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*

First version: August 2020
This version: October 2021

Abstract

Career prioritization in dual-earner couples amplifies pay differences between spouses. Yet, standard decomposition approaches for the gender wage gap do not take this effect into account. We show that, as a consequence, these decompositions assign to small a share of the wage gap to observable differences between men and women. We develop an extended decomposition approach that corrects this shortcoming. In U.S. survey data, we find that our extended decomposition explains considerably more of the wage gap than a standard approach – in line with our theory.

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

Keywords: Gender wage gap; wage-gap decomposition; dual-earner households; discrimination

*We thank conference participants at the 2020 World Congress of the Econometric Society, the 2021 Conference of the International Association for Applied Econometrics, and the 2021 Labor Economics Meeting of the Spanish Economic Association for helpful comments and suggestions. The data used in this article can be obtained from the website of the Panel Study of Income Dynamics (https://simba.isr.umich.edu/default.aspx). The authors have nothing to disclose.
†Correspondence to bredemeier@uni-wuppertal.de.

1 Introduction

The gender wage gap decreased substantially in the second half of the 20th century, but a persistent gap remains, see, e.g., Olivetti and Petrongolo (2016). As shown by, e.g., Blau and Kahn (2017), a considerable part of the wage gap can be related to observable differences between men and women in individual characteristics such as work experience or occupation. However, a substantial part of the wage gap appears to remain unexplained by these observable characteristics. The unexplained part of the gender wage gap has attracted considerable attention in both the academic literature as well as the public discourse.

In empirical decompositions of the gender wage gap, the unexplained part may be due to several reasons. Women in general could be discriminated against, or the labor market could reward unobservable characteristics such as noncognitive skills or certain personality traits in which men and women differ from one another. Another possibility is that the decomposition approach misestimates the importance of observable characteristics because the relationship between wages and characteristics is modeled too restrictively. While the literature has emphasized functional form, we emphasize that the influence of observable characteristics is also misestimated because the decomposition approaches applied in the literature do not capture all mechanisms through which observable characteristics affect wages in reality.[1]

Specifically, we build on the family economics literature and emphasize that existing decomposition approaches overlook the fact that individual characteristics also influence *one's partner's wage*. As a result, these approaches misestimate the impact of observable differences between men and women because they abstract from a channel through which these differences affect the wages of women and men. The literature has long recognized that dual-earner couples regularly have to take decisions whose career to promote, for example, in the context of family migration (e.g., Mincer 1978, Compton and Pollak 2007, Foged 2016, Braun et al. 2021), the choice of employers at a given location (Bredemeier 2019, Petrongolo and Ronchi 2020), or career investment in

[1]Frölich (2007), Mora (2008), and Ñopo (2008) use non-parametric decompositions and Neuman and Oaxaca (2004), Blau and Kahn (2006), Olivetti and Petrongolo (2008), and Oaxaca and Choe (2016) address selection in the Oaxaca-Blinder approach. Our empirical analysis will address these issues, as well as biases arising from omitted wage determinants.

the form of working long hours (Cortés and Tessada 2011, Cortés and Pan 2019).[2] 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. Importantly, this career prioritization implies that observed wages depend both on individual characteristics (such as education, experience, and occupation) but also on the family situation and, hence, the characteristics of the partner. For given individual characteristics, the earnings potential of the 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. Yet, partner characteristics are not accounted for in existing decomposition approaches.

To illustrate the resulting problem, we focus on the Oaxaca-Blinder decomposition approach, which remains the most frequently applied empirical decomposition approach of the gender wage gap. The wage equation in the Oaxaca-Blinder decomposition accounts only for individual characteristics, and the family situation, as captured by the characteristics of the partner, is ignored, inducing a bias in the decomposition.

This bias is due to two effects. First, the parameter estimates on the worker's own characteristics are biased in the standard Oaxaca-Blinder wage equation. Second, in the decomposition of the wage gap, the contribution of the characteristics of the worker's partner is ignored. To correct these shortcomings, the characteristics of the partner should be included in the wage equation . For example, a worker's education should be included on the right-hand side of the worker's own wage equation but also on the right-hand side of the wage equation of the worker's partner to account for the effect of the worker's education on the family's investment into the partner's career. In the decomposition, one would then capture the extent to which women's relative wages are compressed by their husbands' characteristics through career-prioritizing decisions of the family.

[2] In other contexts, taking into account that spouses take labor-market decisions jointly has helped understand better phenomena such as consumption insurance against wage-rate shocks (e.g., Blundell et al. 2016, Autor et al. 2019), female labor supply (e.g., Bick and Fuchs-Schündeln 2017, 2018), time use (e.g., Rogerson and Wallenius 2019, Browning et al. 2020) or the effects of policies such as unemployment insurance (e.g., Ortigueira and Siassi 2013, Choi and Valladares-Esteban 2020) or pension systems (e.g., Nishiyama 2019, Groneck and Wallenius 2020).

The implication of career prioritization (or similar family-economics mechanisms) to include partner characteristics is not limited to the Oaxaca-Blinder approach but applies to all approaches that seek to assign a part of the wage gap to differences in observable characteristics. For example, career prioritization in the family implies that matching-based approaches (e.g., Strittmatter and Wunsch 2021, Meara et al. 2020) should include partner characteristics in the matching process, independent of the specifics of this process. Effectively, one should compare men and women with similar own characteristics *and similar partners* to account appropriately for the role of the family in the determination of wages.

To make our point explicit, we set up a model of dual-earner couples for whom career prioritization stems from a joint location choice because couples having 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 – through the mediator distance to optimal location – negatively on the earnings potential of the individual's partner.

In the model, we allow for observable as well as unobservable determinants of earnings potentials and for gender differences in earnings potentials conditional on characteristics which can be interpreted as discrimination. For illustration, we first consider a model version where an individual's earnings potential depends only on observable characteristics but not on gender or unobservable characteristics. We use this version to show that the standard decomposition yields a share of the wage gap which is supposedly "unexplained" even if the gap is entirely due to differences in observable characteristics. In more general settings, where the model allows for gender-biased wage setting and unobservable determinants of earnings potentials, 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 decomposition by the characteristics of the partner resolves this problem, yielding larger parts of the wage gap assigned to observable characteristics.

Importantly, including 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. To clarify, our point goes beyond simply arguing that *additional* characteristics should be included in the Oaxaca-Blinder approach, but instead relates to the way *how* characteristics that have been isolated as important by the literature should enter the decomposition. Suppose, for example, the wage gap were entirely due to differences in years of schooling. Then, a standard Oaxaca-Blinder decomposition with years of schooling would still label some part of the gap as "unexplained" because differences in years of schooling affect the wage gap twice – through the direct effect of education on earnings potentials and through career prioritization in favor of the better educated partner. The standard approach captures only one of these channels (and in a biased way).

We apply our improved decomposition to U.S. data from the Panel Study of Income Dynamics (PSID). In line with the literature, standard Oaxaca-Blinder decompositions explain roughly half of the gap and hence suggest that a substantial part of the wage gap is unrelated to the included characteristics such as human-capital variables and job information. Our extended decompositions systematically explain larger shares of the wage gap as a consequence of gender differences in observable characteristics.For some years, the extended decomposition explains up to 100% of the wage gap.

We corroborate 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 do not play a role. In fact, we find that, for these groups, standard Oaxaca-Blinder decompositions attribute substantially larger shares of the gender wage gap to observable characteristics than it does for men and women living in dual-earner couples.

