

IZA DP No. 9036

Average Wage Gaps and Oaxaca-Blinder Decompositions

Tymon Słoczyński

May 2015

Average Wage Gaps and Oaxaca–Blinder Decompositions

Tymon Słoczyński
*Warsaw School of Economics
and IZA*

Discussion Paper No. 9036
May 2015

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Average Wage Gaps and Oaxaca–Blinder Decompositions^{*}

In this paper I develop a new version of the Oaxaca–Blinder decomposition whose unexplained component recovers a parameter which I refer to as *the average wage gap*. Under a particular conditional independence assumption, this estimand is equivalent to the average treatment effect (ATE). I also provide treatment-effects reinterpretations of the Reimers, Cotton, and Fortin decompositions as well as estimate average wage gaps, average wage gains for men, and average wage losses for women in the United Kingdom. Conditional wage gaps increase across the wage distribution and therefore, on average, male gains are larger than female losses.

JEL Classification: C21, J31, J71

Keywords: decomposition methods, gender wage gaps, glass ceilings, treatment effects

Corresponding author:

Tymon Słoczyński
Department of Economics I
Warsaw School of Economics
ul. Madalinskiego 6/8 p. 228
02-513 Warszawa
Poland
E-mail: tymon.sloczynski@gmail.com

^{*} A previous version of this paper circulated under the title “Population Average Gender Effects” (IZA DP No. 7315). I am grateful to Krzysztof Karbownik, Michał Myck, and two anonymous referees for detailed comments on earlier drafts of this paper as well as to Arun Advani, Joshua Angrist, Anna Baranowska-Rataj, Thomas Crossley, Steven Haider, Patrick Kline, Mateusz Myśliwski, Ronald Oaxaca, Jörg Schwiebert, Gary Solon, Adam Szulc, Joanna Tyrowicz, Glen Waddell, Rudolf Winter-Ebmer, Jeffrey Wooldridge, and seminar and conference participants at many institutions for further comments and discussions. I acknowledge financial support from the National Science Centre (grant DEC-2012/05/N/HS4/00395), the Foundation for Polish Science (a START scholarship), and the “Weź stypendium—dla rozwoju” scholarship program. I would also like to thank the Clifford and Mary Corbridge Trust, the Cambridge European Trust, and the Faculty of Economics at the University of Cambridge for financial support which allowed me to undertake graduate studies at the University of Cambridge where this project was started. Data used in this paper come from the UK Data Archive at the University of Essex. It bears no responsibility, however, for data analysis or interpretation presented in this paper.

1 Introduction

Since the seminal contributions of Oaxaca (1973) and Blinder (1973), one of the most extensive strands of the literature in empirical labor economics has aimed at decomposing wage gaps into components attributable to group composition and net effects of group membership, often referred to as the explained component and the unexplained component, respectively. Recently, several researchers (Barsky *et al.*, 2002; Black *et al.*, 2006, 2008; Melly, 2006; Fortin *et al.*, 2011; Kline, 2011) have noted that the unexplained component in the most basic version of the Oaxaca–Blinder decomposition can sometimes be interpreted as the average treatment effect on the treated (ATT).¹ In this paper I extend this literature by deriving a new version of the Oaxaca–Blinder decomposition whose unexplained component can be interpreted as the average treatment effect (ATE), which is likely to be the primary object of interest in various empirical contexts. Because the potential outcome model (see, *e.g.*, Holland, 1986; Imbens and Wooldridge, 2009) is rarely invoked in the wage gap literature, I usually refer to this object as *the average wage gap*—an equivalent parameter which lacks a causal interpretation.

It is important to note that this new decomposition is distinct from previous versions of the so-called “generalized Oaxaca–Blinder decomposition” (Reimers, 1983; Cotton, 1988; Neumark, 1988; Oaxaca and Ransom, 1994; Fortin, 2008), although it easily fits into this class of models. Different members of this class are defined by the choice of *the comparison wage structure*—a counterfactual wage setting function with which all actual wages are compared. In this paper I study whether the average wage gap can be recovered with some version of the generalized Oaxaca–Blinder decomposition. I derive such a new model which uses a linear combination of the regression coefficients for both subpopulations (advantaged and disadvantaged workers) as the comparison wage structure. However, these coefficients are weighted in a nonstandard way, namely the population proportion of advantaged workers is used to weight the coefficients for disadvantaged workers, and vice versa. Clearly, such a weighting procedure may at first look counterintuitive.² Nevertheless, within the framework of this paper the role of each group’s wage structure is to serve as the counterfactual for the other group, and therefore we should

¹Other contributions to the decomposition literature have recently concentrated on semi- and nonparametric analogues of traditional Oaxaca–Blinder decompositions (Barsky *et al.*, 2002; Black *et al.*, 2006, 2008; Frölich, 2007; Mora, 2008; Nõpo, 2008) and extensions to other distributional statistics besides the mean (Juhn *et al.*, 1993; DiNardo *et al.*, 1996; Machado and Mata, 2005; Melly, 2005; Firpo *et al.*, 2007; Chernozhukov *et al.*, 2013).

²Note that a similar model had already been used by Duncan and Leigh (1985) in an application to union wage premiums, but such an approach was criticized—as “not a very intuitive procedure”—by Oaxaca and Ransom (1988).

indeed put more weight on the coefficients for the smaller group in order to recover the average wage gap. Note that a similar intuition can also be applied to provide a reinterpretation of the Reimers (1983), Cotton (1988), and Fortin (2008) decompositions. Each of these decompositions is easily shown to recover some generally uninteresting weighted average of conditional wage gaps.

As demonstrated by Fortin *et al.* (2011), identification of the explained component and the unexplained component requires a set of assumptions: simple counterfactual treatment, overlapping support, and conditional independence/ignorability. The assumptions of a simple counterfactual treatment and conditional independence together imply that the conditional wage distribution remains invariant to manipulations of the marginal distribution of covariates (“invariance of conditional distributions”). Therefore, it is possible to construct a counterfactual wage distribution which would be observed if disadvantaged workers were paid according to the wage setting function of advantaged workers, and vice versa. Then, the unexplained component can be interpreted as a treatment effect provided that the potential outcome model is also invoked—which might not be desirable, however, in the wage gap literature.

In this paper I also provide an empirical example which uses my new decomposition as well as other econometric methods to study gender wage gaps with the UK Labour Force Survey (LFS) data, for each year from 2002 to 2010. In particular, the standard parametric approach is complemented by normalized reweighting and a combination of stratification and different versions of the Oaxaca–Blinder decomposition. Importantly, I provide separate estimates of the average wage gap, the average wage gain for men, and the average wage loss for women. This is the first paper to clarify the distinction between these parameters and provide separate estimates for each of them.

2 Theory

2.1 Framework and Notation

Consider a population which is divided into two mutually exclusive groups, indexed by $d_i \in \{0, 1\}$ and referred to as the advantaged group ($d_i = 1$) and the disadvantaged group ($d_i = 0$). For each unit i , we also observe a (log) wage, y_i , and a row vector of covariates, X_i . In that case, $E[y_i \mid X_i = x, d_i = 1]$ is the expected (log) wage of an advantaged worker with observed characteristics $X_i = x$ and $E[y_i \mid X_i = x, d_i = 0]$ is the expected (log) wage of a disadvantaged worker with these characteristics. Moreover, define *the conditional wage gap* to be $\tau(x) = E[y_i \mid X_i = x, d_i = 1] - E[y_i \mid X_i = x, d_i = 0]$, *i.e.* the

gap between the expected (log) wages of an advantaged worker and a disadvantaged worker with $X_i = x$. Dependent on the question we wish to answer, we may average $\tau(X_i)$ over the whole population, over the subpopulation of advantaged workers or over the subpopulation of disadvantaged workers. Define *the average wage gap* to be:

$$\tau_{gap} = E[\tau(X_i)]. \quad (1)$$

Within the framework of a potential outcome model, and under additional assumptions, this parameter is equivalent to the average treatment effect. Moreover, define *the average wage gain for advantaged workers* and *the average wage loss for disadvantaged workers* to be:

$$\tau_{gain} = E[\tau(X_i) \mid d_i = 1] \quad \text{and} \quad \tau_{loss} = E[\tau(X_i) \mid d_i = 0], \quad (2)$$

respectively. Similarly, under certain conditions, these parameters can be regarded as equivalents of the average treatment effect on the treated and the average treatment effect on the controls. It is also the case that:

$$\tau_{gap} = P[d_i = 1] \cdot \tau_{gain} + P[d_i = 0] \cdot \tau_{loss}. \quad (3)$$

Thus, a particular weighted average of the average wage gain for advantaged workers and the average wage loss for disadvantaged workers is equal to the average wage gap.

