442	-1.0154	0.8668	-0.9441	0.8668	-1.0154	0.9441
	25	1.0001	2.4996	1.0001	2.3807	1.0001	2.4130	-1.4995	1.3806	-1.4129	1.3806	-1.4995	1.4129
	50	1.4806	5.4876	1.6132	4.3110	1.4899	5.3983	-3.8744	2.8304	-3.7850	2.8211	-3.9977	3.9177
	75	2.4990	8.7710	2.4599	8.7710	2.4350	8.7710	-6.3110	6.2720	-6.3110	6.3360	-6.3360	6.2720
	90	3.0129	8.7710	2.9772	8.7710	2.9223	8.7710	-5.7938	5.7581	-5.7938	5.8487	-5.8487	5.7581

[1] The column "lower" ("upper") refers to the lower (upper) bounds for the corresponding measures.

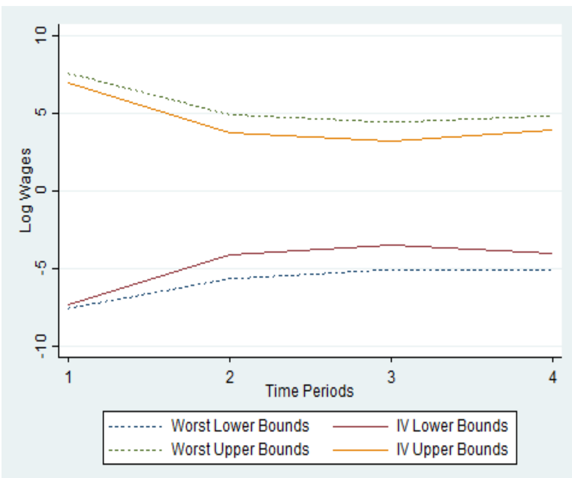
[2] The lower bound (column (7)) for the overall gap is calculated as the difference between the lower bound for blacks and the upper bound for whites. The upper bounds (column (8)) for the overall gap is calculated as the difference between the upper bound for black and the lower bound for whites. The upper and lower bounds for structural and composition effects are similarly calculated. See Equation (4.5) in the text for details.