Our results imply that pay differences between men and women are more strongly related to differences in observable characteristics than suggested by standard decompositions, stressing the role of family decisions for the observed pay gap. Career prioritization within couples amplifies pay differences between men and women and, 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 rule out that the gender wage gap is a result of discrimination against women. Our empirical results indicate that, in most years, the labor market does not yield the same wages form men and women even conditional on their, and their partners', observable characteristics.[3] 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.[4] Moreover, career prioritization amplifies both non-discriminatory and discriminatory differences in earnings potentials. A family observing that women are discriminated against faces incentives to prioritize the husband's career over the wife's even if the two are identical in terms of objective characteristics. Career prioritization can also perpetuate pay differences between men and women because couples take important career-prioritizing decisions early in life, with lifelong consequences for spouses' careers. This is a potential factor behind the persistence of the gender wage gap despite considerable progress in gender equity. Policy might exploit the amplification mechanism of career prioritization as policy measures that improve women's earnings potentials can result in families investing more strongly in women's careers, thereby reinforcing the direct effects on the wage gap.

The remainder of this paper is organized as follows. In Section 2, we present the model and use it to study alternative decomposition approaches. In Section 3, we present our empirical analysis. Section 4 concludes.

[3]This can be due to discrimination or gender differences in non-cognitive abilities or personality traits, which are generally not observed in large cross-sectional studies. There is a significant empirical literature, mostly experimental, on differences between men and women with respect to non-cognitive abilities, personality traits, and preferences, including the willingness to compete (2009, Flory et al. 2015, Buser and Yuan 2019), negotiation styles (Babcock and Laschever 2003, Exley et al. 2020), promotion-seeking (Bosquet et al. 2019) the willingness to take on non-promotable tasks (Babcock et al. 2017), risk aversion (Croson and Gneezy 2009, Dohmen and Falk 2011), self-promotion (Exley and Kessler 2021) and preferences for non-wage job characteristics (Mas and Pallais 2017, Wiswall and Zafar 2018, Le Barbanchon et al. 2020). Several studies (Mueller and Plug 2006, Nyhus and Pons 2012, Le et al. 2011, Reuben et al. 2015, Heinz et al. 2016, Chen et al. 2017, Flinn et al. 2018, Jung et al. 2018, Roussille 2021) have documented that a part of the wage gap can be attributed to such factors, but their quantitative role seems to be limited. Further, there is direct evidence from field experiments for a role of discrimination in hiring and wage-setting processes (e.g., Neumark et al. 1996, Goldin and Rouse 2000, Beaurain and Masclet 2016, Sarsons et al. 2021). While the unexplained wage gap may reflect discrimination or differences in unobservable characteristics, it is unlikely to reveal their influence to the full extent. The effects of some personality traits or preferences may be captured in the explained fraction as long as they run through observable mediators such as occupation choice. Moreover, the unexplained wage gap cannot capture the full extent of possible discrimination in the labor market if some of the explanatory variables are affected by discrimination.

[4]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).

2 Theory

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 briefly discuss related approaches from the literature that have similar implications for wage rates. Thereafter, we analyze the role of career prioritization for empirical decompositions of the gender wage gap, propose an extended decomposition that takes into account career prioritization appropriately and compare it to the standard Oaxaca-Blinder approach.

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 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 (see below), 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

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

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:

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

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.

This 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, 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 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 individual characteristics Z_i,

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

where $g\left(i\right)$ denotes individual i's gender and can take the values m (for male) and f (for female). Z_i is a column vector of individual 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 characteristics, 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 $2(1-\kappa)\sigma^2$.[5] This point of approximation ensures that both, the earnings potential ψ, which reflects individual

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

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.

Applying the approximation gives

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

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 A 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.[6] 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.

Alternative mechanisms. 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

[6]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 $a \in (0,1)$.

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 promotes the designated primary earner's career. This tends to have negative consequences for wage rates paid to the secondary earner.

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 strongly 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 and Albrecht et al. (2018) document that men experience higher wage gains upon switching employers than women, whose firm-to-firm transitions appear motivated by job attributes other than pay. 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 potential 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. This suggests that women's careers came second to their husbands' ones 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.

2.3 Wage-gap decomposition in the model

In the model, gender differences in pay can stem from differences in the characteristics Z and from differences in how earnings potentials depend on characteristics as captured by the coefficients γ and, consequently, β. In order to separate these two sources, the (average) gender wage gap $\Delta = \overline{w}_m - \overline{w}_f$, where $\overline{w}_g$ denotes average log wages by gender, can be decomposed as

$$\Delta = \underbrace{(\beta_{1,m} - \beta_{2,m}) \cdot (\overline{Z}_m - \overline{Z}_f)}_{\Delta|_Z} + \underbrace{(\beta_{1,m} - \beta_{1,f}) \cdot \overline{Z}_f + (\beta_{2,m} - \beta_{2,f}) \cdot \overline{Z}_m}_{\Delta|_\beta}, \qquad (7)$$

where $\overline{Z}_g$ denotes gender-specific average characteristics. The first term on the right-hand side, $\Delta|_Z$, is the wage gap that is due to gender differences in characteristics Z. 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.

2.4 Career prioritization and empirical wage-gap decompositions

We now analyze *empirical* decomposition approaches to the gender wage gap in our model of career prioritization. We will show that the standard Oaxaca-Blinder approach, in which a worker's wage is related only to the worker's own characteristics, assigns too small a share of the wage gap to observable characteristics. We will then show that an extended decomposition where the characteristics of the worker's partner are included in the wage equation solves this problem.

Standard Oaxaca-Blinder decomposition. The first step of the standard Oaxaca-Blinder decomposition is to estimate a log wage equation, separately for men

$(g = m)$ and women $(g = f)$:

$$w_i = b_{0,g(i)} + b_{1,g(i)} \cdot X_i + e_i, \tag{8}$$

where $b_{0,g(i)}$ is a constant, $b_{1,g(i)}$ is a vector of coefficients, X_i is a vector of observable characteristics, and e_i is a residual. In the following, we will first consider the case where the set X of characteristics included in the decomposition is the same as the set of pay-relevant characteristics Z. Thereafter, we consider the case where not all pay-relevant characteristics are included in the decomposition.

The empirical decomposition yields an "explained" part of the gap,

$$\left(\widehat{\Delta}|_X\right)^{std} = \widehat{b}^{std}_{1,m}(\overline{X}_m - \overline{X}_f),$$

where $\widehat{b}^{std}$ indicate estimates, that is assigned to differences in observable characteristics and an "unexplained" part

$$\left(\widehat{\Delta}|_b\right)^{std} = \widehat{b}^{std}_{0,m} - \widehat{b}^{std}_{0,f} + (\widehat{b}^{std}_{1,m} - \widehat{b}^{std}_{1,f})\overline{X}_f$$

that this approach identifies as unrelated to observable characteristics.

