I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13601

Decomposing Gender Wage Gaps: A Family Economics Perspective

Dorothée Averkamp
Christian Bredemeier
Falko Juessen

AUGUST 2020

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13601

Decomposing Gender Wage Gaps: A Family Economics Perspective

Dorothée Averkamp
University of Wuppertal

Christian Bredemeier
University of Wuppertal and IZA

Falko Juessen
University of Wuppertal and IZA

AUGUST 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Decomposing Gender Wage Gaps: A Family Economics Perspective

We show that parts of the unexplained wage gap in standard Oaxaca-Blinder decompositions result from the neglect of the role played by the family for individual wages. We present a simple model of dual-earner households facing a trade-off regarding whose career to promote and show analytically that the standard Oaxaca-Blinder approach overestimates the degree of pay discrimination. Unbiased decompositions can be obtained when the Oaxaca-Blinder wage equation is augmented by the characteristics of the individual's partner. In an empirical application, we find that this extended decomposition explains considerably larger shares of the gender wage gap than does the standard decomposition.

JEL Classification: J31, J16, J12, J71, J24

Keywords: gender wage gap, Oaxaca-Blinder, dual-earner households, discrimination

Corresponding author:
Christian Bredemeier
University of Wuppertal
Schumpeter School of Business and Economics
Gaußstraße 20
42119 Wuppertal
Germany

E-mail: bredemeier@uni-wuppertal.de

1 Introduction

The gender wage gap is not only an important issue for society but also an interesting phenomenon for economists. As Olivetti and Petrongolo (2016) summarize, it decreased substantially in the second half of the 20th century, but a persistent gap remains. Using an Oaxaca-Blinder decomposition (Oaxaca 1973, Blinder 1973), Blau and Kahn (2017) show that a considerable part of the wage gap can be related to observable differences between men and women such as differences in work experience or occupation. However, they also show that a substantial part of the wage gap is unexplained by differences in observable characteristics of workers. The unexplained part of the gender wage gap has attracted considerable attention in both the academic literature as well as the public discourse.

In general, there can be three reasons why Oaxaca-Blinder decompositions do not explain the entire gender wage gap. First, the remaining gap can be due to discrimination of various forms, i.e., taste-based, statistical, or monopsonistic discrimination (see Blinder 1973, Hersch 2007, Blau and Kahn 2006, 2017). Second, there can be differences between men and women in unobservable characteristics that matter for wages, such as the willingness to take risky career choices or to work in competitive environments (Croson and Gneezy 2009, Bertrand 2011, Blau and Kahn 2017). Third, the remaining gap can be due to an incomplete understanding of the labor market. More specifically, there can be observable factors which are so far not considered to be important for wages which leads to a misspecification of the wage equation used in the Oaxaca-Blinder decomposition.

In this paper, we focus on the last point and show that decompositions of the wage gap into an explained and an unexplained part delivered by the standard Oaxaca-Blinder approach are biased due to the neglect of the role played by the family for wage rates. The literature on dual-earner couples has long recognized that couples regularly have to take decisions whose career to promote. Situations where conflicts between careers occur and it is rational for couples to give one career precedence over the other arise, for example, in the context of family migration (e.g., Mincer 1978, Compton and Pollak 2007, Foged 2016, the choice of employers at a given location (Bredemeier 2019, Petrongolo and Ronchi 2020), or career investment in the form of working long hours (Cortés and Tessada 2011, Cortés and Pan 2019). In such situations, dual-earner couples have been found to choose one spouse whose career is favored in major decisions even when this is at the expense of the other spouse's career. This behavior is summarized by the terms career prioritization (Philliber and Vannoy-Hiller 1990, Pixley and Moen 2003, Pixley 2008, 2009) or career hierarchy (Winkler and Rose 2000). An implication of career prioritization is that observed wages depend both on individual characteristics (such as education, experience, and occupation) but also on the family situation and, hence, also on the characteristics of the partner. For given individual characteristics, the earnings potential of the individual's partner determines whether and how strongly the couple prioritizes the individual's career or how often choices are taken that actually harm the individual's career. Hence, the partner's characteristics influence an individual's observed wage rate, conditional on the individual's own characteristics. While the wage equation that is estimated to implement the Oaxaca-Blinder decomposition does account for individual characteristics, the family situation, as captured by the characteristics of the partner, is usually ignored. This induces a bias in the estimated Oaxaca-Blinder decomposition, as the family situation, i.e., relative earnings potentials of the spouses, matters for wages paid to men and women. The bias in the decomposition is due to two effects. First, the parameter estimates of the worker's own characteristics are biased in the standard wage equation. Second, in the decomposition of the wage gap, the contribution of the characteristics of the worker's partner is ignored. The implication for Oaxaca-Blinder decompositions is that the characteristics of the partner should be included in the wage equation to account for different earner roles in the household.

We set up a model of career prioritization in dual-earner couples and use it to illustrate the bias in standard Oaxaca-Blinder decompositions. In our model, career prioritization stems from a joint location choice where couples need to compromise between locations promoting the husband's career and locations promoting the wife's career. For a couple, it is rational to prioritize the career of the spouse with the higher earnings potential and it chooses to live closer to the place which promotes optimally the career of the spouse with the higher earnings potential. As a consequence, the realized wage of an individual depends positively on the individual's own earnings potential and negatively on the earnings potential of the individual's partner. The latter effect reflects that individuals with high-potential partners tend to locate in places not promoting their own careers very strongly.

For illustration, we consider a model version where an individual's wage depends on the individual's earnings potential and location but not on gender. This version without gender discrimination is helpful to isolate the bias in standard Oaxaca-Blinder decompositions that exclude characteristics of the partner. We show that the standard decomposition yields a share of the wage gap which is supposedly "unexplained" even if there were no gender dimension in pay. In a more general setting, where the model accounts for wage discrimination against women, we show that the standard Oaxaca-Blinder approach underestimates the fraction of the wage gap that is due to observable characteristics. We then show that extending the Oaxaca-Blinder decomposition by the characteristics of the partner yields an unbiased decomposition. Importantly, the inclusion of additional (partner) characteristics does not mechanically increase the explained fraction of the gender gap. This only happens if the data are consistent with career prioritization or other mechanisms that induce one's own wage to depend *negatively* on the earnings potential of one's partner.

We apply our improved decomposition to U.S. data from the Panel Study of Income Dynamics (PSID). In a sample of dual-earner households the raw gender wage gap is between 35% (1980) and 23% (2010). In line with the literature, standard Oaxaca-Blinder decompositions explain roughly half of the gap and hence suggest the unexplained part of the gender wage gap to be substantial. Our extended Oaxaca-Blinder decompositions systematically explain larger shares of the wage gap as a consequence of gender differences in observable characteristics than does the standard decomposition. For some years, the extended decomposition explains up to 100% of the wage gap.

Our results imply that differences in labor-market outcomes relate less to pay discrimination by gender *per se* than suggested by a standard Oaxaca-Blinder decomposition. Rather, career prioritization within couples amplifies pay differences between men and women. As a result, the gender gap in actual earnings is larger than the gender gap in earnings potentials. To be clear, this interpretation does not at all rule out that the gender wage gap is a result of discrimination against women. Our empirical results indicate that, in most years, firms offer different wages to men and women even conditional on their characteristics. Further, neither the model nor our empirical analysis is informative about the reasons of gender differences in pay-relevant characteristics. In fact, important determinants of earnings potentials such as career interruptions or occupation choices are plausibly affected by gender roles, stereotypes, or prejudices. For example, empirical evidence shows that female labor supply and hence the accumulation of work experience is affected by gender identity norms (Bertrand et al. 2015) and cultural factors (Blau et al. 2020).

We corroborate our point that the neglect of the family situation is responsible for a substantial part of the supposedly unexplained wage gap by performing standard decompositions for singles and for married individuals without a working partner. For these groups, career prioritization or related aspects specific to dual-earner households cannot play a role and gender-neutral wage setting would imply that there should be no wage gap unrelated to individual characteristics. In fact, we find that, for these groups, standard Oaxaca-Blinder decompositions attribute substantially larger shares of the gender wage gap to observable characteristics of the individual workers than it does for men and women living in dual-earner couples.

The paper is organized as follows. In Section 2, we present a simple model of joint location choice to highlight a link between an individual's wage rate and the characteristics of the partner. We also draw on existing literature and discuss how wages paid by firms depend on the family situation of workers and are therefore dependent on one's own but also on one's partner's characteristics. In Section 3, we show analytically that a standard Oaxaca-Blinder decomposition is subject to a bias. We also show analytically that accounting for the characteristics of the partner eliminates this bias. In Section 4, we present an empirical application using PSID data. Section 5 concludes.

2 Theoretical framework

Recent literature emphasizes the role of the family for wages paid to women and men. The common implication of this literature is that a worker's observed wage does not only depend on the worker's own characteristics but also on the family situation and thereby on the characteristics of the worker's partner. In this section, we present a simple model of joint location choice to make this link explicit. We then discuss related approaches from the literature that have similar implications for wage rates.