It is important to note that without further assumptions $\tau(x)$, τ_{gap} , τ_{gain} , and τ_{loss} cannot be interpreted as causal or counterfactual; they are also identified from the data. As demonstrated by Fortin *et al.* (2011), a counterfactual interpretation can be justified by a set of three additional assumptions: simple counterfactual treatment, overlapping support, and conditional independence/ignorability. These assumptions are discussed below for completeness.

Assumption 1 (Simple Counterfactual Treatment). *The observed wage structure of advantaged (disadvantaged) workers represents a counterfactual wage structure for disadvantaged (advantaged) workers.*

This assumption restricts the analysis to counterfactuals which are based on the observed wage structure for the other group. In other words, the observed wage structure of advantaged workers provides a counterfactual for disadvantaged workers, and vice versa. It is important to note that this assumption rules out the presence of general equilibrium effects, and this might be a substantial restriction in some empirical contexts.

Assumption 2 (Overlapping Support). Define ϵ_i to be the worker's unobserved characteristics. Then, $0 < P[d_i = 1 \mid X_i = x, \epsilon_i = e] < 1$.

The overlapping support assumption ensures that no combination of observed and unobserved characteristics can be used to identify group membership. This restriction has recently become controversial in the gender wage gap literature, because “a significant number of males exhibit[s] a set of characteristics that have no female counterparts, and these characteristics are highly rewarded in the labor markets” (Ñopo, 2008); clearly, similar problems can also arise in other empirical contexts.

Assumption 3 (Conditional Independence/Ignorability). Let (d_i, X_i, ϵ_i) have a joint distribution. Then, $d_i \perp \epsilon_i \mid X_i$, i.e. the worker's unobserved characteristics are independent of group membership, conditional on observed covariates.

This assumption rules out the presence of unobserved characteristics which would be correlated with both group membership and wages, conditional on observed covariates. Such a requirement is potentially problematic in the case of gender wage gaps, because in most studies there are several omitted variables which have been shown to be correlated with both gender and wages. These unobserved covariates include college major (Brown and Corcoran, 1997; Loury, 1997; Machin and Puhani, 2003; Black *et al.*, 2008), propensity to negotiate wages (Leibbrandt and List, 2014), test scores (Blackburn, 2004), gender role attitudes (Fortin, 2005), absenteeism (Ichino and Moretti, 2009), and various personality traits (Mueller and Plug, 2006; Fortin, 2008; Manning and Swaffield, 2008).³

It is important to note that each of these three assumptions is potentially controversial, but they are still required to disentangle the explained component and the unexplained component (Fortin *et al.*, 2011). If we maintain these assumptions, it becomes possible to construct a counterfactual wage distribution which would be observed if disadvantaged workers were paid according to the wage structure of advantaged workers, and vice versa. This counterfactual experiment provides a meaningful interpretation of τ_{gap} , τ_{gain} , and τ_{loss} . The average wage gap, τ_{gap} , is equal to the difference between mean wages in two counterfactual distributions: in the first distribution, all workers are paid according to the wage setting function of advantaged workers; in the second distribution, all workers are paid according to the wage setting function of disadvantaged workers.

³Of course, some form of endogeneity might also arise if there are unobserved covariates with different correlation patterns. However, as demonstrated by Fortin *et al.* (2011), identification of the explained and unexplained components is not threatened unless the conditional independence assumption is violated.

Similarly, the average wage gain for advantaged workers, τ_{gain} , is equal to the average gap between actual wages of advantaged workers and their counterfactual wages which would be observed if these workers were paid according to the wage structure of disadvantaged workers. Analogously, the average wage loss for disadvantaged workers, τ_{loss} , is equal to the average gap between counterfactual wages of disadvantaged workers which would be observed if these workers were paid according to the wage structure of advantaged workers and their actual wages. Although τ_{loss} might be the most intuitive estimand in some contexts, the decomposition literature has often been concerned with both gains and losses (see, *e.g.*, Fortin, 2008), and therefore τ_{gap} and τ_{gain} are also usually interesting. Especially, the average wage gap—a noncausal equivalent of the average treatment effect—is likely to be the primary object of interest in many empirical studies.

2.2 Oaxaca–Blinder Decompositions

Let the model for outcomes be linear and separable in observed and unobserved characteristics, and allow the regression coefficients to be different for both groups of interest:

$$y_i = X_i\beta_1 + v_{1i} \quad \text{if } d_i = 1 \quad \text{and} \quad y_i = X_i\beta_0 + v_{0i} \quad \text{if } d_i = 0. \quad (4)$$

Also, $E[v_{1i} | X_i, d_i] = E[v_{0i} | X_i, d_i] = 0$. The raw wage gap, $E[y_i | d_i = 1] - E[y_i | d_i = 0]$, can then be decomposed as:

$$\begin{aligned} E[y_i | d_i = 1] - E[y_i | d_i = 0] &= E[X_i | d_i = 1] \cdot (\beta_1 - \beta_0) \\ &\quad + (E[X_i | d_i = 1] - E[X_i | d_i = 0]) \cdot \beta_0, \end{aligned} \quad (5)$$

where the first element, $E[X_i | d_i = 1] \cdot (\beta_1 - \beta_0)$, reflects intergroup differences in regression coefficients, and is often referred to as *the unexplained component*, while the second element, $(E[X_i | d_i = 1] - E[X_i | d_i = 0]) \cdot \beta_0$, reflects intergroup differences in mean covariate values, and is often referred to as *the explained component*. Similarly:

$$\begin{aligned} E[y_i | d_i = 1] - E[y_i | d_i = 0] &= E[X_i | d_i = 0] \cdot (\beta_1 - \beta_0) \\ &\quad + (E[X_i | d_i = 1] - E[X_i | d_i = 0]) \cdot \beta_1. \end{aligned} \quad (6)$$

The difference between Equations 5 and 6 rests upon using alternate comparison coefficients to calculate the explained component as well as measuring the distance between the regression functions, $\beta_1 - \beta_0$, for a different set of covariate values. Moreover, Equations 5 and 6 recover the average wage gain for advantaged workers and the average

wage loss for disadvantaged workers, respectively:

$$\tau_{gain} = E[X_i | d_i = 1] \cdot (\beta_1 - \beta_0) \quad \text{and} \quad \tau_{loss} = E[X_i | d_i = 0] \cdot (\beta_1 - \beta_0). \quad (7)$$

There has been a long-lasting tendency in the decomposition literature to claim that the choice of the comparison group in this context (choosing between Equations 5 and 6) is necessarily ambiguous. The standard response has been to suggest alternative wage structures to solve this comparison group choice problem. Such an approach is referred to as “generalized Oaxaca–Blinder”, and it involves an alternative decomposition:

$$\begin{aligned} E[y_i | d_i = 1] - E[y_i | d_i = 0] &= E[X_i | d_i = 1] \cdot (\beta_1 - \beta^*) + E[X_i | d_i = 0] \cdot (\beta^* - \beta_0) \\ &\quad + (E[X_i | d_i = 1] - E[X_i | d_i = 0]) \cdot \beta^*, \end{aligned} \quad (8)$$

where β^* is the set of comparison coefficients, typically referred to as the “nondiscriminatory” or “competitive” wage structure. Note that if $\beta^* = \beta_1 = \beta_0$, then there is no unexplained component, because $\beta_1 = \beta_0$ implies that all workers are paid according to the same wage structure. Also, some authors provide a simple interpretation of the two elements of the unexplained component in Equation 8. Namely, $E[X_i | d_i = 1] \cdot (\beta_1 - \beta^*)$ is sometimes interpreted as “the amount by which ... productivity characteristics [of advantaged workers] are overvalued”, in which case $E[X_i | d_i = 0] \cdot (\beta^* - \beta_0)$ is interpreted as “the amount by which ... productivity characteristics [of disadvantaged workers] are undervalued” (Cotton, 1988). These two objects, in general, do not overlap with τ_{gain} and τ_{loss} . However, for some choices of β^* , they are very closely related.