Extended decomposition. Our extended decomposition accounts for the role of the family for individual wage rates in dual-earner households. Specifically, we account for the characteristics of the individual's partner and estimate

$$w_i = b_{0,g(i)} + b_{1,g(i)} \cdot X_i + b_{2,g(i)} \cdot X_{-i} + e_i, \tag{9}$$

which yields an explained gap of

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

and an unexplained gap of

$$\left(\widehat{\Delta}|_b\right)^{ext} = \widehat{b}^{ext}_{0,m} - \widehat{b}^{ext}_{0,f} + (\widehat{b}^{ext}_{1,m} - \widehat{b}^{ext}_{1,f})\overline{X}_f + (\widehat{b}^{ext}_{2,m} - \widehat{b}^{ext}_{2,f})\overline{X}_m.$$

If the set of observable characteristics X in the decomposition includes all characteristics Z relevant for earnings potentials ψ, the extended decomposition identifies correctly the shares of the gender wage gap which are caused by differences in these characteristics and by differences in coefficients ($\Delta|_Z$ and $\Delta|_\beta$), respectively. This is not surprising since the wage equation in the extended decomposition (9) is then identical to the data-generating wage equation (6). By contrast, the standard decomposition misestimates the importance of differences in characteristics *even if the wage equation accounts for all variables Z which are relevant for earnings potentials and wages* because it fails to account for the career-prioritization channel through which these variables impact on gender-specific wages. We will now demonstrate this point.

2.5 Comparing the decompositions

We start with the case where the set of variables accounted for in the empirical decomposition, X, includes all pay-relevant characteristics in the set Z. Later, we will turn to the case where not all pay-relevant characteristics can be accounted for in the empirical decomposition. For simplicity, we restrict the set of characteristics in Z to a single observable characteristic, x. We consider the case where both decomposition approaches account for this characteristic, albeit in different ways. For simplicity, we assume that the characteristic is measured in a way that it increases earnings potentials, $\gamma_{x,g} > 0$ (a classic example is education) and that some part of the gender wage gap can in fact be attributed to this characteristic, i.e., $\overline{x}_m > \overline{x}_f$.

Model version with gender-neutral wage setting. First, we simplify the model further and consider a 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 (7) in this example simplifies to

$$\Delta = \overline{w}_m - \overline{w}_f = (\beta_1 - \beta_2) \cdot (\overline{x}_m - \overline{x}_f) = \Delta|_x. \tag{10}$$

Hence, in this example, the entire wage gap is due to gender differences in the observable characteristic x and there is no unexplained wage gap *by construction.* We now apply the standard Oaxaca-Blinder decomposition to this example which yields the explained part of the gap

$$\left(\widehat{\Delta}|_x\right)^{std} = \left(\beta_1 + \beta_2 \cdot \frac{\text{cov}(x_m, x_f)}{\text{var}(x_m)}\right) \cdot (\overline{x}_m - \overline{x}_f).$$

Importantly, the explained gap derived from the standard Oaxaca-Blinder decomposition is in general not equal to the true explained gap, $\Delta|_x = (\beta_1 - \beta_2) \cdot (\overline{x}_m - \overline{x}_f)$. As long as $\text{cov}(x_i, x_{-i})/\,\text{var}(x_i) > -1$, i.e., as long as there is no pronounced *negative* assortative mating, we have

$$\left(\widehat{\Delta}|_x\right)^{std} < \Delta = \Delta|_x,$$

and

$$\left(\widehat{\Delta}|_\beta\right)^{std} > 0 = \Delta|_\beta.$$

Thus, although the entire wage gap in this model version is caused by gender differences in the observable characteristics ($\Delta = \Delta|_x$ and thus $\Delta|_\beta = 0$), a standard Oaxaca-Blinder decomposition labels some part of the gap “unexplained”.

Model version with gender-biased wage setting. We now consider the case where the model accounts for gender-biased wage setting, i.e., $\gamma_m \neq \gamma_f$. Differences in the true coefficients are frequently interpreted as discrimination. In our model, this can arise from discrimination in the labor market or in the family, i.e., from earnings potentials differing by gender conditional on characteristics or from the family tending to prioritize one member’s career based on gender rather than characteristics.

The standard Oaxaca-Blinder wage regression yields an explained gender wage gap of

$$\widehat{\Delta}|_x = \widehat{b}^{std}_{1,m} \cdot (\overline{x}_m - \overline{x}_f) = \left(\beta_{1,m} + \beta_{2,m} \cdot \frac{\text{cov}\,(x_m, x_f)}{\text{var}\,(x_m)}\right) \cdot (\overline{x}_m - \overline{x}_f)\,.$$

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 (7). Our model implies $\beta_{2,m} < 0$ due to career prioritization and hence the estimated explained gap is smaller than the true explained gap,

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

as long as $\text{cov}\,(x_m, x_f)\,/\,\text{var}\,(x_m) > -1$. Hence, a standard Oaxaca-Blinder decomposition assigns too small a share of the wage gap to differences in the characteristics and in turn identifies too large a share of the wage gap as unexplained.

By contrast, applying the extended decomposition gives the estimated coefficients $\widehat{b}^{ext}_{1,m} = \beta_{1,m}$, $\widehat{b}^{ext}_{2,m} = \beta_{2,m}$, $\widehat{b}^{ext}_{1,f} = \beta_{1,f}$, and $\widehat{b}^{ext}_{2,f} = \beta_{2,f}$ as well as the estimated explained gap as

$$\widehat{\Delta}|_x = \widehat{b}^{ext}_{1,m}(\overline{x}_m - \overline{x}_f) + \widehat{b}^{ext}_{2,m}(\overline{x}_f - \overline{x}_m) = (\beta_{1,m} - \beta_{2,m})\,(\overline{x}_m - \overline{x}_f)$$

and corresponds to the true explained gap $\Delta|_x$, see (7). 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 $\Delta|_\beta$.

Decomposing differences between the decompositions. The difference in the explained gap between our extended and the standard decomposition can be decomposed into two systematic mechanisms: 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,

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

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 $\operatorname{cov}(x_m, x_f) > 0$ (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,

$$\beta_{2,m} < 0 \Leftrightarrow (\Delta|_x)^{ext} - (\Delta|_x)^{std} > 0. \tag{11}$$

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

$$\beta_{2,m} < 0 \Leftrightarrow \left(\widehat{b}^{ext}_{1,m} - \widehat{b}^{std}_{1,m}\right) \cdot (\overline{x}_m - \overline{x}_f) = -\beta_{2,m} \cdot \frac{\operatorname{cov}(x_m, x_f)}{\operatorname{var}(x_m)} \cdot (\overline{x}_m - \overline{x}_f) > 0, \tag{12}$$

where the last step uses the estimate from the standard Oaxaca-Blinder approach and

$$\beta_{2,m} < 0 \Leftrightarrow \widehat{b}^{ext}_{2,m} \cdot (\overline{x}_f - \overline{x}_m) = -\beta_{2,m} \cdot (\overline{x}_m - \overline{x}_f) > 0. \tag{13}$$

This implies that 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 (12) and (13) of the difference between the explained gaps are predicted to be positive. Further, (12) and (13) show that including 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., $\beta_{2,m} < 0$).

Allowing for unobservable wage determinants. In Appendix B, we extend the analysis to the case where not all wage determinants are observable to the econometrician, as frequently the case in observational data (a classic example is ability). We show that, also in a setting where some wage determinants are unobservable, and under mild assumptions regarding the covariances between characteristics, our extended decomposition assigns a larger part of the wage gap to observable characteristics than does the standard decomposition.

3 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.[7]

3.1 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.[8] Blau and Kahn

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

[8]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).

(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.[9]

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.[10] 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.[11] For our extended decomposition, we augment the wage equation by the partner's edu-

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

[10] 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.

[11] Note that the geographical controls play a different role than the mediator "location" which we stress in our model. Rather than *whether* one lives in a metropolitan area, the mechanism we highlight emphasizes *which* metropolitan area one lives in or where *inside* a metropolitan area a couple locates.

cation, experience, and job information.[12]

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 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), Artuç and McLaren (2015), and 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 largely 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). In our baseline set-up, we include experience in the set of control variables which preserves direct comparability to Blau and Kahn (2017) and facilitates the interpretation of the unexplained gap.[13] Finally, union coverage is mostly determined by the choice of employer and hence a joint decision of the couple

[12] 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.

