Canadian Labour Economics Forum

WORKING PAPER SERIES

Pay Gaps and Outside Wages: The U.S. Gender Wage Gap 1980-2010

Chris Bidner (Simon Fraser University)
Ben Sand (York University)

CLEF WP #82

Pay Gaps and Outside Wages:

The U.S. Gender Wage Gap 1980-2010*

Chris Bidner[†] and Ben Sand[‡]

March 17, 2023

Abstract

We consider the role of differences in *outside wages*—the wages earned by similar workers in other jobs—in shaping pay differentials, focusing on the gender dimension in the U.S. during 1980-2010. Using instruments that exploit differential exposure to common industry wage shocks, we find a substantial role for outside wages. Differences in outside wages account for one-half of the level of, and trends in, the unexplained gender wage gap. Our results offer a reinterpretation of trends in the gender wage gap, and suggest that standard wage decompositions are systematically misleading.

1 Motivation

A variety of economic models imply that a worker's wage is causally affected by the wages earned by similar workers in other jobs. In this paper we take

*Previously circulated as "Industrial structure and the gender gap: Evidence from U.S. cities". We are grateful for the valuable feedback received during numerous conferences and seminars, as well as discussions with Paul Beaudry, David Card, Kelly Foley, Nicole Fortin, David Green, Kevin Lang, Thomas Lemieux, Enrico Moretti, and Krishna Pendakur. All errors are our own.

[†] Department of Economics, Simon Fraser University. cbidner@sfu.ca

[‡] Department of Economics, York University. bmsand@yorku.ca

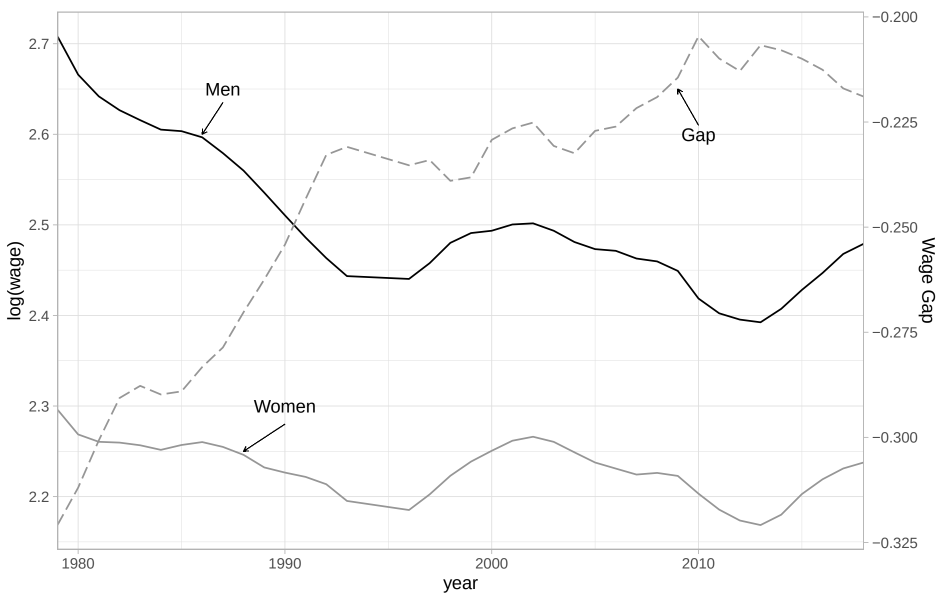
seriously the possibility that a worker's wage is influenced by the average wage earned by those of the same gender—what we call their *outside wage*. Our interest lies in a natural consequence of this: the possibility that wage gaps, even between otherwise identical workers, are sensitive to differences in outside wages. Given that the wage spillovers inherent in such settings have consequences for interpreting past trends and for crafting future policy, our goal is to quantify the relevance of this possibility.

Our specific focus is on the within-industry gender wage gap among non-college workers in the U.S. between 1980 and 2010. This setting is particularly useful for our purposes for two reasons. First, this period saw significant changes in the within-industry gender wage gap. Figure 1 shows how this gap, similar to the unadjusted gender wage gap (Blau and Kahn (2006)), narrowed significantly over this period and particularly in the 1980s. Second, the period saw structural changes that had a significant impact on outside wages. The impact of forces such as globalization, deregulation, de-unionization, and routinization was particularly pronounced for less-educated workers, and especially so for men (Binder and Bound (2019)). Figure 1 also shows how the real wages of non-college workers fell for both men and women, especially during the 1980s, but that the decline was much larger for men.

Our approach focuses on the variation in outside wages that stem from an economy's industrial structure. The key idea is that an economy's array of industrial advantages will generally imply quite different outside wages for men and women. For instance, the characteristics of an economy's steel sector will influence the outside wages of workers in the retail sector, but much more so for men owing to their greater propensity to work in the steel sector. If wages in the steel sector decline, say because of greater import competition, then so too will the quality of outside wages of males relative to females in the retail sector, potentially affecting the gender wage gap in retail. This effect will be stronger in periods where wages in steel declined more rapidly (e.g. in the 1980s) and in places where the steel sector is more prominent (e.g. Pittsburgh).[1] We generalize this argument in Section 2 where we construct

[1]Note that this argument concerns industry-specific *wages*, and does not rely on changes

Figure 1: Gender Wage Gap Narrows Because Male Wages Fall More



Notes: The figure uses data from Outgoing Rotation Group (ORG) Supplements of the Current Population Survey. The wage series are calculated as the average log wage for each gender-education group using fixed weights, where the fixed weights hold the within-group age and sub-education composition constant over time. The plotted series are smoothed using a three-year moving average.

a measure of gender differences in exposure to high-paying industries. We present a series of stylized facts showing how, consistent with our argument, this measure predicts within-industry gender wage gaps across time and space.

The stylized facts in section 2 suggest the relevance of outside wages and motivate a more careful analysis. Our empirical strategy is derived from a class of economic models in which workers of a given type (gender) earn an equilibrium wage that depends on their industry *and* the average wage among workers of the same type. We address the obvious endogeneity issue in mul-

in industry-specific *employment*. Indeed, as will become clear, our empirical strategy discards variation stemming from changes in employment. Our argument (and results) thus do not rely on the well-documented decline in employment in male-dominated industries (e.g. Autor et al. (2013)). See Beeson et al. (2001) for a discussion of the causes and consequences of the decline in steel wages experienced in the U.S. in the 1980s.

tiple ways. First, we exploit variation across gender, industries, local labor markets (commuting zones, CZs), and time. This allows us to control for any unobserved characteristics that affect average wages of gender×CZ×year cells that arise at the level of industry×gender×year, industry×gender×CZ, and industry×CZ×year. In practice, we ask whether places that saw large changes in within-industry gender wage gaps also tend to be places that saw large changes in the raw gender wage difference. This still leaves us with a mechanical endogeneity problem due to unobserved place-specific drivers of changes in within-industry gender pay gaps. We deal with this by using the inner product structure of average wages to propose a range of Bartik-style instruments. These instruments predict the change in a location's gender gap in outside wages without using variation stemming from a change or gender difference in any location characteristic. Instead, the instruments predict gender differences in exposure to common national-level industry wage shocks using the location's lagged industry structure and the national-level gender composition of each industry. Finally, we also use a series of controls in an attempt to capture potential confounders, and we probe the robustness of our findings in various ways (in section 5.2).

Our results suggest that gender differences in outside wages have a substantial impact on within-industry gender wage differentials. We find that gender differences in outside wages are responsible for around 50% of the level of, and change in, the within-industry wage gap in any given year. Furthermore, our results suggest important general equilibrium effects whereby any factor that has a direct effect on the raw gender wage differential will also have implications for outside wages and therefore will have an indirect effect on within-industry gender wage differentials. We estimate that direct effects are accompanied by additional indirect effects around two-thirds as large. By ignoring such indirect effects, a standard Blinder-Oaxca decomposition will systematically inflate the unexplained component and will suggest an effect size that is only 60% of the actual total effect.

Our work contributes to an enormous literature concerned with understanding wage differentials, and in particular that work concerned with gender wage differentials.[2] This literature overwhelmingly focuses on internal conditions and emphasizes women's gains.[3] Our work contrasts with this literature in two main ways. The first is our emphasis on external conditions. Existing explanations are surely relevant, and we take great care to control for them in our empirical strategy. Still, we find that outside wages are at least as relevant for understanding the gender wage gap as all internal conditions combined. The second is our emphasis on men's losses.[4] Figure 1 clearly shows that the gender wage gap narrows the most in periods where men's wages *fall more* than women's. Our results suggest that this is less about women 'swimming upstream' (Blau and Kahn (1997)) and more about men being more strongly tethered to those most exposed to the current. That is, changing economic conditions had a detrimental direct effect on the wages of some workers, mainly men. This reduced mens' outside wages, thereby inducing an indirect effect that tended to lower the wages of all men.

Our paper more directly contributes to a body of work that highlights the role of external conditions in generating wage differentials. For instance, the literature has shown that wage differentials can arise when otherwise identical workers experience different conditions in other jobs; e.g. with respect to prejudice (Black (1995), Flabbi (2010)) and productivity (Albrecht and Vroman (2002)). Our framework incorporates these possibilities as special cases of sources of outside wage differentials.

Our work complements recent studies that use rich employee-employer matched data to examine the role of outside options. These data, combined

[2] There are numerous excellent surveys, including Goldin (2014), Blau and Kahn (2016, 2000), Bertrand (2011), Altonji and Blank (1999) and Olivetti and Petrongolo (2016).

[3] For instance, the literature stresses factors such as the extent of discrimination (Guryan and Charles (2013)), the level of unobserved skill (Mulligan and Rubinstein (2008)), the return to unobserved skill (Blau and Kahn (1997)), the composition of unobserved skill (Weinberg (2000), Welch (2000), Beaudry and Lewis (2014), Cortes et al. (2016)), changing preferences for job amenities (Flabbi and Moro (2012), Goldin (2014)), changing incentives to invest in human capital (Bailey et al. (2012)), and differences in probabilities of transitioning to unemployment (Bowlus (1997), Bowlus and Grogan (2009)).

[4] To be clear, we are referring to our explanation for the *within-industry* gender wage gap. It is, of course, well-understood that wages have declined in male-dominated industries and that this has a mechanical effect on the raw wage gap (e.g. O'Neill and Polachek (1993)).

with innovative methodologies, have allowed researchers to produce insights via the use of fine-grained measures of the job surplus (Card et al. (2015)) and the outside options (Caldwell and Harmon (2019), Caldwell and Danieli (2022)) available to individual workers. For instance, Caldwell and Danieli (2022) use an outside options index, derived from a frictionless transferable utility matching model with heterogeneous workers and firms, to show that gender differences in outside options in Germany are largely due to differences in the willingness to commute or move. In contrast, the weak data requirements of our approach allow us to analyze times and places where sufficiently rich data do not exist, albeit at a more aggregated level. Our level of analysis has the important advantage of offering a transparent view of the general equilibrium implications of outside wages. For instance, Card et al. (2015) find gender wage differentials in Portugal are subject to a "bargaining" effect whereby women extract less of the surplus available at their firm, and to a "sorting" effect whereby women tend to work at lower-surplus firms. Our approach explicitly connects these two effects: a "sorting" effect reasonably suggests that women have worse outside options at any given firm, thereby *generating* a "bargaining" effect (even if there are no gender differences in bargaining skills per se). Indeed, we find no evidence of gender differences in bargaining skills.

Our work also contributes to the literature concerned with the impact of industrial structure on wages. Whilst much of this literature is concerned with the mechanical effect of industrial composition (e.g. Borjas and Ramey (1995)), we are concerned with the indirect effect arising from the role of industrial structure in shaping outside wages (see Green (2015)). In this respect, our work builds on Beaudry et al. (2012) by considering wage *gaps*, by emphasizing a broader interpretation of the underlying mechanism, by proposing new Bartik-style instruments, and by employing a more direct estimation strategy. Our conceptual framework requires a modest extension to incorporate the central (and empirically relevant) feature that industrial employment distributions vary by gender (e.g. see Olivetti and Petrongolo (2016)). Further, by considering the gender dimension we are able to eliminate their main identification threat by including industry×CZ×year fixed effects.

We present some motivating stylized facts in Section 2. To better understand these observations, we provide a conceptual framework in Section 3 and then use this to develop our empirical strategy in Section 4. We present our results in Section 5 before offering concluding comments in Section 6.

2 Stylized Facts

Do gender differences in outside wages causally affect gender differences in pay within industries? To begin exploring this, we focus on the role of industrial structure in affecting gender differences in outside wages. Specifically, consider the following measure of the gender difference in exposure to high-paying industries:

$$\textsc{ExposureGap}_{ct} \equiv \sum_{j=1}^{J} [\pi_{fjct} - \pi_{mjct}] \cdot d_{jt}, \tag{1}$$

where π_{gjct} is the share of gender $g \in \{f, m\}$ workers employed in industry j in location c at time t and d_{jt} is a measure of industry wages. The term in brackets thus measures the gender difference in exposure to industry j in location c; we refer to industries with positive values as 'female-exposed' and with negative values as 'male-exposed'.

We begin by treating the whole of the U.S. as the relevant economy in order to focus on time series variation. For this purpose we use Current Population Survey (CPS) data on non-college workers from 1979-2018.[5] The data clearly show that $\textsc{ExposureGap}_{ct}$ declined substantially over the 1980-2010 period. We emphasize three points. First, the decline was largely due to male-exposed industries having their high-paying status eroded. Figure 2 shows that male-exposed industries were also high-paying industries in 1980, but experienced the least wage growth in the period to 2010. Second, the specific time path of

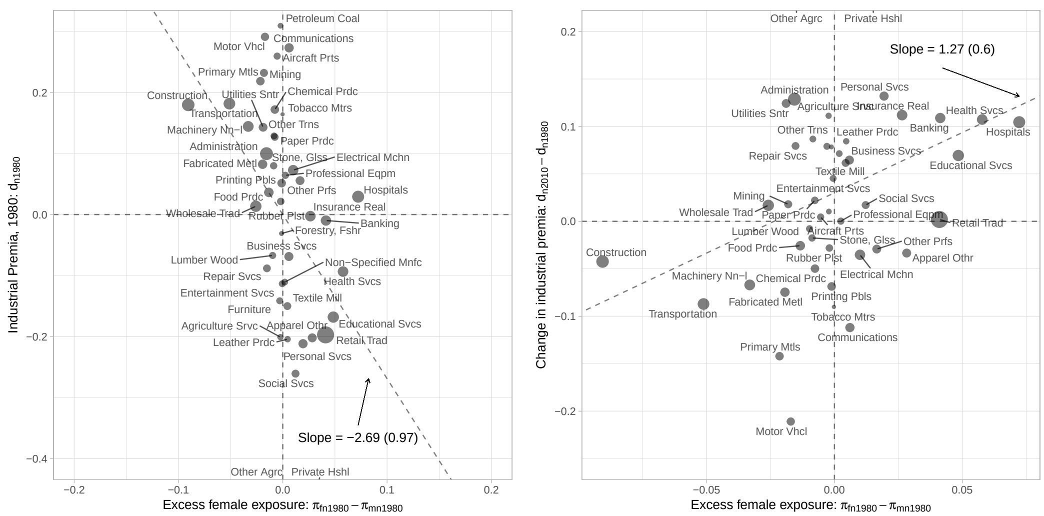
[5]Relative industry wages are estimated coefficients on industry dummies from a log wage regression that also controls for gender, human capital characteristics (education, potential experience), race and occupation. See appendix section C for details.