Several influential papers have been devoted to suggesting alternative sets of comparison coefficients for Equation 8, and these coefficients have often been formulated as $\beta^* = \lambda \cdot \beta_1 + (1 - \lambda) \cdot \beta_0$ where $\lambda \in [0, 1]$ is a weighting factor. If $\lambda = 0$, then disadvantaged workers are used as reference, $\beta^* = \beta_0$, and Equation 8 simplifies to Equation 5. Similarly, if $\lambda = 1$, then advantaged workers are used as reference, $\beta^* = \beta_1$, and Equation 8 simplifies to Equation 6. Alternatively, Reimers (1983) suggested $\lambda = \frac{1}{2}$ and Cotton (1988) suggested $\lambda = P[d_i = 1]$, the population proportion of advantaged workers. Moreover, Neumark (1988) developed a simple model of Beckerian discrimination, and showed that identification of the nondiscriminatory wage structure is ensured, for example, if the utility function of the representative producer is homogeneous of degree zero with respect to labor inputs of advantaged and disadvantaged workers. Such a wage structure can be approximated by regression coefficients in a pooled model which excludes group membership (Neumark, 1988). Although this solution to the comparison group choice problem has been the most popular alternative to the basic Oaxaca–Blinder

decomposition (Weichselbaumer and Winter-Ebmer, 2005), it has been criticized by both Fortin (2008) and Elder *et al.* (2010), since exclusion of the group membership dummy can bias coefficients on other covariates which also affects the unexplained component. Therefore, Fortin (2008) has proposed to use a pooled model including group membership as the comparison wage structure. As noted by Fortin (2008) and Fortin *et al.* (2011), the unexplained component in such a decomposition is equal (by construction) to the coefficient on the group membership dummy in a pooled regression.

2.3 Oaxaca–Blinder and the Average Wage Gap

In this subsection I provide an alternative solution to the comparison group choice problem by developing a new version of the Oaxaca–Blinder decomposition whose unexplained component recovers the average wage gap. This model also uses Equation 8 and a comparison wage structure which is constructed as a specific linear combination of the regression coefficients for both subpopulations of interest.

Proposition 1 (Oaxaca–Blinder and the Average Wage Gap). *If $\beta^* = P[d_i = 0] \cdot \beta_1 + P[d_i = 1] \cdot \beta_0$, then the unexplained component of the Oaxaca–Blinder decomposition in Equation 8 recovers the average wage gap.*

Proof. Combine Equation 8 with $\beta^* = P[d_i = 0] \cdot \beta_1 + P[d_i = 1] \cdot \beta_0$ and reformulate:

$$\begin{aligned}
E[y_i | d_i = 1] - E[y_i | d_i = 0] &= E[X_i | d_i = 1] \cdot (\beta_1 - \beta^*) + E[X_i | d_i = 0] \cdot (\beta^* - \beta_0) \\
&\quad + (E[X_i | d_i = 1] - E[X_i | d_i = 0]) \cdot \beta^* \\
&= E[X_i | d_i = 1] \cdot (\beta_1 - (P[d_i = 0] \cdot \beta_1 + P[d_i = 1] \cdot \beta_0)) \\
&\quad + E[X_i | d_i = 0] \cdot ((P[d_i = 0] \cdot \beta_1 + P[d_i = 1] \cdot \beta_0) - \beta_0) \\
&\quad + (E[X_i | d_i = 1] - E[X_i | d_i = 0]) \cdot \beta^* \\
&= P[d_i = 1] \cdot E[X_i | d_i = 1] \cdot (\beta_1 - \beta_0) \\
&\quad + P[d_i = 0] \cdot E[X_i | d_i = 0] \cdot (\beta_1 - \beta_0) \\
&\quad + (E[X_i | d_i = 1] - E[X_i | d_i = 0]) \cdot \beta^* \\
&= P[d_i = 1] \cdot \tau_{gain} + P[d_i = 0] \cdot \tau_{loss} \\
&\quad + (E[X_i | d_i = 1] - E[X_i | d_i = 0]) \cdot \beta^* \\
&= \tau_{gap} + (E[X_i | d_i = 1] - E[X_i | d_i = 0]) \cdot \beta^*. \quad \square
\end{aligned}$$

Although using the population proportion of advantaged workers to weight the coefficients for disadvantaged workers and using the population proportion of disadvantaged

workers to weight the coefficients for advantaged workers may at first look counterintuitive, each of the wage structures plays a clearly defined role in such a decomposition—it serves as the counterfactual for the other group (see Assumption 1). This is exactly the reason why more weight should be put on the wage structure of the smaller group which is used to provide the counterfactual for the larger one.

Interestingly, this alternative decomposition is equivalent to a flexible linear regression model for the average treatment effect, presented in Imbens and Wooldridge (2009) and Wooldridge (2010). The average treatment effect can be recovered as the coefficient on d_i in the regression of y_i on 1, d_i , X_i , and $d_i \cdot (X_i - E[X_i])$. Imbens and Wooldridge (2009) have noted that such a model can alternatively be written as:

$$\begin{aligned} \tau_{ATE} = & E[y_i | d_i = 1] - E[y_i | d_i = 0] \\ & - (P[d_i = 0] \cdot \beta_1 + P[d_i = 1] \cdot \beta_0) \cdot (E[X_i | d_i = 1] - E[X_i | d_i = 0]), \quad (9) \end{aligned}$$

which is equivalent to the Oaxaca–Blinder decomposition in Proposition 1. Similarly, the unexplained component of the Oaxaca–Blinder decomposition in Equation 5 is equal to the coefficient on d_i in the regression of y_i on 1, d_i , X_i , and $d_i \cdot (X_i - E[X_i | d_i = 1])$ and the unexplained component of the Oaxaca–Blinder decomposition in Equation 6 is equal to the coefficient on d_i in the regression of y_i on 1, d_i , X_i , and $d_i \cdot (X_i - E[X_i | d_i = 0])$.

The logic and approach of Proposition 1 applies also to the well-known versions of the Oaxaca–Blinder decomposition in Reimers (1983), Cotton (1988), and Fortin (2008). It can be easily verified that (i) the unexplained component of the Reimers (1983) decomposition is equal to the arithmetic average of the average wage gain for advantaged workers and the average wage loss for disadvantaged workers; (ii) the unexplained component of the Cotton (1988) decomposition is equal to a weighted average of the average wage gain for advantaged workers and the average wage loss for disadvantaged workers, with reversed weights attached to both these parameters;⁴ and (iii) the unexplained component of the Fortin (2008) decomposition is approximately equal to the same parameter. This last interpretation is based on the similarity between the unexplained component of the Cotton (1988) decomposition and the coefficient on the group membership dummy in a simple linear regression (Elder *et al.*, 2010) and a related reinterpretation of the linear regression estimand in the presence of heterogeneous treatment effects (Słoczyński, 2014). In consequence, whenever we are concerned with heterogeneity in conditional wage gaps, there is good reason to choose the new version of the Oaxaca–Blinder decompo-

⁴In other words, the proportion of disadvantaged workers is used to weight the average wage gain for advantaged workers and the proportion of advantaged workers is used to weight the average wage loss for disadvantaged workers.

sition in Proposition 1 to estimate the average wage gap.⁵ When instead we use the Reimers (1983), Cotton (1988), or Fortin (2008) decompositions, we risk overstating the importance of the smaller group. As an extreme example, assume that the proportion of advantaged workers goes to zero. In that case, in the limit, the unexplained components of the Cotton (1988) and Fortin (2008) decompositions approach the average wage gain for advantaged workers, *i.e.* the group which is nearly absent by assumption.