2.1 A simple model of joint location choice in dual-earner households

We consider couple households with members indexed by i that have to decide over location. Location is a continuous variable $r \in (0, 1)$. An individual's ideal location, i.e., the location where (s)he can earn the highest wage is denoted by a_i. For every individual, this variable is drawn from a distribution $f(a)$ with mean μ and variance σ^2. The correlation between the ideal locations of partners in a couple is denoted by κ.

The wage $W_{i,r}$ of individual i in location r consist of two elements,

$$W_{i,r} = \psi_i z_{i,r}, \tag{1}$$

where ψ_i denotes the earnings potential of individual i, reflecting individual characteristics such as education and experience, and $z_{i,r}$ is a location-worker match variable. $W_{i,r}$ is the highest wage offered to individual i by firms located in location r.

We assume that the location-worker match variable is given by

$$z_{i,r} = \left(1 - (r - a_i)^2\right). \tag{2}$$

If individual i is at its ideal location, $r = a_i$, the individual achieves her full earnings potential. If the individual is in a location that differs from her ideal one, there is a wage penalty captured by $(r - a_i)^2$. The strength of this penalty depends on the distance between the actual location and the ideal one. This specification captures for instance spatial correlation in the industry mix in different regions.

Notation of the household structure in the model is as follows: individual i lives in household I, together with individual $-i$. At location r, couple I receives utility $u(c_{I,r})$ from household consumption $c_{I,r}$, with derivatives $u' > 0$ and $u'' < 0$. The couple's budget constraint at location r is given by

$$c_{I,r} = W_{i,r} + W_{-i,r}. \tag{3}$$

The couple's decision problem is to maximize $u(c_{I,r})$ subject to (1), (2), and (3) by choosing the optimal location for the couple household, which by substituting in the constraints reads

$$\max_r \; u\left(\psi_i \left(1 - (r - a_i)^2\right) + \psi_{-i} \left(1 - (r - a_{-i})^2\right)\right).$$

The first-order condition is

$$u'(c_{I,r}) \cdot (-2\psi_i (r - a_i) - 2\psi_{-i} (r - a_{-i})) = 0$$

so that the optimal location for the couple is

$$r_I^* = \frac{\psi_i}{\psi_i + \psi_{-i}} a_i + \frac{\psi_i}{\psi_i + \psi_{-i}} a_{-i}. \tag{4}$$

The household chooses its location as a weighted average of the ideal locations of its members. The weights are given by the relative earnings potentials of the two partners. The higher the earnings potential of either member, the closer the household moves to this member's ideal location.

Now consider log wage rates, $w_{i,r} = \log W_{i,r}$,

$$w_{i,r} = \log \psi_i + \log z_{i,r} = \log \psi_i + \log \left(1 - (r - a_i)^2\right),$$

and substitute the optimal location r_I^* from (4) to obtain equilibrium log wages w_i:

$$w_i = \log \psi_i + \log \left(1 - \left(\frac{\psi_{-i}}{\psi_i + \psi_{-i}} (a_{-i} - a_i)\right)^2\right), \tag{5}$$

The latter term can be interpreted as the discount from the full earnings potential when the individual is not living at her ideal location. When living at one's ideal location, one earns the full amount ψ_i but this is only the case if both partners happen to have the identical ideal location, $a_i = a_{-i}$. Whenever $a_i \neq a_{-i}$, the household chooses a location that is suboptimal for either partner and both spouses do not realize their full earnings potentials.

Our simple model of joint location choice implies that, for any given difference in ideal locations a_i and a_{-i} (which an econometrician cannot observe), the penalty term depends on the partner's share in full earnings potentials $\psi_{-i}/\left(\psi_i + \psi_{-i}\right)$. The higher the partner's share (hence, the lower one's own share), the farther away one lives from one's ideal location and the higher is hence the wage penalty. Thus, in this model, the observed wage rate of an individual does not only depend on the individual's own characteristics but also on the wage potential of the individual's partner. In particular, a higher earnings potential of the partner leads to a *lower* realized wage rate for oneself.

2.2 Linking equilibrium wages to observable characteristics

To perform an Oaxaca-Blinder wage-gap decomposition in the model, we need to link earnings potentials ψ to observable characteristics of the workers and linearize the wage equation. We express earnings potentials as a function of observable characteristics X_i,

$$\log \psi_i = \gamma_{g(i)} X_i,$$

where $g\left(i\right)$ denotes individual i's gender and can take the values m (for male) and f (for female). X_i is a column vector of observable characteristics of individual i and $\gamma_{g(i)}$ is a row vector of parameters. In general, the mapping from characteristics to earnings potentials can be gender-specific (such that $\gamma_m \neq \gamma_f$) which allows us to capture discrimination.

To obtain a log-linear relation between wages and observables, we apply a first-order Taylor approximation of the equilibrium wage equation (5) around a symmetric situation with $\psi_i = \psi_{-i} = \psi$, where ψ is the mean earnings potential in the economy, and values a_1 and a_2 for a_i and a_{-i}, respectively, that lead to the penalty term $\left(a_{-i} - a_i\right)^2$ in the wage equation (5) taking its expected value $2(1-\kappa)\sigma^2$.[1] This point of approximation ensures that both, the earnings potential ψ, which reflects individual characteristics, and the log wage w take their average values. It can thus be understood as the centroid of a regression of log wages on the individual characteristics embodied in the earnings potential ψ.

[1] The expected value of $\left(a_{-i} - a_i\right)^2$ is $\mathrm{E}\left(a_i - a_{-i}\right)^2 = \mathrm{E}\left(a_i^2 - 2a_i a_{-i} + a_{-i}^2\right) = 2\,\mathrm{E}\left(a_i^2\right) - 2\,\mathrm{E}\left(a_i a_{-i}\right) = 2\left(\mathrm{E}\left(a^2\right) - \mathrm{E}\left(a\right)^2 - \mathrm{cov}(a_i, a_{-i})\right) = 2(\mathrm{var}\left(a\right) - \mathrm{cov}(a_i, a_{-i})) = 2(\sigma^2 - \kappa\sigma^2) = 2(1-\kappa)\sigma^2$.

Applying the approximation gives

$$w_i \approx \beta_0 + \beta_{1,g(i)} X_i + \beta_{2,g(i)} X_{-i} + \varepsilon_i, \tag{6}$$

where

$$\begin{aligned}
\beta_0 &= \log \psi - \log \left(1 - \frac{1}{2}(1-\kappa)\sigma^2\right), \\
\beta_{1,g(i)} &= \left(1 + \frac{1}{\sqrt{2}} \cdot \frac{(1-\kappa)\sigma^2}{1-(1-\kappa)\sigma^2}\right) \cdot \gamma_{g(i)}, \\
\beta_{2,g(i)} &= -\frac{1}{\sqrt{2}} \cdot \frac{(1-\kappa)\sigma^2}{1-(1-\kappa)\sigma^2} \cdot \gamma_{g(-i)},
\end{aligned}$$

and

$$\varepsilon_i = \frac{\sqrt{2(1-\kappa)\sigma^2}}{2-(1-\kappa)\sigma^2} \left(a_{-i} - a_i\right),$$

see Appendix for a derivation. Condition (6) can be read as a regression equation. In a regression of the log wage on the worker's own characteristics and the partner's characteristics, β_0 is a constant, $\beta_{1,g(i)}$ and $\beta_{2,g(i)}$ are vectors of coefficients, and ε_i is a (mean-zero) residual since ideal locations a_i and a_{-i} cannot be observed by the econometrician. Note that the entries in $\beta_{1,g}$ tend to have the opposite sign compared to their counterparts in $\beta_{2,g}$. Consider, for example, a characteristic that is wage promoting for both men and women (i.e., for which the corresponding entries in γ_m and γ_f are positive). For this characteristic, the associated entry in $\beta_{1,g}$ is positive whereas the associated entry in $\beta_{2,g}$ is negative.[2] Hence, a characteristic of a worker influences the worker's own wage and the wage of the worker's partner in opposite directions. In our model, this relation is due to career prioritization.

2.3 Similar mechanisms in the literature

The key implication of our model is that observed wages do not only depend on individual characteristics but also on the characteristics of the partner. This implication can also be derived from other approaches that in general emphasize the role of the family for wages paid to women and men. Our model is similar to Foged (2016) who also provides a model of the joint location choice of dual-earner households but focuses on the extensive-margin choice *whether* to move to another location rather than the intensive-margin choice *where* to locate. Also in Foged (2016), wages depend on location and it is rational for a household to decide on a location that rather promotes the designated primary earner's career. This tends to have negative consequences for wage rates paid to the secondary earner.

[2] The term $(1-\kappa)\sigma^2/(1-(1-\kappa)\sigma^2)$ is weakly positive because $0 \leq \kappa \leq 1$ and $0 \leq \sigma^2 \leq 0.25$ as $a \in (0,1)$.