ExposureGap$_{ct}$ shows a striking resemblance to the time path of the within-industry gender pay gap. Figure 3 shows, for instance, how the relatively steep narrowing of the pay gap in the 1980s is accompanied by a relatively steep narrowing of ExposureGap$_{ct}$ in this period. Third, we emphasize that there is no mechanical relationship between these two series; e.g. the fact that male-exposed industries lost their large pay advantage (for both men and women) has no necessary connection with pay differences within industries. In order for the close relationship to be spurious it must be that there is some time-varying omitted factor correlated with both the relative pay of women within each industry and with women's relative exposure to high-paying industries; e.g. if women moved into high-paying industries as a result of secular declines in discrimination. To deal with this potential issue, we now turn to cross-sectional variation.

We now explore the relationship between within-industry wage gaps and ExposureGap$_{ct}$ across local labour markets within a given year using Census and ACS data. Here too the data reveal a strong positive relationship. We again emphasize three points. First, within-industry gender pay gaps vary substantially across local labour markets. Second, this pay gap is highly correlated with ExposureGap$_{ct}$. Figure 4 Panels (a) and (b) show the relationship in both 1980 and 2010.[6] Third, we emphasize that there is no mechanical relationship between the two variables: e.g. there is no necessary reason why locations that are heavy in high-paying and male-exposed industries are also the locations where men tend to be paid more than women within industries. The strong cross-sectional relationship is spurious if there is an omitted location-specific factor correlated with both women's relative pay within industries and women's relative exposure to high-paying industries; e.g. discrimination may be fostered by a preponderance of industries that are simultaneously male-exposed and high-paying. To address this potential issue we reintroduce the time dimension.

[6] A commuting zone's gender pay gap is obtained by taking the coefficient on female×city interactions from a log wage regression that also controls for human capital variables, race, occupation, and industry. See appendix section B.1 for further details.

Figure 2: Male Jobs Were Good, Female Jobs Improved



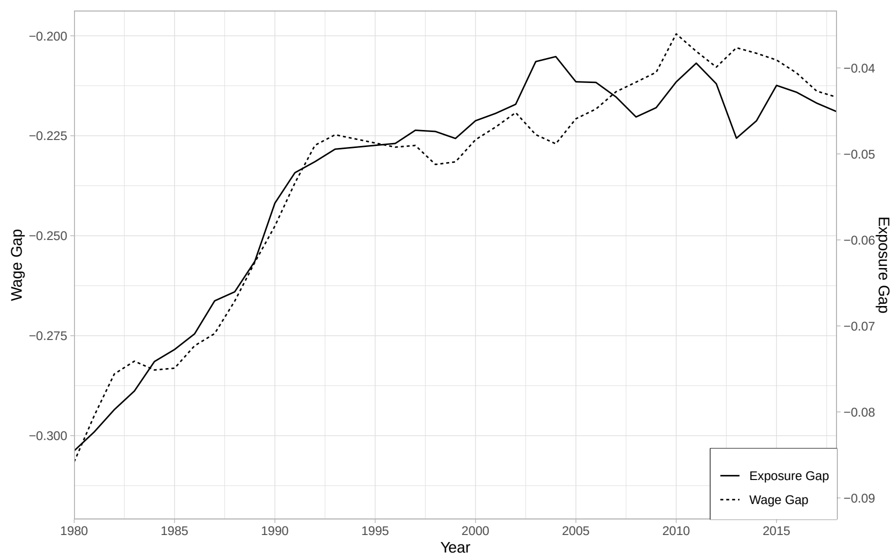
(a) Male Jobs are Good Jobs, 1980 (b) Wages Rose Most in Female Jobs

Notes: The figure uses data from Outgoing Rotation Group (ORG) Supplements of the Current Population Survey. Industry categories are aggregated into 45 industry groups (details are provided in Appendix B). Relative industrial wages, d_{nt}, are calculated as the coefficients on a full set of industry dummies in a regression of log hourly wages on education, age, race, and aggregated occupation dummies (additional details provided in in Appendix C)). The π_{fnt} denote the share of hourly employment of gender, ($g \in \{\underline{f}\text{emale}, \underline{m}\text{ale}\}$), industry ($n$), and year ($t$). The size of the marker indicates the relative size of the industry in terms of share of total hours of employment. The dashed line indicates a linear fit with the standard error in parenthesis.

We now consider cross-sectional variation in the *change* in the gaps to control for all time-specific and location-specific factors. Figure 4 Panel (c) shows a strong positive correlation between a location's change in within-industry wage gap and change in $\textsc{ExposureGap}_{ct}$. The figure also shows how the largest change in $\textsc{ExposureGap}_{ct}$ occurred in rust belt locations, such as Pittsburgh, where industries that are both male-exposed and saw low-wage-growth, such as steel, are over-represented. We emphasize that being male-exposed is not sufficient: we see the smallest change in $\textsc{ExposureGap}_{ct}$ in places, such as Chattanooga, where industries that are both male-exposed and saw high wage-growth, such as utilities, are over-represented.

The stylized facts laid out in this section use various sources of variation

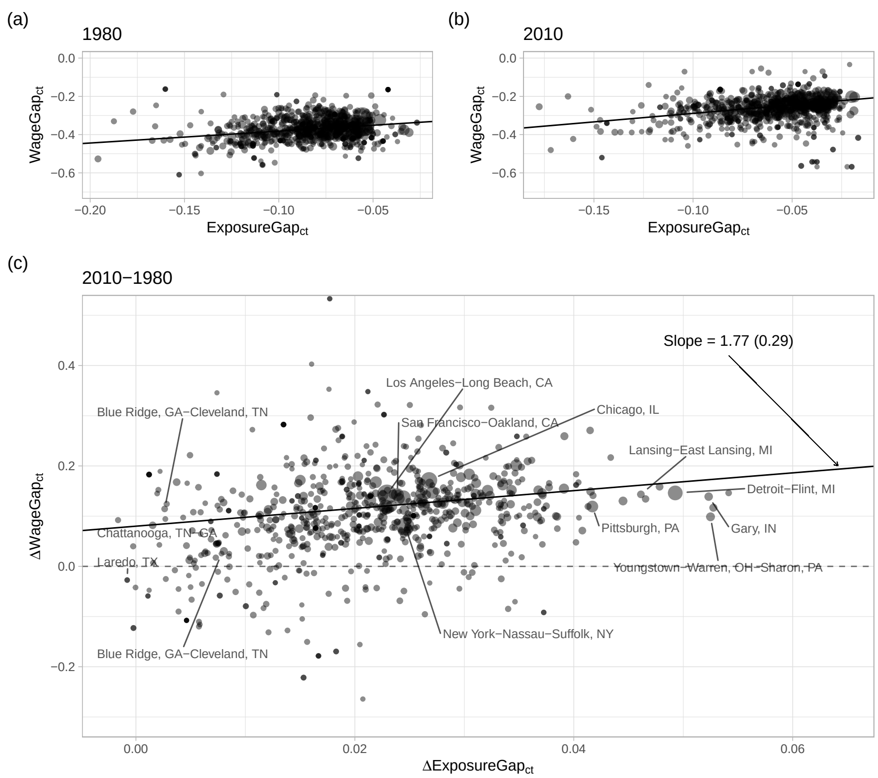
Figure 3: Gender Wage and Exposure Gaps



Notes: The figure uses data from Outgoing Rotation Group (ORG) Supplements of the Current Population Survey. The Gender Wage Gap is the coefficient on a gender dummy in a regression of wages on education, age, race, occupation controls and a full set of industry dummy variables (Details in Appendix C). The Exposure Gap is calculated as $\textsc{ExposureGap} = \sum_n (\pi_{mnt} - \pi_{fnt}) \cdot d_{nt}$, where d_{nt} are the coefficients on the industry dummy variables and π_{gnt} denote gender-industry employment shares in time t.

to demonstrate a tight connection between gender pay differences within industries and gender differences in exposure to high paying industries. The following section is devoted to better understanding this relationship, including drawing out broader general equilibrium implications, identifying the a key parameter of interest, proposing a methodology to estimate this parameter, and identifying threats to a causal interpretation of the resulting estimate.

Figure 4: Gender Wage and Exposure Gaps across Local Labor Markets



Notes: The figure uses data from the US Census and ACS. The y-axis denotes the change in the regression adjusted, within-industry, commuting-zone level gender gap between 1980 and 2010, the x-axis denotes the change in the commuting zone-level $\text{ExposureGap}_{ct} = \sum_n (\pi_{mcn1980} - \pi_{fcn1980}) \cdot \Delta d_{nt}$. The marker size is proportional to the size of the commuting zone in 1980. See appendix section B.1 for further details.

3 Conceptual Framework

3.1 Basic Set Up

We begin by considering a particular economy at a particular date. A worker of type $g \in \{f, m\}$ in sector $n \in \{1, ..., N\}$ earns a wage, w_{gn}, that is shaped by factors internal to the worker-firm relationship (e.g. productivity, discrimination, rents) denoted ψ_{gn}, as well as the level of outside wages, denoted ρ_g. In particular,

$$w_{gn} = \psi_{gn} + \delta \cdot \rho_g, \tag{2}$$

where $\delta \in [0, 1)$ is a key parameter examined further below. The level of outside wages is given by

$$\rho_g \equiv \sum_j \xi_{gj} \cdot \tilde{w}_{gj} \tag{3}$$

where $\tilde{w}_{gj}$ is the expected wage for type g workers in sector j, and ξ_{gj} is the *exposure* of type g workers to such jobs, where $\xi_{gj} \geq 0$ and $\sum_j \xi_{gj} = 1$. Equilibrium wages are those that satisfy (2), (3), and the rational expectations condition $\tilde{w}_{gn} = w_{gn}$.

This basic structure arises from a broad class of models. In Section A of the appendix we provide two dynamic models with search frictions in which wages depend on outside wages because they represent outside options. The first, in the spirit of Pissarides (2000), is a search and matching model where outside options matter because of wage bargaining. The second, in the spirit of Shapiro and Stiglitz (1984), is an efficiency wage model where outside options matter because of incentives to shirk. The key common features are (i) productivity varies across sectors and (potentially) worker type, and (ii) the exposure vector differs by worker type.[7] These models also endogenize exposure as a result of type-specific search proclivities interacting with firms' incentives to open

[7] See Le Barbanchon et al. (2020) and Sorkin (2017) for analyses of gender differences in search.

vacancies.

The basic structure also arises in an entirely different class of models that emphasize social comparisons. In these models outside wages are interpreted as a reference wage that workers compare their wage to. Such comparisons matter because they affect productivity via aggreivement or declining morale (e.g. Akerlof and Yellen (1990), Breza et al. (2017)). Regardless of the specifics, this sort of structure implicitly underlies empirical work that emphasizes wage spillovers, for instance from union negotiations (Lewis (1963)) and public sector wages (Babcock et al. (2005)). Our goal in this paper is not to distinguish between the various possible underlying models, but rather to lay out a clear framework that is sufficiently generic to provide a foundation for establishing whether outside wages matter quantitatively.

To take the model to the data, we need to place structure on the exposure terms. In line with the equilibrium outcomes of the models presented in Section A and with our approach in Section 2, we equate exposure with employment shares. That is, we set $\xi_{gn} = \pi_{gn}$, where π_{gn} is the share of type g workers employed in industry n.

3.2 What does the model say about Pay Gaps?

We now draw out some implications for understanding pay gaps across worker types. From (2), the average wage for type g workers is

$$\bar{w}_g = \sum_j \pi_{gj} \cdot \psi_{gj} + \delta \cdot \rho_g \tag{4}$$

where π_{gj} are employment shares. A standard Blinder-Oaxaca decomposition (Fortin et al. (2011)) can be used to decompose the difference in average wages, $\bar{w}^* \equiv \bar{w}_f - \bar{w}_m$, into an explained component, E, and an unexplained component, U:

$$\bar{w}^* = \underbrace{\sum_j \pi_j^* \cdot \bar{\psi}_j}_{E} + \underbrace{\overbrace{\sum_j \bar{\pi}_j \cdot \psi_j^*}^{U_{\text{in}}} + \overbrace{\delta \cdot \rho^*}^{U_{\text{out}}}}_{U} \tag{5}$$

where $\rho^* \equiv \rho_f - \rho_m$ is the gender difference in outside wages, $\{\psi_j^*, \pi_j^*\}$ are gender differences in internal conditions and employment shares in sector j, and $\{\bar{\psi}_j, \bar{\pi}_j\}$ are gender-neutral internal conditions and employment shares in sector j.[8]

Here E reflects that part of the wage difference that is explained by the gender difference in employment distributions. The unexplained component, U, contains two sub-components; U_{in} is standard and reflects average gender differences in internal conditions whereas U_{out} is our focus as it reflects the difference in outside wages. It is straightforward to show that

$$U_{\text{out}} = \frac{\delta}{1-\delta} \cdot \left[\sum_j \xi_j^* \cdot \bar{\psi}_j + \sum_j \bar{\xi}_j \cdot \psi_j^* \right] = \frac{\delta}{1-\delta} \cdot \left[E + U_{\text{in}} \right], \tag{6}$$

where the final equality is due to the equating of exposure and employment shares. Equation (6) gives a clear view of general equilibrium implications arising from a multiplier effect that is familiar from the social interactions literature (Manski (1993); Moffitt (2001)), and in the context of industrial structure and wage levels, is central in Beaudry et al. (2012) and lucidly discussed in Green (2015).

Changes in internal conditions or employment distributions, $\{\psi_{gn}, \pi_{gn}\}$, will have a *direct* effect on the wage gap operating through E or U_{in}. But such direct effects will, in turn, affect differences in outside wages and thus also have an *indirect* effect on the wage gap. This indirect effect operates through U_{out}, and (6) tells us that the indirect effect will be $\delta/(1-\delta)$ times the magnitude of the direct effect. Equivalently, δ tells us the proportion of the total effect that is due to the indirect channel. The Blinder-Oaxaca decomposition will

[8] Specifically, $\bar{\psi}_j = \alpha \cdot \psi_{fj} + (1-\alpha) \cdot \psi_{mj}$ and $\bar{\pi}_j = (1-\alpha) \cdot \pi_{fj} + \alpha \cdot \pi_{mj}$ for some $\alpha \in \mathbb{R}$.

make misleading predictions about the effect of changes in internal conditions or employment distributions if the indirect channel we emphasize is ignored. Specifically, the decomposition will fail to predict a proportion δ of the actual change.

Finally, (6) can be used to back out $(U_{\text{out},t}, U_{\text{in},t})$ from knowledge of (E_t, U_t) and a value of δ. Doing so allows us to evaluate the relevance of outside wages for understanding gender wage gaps. In particular:

$$U_{\text{out},t} = \delta \cdot [E_t + U_t], \quad U_{\text{in},t} = U_t - U_{\text{out},t}. \tag{7}$$

This allows us to quantify the proportion of the unexplained gap that becomes explained by outside wages, $U_{\text{out},t}/U_t$, and the proportion of the change in the unexplained gap that is explained by changes in outside wages, $\Delta U_{\text{out},t}/\Delta U_t$. We report on estimates of these series in Section 5.3. But now we turn to estimating δ.