On the other hand, it should be noted that these interpretations of the Reimers (1983), Cotton (1988), and Fortin (2008) decompositions are based on the assumption of a simple counterfactual treatment (Assumption 1), while this assumption has not been invoked in any of these papers. More precisely, each of these papers has attempted to account for the presence of general equilibrium effects—which are ruled out by Assumption 1—and to derive a comparison wage structure which would be observed if wage discrimination ceased to exist. It is very difficult, however, to correctly guess the form of this “nondiscriminatory” or “competitive” wage structure—and Reimers (1983), Cotton (1988), and Fortin (2008) have not provided any theoretical models to rationalize their choices. In this situation we might perhaps prefer to invoke the assumption of a simple counterfactual treatment instead of relying on the general-equilibrium approach—in which case the Reimers (1983), Cotton (1988), and Fortin (2008) decompositions would be problematic.

2.4 A Semiparametric Extension

If we indeed decide to aim at estimating τ_{gap} , τ_{gain} , and τ_{loss} , then any of the standard estimators of the average treatment effect and the average treatment effect on the treated can be used to estimate τ_{gap} and τ_{gain}/τ_{loss} , respectively, and we can safely assume that the better an estimator is for various average treatment effects, the better it is also for various averages of conditional wage gaps (for a similar discussion, see Fortin *et al.*, 2011).

Indeed, recent applications have used reweighting (Barsky *et al.*, 2002), matching on covariates (Black *et al.*, 2006, 2008; Ñopo, 2008), propensity score matching (Frölich, 2007), and regression trees (Mora, 2008) to study intergroup differences in various outcomes. In this subsection I demonstrate how various averages of conditional wage gaps can also be estimated using a combination of stratification (on the propensity score) and the Oaxaca–Blinder decomposition. This new estimation method is very similar to combining stratification and linear regression. Although this latter estimator is well established (see, *e.g.*, Dehejia and Wahba, 1999; Imbens and Wooldridge, 2009), it does not allow for

⁵Note that there is persuasive evidence on such heterogeneity which has been accumulated in the literature on glass ceilings (see, *e.g.*, Albrecht *et al.*, 2003; Arulampalam *et al.*, 2007; de la Rica *et al.*, 2008; Chzhen and Mumford, 2011).

within-strata treatment effect heterogeneity. Since the Oaxaca–Blinder decomposition can be regarded as a version of flexible OLS (in the sense that it allows for heterogeneity in effects), this shortcoming can be addressed by a combination of stratification and the Oaxaca–Blinder decomposition. On the other hand, it needs to be noted that this estimation method only allows for a very specific form of effect heterogeneity—linear in X_i .

Such an estimator of the average wage gap requires a first-step estimation of the propensity score, *i.e.* the conditional probability that a sample member belongs to the advantaged group given his or her covariate values. The estimated propensity score is then used to divide the whole sample into J strata. For each stratum j , a separate estimate of the average wage gap, $\hat{\tau}_{gap, j}$, is obtained using the decomposition in Proposition 1. Also, the estimated variance, $\hat{V}_{gap, j}$, can be computed, and these estimates are averaged using a procedure in Imbens and Wooldridge (2009):

$$\hat{\tau}_{gap} = \sum_{j=1}^J \left(\frac{n_{j0} + n_{j1}}{n} \right) \cdot \hat{\tau}_{gap, j} \quad \text{and} \quad \hat{V}_{gap} = \sum_{j=1}^J \left(\frac{n_{j0} + n_{j1}}{n} \right)^2 \cdot \hat{V}_{gap, j}. \quad (10)$$

Similarly, an analogous estimator of the average wage gain for advantaged workers involves estimating Equation 5 in each of the strata, while the average wage loss for disadvantaged workers can be estimated with the use of the Oaxaca–Blinder decomposition in Equation 6. These within-strata estimates are averaged to obtain:

$$\hat{\tau}_{gain} = \sum_{j=1}^J \left(\frac{n_{j1}}{n_1} \right) \cdot \hat{\tau}_{gain, j} \quad \text{and} \quad \hat{V}_{gain} = \sum_{j=1}^J \left(\frac{n_{j1}}{n_1} \right)^2 \cdot \hat{V}_{gain, j}, \quad (11)$$

$$\hat{\tau}_{loss} = \sum_{j=1}^J \left(\frac{n_{j0}}{n_0} \right) \cdot \hat{\tau}_{loss, j} \quad \text{and} \quad \hat{V}_{loss} = \sum_{j=1}^J \left(\frac{n_{j0}}{n_0} \right)^2 \cdot \hat{V}_{loss, j}. \quad (12)$$

These estimators require fewer functional form assumptions compared with the fully parametric approach, and such a property might be important to ensure robustness in the presence of nonlinearities in the existing wage structures (see, *e.g.*, Barsky *et al.*, 2002). They can also provide a useful alternative outside the decomposition context, and can be used to estimate the average treatment effect or the average treatment effect on the treated. These estimators are tested in an empirical application in the next section together with parametric Oaxaca–Blinder decompositions and normalized reweighting, which has performed very well in a recent Monte Carlo study by Busso *et al.* (2014).

3 An Empirical Application

In this section various versions of the Oaxaca–Blinder decomposition are compared in an application to the UK gender wage gap. The source of data is the Quarterly Labour Force Survey (LFS), 2002Q1–2010Q4 (Office for National Statistics, Various years). In this application I append all the quarterly datasets and restrict the resulting sample in such a way that yearly representative results, for each year from 2002 to 2010, can be produced.

I further restrict the 2002–2010 LFS dataset to those individuals who are at least 18 years old, have reported nonzero average gross hourly pay, and do not have missing information on any of the control variables specified below. The outcome variable is the log hourly wage. The set of control variables includes polynomials in age, tenure, and potential experience as well as dummies for marital status (5 categories), ethnic origin (11 categories), country of residence (4 categories), occupation (9 categories), and public/private sector. For current students, potential experience is coded as 0. For individuals with no education, potential experience is coded as the number of years since the age of fifteen.