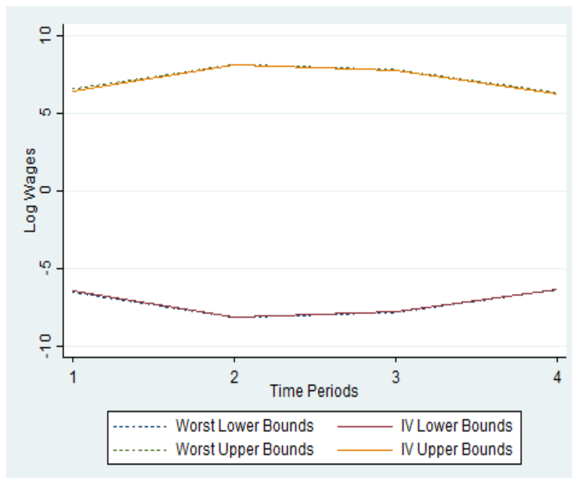
(a) $\tau = .10$



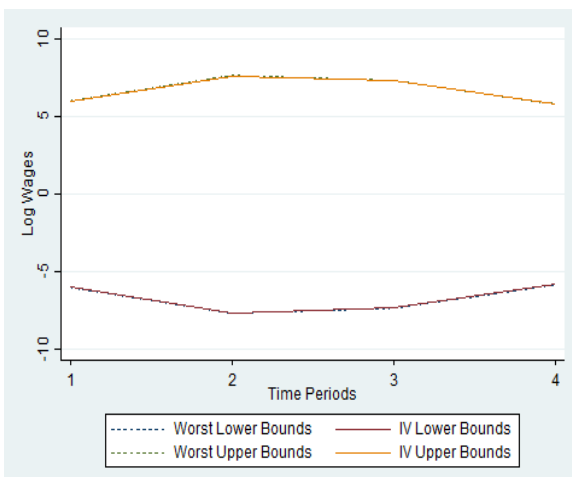
(b) $\tau = .25$



(c) $\tau = .50$

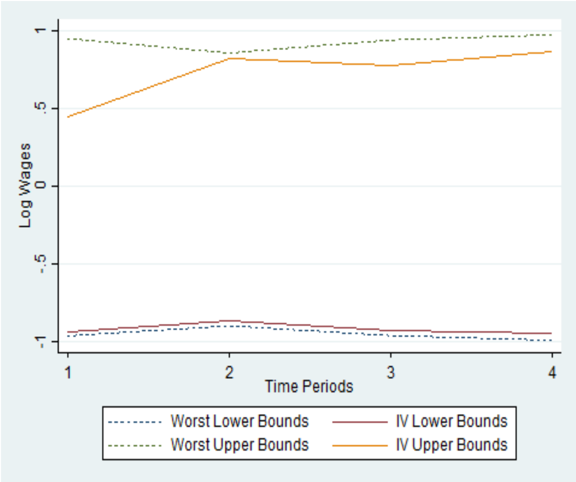


(d) $\tau = .75$

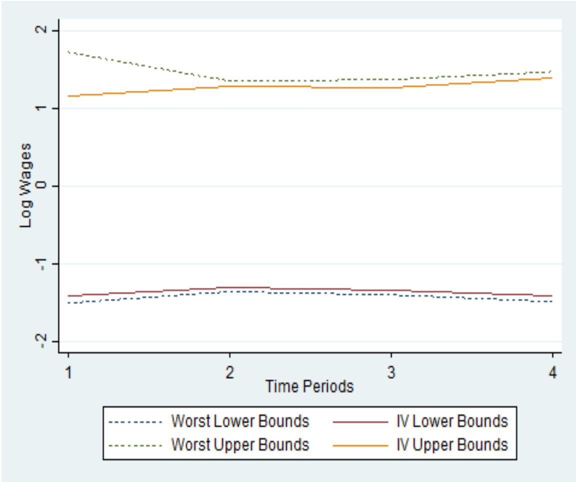


(e) $\tau = .90$

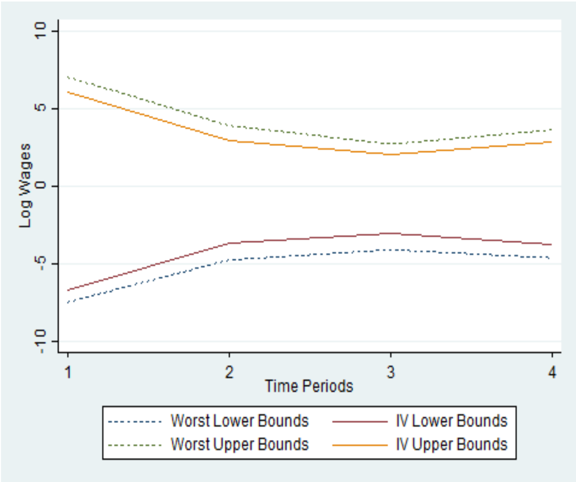
Figure 1: Bounds on Composition Effects



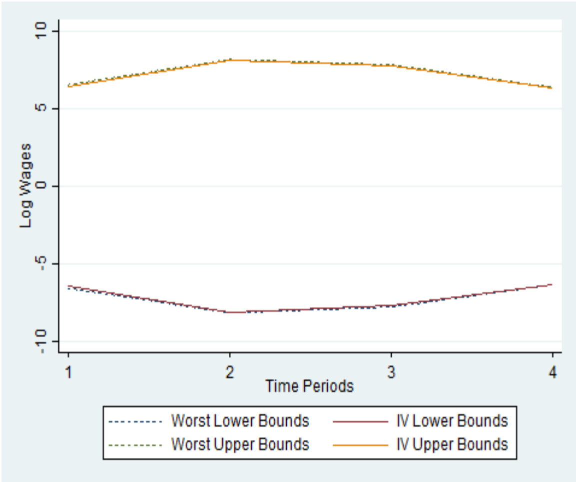
(a) $\tau = .10$



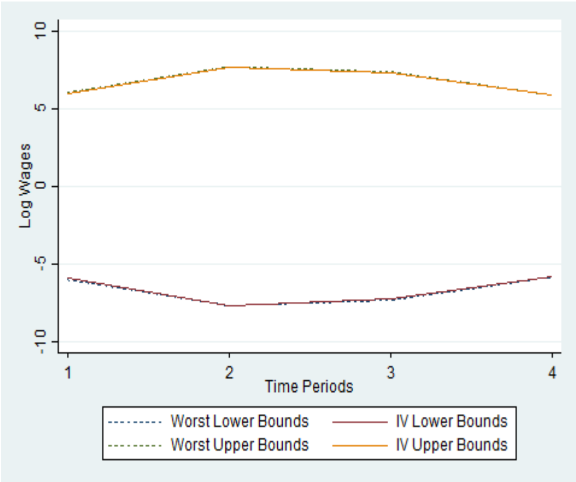
(b) $\tau = .25$



(c) $\tau = .50$



(d) $\tau = .75$



(e) $\tau = .90$

Figure 2: Bounds on Structural Effects