3.3 Estimating δ

We treat the above analysis as applying to a particular location $c \in \{1, ..., C\}$ at a particular point in time, $t \in \{1, ..., T\}$. It is straightforward to show that equilibrium wages satisfy the following equilibrium relationship:

$$w_{gnct} = \psi_{gnct} + \delta \cdot \bar{w}_{gct}, \tag{8}$$

where $\bar{w}_{gct} \equiv \sum_j \pi_{gjct} \cdot w_{gjct}$ is the average wage of type g workers in their economy. For each $t \geq 2$, taking the differencing across worker type and time gives:

$$\Delta w^*_{nct} = \Delta\psi^*_{nct} + \delta \cdot \Delta\bar{w}^*_{ct}. \tag{9}$$

To derive an estimating equation, decompose $\Delta\psi^*_{nct}$ into an industry component, a mean zero location component, and an orthogonal residual component:

$$\Delta\psi^*_{nct} = \zeta_{nt} + \zeta_{ct} + \nu_{nct} \tag{10}$$

where $\sum_c \zeta_{ct} = \sum_n \nu_{nct} = \sum_c \nu_{nct} = 0$ for all $t \geq 2$.[9] To be sure, ζ_{ct} captures location-specific factors that affect the change in the gender difference in internal conditions *within* industries. For instance, ζ_{ct} captures location-specific changes in selection, discriminatory attitudes, and adoption of gender-biased technology. Using (10) in (9) gives:

$$\Delta w^*_{nct} = \zeta_{nt} + \delta \cdot \Delta\bar{w}^*_{ct} + \varepsilon_{nct}, \tag{11}$$

where $\varepsilon_{nct} \equiv \zeta_{ct} + \nu_{nct}$. This forms the basis of an estimating equation whereby ζ_{nt} is captured industry×time fixed effects and ε_{nct} is treated as the error term. It is clear that OLS is inappropriate because ζ_{ct} will have a mechanical impact on $\Delta\bar{w}^*_{ct}$. Furthermore, there are various reasons to suspect a non-mechanical correlation between the two.[10] To address this, we use instrumental variables.

3.3.1 Instrumental Variables

Our main instrument borrows from the exposure literature.[11] It is a Bartik-style instrument that, in essence, represents a *predicted* change in $\textsc{ExposureGap}_{ct}$. The only location-specific characteristic that enters into the prediction is the initial industrial structure. In particular, our 'index' instrument is:

$$Z^{\text{Index}}_{ct} \equiv \sum_j \hat{\pi}^*_{jc,t-1} \cdot \Delta d_{jt}, \tag{12}$$

[9] That is, define $\zeta_{nt} \equiv (1/C) \cdot \sum_c \Delta\psi^*_{nct}$, $\zeta_{ct} \equiv (1/N) \cdot [\sum_n \Delta\psi^*_{nct} - (1/C) \cdot \sum_c \sum_n \Delta\psi^*_{nct}]$, and $\nu_{nct} \equiv \Delta\psi^*_{nct} - \zeta_{nt} - \zeta_{ct}$.

[10] For instance, ζ_{ct} may be systematically related to changes in a location's industrial structure, changes in the the array of location-specific industrial advantages, gender differences in employment distributions (or the change in this), and so on.

[11] For example, Adão et al. (2019); Beaudry et al. (2012, 2018); Borusyak et al. (2020); Goldsmith-Pinkham et al. (2020)

where Δd_{jt} is a national-level measure of wage growth in industry j and $\hat{\pi}^*_{jc,t-1} \equiv \hat{\pi}_{fnc,t-1} - \hat{\pi}_{mnc,t-1}$ is a location-specific predicted base period gender difference in exposure to industry j. In particular, letting $\text{Empl.}_{nc,t-1}$ denote location c's total employment in industry n at date $t-1$, we define

$$\hat{\pi}_{gnc,t-1} \equiv \frac{\text{Empl.}_{nc,t-1} \cdot \varphi_{gn,t-1}}{\sum_j \text{Empl.}_{jc,t-1} \cdot \varphi_{gj,t-1}}$$

where $\varphi_{gn,t-1} \equiv \text{Empl.}_{gn,t-1}/\text{Empl.}_{n,t-1}$ is the national proportion of workers in industry n at date $t-1$ that are of type $g \in \{f, m\}$.

Unlike $\Delta \bar{w}^*_c$, this instrument is not mechanically correlated with ζ_c. In fact, the instrument predicts the change in a location's gender pay gap without using changes or gender differences in *any* location characteristic. A sufficient condition for Z^{Index}_{ct} to provide consistent estimates is that cross-location differences in lagged industrial composition are uncorrelated with ζ_{ct}–a condition emphasized in Goldsmith-Pinkham et al. (2020) and Beaudry et al. (2012, 2018).

Note that Z^{Index}_{ct} is constructed by combining *national*-level changes industrial wages weighted by the gender difference in location-level exposure. The location-level exposure is given by information on the $t-1$ size of the industry at the local level and *national*-level proportion of each gender in an industry. Thus, *all* of the cross-sectional variation in the instrument comes from differences across commuting zones in lagged industrial structure. Given that all of our specifications contain industry-by-year fixed effects, the identifying variation we are using is across location, within-industry variation in gender gaps. The implication is that instrument validity concerns the *cross-commuting zone* correlation between our instrument and the error term in (15).

The credibility of our index instrument can be examined by unpacking it in various ways. First, we can unpack along the gender dimension by writing $Z^{\text{Index}}_{ct} = Z^{\text{Index}:f}_{ct} - Z^{\text{Index}:m}_{ct}$ where

$$Z^{\text{Index}:g}_{ct} \equiv \sum_j \hat{\pi}_{gjc,t-1} \cdot \Delta d_{jt} \tag{13}$$

for $g \in \{f, m\}$. This allows us to examine the model's implication that $Z_{ct}^{\text{Index}:f}$ and $Z_{ct}^{\text{Index}:m}$ have equal but opposite effects on $\Delta\bar{w}_c^*$. Including both instruments also allows for an over-identification test.

Second, we can unpack the index instrument along the industry dimension. For each date t, we partition the set of industries into three groups, $\{\mathcal{N}_{Lt}, \mathcal{N}_{Mt}, \mathcal{N}_{Ht}\}$. Here $\mathcal{N}_{Lt}$ are the 'wage-decline' industries, containing the third of industries with the lowest change in industry wages over the preceding period, $\mathcal{N}_{Ht}$ are the 'wage-growth' industries, containing the third of industries with the highest change in industry wages over the preceding period, and $\mathcal{N}_{Mt}$ contains the remaining middle third of industries. For each $k \in \{L, M, H\}$ we define our 'exposure' instrument as:

$$Z_{k:ct}^{\text{Exposure}} \equiv \eta_{kt} \cdot \sum_{j \in \mathcal{N}_{kt}} \hat{\pi}_{jc,t-1}^*, \tag{14}$$

where $\eta_{kt} > 0$ is a location-independent scaling term.[12] That is, $Z_{H:ct}^{\text{Exposure}}$ tells us the location's initial gender difference in exposure to wage-growth industries, whereas $Z_{L:ct}^{\text{Exposure}}$ tells us the location's initial gender difference in exposure to wage-decline industries.

The exposure instruments offer several unique advantages relative to the standard Bartik-type instrument. First, the exposure instruments allow for a transparent evaluation of the exclusion restriction since they help disentangle the contribution of location and industry characteristics. Second, they also allow us to verify that $Z_{H:ct}^{\text{Exposure}}$ and $Z_{L:ct}^{\text{Exposure}}$ have opposing effects on $\Delta\bar{w}_{ct}^*$. Third, this set of instruments provides a standard over-identification test that tests whether differential gender exposure to wage-decline or wage-growth industries has similar impacts on within-industry gender pay gaps. Finally, it allows us to emphasize the spillover mechanism by examining the effect of outside wages on the subset of workers that do not belong to those industries that are used to construct the instrument.

[12] The scaling term does not affect identification but allows for easier comparisons of first-stage coefficients across k and t. In practice, we take η_{kt} to be the absolute value of the slope coefficient from a regression of $Z_{k:ct}^{\text{Index}} \equiv \sum_{j \in \mathcal{N}_{kt}} \hat{\pi}_{jct}^* \cdot \Delta d_{jt}$ on $z_{k:ct}^{\text{Exposure}} \equiv \sum_{j \in \mathcal{N}_{kt}} \hat{\pi}_{jct}^*$ run separately by k and t.

4 Data and Empirical Implementation

4.1 Data

Our main analysis uses data from the U.S. Census Public Use Micro-Samples (PUMS) for 1980, 1990 and 2000. For 2010 we aggregate the 2009-2010-2011 American Community Surveys. We focus on individuals between the ages of 22 and 54 with at least one year of potential labor market experience. Our location variable is commuting zone (CZ) with consistent geographic definitions from Dorn (2009). We use an industry coding that is consistent across Censuses and based on an aggregation of 1990 industry definitions into 45 industrial groups.

Our measure of w_{gnct} is constructed from the coefficients on a complete set of gender×industry×CZ×year fixed effects in individual wage regressions that also flexibly control for worker characteristics.[13] Employment shares, π_{gnct}, are constructed by using the observed distribution of hours worked across industries. Further details are provided in Appendix B.

4.2 Estimation

Our main estimating equation, based on (11), is

$$\Delta w^*_{nct} = \zeta_{nt} + \delta \cdot \Delta \bar{w}^*_{ct} + X'_{c,t-1}\alpha + \nu_{nct} \tag{15}$$

where ζ_{nt} are industry×time fixed effects, $X_{c,t-1}$ are controls, and ν_{nct} is an error term analogous to ε_{nct}. We cluster standard errors at the State level.

Our instruments, described in section 3.3, require a national-level measure of wage growth by industry (Δd_{jt}). To get these we regress w_{gnct} on a set of gender×CZ×year and industry×year fixed effects, and then take the time

[13]These regressions include controls for a quartic in potential experience; hispanic, black, and immigration dummies; an indicator for whether an individual is observed in a CZ located in their birth state–all interacted with education (three categories in our non-college sample)–and four occupation dummies. All covariates have coefficients that vary by gender. This procedure is repeated separately for each census year.

difference in the latter. In constructing the exposure instruments, we rank industries by Δd_{jt} for each t and group them by terciles.[14]

Following Goldsmith-Pinkham et al. (2020) and Borusyak et al. (2020), our set of controls X_{ct-1} include CZ characteristics that are potentially correlated with the instruments (lagged industrial structure in particular) and the error term (changes in within-industry gender gaps in internal conditions). Our baseline controls include log commuting zone size, the unemployment rate of each gender, the fraction of foreign born, because these systematically vary with industrial structure (e.g. Lewis (2011)) and are potentially correlated with changes in discriminatory attitudes or labor force participation. Our baseline controls also include a measure of aggregate education since this predicts the adoption of technologies that raise the relative productivity of women (Beaudry and Lewis (2014)).[15]

5 Results

5.1 Main Results

Table 1 presents the results from our main specification (15). The first two columns present OLS estimates of δ for completeness. The remaining columns present 2SLS estimates, along with first-stage estimates and diagnostics in the lower panel. A brief inspection reveals that, as anticipated, instrumenting

[14]The set of industries in each tercile can vary by decade. For instance, mining is in $\mathcal{N}_{Ht}$ in the 2000s but was in $\mathcal{N}_{Lt}$ in the 1980s. Some industries, such as Transportation and Machinery are in $\mathcal{N}_{Lt}$ in all three decades and others, such as Health Services, are in $\mathcal{N}_{Ht}$ in all three decades.

[15]Beaudry and Lewis (2014) argue that cities that adopted technology (e.g. the PC) relatively extensively during the 1980s and 1990s experienced a narrowing of their adjusted gender pay gap (as such technologies are argued to benefit brains over brawn; Weinberg (2000)). They show how PC adoption was more extensive in cities that were more educated in 1980. We follow Beaudry and Lewis (2014) in measuring city education as the log ratio of college to high school equivalents in 1980. Beaudry and Lewis (2014) follow Card (2009) in treating those with less than high school as contributing 70% of a high school worker, and those with some post-secondary are treated as contributing 60% of a high school worker and 40% of a college worker. Our results are not sensitive to this particular measure–alternatives such as BA share give very similar results.

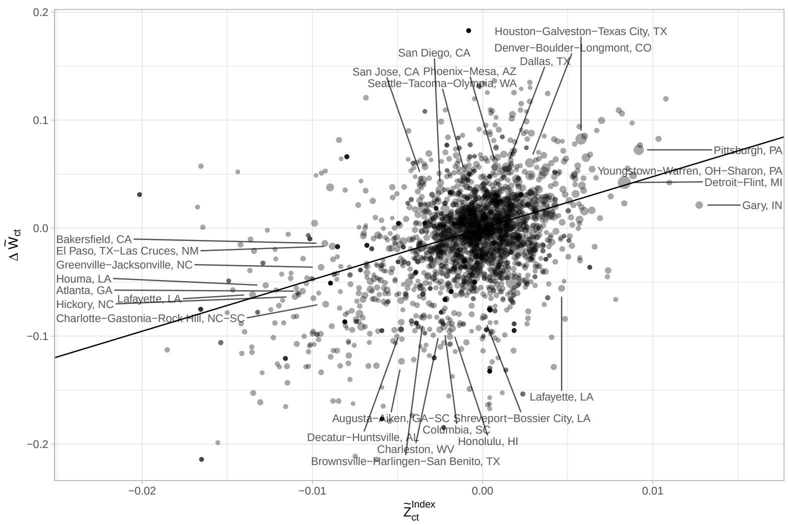
corrects for the mechanical upward bias in OLS: the 2SLS estimates are around 30-50% smaller.

Table 1: Main Results

	OLS		Index			Exposure		Decline	Growth
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\Delta\bar{w}^*_{ct}$	0.75**	0.75**	0.49**	0.44**	0.45**	0.47**	0.38**	0.34*	0.43**
	(0.030)	(0.029)	(0.084)	(0.093)	(0.093)	(0.10)	(0.13)	(0.18)	(0.20)
Obs.	27629	27629	27629	27629	27629	27629	27629	17716	17485
R^2	0.534	0.534	0.047	0.046	0.047	0.046	0.043	0.041	0.041
Fixed Effects:									
Ind.× Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls									
Baseline		Yes		Yes	Yes		Yes	Yes	Yes
First-Stage:									
Z^{Index}_{ct}			4.77**	4.29**					
			(0.64)	(0.65)					
$Z^{\text{Index}}_{ct,\text{female}}$					3.61**				
					(0.84)				
$Z^{\text{Index}}_{ct,\text{male}}$					-4.24**				
					(0.66)				
$Z^{\text{Exposure}}_{H:ct}$						4.67**	3.60**	4.60**	
						(0.93)	(0.98)	(1.04)	
$Z^{\text{Exposure}}_{L:ct}$						-3.08**	-3.37**		-4.73**
						(1.51)	(1.28)		(1.37)
F-Stat.			56.37	43.85	27.93	15.61	9.76	19.67	11.95
p-val			0.00	0.00	0.00	0.00	0.00	0.00	0.00
Over-id. p-val					0.37	0.14	0.45		

Notes: This table displays results from the estimation of equation (15) via OLS (columns 1 and 2) and 2SLS (columns 3 - 8) using US Census and ACS from 1980-2010. Standard errors, in parentheses, are clustered at the state level. (*) and (**) denote significance at the 10% and 5% level, respectively. The dependent variable is the decadal change in in the CZ-industry gender gap. Baseline controls include the start of the period log CZ size, unemployment rate of each gender, fraction of foreign born, and the ratio college to non-college workers. Regressions weighted by the start of period size of the industry-CZ. First-stage coefficients and diagnostics are reported in the bottom panel.