Bredemeier (2019) shows that wages are affected by earner roles in the household through the choice of which employer to work for. In his model, there is a trade-off between pay and non-pay attributes of jobs and high earnings of the partner reduce the importance of the pay dimension in one's own employer choice. As a result, designated secondary earners weigh non-pay job attributes rather heavily when choosing employers and the wage sensitivity of an individual's job choice depends positively on the share that the individual contributes to household income. Firms with monopsonistic power on the labor market exploit this and pay lower wages to individuals married to partners with high earnings potentials. Relatedly, Petrongolo and Ronchi (2020) provide evidence that women more often than men trade off better earnings for non-pay job attributes such as shorter commutes or flexible work schedules. Arguably, the importance of these attributes reflects women's role as the primary child-care provider in most households – which can be expected to be more pronounced the higher is the husband's earnings potentials relative to the wife's one.

Cortés and Tessada (2011) and Cortés and Pan (2019) propose another channel for the link between an individual's wage rate and the respective partner's characteristics. In occupations where wages are highest, individuals have to work long hours to have a successful career. For the family, the cost of supplying long working hours is convex, i.e., working long hours is more costly if one's partner is already working long hours, for example due to child-care obligations. Then, the optimal time allocation mostly promotes the designated primary earner's career while designated secondary earners may forego important investments into their careers. Cortés and Tessada (2011) show that a decrease in the costs of services that are close substitutes to household production increases the labor supply of highly skilled women. The effect is strongest in occupations where success is related to working longer hours.[3] Thus, the results indicate that households did not prioritize women's careers before the cost reduction. Hence, restrictions on affordable household help and the resulting optimal time allocation between spouses reveal the link between wages and an individual's role in the family.

3 Standard and extended Oaxaca-Blinder decompositions

The usual approach to decompose the gender wage gap into a part explained by differences in observable characteristics and an unexplained part is the Oaxaca-Blinder decomposition (Oaxaca 1973, Blinder 1973). In this section, we first recapitulate the standard Oaxaca-

[3]Cortés and Pan (2019) show that facilitated access to substitutes for household production also increases women's entry into these occupations.

Blinder decomposition and show analytically that the standard approach is subject to a bias when observed wages depend on the characteristics of the partner as, for instance, in the model of career prioritization presented in Section 2. We then show analytically that accounting for the characteristics of the partner yields an unbiased decomposition.

3.1 Bias in the standard Oaxaca-Blinder decomposition

The first step of the Oaxaca-Blinder decomposition is to estimate a log wage equation, separately for men ($g = m$) and women ($g = f$):

$$w_i = b_{0,g(i)} + b_{1,g(i)} \cdot X_i + \epsilon_i, \tag{7}$$

where $b_{0,g(i)}$ is a constant, $b_{1,g(i)}$ is a vector of coefficients, and ϵ_i is a residual. The (average) gender wage gap $\Delta = \overline{w}_m - \overline{w}_f$, where $\overline{w}_g$ denotes average log wages by gender, is then decomposed as

$$\begin{aligned}\Delta = \overline{w}_m - \overline{w}_f &= \widehat{b}_{0,m} - \widehat{b}_{0,f} + \widehat{b}_{1,m}\overline{X}_m - \widehat{b}_{1,f}\overline{X}_f \\ &= \underbrace{\widehat{b}_{1,m} \underbrace{(\overline{X}_m - \overline{X}_f)}_{\substack{\text{Differences} \\ \text{in covariates}}}}_{\widehat{\Delta}|_X \text{ (“explained”)}} + \underbrace{\overline{X}_f \underbrace{(\widehat{b}_{1,m} - \widehat{b}_{1,f}) + \widehat{b}_{0,m} - \widehat{b}_{0,f}}_{\substack{\text{Differences} \\ \text{in coefficients}}}}_{\widehat{\Delta}|_b \text{ (“unexplained”)}},\end{aligned} \tag{8}$$

where $\overline{X}_g$ denotes gender-specific average characteristics. The first term on the right-hand side of (8), $\widehat{\Delta}|_X = \widehat{b}_{1,m}(\overline{X}_m - \overline{X}_f)$, is the estimated explained part of the gap and gives the impact of gender differences in covariates, evaluated by the male wage equation. The second term on the right-hand side of (8), $\widehat{\Delta}|_b = \overline{X}_f(\widehat{b}_{1,m} - \widehat{b}_{1,f}) + \widehat{b}_{0,m} - \widehat{b}_{0,f}$, is the estimated unexplained part of the gap and captures the differences in estimated coefficients between men and women. It can be interpreted as the difference in how the male wage equation *would* value the characteristics of women and how the female equation actually values them (Blinder 1973).[4]

Oaxaca-Blinder decompositions in the model. In order to illustrate the bias in Oaxaca-Blinder decompositions of the wage gap, we assume that the vector X consists of just one pay-relevant observable characteristic. For simplicity, we consider a characteristic that is wage enhancing, i.e., $\gamma_m > 0$ and $\gamma_f > 0$, e.g., experience.

[4]While one could also calculate the explained wage gap using the coefficients of the female wage equation, it has become standard to use the coefficients from the male wage equation because, supposedly, they reflect the discrimination-free returns to characteristics.

Using the relation between wages and observables derived from our theoretical model, (6), the gender wage gap in the model is

$$\begin{aligned}\Delta = \overline{w}_m - \overline{w}_f =& \beta_{1,m}\overline{X}_m + \beta_{2,m}\overline{X}_f - \beta_{1,f}\overline{X}_f - \beta_{2,f}\overline{X}_m \\ =& \left(\beta_{1,m} - \beta_{2,f}\right)\overline{X}_m - \left(\beta_{1,f} - \beta_{2,m}\right)\overline{X}_f,\end{aligned} \tag{9}$$

which can be decomposed as

$$\Delta = \underbrace{(\beta_{1,m} - \beta_{2,m}) \cdot (\overline{X}_m - \overline{X}_f)}_{\Delta|_X} + \underbrace{(\beta_{1,m} - \beta_{1,f}) \cdot \overline{X}_f + (\beta_{2,m} - \beta_{2,f}) \cdot \overline{X}_m}_{\Delta|_\beta}. \tag{10}$$

The first term on the right-hand side, $\Delta|_X$, is the wage gap that is due to gender differences in the observable characteristic X. It comprises both the effect that these characteristics exert on one's own wage and the one that they exert on one's partner's wage. The second term, $\Delta|_\beta$, is the wage gap that is due to gender-specific wage setting, or discrimination – it is zero when the coefficients are the same for both genders.

Model version without discrimination. For purposes of illustration, we first consider a model version where wage setting is gender neutral, i.e., for a given characteristic X_i, ideal location a_i, and actual location r_i, the wage does not depend on gender. In this situation, we have

$$\gamma_m = \gamma_f$$

and, consequently,

$$\beta_{1,m} = \beta_{1,f} = \beta_1$$

as well as

$$\beta_{2,m} = \beta_{2,f} = \beta_2.$$

The true wage gap (10) in this example simplifies to

$$\Delta = \overline{w}_m - \overline{w}_f = (\beta_1 - \beta_2) \cdot (\overline{X}_m - \overline{X}_f) = \Delta|_X. \tag{11}$$

Note that, in this example, there is no unexplained wage gap *by construction.* We now apply the standard Oaxaca-Blinder decomposition to this example. In the standard approach, one estimates gender-specific wage regressions without partner characteristics, (7), instead of the true model (6). The estimate for $b_{1,m}$ from the standard approach is

$$\widehat{b}_{1,m} = \beta_1 + \beta_2 \cdot \frac{\text{cov}(X_m, X_f)}{\text{var}(X_m)}$$

and, according to (8), the standard Oaxaca-Blinder approach yields the explained part of the gap

$$\widehat{\Delta}|_X = \left(\beta_1 + \beta_2 \cdot \frac{\operatorname{cov}(X_m, X_f)}{\operatorname{var}(X_m)}\right) \cdot (\overline{X}_m - \overline{X}_f).$$

The *true* explained gap is given by $\Delta|_X = \Delta = (\beta_1 - \beta_2) \cdot (\overline{X}_m - \overline{X}_f)$, see (11). Importantly, the explained gap derived from the standard Oaxaca-Blinder decomposition is in general not equal to the true explained gap.[5]

The standard Oaxaca-Blinder decomposition suffers from two problems. First, the estimated coefficients in the wage equation are potentially biased due to the omission of the partner characteristics, so that the contribution of the worker's own characteristics is not measured correctly. Second, the decomposition misses the differences in the omitted covariates so that the contribution of the characteristics of the worker's partner are not accounted for. This implies that, although the entire wage gap in this model version is caused by gender differences in characteristics, a standard Oaxaca-Blinder decomposition labels some part of the gap "unexplained" and hence seemingly detects "discrimination".

Discrimination in the model. We now consider the case where the model accounts for discrimination, i.e., $\gamma_m \neq \gamma_f$, and show that a standard Oaxaca-Blinder decomposition assigns too large a share of the wage gap to discrimination.