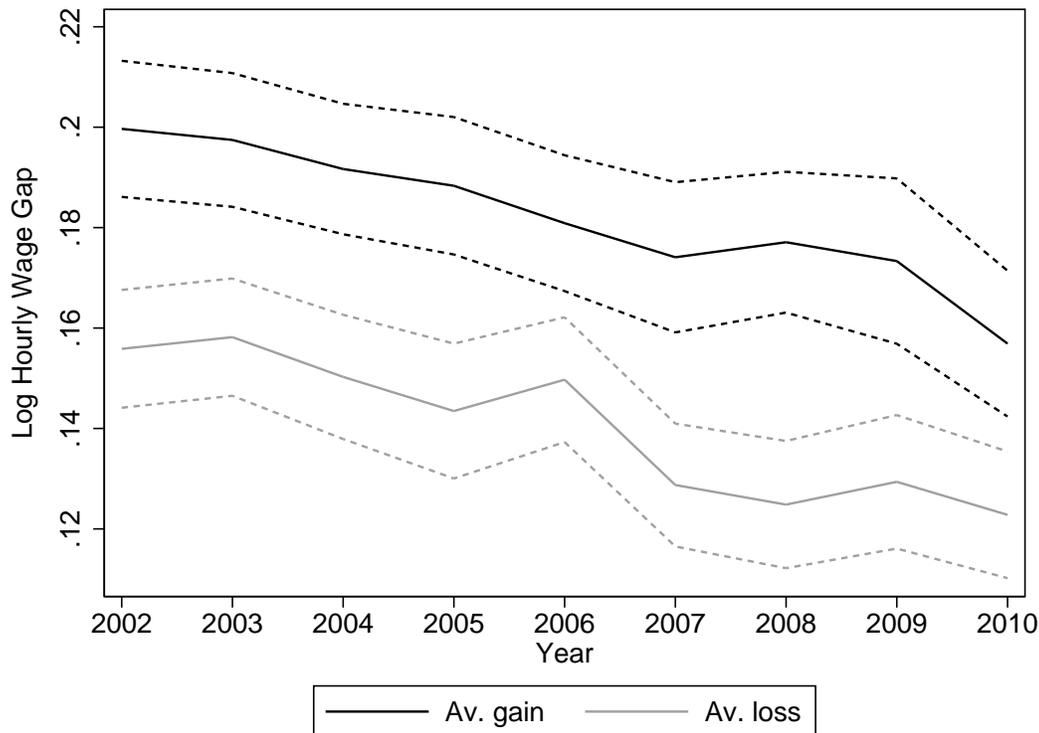
In all instances, the estimation is done using weighted least squares, with weights recommended for income data. In particular, I estimate β_1 and β_0 in Equation 4 by WLS, and apply the results to construct estimates of unexplained components of various versions of the Oaxaca–Blinder decomposition (Equations 7 and 8). Whenever I need to estimate $E[X_i | d_i = 1]$, $E[X_i | d_i = 0]$, $P[d_i = 1]$, or $P[d_i = 0]$, I use (weighted) sample means. Because β^* in the Neumark (1988) and Fortin (2008) decompositions cannot be easily constructed from β_1 and β_0 , I estimate it directly, again using weighted least squares.⁶

Table 1 presents descriptive statistics for outcome and selected control variables in all the yearly samples and Table 2 presents estimates of unexplained components of various versions of the Oaxaca–Blinder decomposition (including $\hat{\tau}_{gap}$, $\hat{\tau}_{gain}$, and $\hat{\tau}_{loss}$) as well as yearly measures of the raw wage gap. All these measures have generally been falling between 2002 and 2010, and have reached their minimum values in 2010. For example, the estimate of the raw wage gap has fallen from 23.06 to 16.86 log points, while $\hat{\tau}_{gain}$ and $\hat{\tau}_{loss}$ have fallen from 19.97 to 15.69 and from 15.59 to 12.28 log points, respectively.

There are several other empirical regularities which are worth mentioning. The average wage gain for men is larger than the average wage loss for women in a very robust way. This difference was as large as 3–5 log points in all years and always statistically

⁶All the applications of the Oaxaca–Blinder decomposition presented in this paper use the `oaxaca` command in Stata (Jann, 2008). The new version of the Oaxaca–Blinder decomposition in Proposition 1 is also easily implementable using this command. If local macros `prf` and `xvars` contain the proportion of women in the estimation sample and the list of control variables, respectively, then this decomposition can be applied as `oaxaca lnwage 'xvars', by(female) weight('prf')`. Note that the `oaxaca` command calculates standard errors which account for the stochastic nature of sample means of control variables.

Figure 1: A Comparison of the Average Wage Gain for Men and the Average Wage Loss for Women in the UK Labor Market

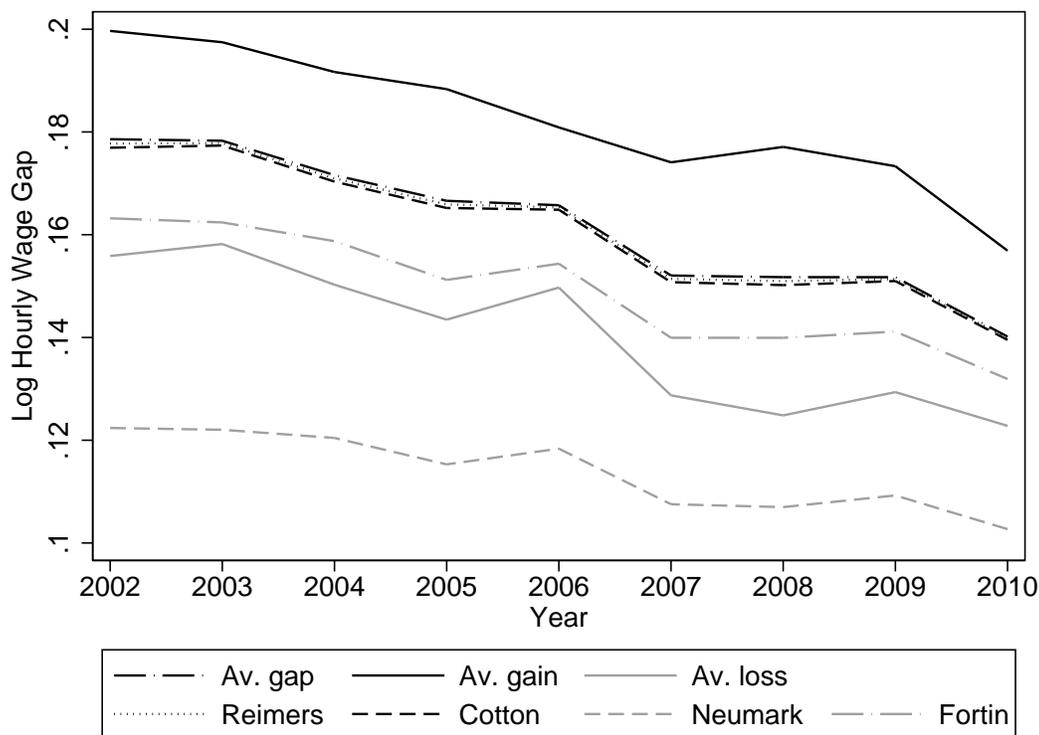


significant at the 1% level.⁷ Such a phenomenon has already been documented in the wage gap literature, but without a clear interpretation (see, *e.g.*, Fortin *et al.*, 2011). More precisely, this empirical regularity was typically understood as an “indeterminacy” of Oaxaca–Blinder decompositions. Such a claim is untenable, however, within the framework of this paper, in which each of these parameters— τ_{gain} and τ_{loss} —has a different interpretation. If τ_{gain} is significantly larger than τ_{loss} , then men gain typically more in comparison with similar women than women lose in comparison with similar men. Because men are located, on average, higher in the wage distribution than women, such a phenomenon means that conditional wage gaps tend to increase with wages. Indeed, the coefficient estimates for a series of quantile regressions suggest that in most years wage gaps at the top of the UK wage distribution (90th centile) were at least 3 log points larger than at the bottom of this distribution (10th centile).⁸ Both time series of estimates ($\hat{\tau}_{gain}$

⁷All the tests of statistical significance referred to in this section are based on bootstrap standard errors of the estimated differences between the unexplained components of various versions of the Oaxaca–Blinder decomposition (with 1,000 resamples).

⁸One can also suspect the existence of the so-called glass ceiling effect. However, the results of the quantile regressions provide mixed evidence on the existence of a glass ceiling in the UK labor market. Although wage gaps generally increase throughout the wage distribution, their acceleration in this distri-

Figure 2: A Comparison of Unexplained Components of Various Versions of the Oaxaca–Blinder Decomposition—in an Application to the UK Labor Market



and $\hat{\tau}_{loss}$) are also plotted in Figure 1, together with 95% confidence intervals.