Columns (3)-(5) present 2SLS results using our index instrument. Column (4) adds the baseline controls to column (3), and column (5) uses the gender-specific version of the index instrument. The point estimates are similar, with $\hat{\delta} \in [0.44, 0.49]$, all statistically significant at the 1% level. The first-stage



Notes: The figure uses data from the US Census and ACS. The y-axis denotes the $\Delta\bar{w}^*_{ct}$ after adjusting for year effects, the x-axis denotes Z^{Index}_{ct} after adjusting for year affects. Each point is a commuting zone-year observation.The marker size is proportional to the size of the commuting zone in 1980. Slope of the regression line is given if the bottom panel of Table 1 column (3).

Figure 5: Two-stage least squares: Residualized visualization of first-stage

results in the bottom panel show that the instrument has strong predictive power. To explore this deeper, we focus on column (3) and display a residualized (for year effects) first-stage in Figure 5. We see that several of the CZs with a high value for Z^{Index}_{ct} during our sample period, suggesting they saw outside wages shift in favor of women, are rust-belt commuting zones. On the other hand, low values of Z^{Index}_{ct} predict a relative deterioration in outside wages for women. Several of these CZs have industry in oil, natural gas and agriculture. Furthermore, column (5) shows that the gender-specific instruments are statistically significant, are of the predicted sign, and have similar magnitudes. This specification also allows us to formally test whether improvements in women's or men's outside wages have an equal impact on the changes in within-industry gender gaps. The over-id test (Hansen's-J p-value) is reported in the last row, and fails to reject at standard significance levels.

Columns (6) to (9) of Table 1 use our exposure instruments. Column (7) adds baseline controls to column (6), whereas columns (8) and (9) use a specification intended to emphasize spillovers. In column (8) we use an instrument based on differential gender exposure to 'wage-growth' industries ($Z_{H:ct}^{\text{Exposure}}$) to examine the impact on within-industry gender wage gaps in all other industries. Column (9) does the same for 'wage-decline' industries. These estimates of δ are similar to, but slightly lower than, previous columns. The estimates are statistically significant at the 1% level in columns (6)-(7), at the 10% level in column (8), and at the 5% level in column (9).

In terms of first-stage results, each instrument takes on the expected sign and is of similar magnitude. The over-id tests in columns (6) and (7) suggest that gender gaps in exposure to 'wage-decline' and 'wage-growth' industries have similar impacts (in absolute value) on within-industry gender gaps. That is, within-industry wage gaps are predicted to decline in CZs with excess male exposure to declining-wage industries or with excess female exposure to increasing-wage industries.

We take $\hat{\delta} = 0.4$ to be a conservative midpoint of these results, but even this indicates a large role for outside wages. As discussed in section 3.2, this estimate suggests standard decompositions will fail to predict around 40% of the effect of changes to internal conditions or employment distributions. In other words, the magnitude of our indirect effect will be around 67% ($0.4/(1-0.4)$) of the magnitude of any direct effect.

5.2 Robustness and Further Results

5.2.1 Additional Controls

In Table 2 we probe the robustness of our main results to additional sets of controls. Columns (1)-(4) use the index instrument and columns (5) to (9) use the exposure instruments; columns (1) and (5) reproduce our estimates with baseline controls from Table 1.

The first set of controls capture the labour market participation of women in the CZ. The goal is to address concerns that our results are due to CZ-specific changes in labor force participation that drive changes in within-industry gender pay gaps (the error term) and systematically vary with lagged

Table 2: Robustness: Additional Controls

	Index				Exposure				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\Delta\bar{w}^*_{ct}$	0.44**	0.46**	0.42**	0.41**	0.38**	0.42**	0.35**	0.34*	0.40**
	(0.093)	(0.088)	(0.089)	(0.14)	(0.13)	(0.12)	(0.13)	(0.18)	(0.15)
Obs.	27629	27629	27629	27629	27629	27629	27629	27629	27629
R^2	0.046	0.048	0.046	0.038	0.043	0.046	0.042	0.035	0.045
Fixed Effects:									
Ind.× Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City				Yes				Yes	
Controls									
Baseline	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Female Empl.		Yes	Yes	Yes		Yes	Yes	Yes	Yes
Tasks			Yes	Yes			Yes	Yes	Yes
Exposure									Yes
First-Stage:									
Z^{Index}_{ct}	4.29**	4.23**	4.07**	4.29**					
	(0.65)	(0.54)	(0.55)	(0.54)					
$Z^{\text{Exposure}}_{H:ct}$					3.60**	3.72**	3.42**	4.17**	3.47**
					(0.98)	(0.86)	(0.88)	(0.95)	(0.88)
$Z^{\text{Exposure}}_{L:ct}$					-3.37**	-3.07**	-3.25**	-3.78**	2.25
					(1.28)	(1.16)	(1.16)	(1.36)	(1.73)
F-Stat.	43.85	61.82	55.44	64.14	9.76	14.70	12.77	14.24	7.73
p-val	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Over-id. p-val					0.45	0.40	0.43	0.31	0.03

Notes: This table displays results from the estimation of equation (15) via 2SLS using US Census and ACS from 1980-2010. Standard errors, in parentheses, are clustered at the state level. (*) and (**) denote significance at the 10% and 5% level, respectively. The dependent variable is the decadal change in in the CZ-industry gender gap. First-stage coefficients and diagnostics are reported in the bottom panel.

industrial structure (the instruments' identifying variation). Specifically, we control for a cubic in CZ's lagged female participation rate and the lagged proportion of the CZ labor force that is female.

The second set of controls capture employment vulnerability in the CZ. Specifically, we control for lagged share of manufacturing employment, share of employment in routine occupations, and the employment-weighted mean of an occupation 'offshoreability index'.[16] The goal is to address concerns that a CZ's lagged industrial structure (the instruments' identifying variation) is systematically related to the CZ's gender difference in employment vulnerability, which in turn is responsible for changes in within-industry gender gaps (the error term) because, for instance, displaced workers are less productive in other sectors.

The third set of controls are CZ fixed effects. These control for all time-invariant CZ characteristics that affect the change in within-industry gender wage gaps. Equivalently, they allow for CZ-specific time trends in within-industry gender gaps; e.g. due to CZ-specific trends in discrimination, preferences for work, gender-biased technology adoption, or CZ-specific exposure to secular changes such as increased globalization.

These three groups of additional controls are added in columns (2)-(4) for the index instrument and in columns (6)-(9) for the exposure instrument. None of these controls has an appreciable effect on the coefficient estimate, its statistical significance, or the first-stage diagnostics.

In column (8) we control for the *total* local employment share of 'wage-decline' and 'wage-growth' industries. Thus, our exposure instruments identify

[16]The latter two variables are constructed as in Autor et al. (2013). In particular, routine-intensive occupations are a set of occupations that perform tasks that are vulnerable to computerization (Autor et al., 2003). Routine occupations are constructed such that they account for one-third of US employment in 1980, and our control variable captures the fraction of employment in these occupations at the commuting zone level. The offshorability index measures the average degree to which the occupations in a commuting zone are susceptible to offshoring, proxied by the extent to which occupations do not require face-to-face contact with workers in the US. Our measure is constructed as in Autor et al. (2013), who standardized offshorability index to a mean of zero and standard deviation of 10 in 1980. We then take the employment-weighted mean of the index across occupations within a commuting zone.

only off of across commuting zone variation in gender employment *differences* within $\{\mathcal{N}_{Lt}, \mathcal{N}_{Ht}\}$-industry groups, holding the total size (exposure) constant. While this weakens the first-stage somewhat, the second-stage point estimate is not significantly different from other point estimates in Table 2.

5.2.2 Understanding Identifying Variation: Rotemberg weights

As suggested in Goldsmith-Pinkham et al. (2020), we now turn to computing the "Rotemberg" weights associated with our Index instrument, Z_{ct}^{Index}. These weights indicate the influence of a particular industry on the 2SLS estimation,[17] and allow us to assess whether we are using the variation stressed by our model. In particular, our model effectively assigns to each industries a 'treatment intensity' of $\hat{\pi}^*_{j,t-1} \cdot \Delta d_{jt}$. Whilst we explain our explanation of the declining gender wage gap in terms of industries that are male-exposed and wage-decline, the mechanism does not apply exclusively to such industries. Any industry with extreme values of treatment intensity are expected to be influential.

Table D.1 of Appendix D.1 presents the Rotemberg weight for each industry, aggregated across time periods. As is common in such applications, the distribution of the Rotemberg weights is highly skewed, with the top five industries accounting for 58 percent of the positive weight in our estimates. These industries include: motor vehicles and equipment, apparel and textile products, mining, construction, and primary metals. Two of these are male-exposed and wage-decline (motor vehicles and primary metals), two are male-exposed and wage-growth (mining and construction), and the remaining one is female-exposed and wage-decline (apparel). Only slightly lower on the list is health services, which is female-exposed and wage-growth. Table D.1 also documents that the identifying variation comes mainly from the 1980s (about two-thirds) and the 2000s (about one-third), while less than one percent comes

[17]Bartik-style instruments can be decomposed into weighted combinations of just-identified estimates, each using a single baseline exposure as an instrument. The 2SLS estimates derived from a Bartik-style instrument are a weighted average of the just-identified estimates. The individual weights, known as "Rotemberg weights," capture the relative importance of each industry's exposure toward the overall identifying variation.

from the 1990s. Interestingly, this corresponds to a period of both stagnant gender wage and exposure gaps, as documented in Figure 3.

Figure 6 documents the properties of influential industries more broadly. Panel (a) shows how industries with more extreme gender exposure gaps tend to be more influential; e.g. 'motor vehicles' and 'mining' both have excess male exposure, whereas 'apparel' and 'health services' both have excess female exposure. Panel (b) shows that industries with more extreme wage shocks tend to be more influential; e.g. 'motor vehicles' and 'apparel' both saw large declines in wages, whereas 'mining' and 'health services' saw large increases in wages. Panel (c) combines these dimensions and shows that industries with extreme values of treatment intensity tend to be more influential; e.g. 'motor vehicles' and 'health services' both have strong positive intensities (suggesting they act to narrow the wage gap), whereas 'mining' and 'apparel' both have strong negative intensities (suggesting they act to expand the wage gap). Finally, panel (d) shows that influential industries also tend to have larger cross-location variation in treatment intensity. To summarize, while many male-exposed industries saw wage declines, our identification comes from a broader range of industries with characteristics that our framework emphasizes.

After identifying the industries that play a larger role in identification, we probe the robustness of our estimates to alternative constructions of our shift-share instrument, Z_{ct}^{Index}. First, we assess the sensitivity of our estimates to high-Rotemberg weight industries by removing the Top-5 weight industries from our index. Second, Goldsmith-Pinkham et al. (2020) recommend fixing the shares in the shift-share instruments to some base year. To assess the sensitivity of our results to fixed-shares, we construct an analogous instrument fixing the CZ industrial structure to their 1980 levels.[18] Table D.4 of Appendix D.2 contains the results of this robustness exercise. Our main estimates are robust to either the fixed or time-varying share construction of the shift-share instrument, and excluding very influential industries from Z_{ct}^{Index}.

[18] A summary of the Rotemberg weights for our fixed-shares instrument can be found in Table D.2 of Appendix D.1 The Top-5 Rotemberg weight industries using fixed-shares coincide closely to those of our baseline instrument; they include: motor vehicles and equipment, apparel and textile products, mining, Justice, and primary metals.

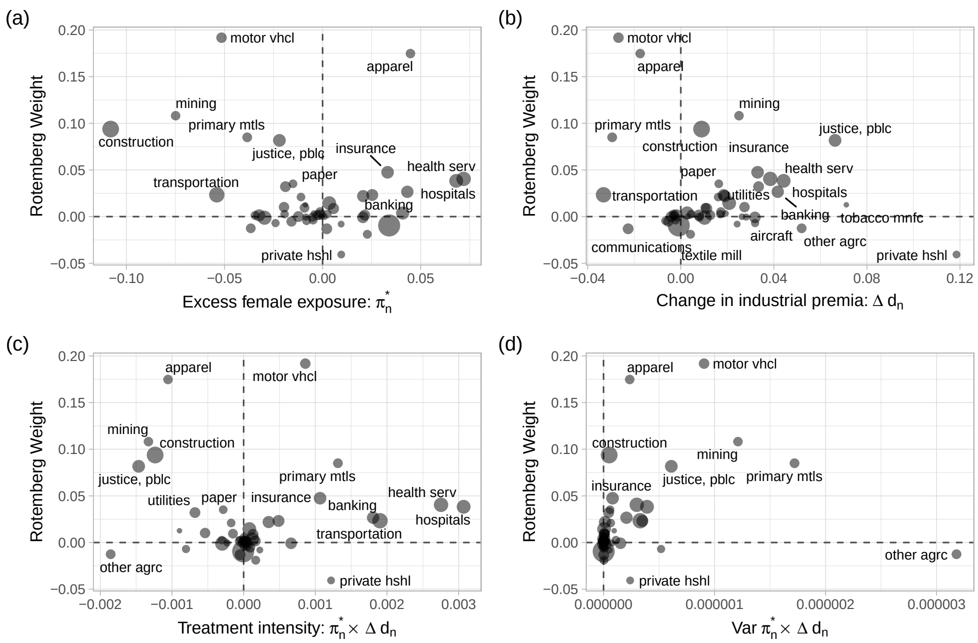


Figure 6: Characteristics of Influential Industries

Notes: The figure uses data from the US Census and ACS. The y-axis denotes the the Rotemberg weights associated with each industry. Each Panel relates the Rotemberg weight to a different aspect of our instrument. Each point is a industry observation aggregated over time. The marker size is proportional to the size of the industry.

5.2.3 Effects by Gender

Our analysis so far has assumed that men and women are equally sensitive to outside wages. This seems counter to a great deal of research which suggests that women negotiate lower wages than men (e.g. Card et al. (2016), Biasi and Sarsons (2021a,b), Roussille (2022)). However, it is not clear whether these findings represent a gender difference in the willingness (or ability) to negotiate or, instead, a gender difference in the quality of outside wages.[19] We now speak to this by estimating δ separately by gender.

We are able to estimate δ by gender using the same variation as our main analysis by estimating a version of (15) without gender differencing applied:

$$\Delta w_{gnct} = \zeta_{gnt} + \zeta_{nct} + \delta_g \cdot \Delta \bar{w}_{gct} + X'_{c,t-1}\alpha + \nu_{gnct}, \tag{16}$$

where ζ_{gnt} are gender$\times$industry$\times$year fixed effects and ζ_{nct} are industry$\times$CZ$\times$year fixed effects. We instrument $\bar{w}_{gct}$ with $Z_{ct}^{\text{Index:}g}$ from equation (13).

In exploring the gender differences in sensitivity to outside wages, we also address selection in two other ways. First, we also examine using only full-time full-year workers.[20] Second, we apply a selection correction based on Mulligan and Rubinstein (2008).[21]

The results are presented in Table 3. We find no evidence of gender differences in the sensitivity to outside wages, suggesting that, on average, it is

[19] Indeed, it could be that an individual's lower willingness to negotiate arises *because* weaker outside wages reduce the expected benefit from the process.

[20] Recall that our main specification includes both full- and part-time workers in our sample and weights observations by the number of hours worked. This approach increases the sample size, but, by treating the wages of all workers equally irrespective of how many hours they supply, the wage distribution will not be representative of the total number of hours worked in the economy. This might be of particular concern when analysing gender differences in wages if women tend to work less than men. As a compromise, we weight each observation by the number of hours worked, which gives more importance to full-time workers.