[13] To shed light on mechanisms behind the wage gap, we have also considered a specification without experience. In this specification, the explained wage gap is reduced but by less than the wage differences that the model including experience assigns to this factor. This indicates that experience is both a mediator of some other included wage determinants or either a wage determinant in itself or a mediator of unobservable factors such as discrimination. Quantitatively, the role of experience as a mediator of other observable determinants seems to be limited, amounting to less than one fifth of the wage differences assigned to differences in experience. This supports our handling of experience as a characteristic in the baseline specification.

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.[14] 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, 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 2.5.[15]

3.2 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

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

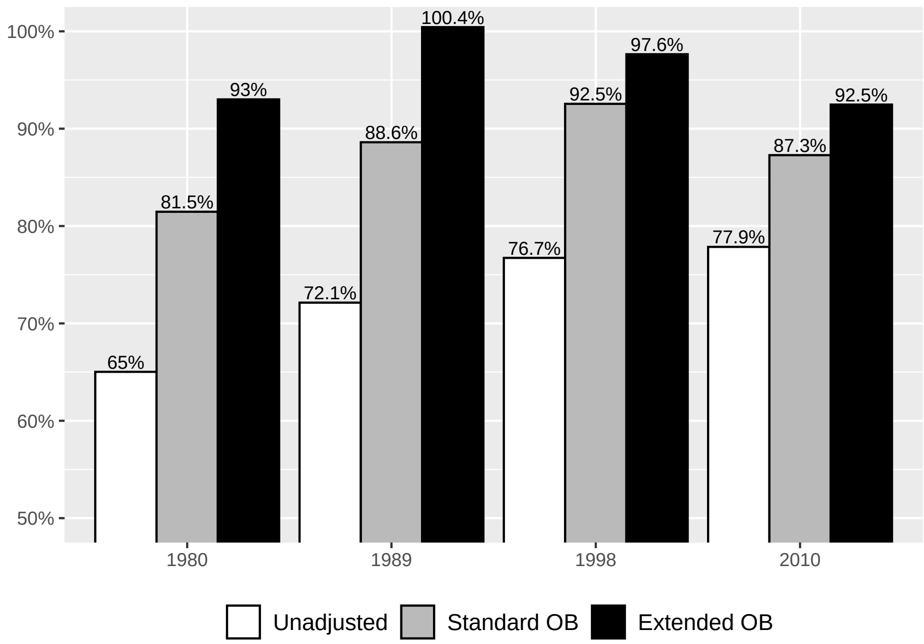
[15] The correlation in full-time experience is mostly driven by the high correlation in spouse's age. The conditional correlation is relatively small.

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.

Figure 1: Comparison of standard Oaxaca-Blinder (OB) decomposition and extended decomposition using partner characteristics, dual-earner sample: Log female to male wage ratio, unadjusted and adjusted for covariates.



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

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 analytical example presented in Section 2.4. 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 (11). 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 (11), and hence corroborates the presence of career prioritization. For 1989, 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 also in line with our theoretical model. While unobservable factors such as discrimination or differences in noncognitive skills do contribute to the wage gap to some extent, a standard Oaxaca-Blinder decomposition understates substantially the extent to which the wage gap is related to observable characteristics.

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. One possible interpretation is that the closure of the wage gap between 1980 and 2010 may 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.[16] 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%.

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

[16]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.

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.

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.

Table 4: Comparison of standard and extended decomposition.

	1980	1989	1998	2010
Increase in explained gap	0.131	0.125	0.054	0.058
$(\Delta\|_X)^{ext} - (\Delta\|_X)^{std}$	(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.

When the labor market discriminates against women, also couples where husband and wife have identical characteristics have 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 occupation).[17] 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

[17] 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.

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 (12) and (13).

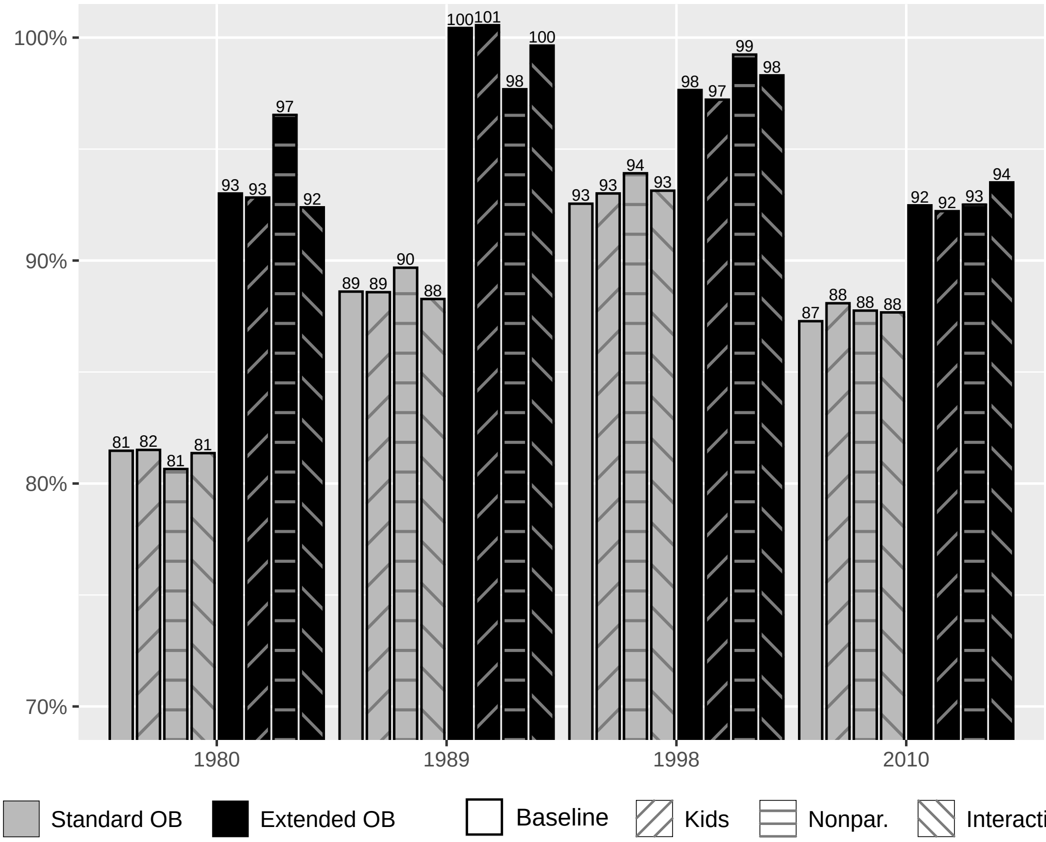
3.3 Sensitivity

We have performed a number of sensitivity checks to corroborate the robustness of our main results. Figure 2 summarizes the adjusted wage ratios ($1/\exp(\widehat{\Delta}|_b)$) obtained in various sensitivity analyses. As in Figure 1, gray bars refer to the standard decomposition and black bars refer to the extended decompositions. The different specifications are represented by a different hatching of the bars. The results of the baseline specification are repeated, without hatching, for convenience.