At the same time, the Reimers (1983) and Cotton (1988) estimates of the unexplained component as well as the Oaxaca–Blinder estimates of τ_{gap} lie always (by construction) in between the corresponding estimates of τ_{gain} and τ_{loss} . The Reimers (1983), Cotton (1988), and τ_{gap} estimates are, however, not only bounded by $\hat{\tau}_{gain}$ and $\hat{\tau}_{loss}$, but also exactly equal (again by construction) to their weighted average, with weights equal to sample proportions of men and women (τ_{gap}), reversed sample proportions of both groups (Cotton, 1988), and 0.5 and 0.5 (Reimers, 1983). Moreover, the Neumark (1988) estimates of the unexplained component are always significantly lower (again, at the 1% level) than any other estimates (for a theoretical explanation, see Elder *et al.*, 2010).

All of these time series of estimates, together with $\hat{\tau}_{gain}$, $\hat{\tau}_{loss}$, and Fortin (2008), are also plotted in Figure 2. As evident in Figure 2, the differences between $\hat{\tau}_{gap}$ and the Reimers (1983) and Cotton (1988) estimates are quantitatively quite small. Note, however, that this is an artifact of nearly equal sample proportions of men and women in

bution’s upper tail can rarely be regarded as “sharp”. See Arulampalam *et al.* (2007) for related evidence on glass ceilings across EU countries as well as Chzhen and Mumford (2011) for a more recent study of the United Kingdom.

the LFS data. If these proportions were significantly different (*e.g.*, when analyzing ethnic/racial wage gaps, or union wage premiums, or gender wage gaps *within* a predominantly male/female occupation), the estimates would tend to diverge.

Also, note that these differences, however small, are always statistically significant. Further evidence on the differences between various Oaxaca–Blinder decompositions is provided in Table 3 by presenting bootstrap standard errors (based on 1,000 resamples) of the estimated differences between $\hat{\tau}_{gap}$ and other Oaxaca–Blinder estimators of the unexplained component, and testing the statistical significance of these differences. All these differences are statistically significant at the 5% level in all years (and, in most cases, also at the 1% level). We get the expected result that $\hat{\tau}_{gap}$ is always significantly smaller than $\hat{\tau}_{gain}$ and larger than $\hat{\tau}_{loss}$. Also, consistent with the evidence in Elder *et al.* (2010), $\hat{\tau}_{gap}$ is always significantly larger than the unexplained component of the Neumark (1988) decomposition, and this difference is quantitatively quite large (4–6 log points). Importantly, however, Table 3 reestablishes the empirical relevance of this paper’s claim that the Reimers (1983), Cotton (1988), and Fortin (2008) decompositions might estimate weighted averages of conditional wage gaps which are difficult to interpret. In all years these decompositions overstate the importance of $\hat{\tau}_{loss}$ (the effect on the smaller subpopulation) in estimating average wage gaps, and consequently $\hat{\tau}_{gap}$ is always significantly larger than the unexplained components of these decompositions (because $\hat{\tau}_{gain} > \hat{\tau}_{loss}$).⁹ In other words, using the Reimers (1983), Cotton (1988), or Fortin (2008) decompositions would negatively bias the estimate of the average wage gap in an application to the UK labor market, although these biases can be quite small in the present application.

As a robustness check, Table 4 presents a comparison of various estimators of τ_{gap} , τ_{gain} , and τ_{loss} , both parametric (Oaxaca–Blinder) and semiparametric (stratification and Oaxaca–Blinder, normalized reweighting). Importantly, all the qualitative results on the relationship between the average wage gain for men and the average wage loss for women

⁹At first, this might be seen as surprising, given the small differences between $\hat{\tau}_{gap}$ and selected other estimates. Take the difference between $\hat{\tau}_{gap}$ (17.86 log points) and the Reimers (1983) estimate (17.78 log points) in 2002 as an example. Both these parameters are estimated with an error of 0.56 log points, so it is natural to expect that the difference of 0.08 log points between them is not statistically significant. As shown in Table 3, however, this is not the case, since this difference is actually estimated with a very small error of 0.02. Why is the initial intuition wrong? First, note that these estimates are (by construction) positively correlated over repeated samples. When $\hat{\tau}_{gap}$ is relatively large, the Reimers (1983) estimate is also likely to be large, but probably smaller than $\hat{\tau}_{gap}$ (since $\hat{\tau}_{gain} > \hat{\tau}_{loss}$). Second, note that the null hypothesis of $P[d_i = 1] \cdot \tau_{gain} + P[d_i = 0] \cdot \tau_{loss} = \frac{1}{2} \cdot \tau_{gain} + \frac{1}{2} \cdot \tau_{loss}$ can be equivalently written as $(P[d_i = 1] - \frac{1}{2}) \cdot (\tau_{gain} - \tau_{loss}) = 0$. That the second factor $(\tau_{gain} - \tau_{loss})$ is significantly larger than zero is evident from Figure 1. Using a one-proportion z-test and data in Table 1, the equality of $P[d_i = 1]$ and $\frac{1}{2}$ can also be rejected (in favor of $P[d_i = 1] > \frac{1}{2}$). Of course, a similar logic applies also to other tests of statistical significance in Table 3.

are confirmed. Also, the trends in the estimated wage gaps are very similar. This pattern is consistent with the recent evidence in Nopo (2008) who has concluded that the linearity assumption is not particularly problematic when decomposing gender wage gaps.

4 Summary

In this paper I have argued for a decomposition framework in which current wages of advantaged workers are compared with current wages of similar disadvantaged workers (and vice versa), and not with an hypothesized “nondiscriminatory” wage structure. To provide a tool for such comparisons, I have derived an alternative solution to the comparison group choice problem which is fundamental in Oaxaca–Blinder, while this new decomposition can be used to estimate the average wage gap and the average treatment effect. I have also pointed out that several other versions of the Oaxaca–Blinder decomposition (Reimers, 1983; Cotton, 1988; Fortin, 2008) can produce misleading results in empirical applications, since each of them is likely to overstate the importance of the smaller of the subpopulations (*e.g.*, men and women, treated and controls, union and nonunion workers) when estimating various averages of conditional wage gaps or treatment effects.

This approach has also been illustrated empirically in an application to UK gender wage gaps. Using data from the Quarterly Labour Force Survey (LFS) I have estimated the average wage gap, the average wage gain for men, and the average wage loss for women for the UK working population, for each year from 2002 to 2010. The major empirical finding of this study is that men gain typically more in comparison with similar women than women lose in comparison with similar men (τ_{gain} is larger than τ_{loss}). This phenomenon is explained by the fact that conditional wage gaps tend to increase with wages and this is indeed the case in the UK labor market.

Future work might involve establishing formal conditions under which causal effects of gender and other immutable characteristics can be identified and estimated (for recent discussions, see Kunze, 2008, Greiner and Rubin, 2011, and Huber, 2014) and, as already suggested by Fortin *et al.* (2011), improving the economic structure behind decomposition methods. It is also essential to understand the links between the decomposition methods and the treatment effects framework. Following an important review in Fortin *et al.* (2011), this paper has attempted to take this ongoing discussion one step further by providing a new (Oaxaca–Blinder) model for the average wage gap which—under a particular causal framework—corresponds to the average treatment effect.