[21] Mulligan and Rubinstein (2008) use a Heckman-two step selection correction estimator to estimate selection corrected pay gaps. While exclusion restrictions for this type of estimator are notoriously hard to come by, we follow Mulligan and Rubinstein (2008) in using the presence of young children at home along with family structure. Further details can be found in Appendix D.5.

differences in outside wages–not in willingness or ability to negotiate–that are responsible for the gender difference in negotiated wages. Columns (1) and (2) compare estimates across our sample and a full-time full-year sample, whereas Columns (3) and (4) do the same only with the selection correction applied. In all cases there is a negligible gender difference in the estimate of δ, with the magnitude and statistical significance of the estimates being very similar to our main results.

Table 3: Gender-Specific Estimates and Selection Correction

	No Selection Correction		Selection Corrected	
	(1) All	(2) FTFY	(3) All	(4) FTFY
$\Delta\bar{w}_{ct}^{\text{Men}}$	0.42**	0.46**	0.41**	0.43**
	(0.12)	(0.12)	(0.12)	(0.13)
$\Delta\bar{w}_{ct}^{\text{Women}}$	0.39**	0.47**	0.37**	0.43**
	(0.16)	(0.18)	(0.16)	(0.19)
Obs.	55258	41062	55258	41062
R^2	0.049	0.059	0.050	0.059
Fixed Effects:				
Gender $\times$ Ind.$\times$ Year	Yes	Yes	Yes	Yes
Ind. $\times$ CZ $\times$ Year	Yes	Yes	Yes	Yes
Controls				
Baseline	Yes	Yes	Yes	Yes
Test $\Delta\bar{w}_{ct}^{\text{Men}} = \Delta\bar{w}_{ct}^{\text{Women}}$				
p-val.	.642	.917	.473	.995
First-Stage:				
F-Stat. ($\Delta\bar{w}_{ct}^{\text{Men}}$)	14.4	11.6	16.1	13.0
F-Stat. ($\Delta\bar{w}_{ct}^{\text{Women}}$)	19.1	17.3	22.5	19.6

Notes: This table displays results from the estimation of equation (15) via 2SLS using US Census and ACS from 1980-2010. Standard errors, in parentheses, are clustered at the state level. (*) and (**) denote significance at the 10% and 5% level, respectively. The dependent variable is the decadal change in in the commuting zone-industry gender gap. Baseline controls include the start of the period log commuting zone size, unemployment rate of each gender, fraction of foreign born, and the ratio college to non-college workers. Regressions weighted by the start of period size of the CZ-industry. First-stage diagnostics are reported in the bottom panel.

5.2.4 Sub-sample Analysis

Our investigation has focused on workers without a college degree, but the proposed mechanism may apply differently to various sub-populations. For instance, wages could be more sensitive to outside wages for workers whose wages are more sensitive to current market conditions, such as lower-educated and younger workers (Schwandt and von Wachter, 2019; von Wachter, 2020), or for workers that are more mobile, such as singles.

Table D.6 of Appendix D.4 contains estimates of equation (15) for several different sub-populations. The results for high school or less are similar to our main estimates from Table 1. For those with at least a college education, the estimate of δ is small and not significantly different from zero, but is also imprecisely estimated and suffers from a weak instrument problem.[22] The estimate for all education groups combined is slightly lower than in our baseline sample of non-college workers. Finally, we find no substantial differences when comparing the estimates for sub-samples based on age or marital status.

5.2.5 Sensitivity to Labor Market and Industry Definitions

The empirical implementation of our theoretical framework required decisions on the handling of the data. In particular, our framework focuses on local labour markets, but these can be empirically defined in several ways. As a baseline, and following much of the literature after Autor and Dorn (2013), we use commuting zones to proxy for local labour markets. Likewise, our baseline industrial classification uses 45 industries, which we chose to balance the size of commuting zone-industry cells with industrial detail. In Appendix D, we show that our results are not sensitive to alternative definitions of local labor markets or more disaggregated industrial classifications.

[22] When focusing on this education group we lose a large number of observations and are left with just 8,234 commuting zone-industry cells. Additionally, the strength of first-stage falls; the F-statistic on the joint significance of each instrument is 3.36, indicating a weak instrument problem.

5.3 Implications for Time Series Evidence

In Section 3.2 we discussed how, given a value of δ, we can decompose the gender difference in wages at each date into a part that is explained by gender differences in employment distributions, E_t, and the remaining unexplained part, U_t, that itself is divided into a part that is now explained by differences in outside wages, $U_{\text{out},t}$, and the part that remains unexplained and is attributable to differences in internal conditions, $U_{\text{in},t}$. Figure 7 plots these three series using CPS data and setting $\delta = 0.4$.

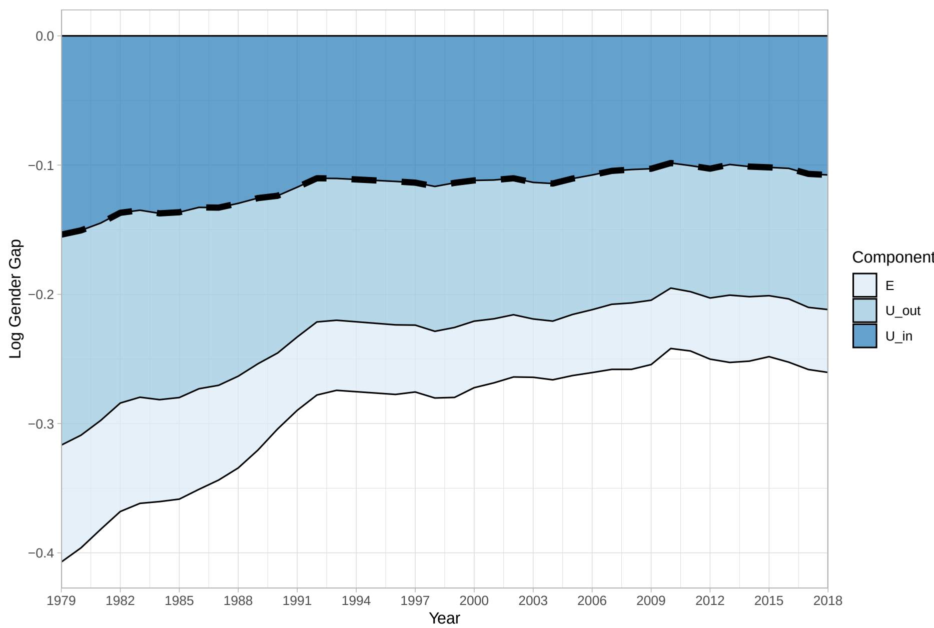


Figure 7: Gender Gap Decomposition

Notes: The figure uses data from Outgoing Rotation Group (ORG) Supplements of the Current Population Survey. Industry categories are aggregated into 45 industry groups (details are provided in Appendix C). Panel (a) reports the average within-industry gender gap as well as the decomposition into explained (gender differences in industrial employment) and unexplained (gender gaps within-industries) components. Panel (b) further decomposes the unexplained component based on equation (6) and an estimate of δ of 0.40.

There are two main points illustrated by the figure. First, the role of outside wages is substantial: $U_{\text{out},t}$ represents around 50% of the total unexplained component, U_t. Further, this proportion is relatively stable across time. The existing literature has offered a wide variety of explanations for within-industry gender wage differentials: taste-based discrimination, selection, technology, inflexible job conditions, personality traits, etc. These are all versions of gender differences in internal conditions and thus would be included in $U_{\text{in},t}$. Our results suggest that gender difference in outside wages is at least as relevant as the gender difference in all such internal factors combined. Second, the time pattern of $U_{\text{out},t}$ plays a large role in generating the well-known time pattern of the unexplained gap, U_t. Slightly more than half of the change in U_t is due to the change in $U_{\text{out},t}$.

6 Conclusions

Our broad goal in this paper was to explore whether gender differences in the quality of outside wages are relevant for understanding within-industry gender pay differences. Our results suggest that such factors are indeed very important–this 'indirect' effect of changes in internal conditions and employment distributions is estimated to be around 67% as large as the direct effect of such changes. Standard decompositions will fail to predict 40% of the total effect. Differences in outside wages explain one-half of the unexplained wage gap, and account for around one-half of the time pattern in the gender wage gap.

In one sense, our findings have a negative implication: the celebrated 'gains' of women during the 1980s may be more accurately characterized as the 'losses' of men, owing to declining male outside wages during this period. But, in another sense, the implication is very positive: the future impact of technological changes on the gender gap is likely to be substantially understated. For instance, Goldin (2014) argues that gender wage equity will require a change in the nature of work, specifically to accommodate greater temporal flexibility. Our analysis suggests that the impact of such a change is under-appreciated– as technologies alter the nature of work in ways that boost the representation and pay of women in some sectors, there will be the sort of direct effect highlighted by Goldin (2014). However, we can expect a further indirect effect, with around two-thirds of the magnitude, because of the associated shifts in outside wages.

This latter aspect highlights policy implications that go beyond equal pay legislation: we estimate that the within-industry pay gap would halve if gender differences in outside wages were eliminated. However, to achieve this there is clearly a need for future research aimed at explaining *why* some industries are more 'male-exposed' than others. Furthermore, the optimal policy response is sensitive to which of the underlying models is at play. For instance, attempts to ensure equal pay for equal work (internal conditions) will have less impact if outside wages matter because they shape incentives to work hard and more impact if they matter because they represent outside options in the bargaining process. Future work in this direction is therefore valuable. The framework could be extended in interesting ways, for instance, by considering systematic biases in beliefs about outside wages (Jäger et al. (2021)). The application to the gender wage gap is particularly valuable since gender pay gaps influence a range of other socially important outcomes such as marriage rates, labor force participation, and domestic violence (Bertrand et al. (2015), Aizer (2010)). Yet, we reiterate that our analysis can be easily applied to any other pay gap of interest, such as racial, native and skill pay gaps. We also hope this sort of analysis will prove useful in other countries and time periods.

References

Adão, R., M. Kolesár, and E. Morales (2019, 08). Shift-Share Designs: Theory and Inference. *The Quarterly Journal of Economics 134*(4), 1949–2010.

Aizer, A. (2010). The gender wage gap and domestic violence. *The American Economic Review 100*(4), 1847–1859.

Akerlof, G. A. and J. L. Yellen (1990, 05). The Fair Wage-Effort Hypothesis and Unemployment . *The Quarterly Journal of Economics 105*(2), 255–283.

Albrecht, J. and S. Vroman (2002, February). A Matching Model with Endogenous Skill Requirements. *International Economic Review 43*(1), 283–305.

Altonji, J. G. and R. M. Blank (1999, October). Race and gender in the labor market. In

O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics*, Volume 3 of *Handbook of Labor Economics*, Chapter 48, pp. 3143–3259. Elsevier.

Autor, D., F. Levy, and R. Murnane (2003, November). The skill content of recent technological change: An empirical exploration. *The Quarterly Journal of Economics 118*(4), 1279–1333.

Autor, D. H. and D. Dorn (2013, aug). The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market. *American Economic Review 103*(5), 1553–1597.

Autor, D. H., D. Dorn, and G. H. Hanson (2013, October). The china syndrome: Local labor market effects of import competition in the united states. *American Economic Review 103*(6), 2121–68.

Babcock, L., J. Engberg, and R. Greenbaum (2005). Wage spillovers in public sector contract negotiations: The importance of social comparisons. *Regional Science and Urban Economics 35*(4), 395–416.

Bailey, M. J., B. Hershbein, and A. R. Miller (2012, July). The opt-in revolution? contraception and the gender gap in wages. *American Economic Journal: Applied Economics 4*(3), 225–54.

Beaudry, P., D. A. Green, and B. Sand (2012). Does industrial composition matter for wages? a test of search and bargaining theory. *Econometrica 80*(3), 1063–1104.

Beaudry, P., D. A. Green, and B. M. Sand (2018, September). In search of labor demand. *American Economic Review 108*(9), 2714–57.

Beaudry, P. and E. Lewis (2014). Do male-female wage differentials reflect differences in the return to skill? cross-city evidence from 1980-2000. *American Economic Journal: Applied Economics 6*(2), 178–94.

Beeson, P., L. Shore-Sheppard, and K. Shaw (2001). Industrial change and wage inequality: Evidence from the steel industry. *Industrial and Labor Relations Review 54*(2A), 466–483.

Bertrand, M. (2011). Chapter 17 - new perspectives on gender. In D. Card and O. Ashenfelter (Eds.), *Handbook of Labor Economics*, Volume 4, Part B, pp. 1543 – 1590. Elsevier.

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

Biasi, B. and H. Sarsons (2021a, 08). Flexible Wages, Bargaining, and the Gender Gap*. *The Quarterly Journal of Economics 137*(1), 215–266.

Biasi, B. and H. Sarsons (2021b, May). Information, confidence, and the gender gap in bargaining. *AEA Papers and Proceedings 111*, 174–78.

Binder, A. J. and J. Bound (2019, May). The declining labor market prospects of less-educated men. *Journal of Economic Perspectives 33*(2), 163–90.

Black, D. A. (1995). Discrimination in an equilibrium search model. *Journal of Labor Economics 13*(2), 309–334.

Blau, F. D. and L. M. Kahn (1997, January). Swimming upstream: Trends in the gender wage differential in 1980s. *Journal of Labor Economics 15*(1), 1–42.

Blau, F. D. and L. M. Kahn (2000). Gender differences in pay. *Journal of Economic Perspectives 14*(4), 75–99.

Blau, F. D. and L. M. Kahn (2006, October). The u.s. gender pay gap in the 1990s: slowing convergence. *Industrial and Labor Relations Review 60*(1), 45–66.

Blau, F. D. and L. M. Kahn (2016, January). The gender wage gap: Extent, trends, and explanations. Working Paper 21913, National Bureau of Economic Research.

Borjas, G. J. and V. A. Ramey (1995, November). Foreign competition, market power, and wage inequality: Theory and practice. *Quarterly Journal of Economics*, 1075–110.

Borusyak, K., P. Hull, and X. Jaravel (2020). Quasi-experimental shift-share research designs.

Bowlus, A. J. (1997, October). A Search Interpretation of Male-Female Wage Differentials. *Journal of Labor Economics 15*(4), 625–57.

Bowlus, A. J. and L. Grogan (2009, April). Gender wage differentials, job search, and part-time employment in the UK. *Oxford Economic Papers 61*(2), 275–303.

Breza, E., S. Kaur, and Y. Shamdasani (2017, 10). The Morale Effects of Pay Inequality . *The Quarterly Journal of Economics 133*(2), 611–663.

Caldwell, S. and O. Danieli (2022). Outside Options in the Labor Market. Technical report.

Caldwell, S. and N. Harmon (2019). Outside Options, Bargaining, and Wages: Evidence from Coworker Networks. Technical report.

Card, D. (2009, May). Immigration and inequality. *American Economic Review 99*(2), 1–21.

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2016, November). Firms and labor market inequality: Evidence and some theory. Working Paper 22850, National Bureau of Economic Research.

Card, D., A. R. Cardoso, and P. Kline (2015, 10). Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women. *The Quarterly Journal of Economics 131*(2), 633–686.

Cortes, G. M., N. Jaimovich, and H. E. Siu (2016). The end of men and rise of women in the high-skilled labor market. Working paper, University of British Columbia.

Dorn, D. (2009, September). Essays on inequality, spatial interaction, and the demand for skills. Technical Report 3613, Dissertation University of St. Gallen.

Flabbi, L. (2010, 08). Gender Discrimination Estimation In A Search Model With Matching And Bargaining. *International Economic Review 51*(3), 745–783.