Children. One variable that so far has been neglected in our analysis but that is relevant in the gender-wage nexus is the number of children in the household. Adda et al. (2017), Lundborg et al. (2017), and Kleven et al. (2019) have documented a substantial decline in mothers' relative wages after child birth which Laun and Wallenius (2021) have linked to foregone human-capital accumulation. Using the Blau and Kahn (2017) sample, Cortés and Pan (2020) have shown that a substantial part of the wage gap that remains unexplained by standard Oaxaca-Blinder decompositions is related to the presence of children. In our first sensitivity check, we have therefore added the number of children below age 18 living in the household as a regressor in the wage equation. While an Oaxaca-Blinder decomposition cannot assign a fraction of the wage gap to this factor since, in dual-earner households, the number of children is the same for both parents, including the number of children in the wage equation might change the coefficients estimated for the other variables.[18] The upward hatched bars in Figure 2

[18]To be precise, there are differences in the average number of children between men and working women and a decomposition assigns some part of the wage gap to these differences. However, this rather reflects selection of some mothers into non-employment and not the contribution of fertility differences between husbands and

Figure 2: Results of sensitivity checks: Log female to male wage ratio, unadjusted and adjusted for covariates.



Notes: White bars show $1/\exp(\Delta)$. Gray bars show $1/\exp((\widehat{\Delta}|_b)^{std})$. Black bars show $1/\exp((\widehat{\Delta}|_b)^{ext})$. No hatching: baseline specification. Upward hatching: number of children (age<18) as additional covariate. Horizontal hatching: years of schooling, full-time experience, and part-time experience (rounded to full years) as categorial variables. Downward hatching: interaction between years of schooling and years of full-time and part-time experience additionally included.

show that this effect is negligible. In all years, they are a similar to the non-hatched bars representing the baseline specification.

Linearity. A potential shortcoming of the Oaxaca-Blinder approach is its linearity assumption and non-parametric wage equations have been estimated as alternatives (DiNardo et al. 1996, Frölich 2007, Mora 2008, and Ñopo 2008). Our baseline specification of the wage equation follows Blau and Kahn (2017) and is mostly non-parametric as all variables except years of schooling and the experience variables are categorical. To check sensitivity, we have also treated these variables as categorical (experience rounded to full years). The horizontally hatched bars in Figure 2 show that this affects

wives to the gender wage gap.

our results only mildly. Relatedly, the Oaxaca-Blinder approach usually does not account for interactions between wage determinants. For this reason, it might overlook for example the age-specific wage premium to education (Bhuller et al. 2017). The results illustrated by the downward hatched bars in Figure 2 refer to a specification where we included interaction terms of years of schooling with years of full and part time experience. Also here, we find only moderate changes relative to our baseline specification.

In summary, the sensitivity checks corroborate that our extended decomposition explains considerably larger shares of the wage gap than the conventional decomposition (the black bars are all higher than the gray bars).

Selection. Selection of women into employment can induce two biases in the decomposition. First, the true gap in offer wages might be larger than the gap in observed realized wages when, systematically, women with low wage offers opt out of the labor force. Second, the sample of employed women may have different characteristics than a full sample of all women. To account for these potential biases, Oaxaca-Blinder decompositions have been extended by corrections for selection (e.g., Neuman and Oaxaca 2004, Machado 2017, Maasoumi and Wang 2019) while other papers have used information from previous or subsequent employment spells of the same individual (Blau and Kahn 2006, Olivetti and Petrongolo 2008). We take a pragmatic approach and exploit that the coefficients of the male wage equation are not subject to selection of women into the labor force and that the average characteristics of all women (independent of labor-force participation) can be calculated from observables. Hence, we can quantify the wage differences (in log points) that can be related to observable differences between men and all women in couple households – though not a gap in offer wages to which we could relate it (in percent). For this exercise, we extend our baseline sample by those couple households where only the male is working. For non-working women, we use the job information on occupation and industry regarding their last or subsequent employment spell. From this sample, we estimate the male wage equation and multiply the resulting coefficients with the average gender differences in characteristics. The results are shown in Table 5 which also repeats the results from our baseline specifications for convenience. Also in the sample including non-participating women, our

Table 5: Wage differences assigned to observable factors, dual-earner households and households with working husbands.

	1980	1989	1998	2010
Dual-earner households				
Standard decomposition	0.225	0.206	0.188	0.114
Extended decomposition	0.357	0.331	0.241	0.172
Couple household with working male				
Standard decomposition	0.223	0.220	0.192	0.143
Extended decomposition	0.370	0.342	0.289	0.193

Notes: Upper block: baseline sample. Lower block: baseline sample plus households with working male but non-working female. For non-working females, job information refers to the last job.

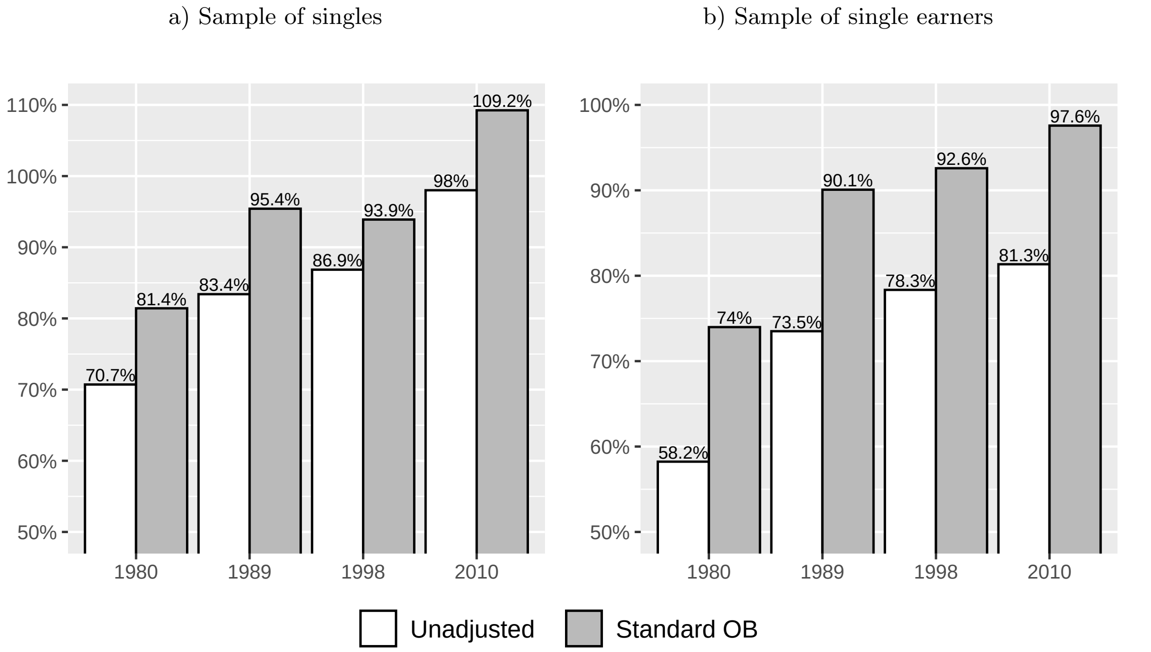
extended approach assigns considerably larger differences in pay to observable factors than does the conventional approach that omits partner characteristics.

3.4 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 3 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. Importantly, the unexplained wage gap between male and female singles is substantially smaller than the one a standard Oaxaca-Blinder decomposition

Figure 3: Standard Oaxaca-Blinder (OB) decomposition in a sample of singles (left) and single earners (right): Log female to male wage ratio, unadjusted and adjusted for covariates.