Table 1: Sample Means of Outcome and Selected Control Variables for Yearly UK LFS Datasets

Year	N	Log wage	Female	Age	Potential experience				Tenure	Public sector	England	Wales	Scotland	Northern Ireland
					Female	Age	Potential experience	Tenure						
2002	31,248	2.10 (.55)	.48 (.50)	39.25 (12.28)	21.67 (13.19)	7.53 (8.24)	.27 (.44)	.84 (.36)	.05 (.21)	.08 (.28)	.03 (.16)			
2003	30,448	2.13 (.55)	.49 (.50)	39.36 (12.34)	21.71 (13.27)	7.60 (8.27)	.28 (.45)	.84 (.37)	.05 (.21)	.09 (.28)	.03 (.16)			
2004	29,424	2.16 (.55)	.48 (.50)	39.29 (12.43)	21.54 (13.33)	7.47 (8.26)	.28 (.45)	.83 (.37)	.05 (.21)	.09 (.29)	.03 (.17)			
2005	26,732	2.21 (.55)	.48 (.50)	39.48 (12.43)	21.65 (13.37)	7.60 (8.30)	.28 (.45)	.84 (.37)	.05 (.21)	.09 (.28)	.03 (.17)			
2006	28,249	2.24 (.56)	.49 (.50)	39.38 (12.60)	21.42 (13.56)	7.52 (8.31)	.28 (.45)	.83 (.37)	.05 (.21)	.09 (.29)	.03 (.17)			
2007	29,628	2.27 (.57)	.49 (.50)	39.51 (12.67)	21.48 (13.63)	7.56 (8.37)	.27 (.45)	.84 (.37)	.04 (.20)	.09 (.28)	.03 (.18)			
2008	28,844	2.31 (.57)	.49 (.50)	39.45 (12.55)	21.41 (13.45)	7.52 (8.31)	.27 (.45)	.83 (.37)	.04 (.20)	.09 (.29)	.03 (.18)			
2009	26,808	2.33 (.57)	.49 (.50)	39.61 (12.54)	21.46 (13.47)	7.77 (8.37)	.29 (.45)	.84 (.37)	.04 (.20)	.09 (.29)	.03 (.17)			
2010	26,201	2.35 (.58)	.49 (.50)	39.53 (12.51)	21.26 (13.43)	7.82 (8.25)	.29 (.45)	.83 (.37)	.04 (.21)	.09 (.28)	.03 (.18)			

NOTE: Standard deviations are in parentheses. Wages are in nominal pounds sterling.

Table 2: A Comparison of Unexplained Components of Various Versions of the Oaxaca–Blinder Decomposition—in an Application to the UK Labor Market

Year	Raw	$\hat{\tau}_{gap}$	$\hat{\tau}_{gain}$	$\hat{\tau}_{loss}$	Reimers	Cotton	Neumark	Fortin
2002	.2306 (.0062)	.1786 (.0056)	.1997 (.0069)	.1559 (.0060)	.1778 (.0056)	.1769 (.0056)	.1224 (.0042)	.1632 (.0055)
2003	.2275 (.0063)	.1783 (.0056)	.1975 (.0068)	.1582 (.0060)	.1778 (.0056)	.1774 (.0056)	.1220 (.0042)	.1624 (.0055)
2004	.2066 (.0065)	.1716 (.0057)	.1917 (.0066)	.1503 (.0063)	.1710 (.0057)	.1703 (.0057)	.1204 (.0044)	.1588 (.0057)
2005	.2064 (.0068)	.1666 (.0061)	.1883 (.0070)	.1435 (.0069)	.1659 (.0061)	.1652 (.0061)	.1153 (.0047)	.1512 (.0061)
2006	.1991 (.0067)	.1657 (.0058)	.1809 (.0069)	.1497 (.0063)	.1653 (.0058)	.1649 (.0058)	.1184 (.0045)	.1544 (.0058)
2007	.1844 (.0067)	.1521 (.0061)	.1741 (.0076)	.1287 (.0062)	.1514 (.0061)	.1508 (.0061)	.1075 (.0047)	.1399 (.0060)
2008	.1936 (.0067)	.1517 (.0060)	.1771 (.0071)	.1248 (.0065)	.1510 (.0060)	.1502 (.0060)	.1070 (.0047)	.1399 (.0060)
2009	.1923 (.0071)	.1517 (.0066)	.1733 (.0084)	.1294 (.0068)	.1514 (.0066)	.1510 (.0066)	.1093 (.0049)	.1411 (.0063)
2010	.1686 (.0073)	.1402 (.0062)	.1569 (.0074)	.1228 (.0064)	.1399 (.0062)	.1395 (.0062)	.1027 (.0050)	.1319 (.0063)

NOTE: Robust standard errors are in parentheses.

Table 3: The Statistical Significance of Differences Between $\hat{\tau}_{gap}$ and Unexplained Components of Other Versions of the Oaxaca–Blinder Decomposition

Year	$\hat{\tau}_{gain}$	$\hat{\tau}_{loss}$	Reimers	Cotton	Neumark	Fortin
2002	-.0211*** (.0031)	.0227*** (.0033)	.0008*** (.0002)	.0017*** (.0003)	.0562*** (.0024)	.0154*** (.0020)
2003	-.0192*** (.0032)	.0201*** (.0033)	.0005*** (.0001)	.0009*** (.0003)	.0563*** (.0022)	.0159*** (.0017)
2004	-.0201*** (.0029)	.0213*** (.0031)	.0006*** (.0002)	.0013*** (.0003)	.0512*** (.0022)	.0128*** (.0018)
2005	-.0217*** (.0032)	.0231*** (.0034)	.0007*** (.0002)	.0014*** (.0004)	.0513*** (.0024)	.0154*** (.0019)
2006	-.0151*** (.0030)	.0160*** (.0032)	.0004*** (.0001)	.0009*** (.0003)	.0474*** (.0023)	.0114*** (.0019)
2007	-.0220*** (.0032)	.0233*** (.0034)	.0006*** (.0002)	.0013*** (.0003)	.0445*** (.0026)	.0121*** (.0022)
2008	-.0254*** (.0031)	.0269*** (.0033)	.0008*** (.0002)	.0015*** (.0004)	.0448*** (.0024)	.0118*** (.0020)
2009	-.0216*** (.0037)	.0224*** (.0038)	.0004** (.0002)	.0008** (.0003)	.0425*** (.0029)	.0106*** (.0025)
2010	-.0167*** (.0032)	.0174*** (.0033)	.0003** (.0001)	.0006** (.0003)	.0375*** (.0022)	.0083*** (.0019)

NOTE: All differences are computed as $\hat{\tau}_{gap} - \hat{\tau}_X$, where $\hat{\tau}_X$ is listed in the column heading. Estimates are listed in Table 2. Bootstrap standard errors of the estimated differences are based on 1,000 resamples. *Statistically significant at the 10% level; **at the 5% level; ***at the 1% level.

Table 4: A Comparison of Various Parametric and Semiparametric Estimators of the Average Wage Gap, the Average Wage Gain for Men, and the Average Wage Loss for Women

Year	$\hat{\tau}_{gap}$		$\hat{\tau}_{gain}$		$\hat{\tau}_{loss}$	
	O-B	Stratification and O-B	O-B	Stratification and O-B	O-B	Stratification and O-B
2002	.1786 (.0056)	.1717 (.0063)	.1997 (.0069)	.1957 (.0086)	.1559 (.0060)	.1488 (.0065)
2003	.1783 (.0056)	.1713 (.0060)	.1975 (.0068)	.1889 (.0079)	.1582 (.0060)	.1551 (.0063)
2004	.1716 (.0057)	.1685 (.0060)	.1917 (.0066)	.1942 (.0075)	.1503 (.0063)	.1444 (.0065)
2005	.1666 (.0061)	.1617 (.0066)	.1883 (.0070)	.1881 (.0075)	.1435 (.0069)	.1373 (.0079)
2006	.1657 (.0058)	.1596 (.0060)	.1809 (.0069)	.1766 (.0075)	.1497 (.0063)	.1442 (.0065)
2007	.1521 (.0061)	.1402 (.0063)	.1741 (.0076)	.1637 (.0084)	.1287 (.0062)	.1188 (.0064)
2008	.1517 (.0060)	.1468 (.0062)	.1771 (.0071)	.1786 (.0079)	.1248 (.0065)	.1175 (.0066)
2009	.1517 (.0066)	.1446 (.0070)	.1733 (.0084)	.1786 (.0097)	.1294 (.0068)	.1133 (.0071)
2010	.1402 (.0062)	.1367 (.0063)	.1569 (.0074)	.1571 (.0079)	.1228 (.0064)	.1180 (.0066)

NOTE: Robust standard errors (O-B, stratification and O-B) and bootstrap standard errors (normalized reweighting) are in parentheses. The bootstrap is based on 1,000 resamples. The propensity score is estimated using a logit model. For stratification and Oaxaca-Blinder samples are divided into five strata of equal size (*i.e.* on the basis of the 20th, 40th, 60th, and 80th centile of the distribution of the estimated propensity score).