Flabbi, L. and A. Moro (2012). The effect of job flexibility on female labor market outcomes: Estimates from a search and bargaining model. *Journal of Econometrics 168*(1), 81 – 95. The Econometrics of Auctions and Games.

Fortin, N., T. Lemieux, and S. Firpo (2011). *Decomposition Methods in Economics*, Volume 4 of *Handbook of Labor Economics*, Chapter 1, pp. 1–102. Elsevier.

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

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020, August). Bartik instruments: What, when, why, and how. *American Economic Review 110*(8), 2586–2624.

Green, D. A. (2015, November). Chasing after good jobs. do they exist and does it matter if they do? *Canadian Journal of Economics 48*, 1215–1265.

Guryan, J. and K. K. Charles (2013, 11). Taste-Based or Statistical Discrimination: The Economics of Discrimination Returns to its Roots. *The Economic Journal 123*(572), F417–F432.

Jäger, S., C. Roth, N. Roussille, and B. Schoefer (2021, December). Worker beliefs about outside options. Working Paper 29623, National Bureau of Economic Research.

Le Barbanchon, T., R. Rathelot, and A. Roulet (2020, 10). Gender Differences in Job Search: Trading off Commute against Wage . *The Quarterly Journal of Economics 136*(1), 381–426.

Lewis, E. (2011, 05). Immigration, Skill Mix, and Capital Skill Complementarity . *The Quarterly Journal of Economics 126*(2), 1029–1069.

Lewis, H. G. (1963). *Unionism and relative wages in the United States: an empirical inquiry.* University of Chicago press.

Manski, C. F. (1993, July). Identification of endogenous social effects: The reflection problem. *Review of Economic Studies 60*(3), 531–42.

Moffitt, R. A. (2001). Policy interventions, low level equilibria, and social interactions. In S. N. Durlauf and H. P. Young (Eds.), *Social Dynamics*. MIT Press.

Mulligan, C. B. and Y. Rubinstein (2008). Selection, investment, and women's relative wages over time. *The Quarterly Journal of Economics 123*(3), 1061–1110.

Olivetti, C. and B. Petrongolo (2016, January). The Evolution of Gender Gaps in Industrialized Countries. IZA Discussion Papers 9659, Institute for the Study of Labor (IZA).

O'Neill, J. and S. Polachek (1993). Why the gender gap in wages narrowed in the 1980s. *Journal of Labor Economics 11*(1), 205–228.

Pissarides, C. (2000). *Equilibrium Unemployment Theory* (Second ed.). Cambridge, MA: MIT press.

Roussille, N. (2022, July). The central role of the ask gap in gender pay inequality. Working paper, MIT.

Schwandt, H. and T. von Wachter (2019). Unlucky cohorts: Estimating the long-term effects of entering the labor market in a recession in large cross-sectional data sets. *Journal of Labor Economics 37*(11926), S161–S198.

Shapiro, C. and J. E. Stiglitz (1984). Equilibrium unemployment as a worker discipline device. *The American Economic Review 74*(3), 433–444.

Sorkin, I. (2017, May). The role of firms in gender earnings inequality: Evidence from the united states. *American Economic Review 107*(5), 384–87.

von Wachter, T. (2020). The Persistent Effects of Initial Labor Market Conditions for Young Adults and Their Sources. *The Journal of Economic Perspectives 34*(4), 168–194.

Weinberg, B. A. (2000, January). Computer use and the demand for female workers. *Industrial and Labor Relations Review 53*(2), 290–308.

Welch, F. (2000, May). Growth in women's relative wages and in inequality among men: One phenomenon or two? *American Economic Review 90*(2), 444–449.

Pay Gaps and Outside Wages:

The U.S. Gender Wage Gap 1980-2010

– Appendix: **For Online Publication** –

Chris Bidner*and Ben Sand†

March 17, 2023

Contents

*Department of Economics, Simon Fraser University. cbidner@sfu.ca
†Department of Economics, York University. bmsand@yorku.ca

A Theory: Microfoundations

Dynamic frictional models generate wages that depend on conditions in the job (match-specific productivity, discrimination, etc.) as well as the value of worker outside wages. For instance outside wages shape surplus division in bargaining, or shape wages required to dissuade shirking.

A.1 Common Set Up

We consider an infinite horizon model in continuous time. There is a unit mass of workers. Each worker has a type $g \in \{f, m\}$, and $\rho_g > 0$ is the measure of type g workers. Workers produce output when matched with a firm. Each firm belongs to one of N sectors. An unemployed worker gets a flow payoff of b_g and encounters a vacancy at rate p. The vacancy belongs to sector n with probability

$$\xi_{gn} \equiv \frac{\lambda_{gn} \cdot \tilde{\rho}_n}{\sum_j \lambda_{gj} \cdot \tilde{\rho}_j} \tag{A.1}$$

where $\tilde{\rho}_n$ is the (endogenous) measure of sector n vacancies and $\lambda_{gn} \geq 0$ are weighting parameters that allow a generalization of matching beyond uniform random. Importantly, these weighting parameters allow us to capture the empirical fact that employment distributions vary greatly with gender.[1] There are many ways to interpret the weighting parameters; they may reflect pure gender discrimination in terms of employment barriers (as opposed to wage levels), they could reflect gender identities associated with particular jobs (Akerlof and Kranton (2000)), or they could arise from the search technology whereby information about job vacancies arrives via social networks (Bayer et al. (2008), Munshi (2003)), and males and females are embedded in different networks (Brashears (2008)).

Thus type g workers encounter vacancies in n at rate $p_{gn} \equiv p\xi_{gn}$ and vacancies in n encounter type g workers at rate $\tilde{p}_{gn} \equiv p\xi_{gn} \cdot (\mu_g/\tilde{\rho}_n)$ where μ_g is the measure of unemployed workers of type g.

A vacancy entails a flow cost of κ_n for the firm. A worker-firm pair separate for exogenous reasons at rate s. The future is discounted at rate r.

A.1.1 Values

Suppose that in equilibrium a match generates utility of u_{gn} for the worker and profit v_{gn} for the firm (these are endogenously determined below). Let (V_{gn}, U_g) denote the equilibrium values for a type g worker that is employed in sector n, and that is unemployed respectively. Let $(\tilde{V}_{gn}, \tilde{U}_n)$ denote the equilibrium values for a firm in sector n when employing a type g worker, and when holding a vacancy respectively. Then, assuming for simplicity that all encounters result in a match, given a vector of vacancy measures, $\{\tilde{\rho}_n\}$, the $(5N + 2)$ values

[1] Consistent with Sorkin (2017), differences in employment distributions will arise because of opportunities rather than preferences.

simultaneously satisfy the following $(5N+2)$ equations:

$$rV_{gn} = u_{gn} + s \cdot [U_g - V_{gn}] \tag{A.2}$$
$$rU_g = b_g + \sum_j p_{gj} \cdot [V_{gj} - U_g] \tag{A.3}$$
$$r\tilde{V}_{gn} = v_{gn} + s \cdot [\tilde{U}_n - \tilde{V}_{gn}] \tag{A.4}$$
$$r\tilde{U}_n = -\kappa_n + \sum_g \tilde{p}_{gn} \cdot [\tilde{V}_{gn} - \tilde{U}_n]. \tag{A.5}$$

The N vacancy measures, $\{\tilde{\rho}_n\}$, are determined by the N free-entry conditions:

$$\tilde{U}_n = 0. \tag{A.6}$$

A.1.2 Steady State

In the steady state, the measure of each type of worker in each sector is constant over time. Let π_{ij} be the steady state employment share so that equating inflows and outflows requires:

$$\mu_g \cdot p \cdot \xi_{gn} = s \cdot (\rho_g - \mu_g) \cdot \pi_{gn}. \tag{A.7}$$

The steady state condition alone delivers two results. First, dividing by ρ_g and summing over n gives the steady state unemployment rate:

$$\frac{\mu_g}{\rho_g} = \frac{s}{s+p}. \tag{A.8}$$

Substituting this back implies that, in the steady state, exposure equals employment shares:

$$\xi_{gn} = \pi_{gn}. \tag{A.9}$$

A.1.3 Conditional Values

Given utilities and profits, $\{u_{gn}, v_{gn}\}$, we can solve for the steady state values as follows.

$$rV_{gn} = \frac{r}{r+s} \cdot u_{gn} + \frac{s}{r+s} \cdot [\frac{r+s}{p+r+s} \cdot b_g + \frac{p}{p+r+s} \cdot \sum_j \pi_{gj} \cdot u_{gj}] \tag{A.10}$$
$$rU_g = \frac{r+s}{p+r+s} \cdot b_g + \frac{p}{p+r+s} \cdot \sum_j \pi_{gj} \cdot u_{gj} \tag{A.11}$$
$$r\tilde{V}_{gn} = \frac{r}{r+s} \cdot v_{gn} \tag{A.12}$$
$$r\tilde{U}_n = 0. \tag{A.13}$$

Firms are clearly willing to match with any worker they encounter. A worker is willing to match with any vacancy they encounter if the smallest value of V_{gn} is no less than U_g. That is, if

$$\min_n \{u_{gn} - b_g\} \geq \frac{p}{p+r+s} \cdot \sum_j \pi_{gj} \cdot [u_{gj} - b_g].$$

We now present two versions of how utilities and profits are determined.

A.2 Bargaining

In the spirit of Pissarides (2000), suppose that a worker of type g in sector n produces an output valued at y_{gn}. Worker flow utility is their wage and firm flow profits are output net of the wage:

$$u_{gn} = w_{gn} \tag{A.14}$$

$$v_{gn} = y_{gn} - w_{gn}. \tag{A.15}$$

The wage is determined by generalized Nash bargaining where ϕ is the relative bargaining power of workers:

$$(1-\phi)\cdot[V_{gn} - U_g] = \phi\cdot[\tilde{V}_{gn} - \tilde{U}_n]. \tag{A.16}$$

Proposition 1 *Under bargaining, equilibrium wages satisfy the 'social interactions' form, (2), where exposure coincides with employment shares*

$$\xi_{gn} = \pi_{gn}, \tag{A.17}$$

the industry effects are

$$\psi_{gn} = \phi\cdot y_{gn} + (1-\phi)\cdot\frac{r+s}{p+r+s}\cdot b_g, \tag{A.18}$$

and the parameter of interest is:

$$\delta = \frac{(1-\phi)\cdot p}{r+s+p}. \tag{A.19}$$

Proof. First, (A.17) comes straight from (A.9). Equations (A.18) and (A.19) are straightforward consequences of (A.10)-(A.16). □

A.3 Efficiency Wages

In the spirit of Shapiro and Stiglitz (1984), suppose that workers must exert costly and non-verifiable effort in order to produce output. Given an effort of e, a worker of type g in sector n produces an output of $y(\theta_{gn}, e)$ where y is differentiable with $y_\theta, y_e > 0$, $y_{\theta e} > 0$, and $y_{ee} \leq 0$. Given effort, e, and a wage, w_{gn}, worker flow utility is $w_{gn} - c(e)$ where $c_e > 0$ and $c_{ee} \geq 0$ (where $y_{ee} < c_{ee}$ so that the surplus-maximizing effort is well-defined). Firm flow profits are $y(\theta_{gn}, e) - w_{gn}$.

Employment is governed by relational contracts. Such a contract specifies a wage and effort, (w, e), and involves the worker being fired if they are caught shirking. A shirking worker is detected at a rate d. Each firm maximizes profits subject to a no-shirking incentive constraint:

$$V(w, e) \geq V^S(w) \tag{A.20}$$

where $V(w,e)$ is the value of working under (w,e) and $V^S(w)$ is the value of shirking. These satisfy

$$rV(w,e) = w - c(e) + s \cdot [U_g - V(w,e)] \tag{A.21}$$
$$rV^S(w) = w + (s+d) \cdot [U_g - V^S(w)]. \tag{A.22}$$

That is, incentive compatibility requires

$$w \geq w_g(e) \equiv \frac{1}{\varphi} \cdot c(e) + r \cdot U_g, \tag{A.23}$$

where

$$\varphi \equiv \frac{d}{r+s+d}. \tag{A.24}$$

The problem for the firm is equivalent to choosing (w,e) to maximize flow payoffs, $y(\theta_{gn}, e) - w$, subject to (A.23). The optimal effort, e_{gn}, satisfies the first-order condition:

$$y_e(\theta_{gn}, e_{gn}) = \frac{1}{\varphi} \cdot c_e(e_{gn}). \tag{A.25}$$

This implies that e_{gn} is increasing in θ_{gn}. The optimal wage is the one that makes the incentive constraint bind:

$$w_{gn} = \frac{1}{\varphi} \cdot c(e^*_{gn}) + r \cdot U_g. \tag{A.26}$$

Subtract w_{gn}/φ from both sides and re-arrange to get

$$u_{gn} \equiv w_{gn} - c(e_{gn}) = (1-\varphi) \cdot w_{gn} + \varphi \cdot r \cdot U_g. \tag{A.27}$$

Proposition 2 *Under efficiency wages, equilibrium wages satisfy the 'social interactions' form, (2), where exposure coincides with employment shares*

$$\xi_{gn} = \pi_{gn}, \tag{A.28}$$

the industry effects are

$$\psi_{gn} = \frac{1}{\varphi} \cdot c(e_{gn}) + (1-\delta) \cdot b_g, \tag{A.29}$$

where e_{gn} satisfies (A.25), and the parameter of interest is:

$$\delta = \frac{p}{r+s+d+p}. \tag{A.30}$$

Proof. First, (A.17) comes straight from (A.9). Equations (A.29) and (A.30) are straightforward consequences of (A.26), (A.27), and (A.11). In particular, use (A.27) in (A.11) to get

$$rU_g = \frac{r+s}{p+r+s} \cdot b_g + \frac{p}{p+r+s} \cdot \sum_j \pi_{gj} \cdot (1-\varphi) \cdot w_{gj} + \frac{p}{p+r+s} \cdot \varphi \cdot r \cdot U_g. \tag{A.31}$$

Solve this for rU_g and use in (A.26). □

The intuition here is that a firm has to pay a worker enough to incentivize them to exert the desired effort. The amount that must be paid is increasing in the desired effort (which itself is increasing in the productivity term θ_{gn}) and in the quality of outside options (which takes into account wages and efforts in other firms). That is, if firms in industry 1 experience a productivity boost then they optimally require their workers to exert more effort. This requires they pay them a higher wage. The higher wage covers the disutility of the extra effort but also an additional amount to account for the additional incentive to shirk. Thus, the utility of workers in industry 1 increases. Workers in industry 2 will now require a higher wage in order to dissuade shirking, owing to their improved outside option, and so on.

B Data

B.1 Census Data

The Census data was obtained with extractions done using the IPUMS system (see Ruggles et al. (2010). The files were the 1980 5% State (A Sample), 1990 State, 2000 5% Census PUMS, and the 2009-2010-2011 American Community Surveys. The initial extraction includes all individuals aged 22 - 54 not living in group quarters. All calculations are made using the sample weights provided. We focus on the log of hourly wages, calculated by dividing wage and salary income by annual hours worked (usual hours worked × annual weeks worked). We impute incomes for top coded values by multiplying the top code value in each year by 1.5. Since top codes vary by State in 1990 and 2000 Census and thte 2009-2010-2011 ACS, we impose common top-code values of $140,000$ in 1990, $175,000$ in 2000 and $200,000$ in the ACS.

A consistent measure of education is not available for these Census years. We use indicators based on the IPUMS recoded variable `EDUCREC` that computes comparable categories from the 1980 Census data on years of school completed and later Census years that report categorical schooling only. To calculate potential experience (age minus years of education minus six), we assign group mean years of education from Table 5 in Park (1994) to the categorical education values reported in the 1990 and 2000 Censuses and the ACS.