The standard Oaxaca-Blinder wage regression yields

$$\widehat{b}_{1,m} = \beta_{1,m} + \beta_{2,m} \cdot \frac{\operatorname{cov}(X_m, X_f)}{\operatorname{var}(X_m)} \tag{12}$$

and an explained gender wage gap of

$$\widehat{\Delta}|_X = \widehat{b}_{1,m} \cdot \left(\overline{X}_m - \overline{X}_f\right) = \left(\beta_{1,m} + \beta_{2,m} \cdot \frac{\operatorname{cov}(X_m, X_f)}{\operatorname{var}(X_m)}\right) \cdot \left(\overline{X}_m - \overline{X}_f\right).$$

As a comparison, the gap which is truly due to differences in the characteristic X is $\Delta|_X = (\beta_{1,m} - \beta_{2,m}) \cdot (\overline{X}_m - \overline{X}_f)$, see (9). Our model implies $\beta_{2,m} < 0$ due to career prioritization and hence the estimated explained gap is less than the true explained gap,

$$\widehat{\Delta}|_X < \Delta|_X,$$

unless by coincidence $\operatorname{cov}(X_m, X_f) / \operatorname{var}(X_m) = -1$. Put differently, a standard Oaxaca-Blinder decomposition identifies too large a share of the wage gap as unexplained.

[5] The only case in which those two gaps are equal is $\widehat{b}_{1,m} = \beta_1 - \beta_2$. This would only be the case if there were perfectly negative assortative mating, i.e., if and only if $\operatorname{cov}(X_i, X_{-i}) / \operatorname{var}(X_i) = -1$.

3.2 Extended Oaxaca-Blinder decomposition

The bias in the Oaxaca-Blinder decomposition can be eliminated if the wage equation is augmented by the characteristics of the partner. Consider gender-specific wage regressions of the form

$$w_i = b_{0,g(i)} + b_{1,g(i)} \cdot X_i + b_{2,g(i)} \cdot X_{-i} + \epsilon_i. \tag{13}$$

Estimating (13) gives the estimated coefficients

$$\widehat{b}_{1,m} = \beta_{1,m},\ \widehat{b}_{2,m} = \beta_{2,m},\ \widehat{b}_{1,f} = \beta_{1,f} \text{ and } \widehat{b}_{2,f} = \beta_{2,f}.$$

The estimated explained gap then is

$$\widehat{\Delta}|_X = \widehat{b}_{1,m}(\overline{X}_m - \overline{X}_f) + \widehat{b}_{2,m}(\overline{X}_f - \overline{X}_m) = (\beta_{1,m} - \beta_{2,m})\,(\overline{X}_m - \overline{X}_f)$$

and corresponds to the true explained gap $\Delta|_X$, see (10). The estimated unexplained gap is $\widehat{\Delta}|_b = \Delta - \widehat{\Delta}|_X = (\beta_{1,m} - \beta_{1,f}) \cdot \overline{X}_f + (\beta_{2,m} - \beta_{2,f}) \cdot \overline{X}_m$ and equals the true unexplained gap $\Delta|_\beta$. Thus, accounting for the characteristics of the partner in the wage equation yields an unbiased decomposition.

Comparing the standard and extended Oaxaca-Blinder decompositions. The difference in the explained gap between the extended and the standard decomposition can itself be decomposed into two parts, the difference in the part of the wage gap that is explained by workers' own characteristics that follows from different parameter estimates and the contribution of the characteristics of the worker's partner. Formally,

$$(\Delta|_X)^{ext} - (\Delta|_X)^{std} = \underbrace{\left(\widehat{b}^{ext}_{1,m} - \widehat{b}^{std}_{1,m}\right) \cdot \left(\overline{X}_m - \overline{X}_f\right)}_{\text{change in contribution of own characteristics}} + \underbrace{\widehat{b}^{ext}_{2,m} \cdot \left(\overline{X}_f - \overline{X}_m\right)}_{\text{contribution of partner characteristics}},$$

where $(\Delta|_X)^{std}$ is the explained gap according to the standard Oaxaca-Blinder decomposition and $(\Delta|_X)^{ext}$ denotes the explained gap from our extended approach. For the following, we assume that $\overline{X}_m > \overline{X}_f$ (which is a convention and induces no loss of generality), that $\beta_{1,m} > 0$ (i.e., that some positive part of the gender wage gap can be explained through observable characteristic X), and that $\text{cov}\,(X_m, X_f) > 0$ (at least some positive assortative mating).

Then, only if there is career prioritization, i.e., $\beta_{2,m} < 0$, the extended decomposition explains larger shares of the gender wage gap than the standard decomposition. Formally,

$$\beta_{2,m} < 0 \Leftrightarrow (\Delta|_X)^{ext} - (\Delta|_X)^{std} > 0. \tag{14}$$

This also holds for both components of the change in the explained gap,

$$\beta_{2,m} < 0 \Leftrightarrow \left(\widehat{b}_{1,m}^{ext} - \widehat{b}_{1,m}^{std}\right) \cdot \left(\overline{X}_m - \overline{X}_f\right) = -\beta_{2,m} \cdot \frac{\operatorname{cov}\left(X_m, X_f\right)}{\operatorname{var}\left(X_m\right)} \cdot \left(\overline{X}_m - \overline{X}_f\right) > 0 \tag{15}$$

where the last step uses the estimate from the standard Oaxaca-Blinder approach, (12), and

$$\beta_{2,m} < 0 \Leftrightarrow \widehat{b}_{2,m}^{ext} \cdot \left(\overline{X}_f - \overline{X}_m\right) = -\beta_{2,m} \cdot \left(\overline{X}_m - \overline{X}_f\right) > 0. \tag{16}$$

Thus, career prioritization has several testable implications for Oaxaca-Blinder decompositions. First, the explained part of the gender gap is larger for the extended decomposition than for the standard decomposition, $(\Delta|_X)^{ext} - (\Delta|_X)^{std} > 0$. Second, both elements (15) and (16) of the difference between the explained gaps resulting from the two decompositions are predicted to be positive. Further, (15) and (16) show that the inclusion of additional (partner) characteristics does not mechanically increase the explained fraction of the gender gap. This only happens if the data are consistent with career prioritization (i.e., $\beta_{2,m} < 0$).

4 Empirical analysis

In this section, we apply our extended Oaxaca-Blinder decomposition empirically using data from the Panel Study of Income Dynamics (PSID). The PSID is the most suited U.S. data set for decompositions of the gender wage gap as it has information on actual labor market experience, a key explanatory variable for the gender wage gap. For comparability to the literature, we follow Blau and Kahn (2017) in terms of sample selection, and in the choice and definition of explanatory variables. As Blau and Kahn (2017), we use data for the years 1980, 1989, 1998, and 2010.[6]

4.1 Sample selection, explanatory variables, and descriptive statistics

Sample. We consider different subsamples of the Blau and Kahn (2017) sample, most importantly the subsample of workers living in dual-earner households.[7] Blau and Kahn (2017) select employees between ages 25 and 64 working full-time in the non-farm/non-military sector for at least 26 weeks per year, excluding the self-employed as well as the immigrant and Latino samples.[8]

[6]Earnings in the PSID refer to the previous year. Hence, we use, e.g., the 1981 data to measure wages in 1980.

[7]In later evaluations, we also consider samples of singles (defined as individuals with no partner, neither married nor cohabiting) and single earners (defined as individuals who are the sole earner in their household independent of marital or cohabitation status).

[8]As is standard, full-time is defined as being employed and working at least 35 hours per week.

To construct a sample of workers living in dual-earner households, which is necessary for our extended Oaxaca-Blinder decomposition, we restrict the Blau-Kahn sample to married or cohabiting individuals with employed spouses for whom all relevant variables are observed. For an individual to be included in our dual-earner sample, neither is the partner required to work full-time nor has an hourly wage rate to be observed for the partner. As these requirements have to be met only for the individual himself, our dual-earner sample contains more men than women, mostly because part-time rates are higher for women. Overall, our dual-worker sample includes roughly 50% of the workers in the Blau-Kahn sample.[9] We will show that our dual-earner sample is similar to the Blau-Kahn sample regarding trends in the gender wage gap and in key explanatory variables as well as with respect to results from standard Oaxaca-Blinder decompositions. This is important as it ensures that differences between the results of our extended Oaxaca-Blinder decomposition and the standard decomposition are in fact due to the methodological extension and are not driven by the different samples.