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

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.

4 Conclusion

We have shown that parts of the unexplained gender wage gap in standard 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 developed extended decompositions 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 explain considerable larger shares of the wage gap. Our findings highlight the role of family decisions which amplify pay differences between men and women.

While our analysis has considered marital status and the distribution of spouses' characteristics in marriages as exogenous, studying the interaction between career prioritization and the marriage market as well as its implications for the gender wage gap is an interesting avenue for future research.

References

Adda, J., Dustmann, C., and Stevens, K. (2017). "The career costs of children". *Journal of Political Economy* 125 (2), 293–337.

Albrecht, J., Bronson, M. A., Thoursie, P. S., and Vroman, S. (2018). "The career dynamics of high-skilled women and men: Evidence from Sweden". *European Economic Review* 105, 83–102.

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), 278–294.

Autor, D., Kostøol, A., Mogstad, M., and Setzler, B. (2019). "Disability benefits, consumption insurance, and household labor supply". *American Economic Review* 109 (7), 2613–54.

Babcock, L. and Laschever, S. (2003). *Women Don't Ask: Negotiation and the Gender Divide.* Princeton University Press.

Babcock, L., Recalde, M. P., Vesterlund, L., and Weingart, L. (2017). "Gender differences in accepting and receiving requests for tasks with low promotability". *American Economic Review* 107 (3), 714–747.

Beaurain, G. and Masclet, D. (2016). "Does affirmative action reduce gender discrimination and enhance efficiency? New experimental evidence". *European Economic Review* 90, 350–362.

Bertrand, M., Kamenica, E., and Pan, J. (2015). "Gender identity and relative income within households". *Quarterly Journal of Economics* 130 (2), 571–614.

Bhuller, M., Mogstad, M., and Salvanes, K. G. (2017). "Life-cycle earnings, education premiums, and internal rates of return". *Journal of Labor Economics* 35 (4), 993–1030.

Bick, A. and Fuchs-Schündeln, N. (2017). "Quantifying the disincentive effects of joint taxation on married women's labor supply". *American Economic Review* 107 (5), 100–104.

Bick, A. and Fuchs-Schündeln, N. (2018). "Taxation and labour supply of married couples across countries: A macroeconomic analysis". *Review of Economic Studies* 85 (3), 1543–1576.

Blau, F. D. and Kahn, L. M. (2006). "The U.S. gender pay gap in the 1990s: Slowing convergence". *ILR Review* 60 (1), 45–66.

Blau, F. D. and Kahn, L. M. (2017). "The gender wage gap: Extent, trends, and explanations". *Journal of Economic Literature* 55 (3), 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". *Review of Economics of the Household* 18 (4), 907–958.

Blundell, R., Pistaferri, L., and Saporta-Eksten, I. (2016). "Consumption inequality and family labor supply". *American Economic Review* 106 (2), 387–435.

Bosquet, C., Combes, P.-P., and García-Peñalosa, C. (2019). "Gender and promotions: evidence from academic economists in France". *Scandinavian Journal of Economics* 121 (3), 1020–1053.

Braun, C., Nusbaum, C., and Rupert, P. (2021). "Labor market dynamics and the migration behavior of married couples". *Review of Economic Dynamics* 42, 239–263.

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

Bredemeier, C. and Juessen, F. (2013). "Assortative mating and female labor supply". *Journal of Labor Economics* 31 (3), 603–631.

Browning, M., Donni, O., and Gørtz, M. (2020). "Do you have time to take a walk together? Private and joint time within the household". *Economic Journal* 131 (635), 1051–1080.

Buser, T. and Yuan, H. (2019). "Do women give up competing more easily? Evidence from the lab and the Dutch math olympiad". *American Economic Journal: Applied Economics* 11 (3), 225–252.

Chen, W., Grove, W. A., and Hussey, A. (2017). "The role of confidence and noncognitive skills for post-baccalaureate academic and labor market outcomes". *Journal of Economic Behavior & Organization* 138, 10–29.

Choi, S. and Valladares-Esteban, A. (2020). "On households and unemployment insurance". *Quantitative Economics* 11 (1), 437–469.

Compton, J. and Pollak, R. A. (2007). "Why are power couples increasingly concentrated in large metropolitan areas?" *Journal of Labor Economics* 25 (3), 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), 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), 351–398.

Cortés, P. and Pan, J. (2020). *Children and the Remaining Gender Gaps in the Labor Market.* NBER Working Paper 27980.

Cortés, P. and Tessada, J. (2011). "Low-skilled immigration and the labor supply of highly skilled women". *American Economic Journal: Applied Economics* 3 (3), 88–123.

Croson, R. and Gneezy, U. (2009). "Gender differences in preferences". *Journal of Economic Literature* 47 (2), 448–474.

DiNardo, J., Fortin, N. M., and Lemieux, T. (1996). "Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach". *Econometrica* 64 (5), 1001–1044.

Dohmen, T. and Falk, A. (2011). "Performance pay and multidimensional sorting: productivity, preferences, and gender". *American Economic Review* 101 (2), 556–590.

Dupuy, A. and Galichon, A. (2014). "Personality traits and the marriage market". *Journal of Political Economy* 122 (6), 1271–1319.

Exley, C. L. and Kessler, J. B. (2021). *The Gender Gap in Self-Promotion.* NBER Working Paper 26345.

Exley, C. L., Niederle, M., and Vesterlund, L. (2020). "Knowing when to ask: The cost of leaning in". *Journal of Political Economy* 128 (3), 816–854.

Flinn, C. J., Todd, P. E., and Zhang, W. (2018). "Personality traits, intra-household allocation and the gender wage gap". *European Economic Review* 109, 191–220.

Flory, J. A., Leibbrandt, A., and List, J. A. (2015). "Do competitive workplaces deter female workers? A large-scale natural field experiment on job entry decisions". *Review of Economic Studies* 82 (1), 122–155.

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

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 (2), 359–407.

Gneezy, U., Leonard, K. L., and List, J. A. (2009). "Gender differences in competition: Evidence from a matrilineal and a patriarchal society". *Econometrica* 77 (5), 1637–1664.

Goldin, C. (2014). "A grand gender convergence: Its last chapter". *American Economic Review* 104 (4), 1091–1119.

Goldin, C. and Rouse, C. (2000). "Orchestrating impartiality: The impact of "blind" auditions on female musicians". *American Economic Review* 90 (4), 715–741.

Greenwood, J., Guner, N., Kocharkov, G., and Santos, C. (2014). "Marry your like: Assortative mating and income inequality". *American Economic Review* 104 (5), 348–353.

Groneck, M. and Wallenius, J. (2020). "It Sucks to Be Single! Marital Status and Redistribution of Social Security". *Economic Journal* 131 (633), 327–371.

Guner, N., Kulikova, Y., and Llull, J. (2018). "Marriage and health: Selection, protection, and assortative mating". *European Economic Review* 104, 138–166.

Heinz, M., Normann, H.-T., and Rau, H. A. (2016). "How competitiveness may cause a gender wage gap: Experimental evidence". *European Economic Review* 90, 336–349.

Jung, S., Choe, C., and Oaxaca, R. L. (2018). "Gender wage gaps and risky vs. secure employment: An experimental analysis". *Labour Economics* 52, 112–121.

Kambourov, G. and Manovskii, I. (2009). "Occupational specificity of human capital". *International Economic Review* 50 (1), 63–115.