References

- ALBRECHT, J., BJÖRKLUND, A. & VROMAN, S. (2003). Is there a glass ceiling in Sweden? *Journal of Labor Economics* 21, 145–177.
- ARULAMPALAM, W., BOOTH, A. L. & BRYAN, M. L. (2007). Is there a glass ceiling over Europe? Exploring the gender pay gap across the wage distribution. *Industrial and Labor Relations Review* 60, 163–186.
- BARSKY, R., BOUND, J., CHARLES, K. K. & LUPTON, J. P. (2002). Accounting for the black-white wealth gap: A nonparametric approach. *Journal of the American Statistical Association* 97, 663–673.
- BLACK, D., HAVILAND, A., SANDERS, S. & TAYLOR, L. (2006). Why do minority men earn less? A study of wage differentials among the highly educated. *Review of Economics and Statistics* 88, 300–313.
- BLACK, D. A., HAVILAND, A. M., SANDERS, S. G. & TAYLOR, L. J. (2008). Gender wage disparities among the highly educated. *Journal of Human Resources* 43, 630–659.
- BLACKBURN, M. L. (2004). The role of test scores in explaining race and gender differences in wages. *Economics of Education Review* 23, 555–576.
- BLINDER, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources* 8, 436–455.
- BROWN, C. & CORCORAN, M. (1997). Sex-based differences in school content and the male-female wage gap. *Journal of Labor Economics* 15, 431–465.
- BUSSO, M., DINARDO, J. & MCCRARY, J. (2014). New evidence on the finite sample properties of propensity score reweighting and matching estimators. *Review of Economics and Statistics* 96, 885–897.
- CHERNOZHUKOV, V., FERNÁNDEZ-VAL, I. & MELLY, B. (2013). Inference on counterfactual distributions. *Econometrica* 81, 2205–2268.
- CHZHEN, Y. & MUMFORD, K. (2011). Gender gaps across the earnings distribution for full-time employees in Britain: Allowing for sample selection. *Labour Economics* 18, 837–844.
- COTTON, J. (1988). On the decomposition of wage differentials. *Review of Economics and Statistics* 70, 236–243.

- DE LA RICA, S., DOLADO, J. J. & LLORENS, V. (2008). Ceilings or floors? Gender wage gaps by education in Spain. *Journal of Population Economics* 21, 751–776.
- DEHEJIA, R. H. & WAHBA, S. (1999). Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *Journal of the American Statistical Association* 94, 1053–1062.
- DINARDO, J., FORTIN, N. M. & LEMIEUX, T. (1996). Labor market institutions and the distribution of wages, 1973–1992: A semiparametric approach. *Econometrica* 64, 1001–1044.
- DUNCAN, G. M. & LEIGH, D. E. (1985). The endogeneity of union status: An empirical test. *Journal of Labor Economics* 3, 385–402.
- ELDER, T. E., GODDEERIS, J. H. & HAIDER, S. J. (2010). Unexplained gaps and Oaxaca–Blinder decompositions. *Labour Economics* 17, 284–290.
- FIRPO, S., FORTIN, N. & LEMIEUX, T. (2007). Decomposing wage distributions using recentered influence function regressions. Unpublished.
- FORTIN, N., LEMIEUX, T. & FIRPO, S. (2011). Decomposition methods in economics. In: *Handbook of Labor Economics* (ASHENFELTER, O. & CARD, D., eds.), vol. 4A. Elsevier.
- FORTIN, N. M. (2005). Gender role attitudes and the labour-market outcomes of women across OECD countries. *Oxford Review of Economic Policy* 21, 416–438.
- FORTIN, N. M. (2008). The gender wage gap among young adults in the United States: The importance of money versus people. *Journal of Human Resources* 43, 884–918.
- FRÖLICH, M. (2007). Propensity score matching without conditional independence assumption – with an application to the gender wage gap in the United Kingdom. *Econometrics Journal* 10, 359–407.
- GREINER, D. J. & RUBIN, D. B. (2011). Causal effects of perceived immutable characteristics. *Review of Economics and Statistics* 93, 775–785.
- HOLLAND, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association* 81, 945–960.
- HUBER, M. (2014). Causal pitfalls in the decomposition of wage gaps. *Journal of Business & Economic Statistics* (forthcoming).

- ICHINO, A. & MORETTI, E. (2009). Biological gender differences, absenteeism, and the earnings gap. *American Economic Journal: Applied Economics* 1, 183–218.
- IMBENS, G. W. & WOOLDRIDGE, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature* 47, 5–86.
- JANN, B. (2008). The Blinder–Oaxaca decomposition for linear regression models. *Stata Journal* 8, 453–479.
- JUHN, C., MURPHY, K. M. & PIERCE, B. (1993). Wage inequality and the rise in returns to skill. *Journal of Political Economy* 101, 410–442.
- KLINE, P. (2011). Oaxaca–Blinder as a reweighting estimator. *American Economic Review: Papers & Proceedings* 101, 532–537.
- KUNZE, A. (2008). Gender wage gap studies: Consistency and decomposition. *Empirical Economics* 35, 63–76.
- LEIBBRANDT, A. & LIST, J. A. (2014). Do women avoid salary negotiations? Evidence from a large-scale natural field experiment. *Management Science* (forthcoming).
- LOURY, L. D. (1997). The gender earnings gap among college-educated workers. *Industrial and Labor Relations Review* 50, 580–593.
- MACHADO, J. A. F. & MATA, J. (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of Applied Econometrics* 20, 445–465.
- MACHIN, S. & PUHANI, P. A. (2003). Subject of degree and the gender wage differential: Evidence from the UK and Germany. *Economics Letters* 79, 393–400.
- MANNING, A. & SWAFFIELD, J. (2008). The gender gap in early-career wage growth. *Economic Journal* 118, 983–1024.
- MELLY, B. (2005). Decomposition of differences in distribution using quantile regression. *Labour Economics* 12, 577–590.
- MELLY, B. (2006). Applied quantile regression. PhD dissertation.
- MORA, R. (2008). A nonparametric decomposition of the Mexican American average wage gap. *Journal of Applied Econometrics* 23, 463–485.
- MUELLER, G. & PLUG, E. (2006). Estimating the effect of personality on male and female earnings. *Industrial and Labor Relations Review* 60, 3–22.

- NEUMARK, D. (1988). Employers' discriminatory behavior and the estimation of wage discrimination. *Journal of Human Resources* 23, 279–295.
- OAXACA, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review* 14, 693–709.
- OAXACA, R. L. & RANSOM, M. R. (1988). Searching for the effect of unionism on the wages of union and nonunion workers. *Journal of Labor Research* 9, 139–148.
- OAXACA, R. L. & RANSOM, M. R. (1994). On discrimination and the decomposition of wage differentials. *Journal of Econometrics* 61, 5–21.
- OFFICE FOR NATIONAL STATISTICS (Various years). *Quarterly Labour Force Survey*. UK Data Archive.
- ÑOPO, H. (2008). Matching as a tool to decompose wage gaps. *Review of Economics and Statistics* 90, 290–299.
- REIMERS, C. W. (1983). Labor market discrimination against Hispanic and black men. *Review of Economics and Statistics* 65, 570–579.
- SŁOCZYŃSKI, T. (2014). New evidence on linear regression and treatment effect heterogeneity. Unpublished.
- WEICHSELBAUMER, D. & WINTER-EBMER, R. (2005). A meta-analysis of the international gender wage gap. *Journal of Economic Surveys* 19, 479–511.
- WOOLDRIDGE, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. MIT Press, 2nd ed.