Hourly wage rates and explanatory variables. The hourly wage rate is calculated as annual labor earnings divided by annual hours worked. The preferred specification of the wage equation in Blau and Kahn (2017) uses as explanatory variables the individual's education (years of schooling and dummy variables for bachelor and master degrees) and experience (years of full-time experience, years of part-time experience), race or ethnicity, Census region dummies, a dummy for living in a metropolitan area, as well as variables containing job information, such as industry (15 two-digit groups, 2000 Census classification), occupation (21 two-digit groups, 2000 Census classification), union coverage, and whether the respondent is working for the government. For our extended decomposition, we augment the wage equation by the partner's education, experience, and job information.[10]

Oaxaca-Blinder decompositions do not aim at identifying causal relations between variables but are merely accounting tools used to assess how much pay differences can be related to differences in observable characteristics. In our context, it is nonetheless important to discuss in how far the additional explanatory (partner) variables added to the wage equation in our extended Oaxaca-Blinder decomposition reflect *choices* of the dual-earner couple. Recall that our theoretical mechanism runs from characteristics of the individual spouses to

[9] Our dual-earner sample contains 902 (in 1980), 1311 (in 1989), 1301 (in 1998) and 1179 (in 2010) men; and 668 (in 1980), 990 (in 1989), 1051 (in 1998), 977 (in 2010) women. In the Blau-Kahn sample, there are 2280 (in 1980), 2614 (in 1989), 2454 (in 1998), 2365 (in 2010) men and 1491 (in 1980), 2053 (in 1989), 2161 (in 1998), 2448 (in 2010) women.

[10] The partner's race or ethnicity, region of residence, and metropolitan status are not included due to collinearity to the corresponding information for the individual itself.

wage-relevant (joint) choices of the couple. While almost all of the explanatory variables described above constitute choices, it makes sense to consider most of them characteristics from the perspective of our model. Education is typically chosen before couple households form and is hence not subject to the joint decision making which is key to our mechanism. Empirical evidence shows that industry and occupation are rarely switched and doing so entails substantial costs, see, e.g., Kambourov and Manovskii (2009), Sullivan (2010), Artuç and McLaren (2015), or Cortes and Gallipoli (2018). Thus, individuals' initial choices on industry and occupation, which for most individuals occur before formation of the marriage, are of significant importance during marriage but usually not subject to joint decision making. Arguably, the accumulation of work experience and the lack thereof occurs during the course of the marriage and is to some part a decision of the couple that may take into account anticipated differences in returns to experience. However, one can also argue that career interruptions are mostly caused by child births and the absence of affordable child care and that their distribution within the couple is to a large extent driven by norms (Bertrand et al. 2015, Blau et al. 2020). Union coverage is mostly determined by the choice of employer and hence a joint decision of the couple from the viewpoint of our model. We nevertheless include this variable in the set of explanatory variables in order to maintain full comparability to Blau and Kahn (2017).

Descriptive statistics. The first part of Table 1 shows average log wage rates by gender as well as the gender wage gap for our dual earner sample (Columns (1) through (4)) as well as for the Blau-Kahn sample (Columns (5) through (7)). Both samples display the substantial decrease of the gender wage gap and the slowing down of the convergence in later years (Goldin 2014).

The table also summarizes education and full-time experience by gender for both samples together with developments of other determinants of wages related to job information.[11] Both samples show the well-known reversal of the gender gap in education and women's catching up in terms of full-time experience. Women less often than men work in managerial occupations but more often in professional occupations. In both types of occupations, female shares are increasing over time. Despite their strong representation in professional occupations in general, women are still the minority in the high-paying professional occupations traditionally dominated by men, such as lawyers and doctors. Union coverage rates and gender differences therein are similar in both samples with women being less frequently covered by collective-bargaining agreements than men in early years and similarly often in recent years. Overall,

[11] The underlying categorization of occupations and industries follows Blau and Kahn (2017).

Table 1: Log wages, human capital, and job attributes by gender, year, and sample.

	Dual-earner sample				Blau-Kahn sample		
Year	Men (1)	Women (2)	Difference (3)	Corr(X_i, X_{-i}) (4)	Men (5)	Women (6)	Difference (7)
Log Wage Rates							
1980	3.08	2.65	0.43	0.36	3.08	2.60	0.48
1989	3.09	2.77	0.32	0.39	3.06	2.76	0.30
1998	3.16	2.89	0.27	0.29	3.11	2.85	0.26
2010	3.29	3.04	0.25	0.37	3.24	3.00	0.23
Years of schooling							
1980	13.09	13.05	0.04	0.61	13.13	12.96	0.17
1989	13.65	13.54	0.11	0.55	13.57	13.51	0.05
1998	14.06	14.16	-0.10	0.53	13.93	13.98	-0.06
2010	14.32	14.62	-0.31	0.50	14.32	14.48	-0.16
Bachelor (in %)							
1980	17.25	15.22	2.03	0.33	17.32	15.99	1.33
1989	19.30	16.59	2.71	0.26	20.00	18.09	1.91
1998	23.54	24.01	-0.47	0.25	23.34	22.53	0.81
2010	24.83	26.52	-1.69	0.27	26.24	24.78	1.46
Advanced degree (in %)							
1980	8.33	6.86	1.47	0.32	8.12	6.09	2.04
1989	10.10	8.28	1.81	0.30	9.63	8.37	1.26
1998	11.99	12.24	-0.25	0.31	11.05	10.22	0.83
2010	13.41	17.86	-4.46	0.30	12.90	15.73	-2.83
Years of full-time experience							
1980	21.92	13.08	8.83	0.51	20.32	13.51	6.81
1989	20.46	13.48	6.98	0.43	19.15	14.72	4.44
1998	21.46	15.15	6.31	0.51	19.77	15.93	3.84
2010	18.95	15.06	3.89	0.66	17.80	16.35	1.44
Managerial jobs (in %)							
1980	21.42	8.92	12.50	0.10	21.52	9.18	12.34
1989	22.06	11.85	10.21	0.13	20.87	10.94	9.94
1998	22.55	16.47	6.08	0.11	21.87	15.44	6.43
2010	19.21	16.81	2.40	0.15	18.35	16.20	2.15
Professional jobs (in %)							
1980	17.83	23.17	-5.34	0.26	17.08	21.80	-4.72
1989	19.27	25.08	-5.82	0.23	19.45	26.11	-6.67
1998	21.41	28.48	-7.07	0.18	20.47	26.95	-6.48
2010	21.88	30.13	-8.26	0.25	21.70	31.09	-9.39
"Male" professional jobs (in %)							
1980	14.32	9.53	4.79	0.15	14.60	10.10	4.50
1989	16.37	13.40	2.97	0.17	17.32	14.14	3.18
1998	18.19	13.59	4.60	0.16	17.61	13.16	4.45
2010	18.37	15.04	3.33	0.14	18.59	17.78	0.81
Collective-bargaining coverage (in %)							
1980	34.36	19.97	14.39	0.20	34.51	21.14	13.37
1989	25.25	18.10	7.15	0.27	25.46	19.30	6.16
1998	21.86	20.08	1.78	0.19	21.44	18.09	3.35
2010	17.71	19.44	-1.73	0.24	17.45	18.94	-1.49

Notes: Columns (1), (2), (5), and (6) show gender-specific averages. Columns (3) and (7) show male average minus female average. Column (4) shows correlation between own and partner characteristics in the sample of male workers in dual-earner couples. "Male" professional occupations are professional occupations other than nurses and non-college teachers.

we conclude that the dual-earner sample and the Blau-Kahn sample have similar properties regarding gender gaps in wage determinants and their trends. Table 1 also shows that pay-relevant characteristics are positively correlated between spouses in dual-earner couples. This supports the assortative-mating assumption applied in Section 3.[12]

4.2 Empirical results

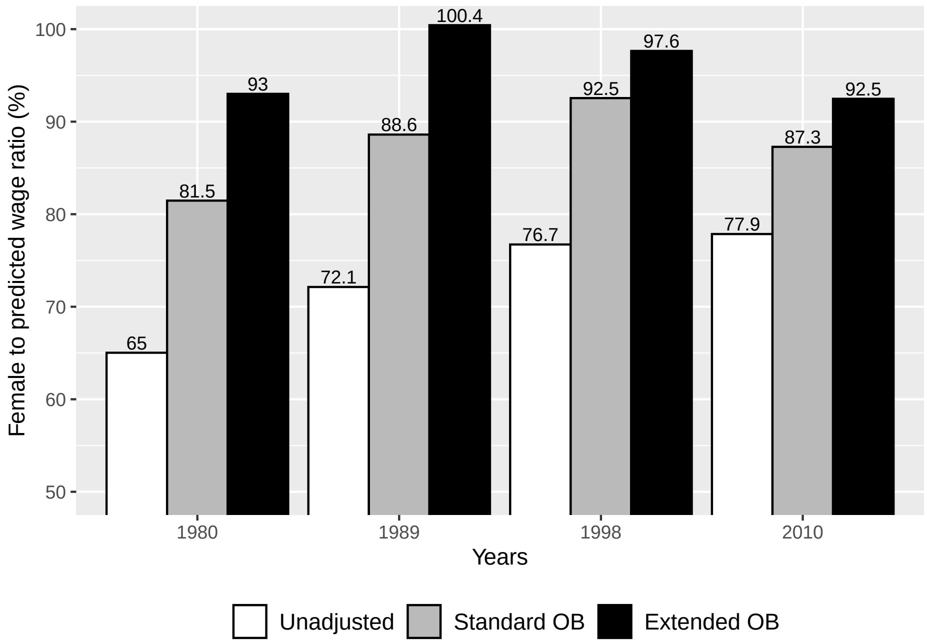
Baseline results. Figure 1 shows the results of Oaxaca-Blinder decompositions in the dual-earner sample. Following Blau and Kahn (2017), we display the inverse exponential of the raw wage gap Δ and of the unexplained wage gap $\widehat{\Delta}|_b$, hence the level of the gap in log points can (approximately) be seen in the figure as the difference between the bars and 100%. The inverse exponential of the raw gap, $1/\exp(\Delta)$, is the unadjusted ratio of women's mean wage rate the one of men. The inverse exponential of the unexplained gap is the adjusted wage ratio, i.e., the ratio of the average wage women actually earn and the average wage women would earn if their characteristics were priced in the same way by the labor market as men's (i.e., if they had the same coefficients as men). The white bars show the unadjusted wage ratios, i.e., correspond to the raw gender wage gaps. The gray bars show the results from the standard Oaxaca-Blinder decomposition. The black bars show the results from our extended approach, where we augment the wage equation by the characteristics of the partner.