Kleven, H., Landais, C., and Søgaard, J. E. (2019). "Children and gender inequality: Evidence from Denmark". *American Economic Journal: Applied Economics* 11 (4), 181–209.

Laun, T. and Wallenius, J. (2021). "Having it all? Employment, earnings, and children". *Scandinavian Journal of Economics* 123 (1), 353–381.

Le, A. T., Miller, P. W., Slutske, W. S., and Martin, N. G. (2011). "Attitudes towards economic risk and the gender pay gap". *Labour Economics* 18 (4), 555–561.

Le Barbanchon, T., Rathelot, R., and Roulet, A. (2020). "Gender differences in job search: Trading off commute against wage". *Quarterly Journal of Economics* 136 (1), 381–426.

Lundborg, P., Plug, E., and Rasmussen, A. W. (2017). "Can women have children and a career? IV evidence from IVF treatments". *American Economic Review* 107 (6), 1611–1637.

Maasoumi, E. and Wang, L. (2019). "The gender gap between earnings distributions". *Journal of Political Economy* 127 (5), 2438–2504.

Machado, C. (2017). "Unobserved selection heterogeneity and the gender wage gap". *Journal of Applied Econometrics* 32 (7), 1348–1366.

Mas, A. and Pallais, A. (2017). "Valuing alternative work arrangements". *American Economic Review* 107 (12), 3722–3759.

Meara, K., Pastore, F., and Webster, A. (2020). "The gender pay gap in the USA: A matching study". *Journal of Population Economics* 33 (1), 271–305.

Mincer, J. (1978). "Family migration decisions". *Journal of Political Economy* 86 (5), 749–773.

Mora, R. (2008). "A nonparametric decomposition of the Mexican American average wage gap". *Journal of Applied Econometrics* 23 (4), 463–485.

Mueller, G. and Plug, E. (2006). "Estimating the effect of personality on male and female earnings". *ILR Review* 60 (1), 3–22.

Neuman, S. and Oaxaca, R. L. (2004). "Wage decompositions with selectivity-corrected wage equations: A methodological note". *Journal of Economic Inequality* 2 (1), 3–10.

Neumark, D., Bank, R. J., and van Nort, K. D. (1996). "Sex discrimination in restaurant hiring: An audit study". *Quarterly Journal of Economics* 111 (3), 915–941.

Nishiyama, S. (2019). "The joint labor supply decision of married couples and the U.S. social security pension system". *Review of Economic Dynamics* 31, 277–304.

Ñopo, H. (2008). "Matching as a tool to decompose wage gaps". *Review of Economics and Statistics* 90 (2), 290–299.

Nyhus, E. K. and Pons, E. (2012). "Personality and the gender wage gap". *Applied Economics* 44 (1), 105–118.

Oaxaca, R. L. and Choe, C. (2016). "Wage decompositions using panel data sample selection correction". *Korean Economic Review* 32 (2), 201–218.

Olivetti, C. and Petrongolo, B. (2008). "Unequal pay or unequal employment? A cross–country analysis of gender gaps". *Journal of Labor Economics* 26 (4), 621–654.

Olivetti, C. and Petrongolo, B. (2016). "The evolution of gender gaps in industrialized countries". *Annual Review of Economics* 8 (1), 405–434.

Ortigueira, S. and Siassi, N. (2013). "How important is intra-household risk sharing for savings and labor supply?" *Journal of Monetary Economics* 60, 650–666.

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

Reuben, E., Sapienza, P., and Zingales, L. (2015). *Taste for competition and the gender gap among young business professionals.* NBER Working Paper 21695.

Rogerson, R. and Wallenius, J. (2019). "Household time use among older couples: Evidence and implications for labor supply parameters". *Quarterly Journal of Economics* 134 (2), 1079–1120.

Roussille, N. (2021). *The central role of the ask gap in gender pay inequality.* Working Paper, University of California, Berkeley.

Sarsons, H., Gërxhani, K., Reuben, E., and Schram, A. (2021). "Gender differences in recognition for group work". *Journal of Political Economy* 129 (1), 101–147.

Strittmatter, A. and Wunsch, C. (2021). *The Gender Pay Gap Revisited with Big Data: Do Methodological Choices Matter?* IZA Discussion Paper 14128.

Wiswall, M. and Zafar, B. (2018). "Preference for the workplace, investment in human capital, and gender". *Quarterly Journal of Economics* 133 (1), 457–507.

Appendix

A 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

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

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

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

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

B Allowing for unobservable wage determinants

To acknowledge the potential importance of wage determinants usually not reported in observational data, we consider in this Appendix the case where there is a second characteristic, y, which is unobservable to the econometrician. The earnings potential of individual i then is

$$\log \psi_i = \gamma_{0,g(i)} + \gamma_{x,g(i)} x_i + \gamma_{y,g(i)} y_i.$$

We denote the variance of a characteristic z as σ_z^2 and its correlation to characteristic $\widetilde{z}$ as $\rho_{z,\widetilde{z}}$. For simplicity, we assume that both characteristics are measured in a way that they increase earnings potentials, $\gamma_{x,g} > 0$, $\gamma_{y,g} \geq 0$ and that a part of the wage gap is actually due to differences in the observable characteristic, i.e., $\overline{x}_m > \overline{x}_f$. Inserting this into the wage equation (6) gives

$$w_i = B_{g(i)} + \beta_{1,g(i)} x_i + \delta_{1,g(i)} y_i + \beta_{2,g(i)} x_{-i} + \delta_{2,g(i)} y_{-i} + \varepsilon_i,$$

where $B_{g(i)} = \beta_0 + \Omega_1 \gamma_{0,g(i)} + \Omega_2 \gamma_{0,g(-i)}$, $\beta_{1,g(i)} = \Omega_1 \gamma_{x,g(i)} > 0$, $\delta_{1,g(i)} = \Omega_1 \gamma_{y,g(i)} > 0$, $\beta_{2,g(i)} = \Omega_2 \gamma_{x,g(-i)} < 0$, and $\delta_{2,g(i)} = \Omega_2 \gamma_{y,g(-i)} < 0$, with $\Omega_1 = 1 + 1/\sqrt{2} \cdot (1-\kappa)\,\sigma^2 / \left(1 - (1-\kappa)\,\sigma^2\right) > 0$ and $\Omega_2 = -1/\sqrt{2} \cdot \left((1-\kappa)\,\sigma^2\right) / \left(1 - (1-\kappa)\,\sigma^2\right) < 0$.

In empirical decompositions, only the observable characteristics can be accounted for. Accordingly, the standard decomposition estimates $w_i = b_{0,g(i)} + b_{1,g(i)} x_i + \epsilon_i$ and the extended decomposition estimates $w_i = b_{0,g(i)} + b_{1,g(i)} x_i + b_{2,g(i)} x_{-i} + \epsilon_i$.