Local labour markets defined by commuting zones are based on Dorn (2009) and Autor and Dorn (2013) and constructed using the code/data generously provided by David Dorn on his website. As a robustness check, we also use 'cities' defined either by CMSAs as described in Doms and Lewis (2006) or SMSAs as described in Beaudry et al. (2012). Code for both definitions was generously provided by Ethan G. Lewis. We further restrict our analysis to the 100 largest cities based on population estimates from 1980.

We use an industry coding that is consistent across Censuses and is based on the IPUMS recoded variable `IND1990`, which recodes census industry codes to the 1990 definitions. We aggregate this variable into 45 detailed industry groups based on standard BLS definitions

We also use an industry coding that is consistent across Censuses and is based on the IPUMS recoded variable `IND1950`, which recodes census industry codes to the 1950 definitions. This generates 144 consistent industries.[2] In addition, we use the industry crosswalks provided by by David Dorn on his website to produce the variable `ind1990dd` which are based on the 1990 Census industry definitions.

[2]See http://usa.ipums.org/usa-action/variableDescription.do?mnemonic=IND1950 for details.

For all analysis using wage data, we further restrict the sample to those (1) currently employed at the time of the census, (2) with positive wage and salary income. In some specifications, we restrict attention to those who are full-time, full year worker, defined as a worker who usually works 35 hours per week and worked at least 40 weeks in the year prior to the Census. The ACS does not contain a continuous measure of weeks worked in the year prior to the survey after 2008. Instead, an interval of weeks worked is provided. We convert this into a continuous measure by taking the mid-point in each interval.

B.2 Merged Outgoing Rotation Group Current Population Survey

MORG CPS data from 1979-2018 are downloaded from the NBER[3] Our initial extractions included all individuals between the ages of 22-54. Prior to 1992, education was reported as the number of completed years. In 1992 and after, education is reported in categories as the highest grade/degree completed. We convert categories to years of completed school in the post-1991 data based on Park (1994). The construction of our wage data closely follows Lemieux (2006). Wage data is based on those who report employment in reference week. In all wage calculations, we set allocated wages to missing. Our hourly wage measure is based on reported hourly wage for those who report hourly payment and not adjusted for topcoding. For workers who are not paid hourly:

1. We use edited weekly earnings. For the years 1984-1986, we use unedited earnings due to the higher topcode value.

2. Adjust topcoded wages by a factor of 1.4.

3. Divide the result by usual hours worked per week.

C Further Details on Construction of Motivating Figures

Motivating Figures 1, 2 and 3 are based on the MORG CPS 1979-2018 sample described in Appendix B.2. Figure 2 requires the calculation of industry shares by gender and industrial premia. For the former, we compute the fraction of hours worked in each by gender and year, and smooth the shares with a 2 year rolling average. We denote the fraction of hours worked in an industry-year by π_{gn} for ($g \in \{\underline{\text{f}}\text{emale}, \underline{\text{m}}\text{ale}\}$) and the 'femaleness' of an industry by their difference. For the latter, we estimate a log wage regression year-by-year, using provided sampling weights multiplied by usual hours worked:

$$\ln W_i = X_i'\beta + d_n + \alpha \cdot \text{female}_i + \epsilon_i, \tag{C.32}$$

where d_n are industry fixed effects (45 categories), female_i is a female indicator, and X_i contains an education indicator (5 categories), age group dummies (6 categories), their interaction, dummies for race (White, black, other), and occupation indicators (4 categories). The estimated coefficients on the industry fixed effects are used as relative industrial wages.[4]

[3] Links are http://www.nber.org/data/cps_may.html and http://www.nber.org/data/morg.html

[4] Figure 2 doesn't require annual data and could be constructed with the Census data we use in our main analysis. The results are remarkably similar and are available from the authors upon request.

Figure 3 plots the time series of the within-industry wage gap and the exposure gap. The within-industry gender wage gap is the estimate of α that comes from the above regression. The exposure gap is $\sum_j(\pi_{fj} - \pi_{mj}) \cdot d_j$, where π_{gn} are calculated as described above.

Figure 4 is based on US Census and ACS data. It is constructed by first estimating equation (15) by year to obtain ω_{gnc}, regression adjusted gender×CZ×industry wages. Using these, we calculate within-industry gender gaps at the CZ-year level: $\omega^*_{nct} \equiv \omega_{fnct} - \omega_{mnct}$ for ($g \in \{\underline{f}\text{emale}, \underline{m}\text{ale}\}$). Next, we extract the CZ level within-industry gender gap by estimating:

$$\omega^*_{nc} = D_{nt} + D_{ct} + \epsilon^*_{nct} \tag{C.33}$$

using the size of the nct-cell as weights. The coefficients on the CZ dummies (D_{ct}) are the regression adjusted, within-industry, commuting-zone level gender gaps. CZ level exposure differences are calculated as in section 4.

D Further Robustness

D.1 Rotemberg Summary

Tables D.1 and D.2 present a summary of Rotemberg weights in the style of Goldsmith-Pinkham et al. (2020) produced with code provided by those authors and adapted to our data structure. Panel A of D.1 reports the share and sum of negative weights, which have a nearly 90 percent positive share. Panel B reports the correlation of the Rotemberg weights (α_n), national level shock (Δ_n), just identified estimates β_n, the just identified first-stage F-statistic (F_n) and the variation of share-gaps across locations $\text{Var}(\pi^*_n)$. Panel C reports variation in the weights across years, showing that nearly all of the variation comes from the 1980s and 2000s. Panel D reports the top ten industries according to the Rotemberg weights. The Δd_n is the national industry premia growth rate, β_n is the coefficient from the just-identified regression. Panel E reports statistics about average estiamtes by the positive and negative Rotemberg weights. Table D.2 replicates this exercise with the fixed-share version of the instrument, discussed in the main text.

Table D.3 reports the correlation between commuting zone level variables and the share-gaps (π^*_{nt}) for the industries with the top-5 highest Rotemberg weights. The table is based on Table A2 of Goldsmith-Pinkham et al. (2020). Each column is a separate regression, pooling three cross sections. The regression is of $t + 1$ characteristics on t share-gaps π^*_{nt}. Each regression also includes a separate intercept for each decade. Standard errors are clustered at the state level. The R^2 of each regression, reported in the final row, show that there is substantial correlation between share-gaps and demographics. In particular, commuting zone size and the fraction of college graduates.

D.2 Instrument Sensitivity

D.2.1 Sensitivity to high-Rotemberg weight industries

To assess the sensitivity of our estimates to high-Rotemberg weight industries, we perform several specification checks in Table D.4. Columns (1)-(4) display two-stage least-squares estimates with different forms of our baseline instrument, Z^{Index}_{ct}. In column (1), we simply remove the Top-5 Rotemberg industries from our index. In column (2), we use only the Top-5 Rotemberg weight industries. In column (3), we use both of these instruments, entered

Table D.1: Summary of Rotemberg Weights

Panel A: Negative and positive weights

	Sum	Mean	Share		
Negative	-0.126	-0.008	0.101		
Positive	1.126	0.034	0.899		

Panel B: Correlations of Industry Aggregates

	α_n	Δd_n	β_n	F_n	$\text{Var}(\pi_n^*)$
α_n	1				
Δd_n	-0.138	1			
β_n	0.018	0.071	1		
F_n	0.168	-0.098	0.077	1	
$\text{Var}(\pi_n^*)$	0.029	0.072	0.515	-0.015	1

Panel C: Variation across years in α_n

	Sum	Mean
1990	0.615	0.013
2000	0.068	0.001
2010	0.317	0.007

Panel D: Top 10 Rotemberg weight industries

	$\hat{\alpha}_n$	Δd_n	$\hat{\beta}_n$	F_n
motor vehicles, equipment	0.192	-0.063	0.567	14.401
apparel and other textiles	0.175	-0.036	0.749	59.731
mining	0.108	0.089	0.083	17.834
construction	0.094	0.019	0.514	9.524
primary metals	0.085	-0.044	0.420	12.253
justice, public order	0.082	0.075	0.740	26.785
insurance and real estate	0.048	0.070	0.547	27.048
health services, except hospitals	0.040	0.045	0.722	6.677
hospitals	0.038	0.061	0.064	1.999
paper and allied products	0.035	0.046	0.595	18.601

Panel E: Estimates of β_n for positive and negative weights

	α-weighted Sum	Share of overall β	Mean
Negative	-0.117	-0.253	0.016
Positive	0.581	1.253	0.562

Notes: This table reports statistics about the Rotemberg weights, based on Goldsmith-Pinkham et al. (2020) and produced using their accompanying software. In all cases, we report statistics about the aggregated weights with normalized growth rates, where we aggregate a given industry across years, as discussed in**?**. Panel A reports the share and sum of negative weights. Panel B reports correlations between the weights (α_n), the national component of growth (Δd_n), the just-identified coefficient estimates (β_n), the first-stage F-statistic of the industry share (F_n), and the variation in the industry shares across locations ($\text{Var}(\pi_n^*)$). Panel C reports variation in the weights across years. Panel D reports the top five industries according to the Rotemberg weights. The Δd_n is the national industry growth rate, $\hat{\beta}_n$ is the coefficient from the just-identified regression, the 95 percent confidence interval is the weak instrument robust confidence interval using the method from Chernozhukhov and Hansen (2008) over a range from -10 to 10, and Ind. Share is the industry share (multiplied by 100 for legibility). Panel E reports statistics about how the values of β_n vary with the positive and negative Rotemberg weights.

Table D.2: Summary of Rotemberg Weights

Panel A: Negative and positive weights					
	Sum	Mean	Share		
Negative	-0.140	-0.009	0.109		
Positive	1.140	0.036	0.891		
Panel B: Correlations of Industry Aggregates					
	α_n	Δd_n	β_n	F_n	$\text{Var}(\pi_n^*)$
α_n	1				
Δd_n	-0.155	1			
β_n	-0.045	0.715	1		
F_n	0.330	-0.106	0.023	1	
$\text{Var}(\pi_n^*)$	-0.007	-0.061	-0.000	-0.056	1
Panel C: Variation across years in α_n					
	Sum	Mean			
1990	0.461	0.010			
2000	0.037	0.001			
2010	0.502	0.010			
Panel D: Top 10 Rotemberg weight industries					
	$\hat{\alpha}_n$	Δd_n	$\hat{\beta}_n$	F_n	
motor vehicles, equipment	0.236	-0.068	0.528	8.202	
justice, public order	0.128	0.074	0.703	34.392	
primary metals	0.121	-0.044	0.495	28.585	
apparel and other textiles	0.119	-0.039	0.704	31.504	
mining	0.115	0.113	0.741	9.209	
health services, except hospitals	0.057	0.040	0.796	17.942	
construction	0.054	0.016	0.400	3.922	
hospitals	0.045	0.062	-0.522	2.388	
insurance and real estate	0.040	0.055	0.514	18.769	
personal services (non-private household)	0.027	0.027	0.597	7.009	
Panel E: Estimates of β_n for positive and negative weights					
	α-weighted Sum	Share of overall β	Mean		
Negative	-0.116	-0.251	1.188		
Positive	0.578	1.251	0.376		

Notes: This table reports statistics about the Rotemberg weights, based on Goldsmith-Pinkham et al. (2020) and produced using their accompanying software. In all cases, we report statistics about the aggregated weights with normalized growth rates, where we aggregate a given industry across years, as discussed in Goldsmith-Pinkham et al. (2020). Panel A reports the share and sum of negative weights. Panel B reports correlations between the weights (α_n), the national component of growth (Δd_n), the just-identified coefficient estimates (β_n), the first-stage F-statistic of the industry share (F_n), and the variation in the industry shares across locations ($\text{Var}(\pi_n^*)$). Panel C reports variation in the weights across years. Panel D reports the top five industries according to the Rotemberg weights. The Δd_n is the national industry growth rate, $\hat{\beta}_n$ is the coefficient from the just-identified regression, the 95 percent confidence interval is the weak instrument robust confidence interval using the method from Chernozhukhov and Hansen (2008) over a range from -10 to 10, and Ind. Share is the industry share (multiplied by 100 for legibility). Panel E reports statistics about how the values of β_n vary with the positive and negative Rotemberg weights.

Table D.3: Relationship between gender-industry share gaps and characteristics

	Mining	Primary Metals	Motor Vehicles	Apparel	Justice
	(1)	(2)	(3)	(4)	(5)
Female employment share	0.22	-0.19*	-0.11	-0.065	-0.060
	(0.35)	(0.11)	(0.22)	(0.10)	(0.099)
Fraction College	-0.26**	0.088**	-0.19**	-0.023**	0.071**
	(0.053)	(0.027)	(0.050)	(0.011)	(0.022)
Fraction Black	0.062**	0.017	0.11**	-0.0036	0.037**
	(0.020)	(0.012)	(0.018)	(0.0085)	(0.013)
Fraction Married	-0.12*	-0.0076	0.078*	0.080**	0.039
	(0.065)	(0.025)	(0.045)	(0.016)	(0.031)
Unemployment Rate	0.28*	-0.17**	-0.18	0.027	-0.0074
	(0.16)	(0.079)	(0.14)	(0.032)	(0.042)
Fraction Forn born	0.057*	0.055**	-0.014	-0.025**	0.090**
	(0.032)	(0.014)	(0.020)	(0.0091)	(0.017)
Log CZ size	0.0094**	-0.0038**	0.0022*	0.0011**	-0.0050**
	(0.0015)	(0.0011)	(0.0012)	(0.00031)	(0.00099)
Female Participation	0.48**	0.0095	-0.013	0.052	0.062
	(0.18)	(0.045)	(0.083)	(0.048)	(0.057)
Constant	-0.44**	0.079**	0.032	-0.086**	-0.026
	(0.069)	(0.022)	(0.043)	(0.017)	(0.029)
Observations	1845	1894	2046	2223	1690
R^2	0.35	0.16	0.28	0.17	0.14

Notes: The dependent variable is the the share-gaps (π^*_{nt}) at the commuting zone level for the industries with the top-5 highest Rotemberg weights. The table is based on Table A2 of Goldsmith-Pinkham et al. (2020). Each column is a separate regression, pooling three cross sections. The regression is of $t+1$ characteristics on t share-gaps π^*_{nt}. Each regression also includes a separate intercept for each decade. Standard errors are clustered at the state level.

separately. While the estimates in column (1) are near our baseline estimates, it is slightly lower than the estimates in column (2). The over-identification test, reported in the last row of Column (3), indicates that the estimates in column (1) and (2) are not statistically different from one another. In column (4), we use the only the Top-5 industries to construct our instrument, but remove these industries in our estimation– which emphasises the spill-over of these industries to gender gaps in other industries.

D.2.2 Sensitivity to fixed-shares

Goldsmith-Pinkham et al. (2020) recommend fixing the shares in the shift-share instruments to some base year. Our baseline empirical work allows the shares vary by decade, which allows for more predictive power in our first-stage. To assess the sensitivity of our results to fixed-shares, we construct an analogous instrument using 1980 shares. Column (5) of Table D.4 displays the fixed-share result analogous to our baseline specification in column (4) of Table 1, and is very similar. A summary of the Rotemberg weights for our fixed-shares instrument can be found in Table D.2 of Appendix D.2. The Top-5 Rotemberg weight industries using fixed-shares coincide closely to those of our baseline instrument; they include: motor vehicles and equipment, apparel and textile products, mining, Justice, and primary metals. For completeness, columns (6)-(9) show the same robustness checks for the fixed-shares instrument as columns (1)-(4).