The white bars show the substantial closure of the gender wage gap during the 1980s and the slowing down of the convergence in later years. The gray bars show that a standard Oaxaca-Blinder decomposition explains a substantial amount of the gender wage gap, as discussed by Blau and Kahn (2017). However, a substantial gap in adjusted wages remains. The adjusted wage ratio stagnates at around 90% from 1989 on. Put differently, a gap of roughly 10 percentage points, which corresponds to between one third and three fifths of the raw gap, remains unexplained by a standard Oaxaca-Blinder decomposition. Note that the results for our dual-earner sample are similar to the ones for the Blau-Kahn sample. Specifically, in their full specification, Blau and Kahn (2017) report adjusted wage ratios of 79.4%, 92.4%, 91.4% and 82.1%, respectively. Thus, moving from the Blau-Kahn sample to our sample of dual-earner households does not affect the results of the standard Oaxaca-Blinder decomposition substantially.

The most important result of our analysis is that, in all years, the adjusted wage ratios using our extended Oaxaca-Blinder decomposition (black bars) are substantially larger than the adjusted wage ratios indicated by the standard approach (gray bars), in line with our

[12]The correlation in full-time experience is mostly driven by the high correlation in spouse's age. The conditional correlation is relatively small.

Figure 1: Comparison of standard Oaxaca-Blinder (OB) decomposition and extended decomposition using partner characteristics, dual-earner sample.



Notes: White bars show $1/\exp(\Delta)$. Gray bars show $1/\exp((\widehat{\Delta}|_b)^{std})$. Black bars show $1/\exp((\widehat{\Delta}|_b)^{ext})$.

analytical example presented in Section 3. In 1989, our extended Oaxaca-Blinder decomposition explains 100% of the gap. For the other years, a small unexplained gap remains but it is considerably smaller than the gap that remains unexplained by the standard decomposition. Thus, accounting for partner characteristics allows to explain a substantially larger part of the gender wage gap.

As discussed before, the inclusion of partner characteristics does not mechanically increase the fraction of the gender wage gap that can be explained using an Oaxaca-Blinder decomposition. Whether this happens depends on the signs of the coefficients in the wage equation. Suppose coefficients on partner characteristics would tend to have the same sign as the respective coefficients on one's own characteristics. Then, including characteristics of the partner into the decomposition would *decrease* rather than increase the explained gap, see (14). Only when the coefficients are mostly of opposite sign, in line with career prioritization, including partner characteristics actually increases the explained part of the gap. The empirical results align well with our model of career prioritization and its predictions for decompositions of the gender wage gap. Given the positive correlation of characteristics within couples (see Table 1), the fact that the inclusion of partner characteristics increases the explained fraction of the wage gap indicates that, in general, the coefficients on one's own characteristics and on one's partner's characteristics in the wage equation are of opposite sign, see (14), and hence corroborates the presence of career prioritization. For 1989, the data are even consistent with the model version without wage discrimination. For this year, we can understand gender differences in wages as simply reflecting gender differences in pay relevant characteristics when we take into account the role of partner characteristics. The results for the other years are in line with our model version that includes wage discrimination against women. For these years, a standard Oaxaca-Blinder decomposition overestimates the degree of discrimination substantially.

Figure 1 also shows that the part of the gap that remains unexplained by the standard Oaxaca-Blinder decomposition (roughly the difference between the gray bars and 100%) declines substantially over time. This seems to suggest that the closure of the wage gap between 1980 and 2010 can to a discernible part be attributed to declining discrimination. This interpretation, however, is not supported by our extended Oaxaca-Blinder decomposition which delivers a roughly constant unexplained gender gap amounting to about 7 percentage points in both 1980 and 2010.

Additional decompositions. We now decompose the explained gender wage gap into the parts explained by different types of variables. Given our focus on the role of partner characteristics for wages, a natural starting point is to distinguish between workers' own characteristics and their partners' characteristics. The results of this decomposition are summarized in Table 2.[13] For every part of the wage gap, the upper number gives the log difference while the lower number in parentheses gives the share of the total wage gap. E.g., for the explained gap, the upper number is $\widehat{\Delta}|_X$ and the lower number in parentheses is $\widehat{\Delta}|_X/\Delta$. Workers' own characteristics explain between 46% and 71% of the gender wage gap according to a standard Oaxaca-Blinder decomposition. In the extended Oaxaca-Blinder decomposition, these shares are raised to 53% to 76%. In line with the predictions of our model, the extended Oaxaca-Blinder decomposition assigns a larger share of the gender wage gap to workers' own characteristics in every considered year. The extended Oaxaca-Blinder decomposition also informs about how much of the wage gap can be assigned to partner characteristics. Also in line with our model, partner characteristics explain a positive part of the wage gap in every considered year and this part amounts to numbers between 11% and 34%.

[13]For simplicity, we subsume race and region in the 'own characteristics' category. These variables do not explain much of the gender wage gap, see Table 3.

Table 2: Decomposition of gender wage gap into parts explained by observable characteristics, standard and extended decomposition.

	1980	1989	1998	2010
Wage gap	0.430	0.327	0.265	0.250
Standard decomposition				
total explained	0.225	0.206	0.188	0.114
	(52%)	(63%)	(71%)	(46%)
unexplained	0.205	0.121	0.077	0.136
	(48%)	(37%)	(29%)	(54%)
Extended decomposition				
total explained	0.357	0.331	0.241	0.172
	(83%)	(101%)	(91%)	(69%)
own characteristics	0.229	0.22	0.201	0.145
	(53%)	(67%)	(76%)	(58%)
partner characteristics	0.128	0.111	0.040	0.027
	(30%)	(34%)	(15%)	(11%)
unexplained	0.072	-0.004	0.024	0.078
	(17%)	(-1%)	(9%)	(31%)

Notes: First line shows log differences, second line (in parentheses) gives percentage of total wage gap.

The contribution of the partner characteristics to the gender wage gap can also be interpreted as a measure of the importance of career prioritization for the wages of men and women. Formally, the contribution of partner characteristics is given by $\beta_{2,m} \cdot (\overline{X}_f - \overline{X}_m)$ and hence measures the reduction in men's wages that would result from their partners having the same characteristics as they themselves. In a counterfactual situation where every man in a dual-earner marriage would be married to a wife whose characteristics are identical to his, the incentive to prioritize the husband's career due to superior characteristics would be shut off. The results in Table 2 strongly indicate that men's wages are fostered by households prioritizing their careers. If their wives had the same characteristics and, hence, incentives for households to prioritize men's careers were smaller, men would earn substantially less. In the early years of our sample period, this channel makes up for more than 10% of men's wages and about one-third of the wage gap. For the year 2010, it still contributes one-tenth of the gender wage gap. For the quantitative interpretation, note that our model predicts that career prioritization induced by differences in characteristics suppresses women's wages to a similar degree as men's wages are promoted. It should further be noted that the thought experiment applied here does not totally shut off career prioritization. When the labor market discriminates against women, also couples where husband and wife have identical characteristics have

Table 3: Detailed decomposition of gender wage gap, standard and extended decomposition.