Estimating the **standard** decomposition yields the slope estimate

$$\widehat{b}_{1,g}^{std} = \beta_{1,g} + \beta_{2,g} \frac{\operatorname{cov}(x_g, x_{-g})}{\operatorname{var}(x_g)} + \iota_{1,g(i)}^{std},$$

where

$$\iota_{1,g}^{std} = \delta_{1,g} \rho_{x_g,y_g} \sigma_{y_g} / \sigma_{x_g} + \delta_{2,g} \rho_{x_g,y_{-g}} \sigma_{y_{-g}} / \sigma_{x_g}$$

summarizes the bias due to the unobservables y_g and x_g and y_g are the vectors of characteristics x and y for individuals of gender g. The estimated explained gap then is

$$\widehat{\Delta}|_x^{std} = \left(\beta_{1,m} + \beta_{2,m} \frac{\operatorname{cov}(x_m, x_f)}{\operatorname{var}(x_g)} + \iota_{1,m}^{std} \right) \cdot (\overline{x}_m - \overline{x}_f).$$

In turn, the **extended** decomposition yields the slope estimates

$$\widehat{b}_{1,g}^{ext} = \beta_{1,g} + \iota_{1,g}^{ext}$$

and

$$\widehat{b}_{2,g}^{ext} = \beta_{2,g} + \iota_{2,g}^{ext},$$

where

$$\begin{aligned}\iota_{1,g}^{ext} =&(\delta_{1,g} \cdot (\rho_{x_g,y_g}\sigma_{y_g}/\sigma_{x_g} - \rho_{x_{-g},y_g}\sigma_{y_g}/\sigma_{x_g}) && \text{(B.1)}\\ &+ \delta_{2,g} \cdot (\rho_{x_g,y_{-g}}\sigma_{y_{-g}}/\sigma_{x_g} - \rho_{x_{-g},y_{-g}}\sigma_{y_{-g}}/\sigma_{x_g}))/(1-\rho^2_{x_m,x_f}) && \text{(B.2)}\end{aligned}$$

and

$$\begin{aligned}\iota_{2,g}^{ext} =&(\delta_{1,g} \cdot (\rho_{x_{-g},y_g}\sigma_{y_g}/\sigma_{x_{-g}} - \rho_{x_m,x_f}\rho_{x_g,y_g}\sigma_{y_g}/\sigma_{x_{-g}}) && \text{(B.3)}\\ &+ \delta_{2,m} \cdot (\rho_{x_{-g},y_{-g}}\sigma_{y_{-g}}/\sigma_{x_{-g}} - \rho_{x_m,x_f}\rho_{x_g,y_{-g}}\sigma_{y_{-g}}/\sigma_{x_{-g}}))/(1-\rho^2_{x_m,x_f}). && \text{(B.4)}\end{aligned}$$

The explained gap from the extended decomposition is

$$\widehat{\Delta}|_x^{ext} = \left(\beta_{1,m} - \beta_{2,m} + \iota_{1,m}^{ext} - \iota_{2,m)}^{ext}\right) \cdot (\overline{x}_m - \overline{x}_f)\,.$$

Comparing the decomposition results. Applying the estimated slope coefficients, the difference between the explained gaps yielded by the standard Oaxaca-Blinder decomposition and our extended decomposition is given by

$$\begin{aligned}\mathrm{d}\widehat{\Delta}|_X =& \left(\widehat{\Delta}|_X\right)^{ext} - \left(\widehat{\Delta}|_X\right)^{std}\\ =& \left(-\beta_{2,m} \cdot \left(1 + \rho_{x_m,x_f} \cdot \frac{\sigma_{x_m}}{\sigma_{x_f}}\right)\right.\\ &+\frac{\rho^2_{x_m,x_f}}{1-\rho^2_{x_m,x_f}} \cdot \frac{\sigma^2_{x_f} + \sigma_{x_m}}{\sigma_{x_m}\sigma_{x_f}} \cdot \left(\delta_{1m}\rho_{x_m,y_m}\sigma_{y_m} + \delta_{2m}\rho_{x_m,y_f}\sigma_{y_f}\right)\\ &\left.-\frac{1}{1-\rho^2} \cdot \frac{\sigma^2_{x_m} + \sigma_{x_f}}{\sigma_{x_f}\sigma_{x_m}} \cdot \left(\delta_{1m}c\rho_{x_f,y_m}\sigma_{y_m} + \delta_{2m}\rho_{x_f,y_f}\sigma_{y_f}\right)\right) \cdot (\overline{x}_m - \overline{x}_f)\,.\end{aligned}$$

The term in the first line on the right-hand side is positive (since $\beta_{2,m}$ is negative) unless σ_{x_m} is very large relative to σ_{x_f} and there is pronounced *negative* assortative mating along observables. The term in the second line is positive unless unobservable wage enhancers of men are *negatively* correlated with observable ones, or unobservable wage enhancers of women are substantially more strongly correlated to men's observable wage enhancers than men's own unobservable wage enhancers. The term in the third line may be negative but it will not dominate the other two terms unless women's observable wage enhancers are substantially more strongly correlated to men's unobservable wage enhancers than to women's own unobservable wage enhancers. The following proposition provides a summary.

Proposition. *There exist negative lower bounds* $\underline{\rho} < 0$, $\underline{\chi}_m < 0$, $\underline{\chi}_f < 0$ *for the extent of assortative mating and the within-gender covariances between observable and unobservable wage determinants as well as positive upper bounds* $\overline{\zeta}_{m,f} > 0$ *and* $\overline{\zeta}_{f,m} > 0$ *for the across-gender covariances between observable and unobservable wage determinants such that, if but not only if* $\rho > \overline{\rho}$, $cov\,(x_m, y_m) > \underline{\chi}_m$, $cov\,(x_f, y_f) > \underline{\chi}_f$, $cov\,(x_m, y_f) < \overline{\zeta}_{m,f}$, *and* $cov\,(x_f, y_m) < \overline{\zeta}_{f,m}$,

$$\left(\widehat{\Delta}|_X\right)^{ext} > \left(\widehat{\Delta}|_X\right)^{std}.$$

Proof. Inserting $\rho = cov\,(x_m, y_m) = cov\,(x_f, y_f) = cov\,(x_m, y_f) = cov\,(x_f, y_m) = 0$ gives $\mathrm{d}\widehat{\Delta}|_X = -\beta_{2,m} > 0$. The unambiguous signs of the partial derivatives $\partial \mathrm{d}\widehat{\Delta}|_X / \partial \rho > 0$ (since $\beta_{2,m} < 0$), $\partial \mathrm{d}\widehat{\Delta}|_X / \partial cov\,(x_m, y_m) > 0$ (since $\delta_{1,m} > 0$), $\partial \mathrm{d}\widehat{\Delta}|_X / \partial cov\,(x_f, y_f) > 0$ (since $\delta_{1,m} > 0$), $\partial \mathrm{d}\widehat{\Delta}|_X / \partial cov\,(x_m, y_f) < 0$ (since $\delta_{2,m} < 0$), and $\partial \mathrm{d}\widehat{\Delta}|_X / \partial cov\,(x_f, y_m) < 0$ (since $\delta_{2,m} < 0$) complete the proof. □

The unobservable wage determinants add some noise to the systematic effects outlined in the main text but they can outweigh them only under extreme conditions. These extreme cases relate to pronounced *negative* assortative mating and unlikely correlations between observables and partner unobservables. Empirically, there is *positive* assortative mating in terms of wage-relevant characteristics, such as education (Bredemeier and Juessen 2013, Greenwood et al. 2014, Guner et al. 2018). For the systematic mechanisms to be dominated by changing biases due to unobservable determinants of income potentials, an individual's unobservable wage enhancers would have to be more strongly correlated to the observable wage enhancers of the partner than to the individual's own observable wage enhancers. Using survey data from the Netherlands which include information on personality traits and risk preferences, Dupuy and Galichon (2014) analyze the extent of assortative mating along these dimensions. While they find significant connections between personality traits of husband and wife, personality traits and risk aversion display low and insignificant connections to the partner's education. This supports that unobservable characteristics, while potentially being relevant for the wage gap, are unlikely to affect the comparison between the two decomposition approaches.