	1980	1989	1998	2010
Wage gap	0.430	0.327	0.265	0.250
Standard decomposition				
total explained	0.225	0.206	0.188	0.114
	(52%)	(63%)	(71%)	(46%)
human capital	0.109	0.087	0.067	0.033
	(25%)	(27%)	(25%)	(13%)
job information	0.107	0.111	0.116	0.079
	(25%)	(34%)	(44%)	(32%)
race and region	0.010	0.008	0.004	0.002
	(2%)	(2%)	(2%)	(1%)
unexplained	0.205	0.121	0.077	0.136
	(48%)	(37%)	(29%)	(54%)
Extended decomposition				
total explained	0.357	0.331	0.241	0.172
	(83%)	(101%)	(91%)	(69%)
own characteristics	0.229	0.220	0.201	0.145
	(53%)	(67%)	(76%)	(58%)
human capital	0.130	0.117	0.080	0.039
	(30%)	(36%)	(30%)	(16%)
job information	0.090	0.097	0.117	0.104
	(21%)	(30%)	(44%)	(42%)
race and region	0.008	0.006	0.003	0.002
	(2%)	(2%)	(1%)	(1%)
partner characteristics	0.128	0.111	0.040	0.027
	(30%)	(34%)	(15%)	(11%)
human capital	0.040	0.077	0.045	0.032
	(9%)	(24%)	(17%)	(13%)
job information	0.088	0.034	-0.004	-0.006
	(20%)	(10%)	(-2%)	(-2%)
unexplained	0.072	-0.004	0.024	0.078
	(17%)	(-1%)	(9%)	(31%)

Notes: First line shows log differences, second line (in parentheses) gives percentage of total wage gap. "human capital": education and experience; "job information": union coverage, industry, occupation, working for government.

incentives to prioritize the husband's career because the husband can earn a higher return on these characteristics. Hence, the contribution of partner characteristics documented in Table 2 quantifies only a part of the wage effects of career prioritization in dual-earner couples.

In the next step, we decompose wage gaps further and distinguish between human-capital variables (education and experience) and job information (union status, industry, and oc-

Table 4: Comparison of standard and extended decomposition.

	1980	1989	1998	2010
Increase in explained gap	0.131	0.125	0.054	0.058
$(\Delta\vert_X)^{ext} - (\Delta\vert_X)^{std}$	(30%)	(38%)	(20%)	(23%)
Change in contribution	0.004	0.015	0.013	0.031
of own characteristics	(1%)	(5%)	(5%)	(12%)
Contribution of	0.128	0.111	0.04	0.027
partner characteristics	(30%)	(34%)	(15%)	(11%)

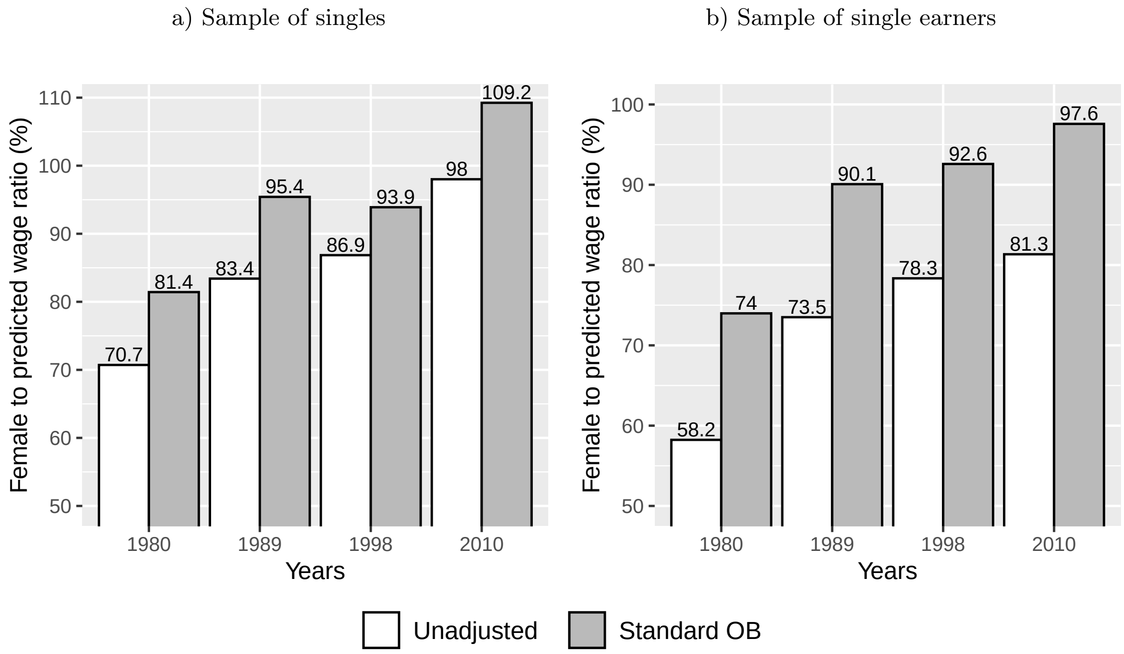
Notes: First line shows change in log differences, second line (in parentheses) gives percentage of total wage gap.

cupation).[14] The results of this decomposition are summarized in Table 3. Similar to the findings of Blau and Kahn (2017), the standard Oaxaca-Blinder decomposition assigns about equal shares of the wage gap to these two dimensions in early years and indicates an increasing importance of job attributes in more recent years. The extended Oaxaca-Blinder decomposition assigns larger shares to workers' own human capital, which is in line with the theoretical model given the strong degree of assortative mating in human capital. Regarding job attributes, the importance assigned by the extended Oaxaca-Blinder decomposition is smaller than the one assigned by the standard decomposition in early years and larger in more recent years. The extended Oaxaca-Blinder decomposition further reveals that both, partners' human capital and partners' job attributes contribute to the gender wage gap, with the importance of partners' job attributes fading in more recent years.

Finally, Table 4 summarizes the differences between the standard and the extended Oaxaca-Blinder decomposition. The extended Oaxaca-Blinder decomposition explains an additional one to two fifths of the gender wage gap. The bulk of this additional explanatory power stems from the direct effect of the included partner characteristics. The indirect effect due to changing coefficients on workers' own characteristics is positive, in line with the model, but rather small compared to the direct effect. Our results are in line with career prioritization, as both components of the change in the explained gap are positive. This indicates that a characteristic that is promoting one's own wage tends to be wage-reducing for one's partner, see (15) and (16).

[14] Given the high correlation of spouses' race and region, these variables—not surprisingly—explain very little of the gender wage gap. They contribute slightly to the gender wage gap due to differences in gender-specific full-time rates across regions and races which affects sample selection.

Figure 2: Standard Oaxaca-Blinder (OB) decomposition in a sample of singles (left) and single earners (right).



Notes: The sample of singles contains 307 (in 1980), 386 (in 1989), 400 (in 1998) and 423 (in 2010) men; and 554 (in 1980), 673 (in 1989), 747 (in 1998) and 869 (in 2010) women. In the single-earner sample, there are 1109 (in 1980), 969 (in 1989), 884 (in 1998) and 952 (in 2010) men and 652 (in 1980), 818 (in 1989), 872 (in 1998) and 1138 (in 2010) women.

4.3 Implications for Bachelor households

Our extended Oaxaca-Blinder decomposition is motivated by joint decision making in dual-earner households and, in our model, we emphasized that joint decision making induces career prioritization. Given that the model mechanism that leads to the bias in a standard Oaxaca-Blinder decomposition is absent for bachelor households or couple households with a single earner, our model implies that a standard Oaxaca-Blinder decomposition should explain larger shares of the gender wage gap in samples of bachelor workers or single earners in general.

To investigate this relation, Figure 2 shows results for singles (defined as individuals with no partner, neither married nor cohabiting, left panel) and single earners (defined as individuals who are the sole earner in their household independent of marital or cohabitation status, right panel). The left panel shows that the standard Oaxaca-Blinder decomposition explains very large shares of the gender wage gap among singles. From 1989 on, it explains more than 90% of the wage gap and in 2010 it explains the entire wage gap. This indicates that wage discrimination by gender is small for these groups. Importantly, the unexplained wage gap between male and female singles is substantially smaller than the one a standard Oaxaca-Blinder decomposition suggests in a sample of dual-earner couples or in a sample of all workers. The right panel reveals a similar pattern for single earners in general. Also here, the standard Oaxaca-Blinder decomposition explains large shares of the gender wage gap, ranging to close to 100%. These results support that a standard Oaxaca-Blinder decomposition underestimates the part of the gender wage gap attributable to observable differences between men and women due to its neglect of the role of partner characteristics for wage rates of workers in dual-earner couples.

5 Conclusion

We have shown that parts of the unexplained gender wage gap in standard Oaxaca-Blinder decompositions result from neglecting the role that partner characteristics play for wage rates in dual-earner couples. We have presented a simple model of location choice to make explicit that observed wage rates depend on the family situation and thereby on partner characteristics. This dependency is ignored in the standard Oaxaca-Blinder approach, so that both, estimated coefficients in the wage equation as well as the decomposition of the wage gap are biased. We have proposed an extended version of the Oaxaca-Blinder decomposition that addresses the bias by accounting for characteristics of the individual's partner.

In a sample of dual earners from the PSID, conventional Oaxaca-Blinder decompositions explain roughly half of the gender wage gap. Our extended Oaxaca-Blinder decompositions with partner characteristics explain considerable larger shares of the wage gap. Our findings suggest that the labor market does not discriminate by gender per se as strongly as conventional methods suggest. Instead, our findings highlight the role of family decisions which amplify pay differences between men and women.

References

Artuç, E. and McLaren, J. (2015). "Trade policy and wage inequality: a structural analysis with occupational and sectoral mobility". *Journal of International Economics* 97.2, pp. 278–294.

Bertrand, M. (2011). "New perspectives on gender". In: *Handbook of labor economics.* Ed. by D. E. Card and O. Ashenfelter. Vol. 4B. Elsevier, North-Holland, pp. 1543–1590.

Bertrand, M., Kamenica, E., and Pan, J. (2015). "Gender identity and relative income within households". *The Quarterly Journal of Economics* 130.2, pp. 571–614.

Blau, F. D. and Kahn, L. M. (2006). "The U.S. gender pay gap in the 1990s: slowing convergence". *ILR Review* 60.1, pp. 45–66.

Blau, F. D. and Kahn, L. M. (2017). "The gender wage gap: extent, trends, and explanations". *Journal of Economic Literature* 55.3, pp. 789–865.

Blau, F. D., Kahn, L. M., Comey, M., Eng, A., Meyerhofer, P., and Willén, A. (2020). *Culture and gender allocation of tasks: source country characteristics and the division of non-market work among US immigrants.* NBER Working Paper 26931.

Blinder, A. S. (1973). "Wage discrimination: reduced form and structural estimates". *The Journal of Human Resources* 8.4, pp. 436–455.

Bredemeier, C. (2019). *Gender gaps in pay and inter-firm mobility.* IZA Discussion Paper 12785.

Compton, J. and Pollak, R. A. (2007). "Why are power couples increasingly concentrated in large metropolitan areas?" *Journal of Labor Economics* 25.3, pp. 475–512.

Cortes, G. M. and Gallipoli, G. (2018). "The costs of occupational mobility: an aggregate analysis". *Journal of the European Economic Association* 16.2, pp. 275–315.

Cortés, P. and Pan, J. (2019). "When time binds: substitutes for household production, returns to working long hours, and the skilled gender wage gap". *Journal of Labor Economics* 37.2, pp. 351–398.

Cortés, P. and Tessada, J. (2011). "Low-skilled immigration and the labor supply of highly skilled women". *American Economic Journal: Applied Economics* 3.3, pp. 88–123.

Croson, R. and Gneezy, U. (2009). "Gender differences in preferences". *Journal of Economic Literature* 47.2, pp. 448–474.

Foged, M. (2016). "Family migration and relative earnings potentials". *Labour Economics* 42, pp. 87–100.

Goldin, C. (2014). "A grand gender convergence: its last chapter". *American Economic Review* 104.4, pp. 1091–1119.

Hersch, J. (2007). "Sex discrimination in the labor market". *Foundations and Trends in Microeconomics* 2.4, pp. 281–361.

Kambourov, G. and Manovskii, I. (2009). "Occupational specificity of human capital". *International Economic Review* 50.1, pp. 63–115.

Mincer, J. (1978). "Family migration decisions". *Journal of Political Economy* 86.5, pp. 749–773.

Oaxaca, R. (1973). "Male-female wage differentials in urban labor markets". *International Economic Review* 14.3, pp. 693–709.

Olivetti, C. and Petrongolo, B. (2016). "The evolution of gender gaps in industrialized countries". *Annual Review of Economics* 8.1, pp. 405–434.

Petrongolo, B. and Ronchi, M. (2020). "Gender gaps and the structure of local labor markets". *Labour Economics* 64, p. 101819.

Philliber, W. W. and Vannoy-Hiller, D. (1990). "The effect of husband's occupational attainment on wife's achievement". *Journal of Marriage and the Family* 52.2, p. 323.

Pixley, J. E. (2008). "Life course patterns of career-prioritizing decisions and occupational attainment in dual-earner couples". *Work and Occupations* 35.2, pp. 127–163.

Pixley, J. E. (2009). "Career Prioritizing in Dual-Earner Couples". In: *Women, feminism, and femininity in the 21st century.* Ed. by B. Mousli and E.-A. Roustang-Stoller. Palgrave Macmillan.

Pixley, J. E. and Moen, P. (2003). "Prioritizing careers". In: *It's about time: Couples and careers.* Ed. by P. Moen. Cornell University Press, pp. 183–200.

Sullivan, P. (2010). "Empirical evidence on occupation and industry specific human capital". *Labour Economics* 17.3, pp. 567–580.

Winkler, A. E. and Rose, D. C. (2000). "Career hierarchy in dual-earner families". *Research in Labor Economics* 19, pp. 147–172.

Appendix: Derivation of the wage equation

We define ϕ such that $\phi^2 = (1-\kappa)\sigma^2$ and $\Lambda_i = \psi_{-i} / \left(\psi_i + \psi_{-i}\right) \cdot (a_{-i} - a_i)$ with derivatives

$$\begin{aligned}
\frac{\partial \Lambda_i}{\partial \psi_i} &= -\frac{\psi_{-i}}{\left(\psi_i + \psi_{-i}\right)^2} \cdot (a_{-i} - a_i) \\
\frac{\partial \Lambda_i}{\partial \psi_{-i}} &= \frac{\psi_i}{\left(\psi_i + \psi_{-i}\right)^2} \cdot (a_{-i} - a_i) \\
\frac{\partial \Lambda_i}{\partial a_i} &= -\frac{\psi_{-i}}{\left(\psi_i + \psi_{-i}\right)} \\
\frac{\partial \Lambda_i}{\partial a_{-i}} &= \frac{\psi_{-i}}{\left(\psi_i + \psi_{-i}\right)}
\end{aligned}$$

In the point of approximation, these expressions evaluate as

$$\Lambda^2 = 1/4 \cdot 2\phi^2 = \frac{1}{2} \cdot \phi^2 \Rightarrow \Lambda = \frac{1}{\sqrt{2}} \cdot \phi,$$

as well as

$$\begin{aligned}
&\frac{\partial \Lambda_i}{\partial \psi_i} = -\frac{\psi}{4\psi^2} \cdot \sqrt{2}\phi = -\frac{\sqrt{2}\phi}{4\psi}, \frac{\partial \Lambda_i}{\partial \psi_{-i}} = \frac{\sqrt{2}\phi}{4\psi}, \\
&\frac{\partial \Lambda_i}{\partial a_i} = -\frac{1}{2}, \text{ and } \frac{\partial \Lambda_i}{\partial a_{-i}} = \frac{1}{2}.
\end{aligned}$$

Applying the approximation gives

$$\begin{aligned}
\log w_i =& \log \psi_i + \log\left(1 - \Lambda_i^2\right) \\
\approx& \log \psi + \log\left(1 - \Lambda^2\right) + \frac{1}{\psi}\left(\psi_i - \psi\right) \\
&- \frac{2\Lambda}{1-\Lambda^2} \cdot \left(\frac{\partial \Lambda_i}{\partial \psi_i}\left(\psi_i - \psi\right) + \frac{\partial \Lambda_i}{\partial \psi_{-i}}\left(\psi_{-i} - \psi\right) + \frac{\partial \Lambda_i}{\partial a_i}\left(a_i - a_1\right) + \frac{\partial \Lambda_i}{\partial a_{-i}}\left(a_{-i} - a_2\right)\right) \\
=& \log \psi + \log\left(1 - \frac{1}{2}\phi^2\right) + \frac{1}{\psi}\left(\psi_i - \psi\right) \\
&- \frac{\sqrt{2}\phi}{1 - \frac{1}{2}\phi^2}\left(\frac{\partial \Lambda_i}{\partial \psi_i}\psi\frac{\psi_i - \psi}{\psi} + \frac{\partial \Lambda_i}{\partial \psi_{-i}}\psi\frac{\psi_{-i} - \psi}{\psi} + \frac{\partial \Lambda_i}{\partial a_i}\left(a_i - a_1\right) + \frac{\partial \Lambda_i}{\partial a_{-i}}\left(a_{-i} - a_2\right)\right) \\
\approx& \log \psi - \log\left(1 - \frac{1}{2}\phi^2\right) + \frac{1}{\psi}\left(\psi_i - \psi\right) \\
&- \frac{\sqrt{2}\phi}{1 - \frac{1}{2}\phi^2}\left(-\frac{\sqrt{2}\phi}{4}\log\left(\psi_i/\psi\right) + \frac{\sqrt{2}\phi}{4}\log\left(\psi_{-i}/\psi\right) - \frac{1}{2}\left(a_i - a_1\right) + \frac{1}{2}\left(a_{-i} - a_2\right)\right) \\
=& \log \psi - \log\left(1 - \frac{1}{2}\phi^2\right) + \frac{1}{\psi}\left(\psi_i - \psi\right) \\
&+ \frac{\phi^2}{2-\phi^2}\log\left(\psi_i/\psi\right) - \frac{\phi^2}{2-\phi^2}\log\left(\psi_{-i}/\psi\right) - \frac{\sqrt{2}\phi}{2-\phi^2}\left(a_i - a_1\right) + \frac{\sqrt{2}\phi}{2-\phi^2}\left(a_{-i} - a_2\right).
\end{aligned}$$

The expression in the last line can be rearranged to condition (6) in the main text.