D.3 Robustness to city and industry definitions

As a baseline, and following much of the literature after Autor and Dorn (2013), we use commuting zones to proxy for local labour markets. These geographical units have the advantage of covering all of the United States, but vary greatly in population density and urbanization. We assess the sensitivity of our results to density or urban/rural divide by splitting our sample of into the 100 largest commuting zones, based on 1980 population, and commuting zones outside the 100 largest. Column (1) and (2) of Table D.5 shows that this the results are nearly identical between larger and smaller commuting zones. In columns (3) and (4), we use as local labour markets "cities" defined by the 1990 definition of Standard Metropolitan Statistical Areas (SMSA) or 1999 definition of Consolidated Metropolitan Statistical Areas (CMSA).[5] CMSA definitions produce larger cities, on average, because these definitions may combine several SMSAs. Working with these alternative definitions of local labour markets has virtually no impact on our estimates.

In the last two columns of Table D.5, we again use commuting zones as our geographical unit, but we change the number of industry groups. Our baseline industrial classification uses 45 industries, which we chose to balance the size of commuting zone-industry cells with industrial detail. Increasing the number of industrial groups results in many smaller CZ-industry cells, but the trade-off of greater detail may provide better measurement of industrial premia. We assess the sensitivity of our results two more disaggregate industrial classifications. First, we use the IPUMS constructed variable `ind1950`, which recodes census industry codes to the 1950 definitions and generates 144 consistent industries groupings. Finally, we use the industrial classifications based on he 1990 Census constructed by David Dorn, which includes

[5]Additional details on construction can be found in Appendix B.1.

Table D.4: Robustness: Instrument Construction

	Index				Fixed Shares Index				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\Delta \bar{w}^*_{ct}$	0.38**	0.51**	0.48**	0.46**	0.45**	0.30*	0.52**	0.46**	0.48**
	(0.17)	(0.11)	(0.091)	(0.092)	(0.080)	(0.16)	(0.094)	(0.080)	(0.099)
Obs.	27629	27629	27629	25533	27629	27629	27629	27629	25293
R^2	0.043	0.050	0.049	0.047	0.047	0.037	0.050	0.048	0.049
Instrument(s)	$Z^{\text{Index}}_{\text{-Top 5},ct}$	$Z^{\text{Index}}_{\text{Top 5},ct}$	$Z^{\text{Index}}_{\text{-Top 5},ct}$ $Z^{\text{Index}}_{\text{Top 5},ct}$	$Z^{\text{Index}}_{\text{Top 5},ct}$	Z^{FS}_{ct}	$Z^{\text{FS}}_{\text{-Top 5},ct}$	$Z^{\text{FS}}_{\text{Top 5},ct}$	$Z^{\text{FS}}_{\text{-Top 5},ct}$ $Z^{\text{FS}}_{\text{Top 5},ct}$	$Z^{\text{FS}}_{\text{Top 5},ct}$
Fixed Effects:									
Ind.× Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls									
Baseline	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Female Empl.					Yes				
First-Stage:									
	3.53**	5.75**	5.44**	5.85**	2.85**	3.16**	3.32**	3.05**	3.35**
	(1.33)	(0.66)	(0.64)	(0.66)	(0.33)	(0.91)	(0.36)	(0.36)	(0.35)
			2.84**					2.48**	
			(1.13)					(0.82)	
F-Stat.	7.08	74.85	56.75	77.51	72.65	11.98	84.80	44.62	89.80
p-val	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Over-id. p-val	.	.	0.47	.	.	.	.	0.13	.

Notes: This table displays results from the estimation of equation (15) via 2SLS using US Census and ACS from 1980-2010. Standard errors, in parentheses, are clustered at the state level. (*) and (**) denote significance at the 10% and 5% level, respectively. The dependent variable is the decadal change in in the CZ-industry gender gap. First-stage coefficients and diagnostics are reported in the bottom panel.

236 industrial groups.[6] These results, presented in columns (5) and (6) of Table D.5, are similar to those reported in Table 1.

Table D.5: Robustness: Sensitivity of Location and Industry Definitions

	Commuting Zones		SMSAs	CMSAs	Industry	
	(1) 100 Largest	(2) Excl. 100 Largest	(3)	(4)	(5) `ind1950`	(6) `ind1990dd`
$\Delta\bar{w}^*_{ct}$	0.47**	0.46**	0.43**	0.49**	0.40**	0.37**
	(0.087)	(0.14)	(0.099)	(0.082)	(0.086)	(0.079)
Obs.	8911	18715	7427	5845	29520	26205
R^2	0.053	0.038	0.051	0.059	0.026	0.021
Fixed Effects:						
Ind.× Year	Yes	Yes	Yes	Yes	Yes	Yes
Controls						
Baseline	Yes	Yes	Yes	Yes	Yes	Yes
First-Stage:						
$Z^{\text{Index}}_{ct,\text{female}}$	2.71**	4.12**	3.40**	4.53**	4.26**	4.18**
	(0.99)	(0.82)	(0.96)	(0.97)	(0.50)	(0.55)
$Z^{\text{Index}}_{ct,\text{male}}$	-3.88**	-4.78**	-4.16**	-5.37**	-4.23**	-4.14**
	(0.84)	(0.48)	(0.60)	(0.67)	(0.51)	(0.51)
F-Stat.	18.64	77.14	47.24	47.52	39.53	35.07
p-val	0.00	0.00	0.00	0.00	0.00	0.00
Over-id. p-val	0.01	0.53	0.84	0.82	0.31	0.95

Notes: This table displays results from the estimation of equation (15) via 2SLS using US Census and ACS from 1980-2010. Standard errors, in parentheses, are clustered at the state (columns 1, 2, 5 and 6), SMSA (column 3), or CMSA (column 4) level. (*) and (**) denote significance at the 10% and 5% level, respectively. The dependent variable is the decadal change in in the commuting zone/city-industry gender gap. Baseline controls include the start of the period log commuting zone/city size, unemployment rate of each gender, fraction of foreign born, and the ratio college to non-college workers. Regressions weighted by the start of period size of the commuting zone/city-industry. First-stage coefficients and diagnostics are reported in the bottom panel.

D.4 Sub-sample Analysis

Column (1) of Table D.6 shows the estimate of δ for those with a High School degree or less, which is very similar to our baseline population estimate in column (5) of Table 1. Column (2) presents the results for those with a college degree or more. When we focus on this smaller subgroup, we lose a great deal of observations and are left with just 8,234 commuting zone-industry cells. Additionally, the strength of first-stage falls; the F-statistic on the joint significance of each instrument is 3.36, indicating a weak instrument problem. The second-stage estimate of δ is small and not significantly different from zero, but is also imprecisely estimated. Given the strength of the first-stage, we are unable to determine whether the

[6]These industry codes can be found on David Dorn's webpage: https://www.ddorn.net/data.htm.

difference in the estimated coefficient between the education subsamples is result of a lack of identification or whether our mechanism fails to apply to college workers. Column (3) shows the estimates for all workers, which are slightly lower than in our baseline sample that consists of workers with less than college.

Columns (5) and (6) of Table D.6 shows results estimated separately by younger (less than 35 years old) and older workers. On the one hand, younger workers may be less sensitive to the local forces emphasised in our framework if younger groups are more mobile. On the other hand, this group maybe more sensitive if younger workers' wages are more directly related to current labor market conditions (Oreopoulos et al., 2012). These results suggest that outside job prospects might have a stronger impact on wages of younger compared to older workers. Finally, we split our sample based on marital status. For instance, one might expect that married women may be more mismatched to the local labour market compared to men, if location decisions are based on the husband. We find no substantiate differences in our estimates between these subsamples.

Table D.6: Robustness: Subsample Analysis

	$\leq$ High School	College	All	Age $<$ 35	Age $\geq$ 35	Married	Single
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\Delta\bar{w}^*_{ct}$	0.42**	0.028	0.32**	0.53**	0.36**	0.42**	0.47**
	(0.10)	(0.39)	(0.090)	(0.088)	(0.085)	(0.077)	(0.13)
Obs.	17067	8234	33660	11402	17183	18941	10127
R^2	0.048	0.005	0.031	0.061	0.042	0.050	0.052
Fixed Effects:							
Ind.× Year	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls							
Baseline	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First-Stage:							
$Z^{\text{Index}}_{ct,\text{female}}$	3.05**	0.30	2.11**	1.55	3.79**	3.66**	0.47
	(0.63)	(0.44)	(0.76)	(1.01)	(0.75)	(0.74)	(1.37)
$Z^{\text{Index}}_{ct,\text{male}}$	-4.28**	-0.94**	-3.41**	-3.90**	-4.01**	-4.09**	-2.87**
	(0.61)	(0.39)	(0.53)	(0.74)	(0.53)	(0.53)	(1.42)
F-Stat.	24.76	3.36	50.55	14.90	34.74	31.98	10.69
p-val	0.00	0.04	0.00	0.00	0.00	0.00	0.00
Over-id. p-val	0.36	0.47	0.41	0.25	0.65	0.20	0.16

Notes: This table displays results from the estimation of equation (15) via 2SLS using US Census and ACS from 1980-2010. Standard errors, in parentheses, are clustered at the state level. The dependent variable is the decadal change in in the CZ-industry gender gap for the indicated subsample. Baseline controls include the start of the period log CZ size, unemployment rate of each gender, fraction of foreign born, and the ratio college to non-college workers. Regressions weighted by the start of period size of the industry-CZ. First-stage coefficients and diagnostics are reported in the bottom panel.

D.5 Selection Correction: Mulligan and Rubinstein (2008)

We implement the Heckman-two-step procedure in the same manner as described in Mulligan and Rubinstein (2008), except for small changes due to the fact that we use a different data source and have a different outcome equation.

1. Following (Mulligan and Rubinstein, 2008), we assume that men's inverse Mill's ratio (IMR) is zero, or that men have no selection bias.

2. We form an variable for the number of children in the household under 5 years old. The categories are zero, 1, 2, 3, and 4 or more.

3. We estimate a first-stage probit equation where the dependent variable is an indicator for a women working with a valid wage. The right hand side of the probit contains all of the demographic variables we use in our baseline procedure described in the text, but additionally includes an indicator for marriage and a marriage-number of children interaction. This latter interaction becomes the exclusion restriction, as in Mulligan and Rubinstein (2008).

4. To be consistent with our second stage, we allow the effects marriage and the marriage-children interaction on working to vary by education level (4 groups) by fully interacting these variables with education. This is a slight departure from Mulligan and Rubinstein (2008).

5. Having estimated the probit, we construct an estimated IMR for women, and include this in our first-stage regression adjustment procedure described in the text. When including the estimated IMR and a marriage indicator, we allow the coefficients on these variables to vary by education. This is a slight generalization of Mulligan and Rubinstein (2008), who restrict the selection effect to be the same for education groups, but is consistent with the arguments of Machado (2012) who suggest that imposing the same selection function for all women is overly restrictive.

Our first-stage results with the inclusion of the IMR ratio are generally consistent with Mulligan and Rubinstein (2008). In particular, we find that the coefficient on the IMR switches signs from negative in the early Census years to positive in the most recent, particularly for those with less education. The fact that the inclusion of the IMR does not impact our point estimates of δ may not be surprising since we rely on quite different variation compared to Mulligan and Rubinstein (2008), and our control for city-level participation does not have much impact. It should also be noted that there is an ongoing debate about the Mulligan and Rubinstein (2008) claim that the narrowing of the gender gap is completely explained by selection. Machado (2012) and Herrmann and Machado (2012) question the identification strategy of Mulligan and Rubinstein (2008), suggesting that imposing the same selection rule for all women is overly restrictive and that family structure is correlated to pre-labor market characteristics. Machado (2012) shows that the gender gap declines for a sub-sample of 'always working' women (women with children who would likely work if they did not have children) and Chandrasekhar et al. (2012) use a bounding technique to show that a decline in the gender gap cannot be rejected in the Mulligan and Rubinstein (2008) data. Bar et al. (2015) argue that Mulligan and Rubinstein (2008) results are biased and overstate the importance of selection in explaining the decline in the gender gap.

E Decomposition of Exposure Gaps

E.1 Decomposing Exposure Differences Over Time

If the exposure gap plays such an important role, then one would naturally like to better understand the drivers of this variable's evolution. To this end, we decompose the gender exposure gap to get a sense of the relative importance of changing employment distributions and changing industrial wages. As described in the Introduction, given a vector of date t relative industry wages d_t and a vector of national employment distributions by gender, π_{gt}, the national gender exposure gap is

$$\textsc{ExposureGap}_t \equiv \underbrace{[\pi_{ft} - \pi_{mt}]}_{\pi_t^*} \cdot d_t. \tag{E.34}$$

To help interpret changes in E_t recall that a typical element of π_t^*–i.e. $\pi_{fnt} - \pi_{mnt}$–is a measure of the excess exposure of women to industry n in year t. We can decompose the evolution of $\textsc{ExposureGap}_t$ into parts due to changes in industry wages, to changes in employment distributions ("exposure"), and to the joint change in these ("interaction"). Specifically, pick some fixed year, $\bar{t}$, and decompose:

$$\Delta\textsc{ExposureGap}_t = \underbrace{[\pi_t^* - \pi_{\bar{t}}^*] \cdot d_{\bar{t}}}_{\Delta\textsc{Exposure}} + \underbrace{\pi_{\bar{t}}^{\text{gap}} \cdot (d_t - d_{\bar{t}})}_{\Delta\textsc{Wages}} + \underbrace{[\pi_t^* - \pi_{\bar{t}}^*] \cdot [d_t - d_{\bar{t}}]}_{\textsc{Interaction}}. \tag{E.35}$$

The ΔExposure component tells us the change in the exposure gap had only employment distributions changed, the ΔWages component tells us the change in the exposure gap had only wages changed, and the interaction component tells us the additional consequences of both changing. From another perspective, the ΔWages component captures the extent to which relative industrial wages shifted toward initially female jobs. This is the variation used by our instrument, Z_{gct} discussed in section 4.2. The ΔExposure component captures the extent to which jobs that became more female were initially good jobs. The interaction component captures the extent to which jobs that became more female were 'improving' jobs (i.e. those experiencing a relatively large improvement in relative wages).

Since the ΔExposure component is the only component to use a time-invariant set of industry wages, only it can be further decomposed to give a sense of the evolution of wage *levels*. Specifically, we can measure whether a larger exposure component is due to women moving into initially better jobs or due to men moving into initially worse jobs. The decomposition is:

$$[\pi_t^* - \pi_{\bar{t}}^*] \cdot d_{\bar{t}} = \underbrace{[\pi_{ft} - \pi_{f\bar{t}}] \cdot d_{\bar{t}}}_{\Delta\textsc{Exposure}_{\text{female}}} - \underbrace{[\pi_{mt} - \pi_{m\bar{t}}] \cdot d_{\bar{t}}}_{\Delta\textsc{Exposure}_{\text{male}}} \tag{E.36}$$

The components of decomposition (E.35) are displayed in Figure E.1, using the CPS data with $\bar{t} = 1979$. We see that wage changes have played a very important role, especially since the mid 1980s. Both ΔWages and ΔExposure contributed to the narrowing of the exposure gap between the mid-1980s and mid-1990s, but the effect of ΔWages was twice as large. The interaction effect played no role during this episode.

Interestingly, the slowdown in the narrowing of the exposure gap (and thus the pay gap) since the mid-1990s seems entirely due to a slowdown in the ΔExposure component and an

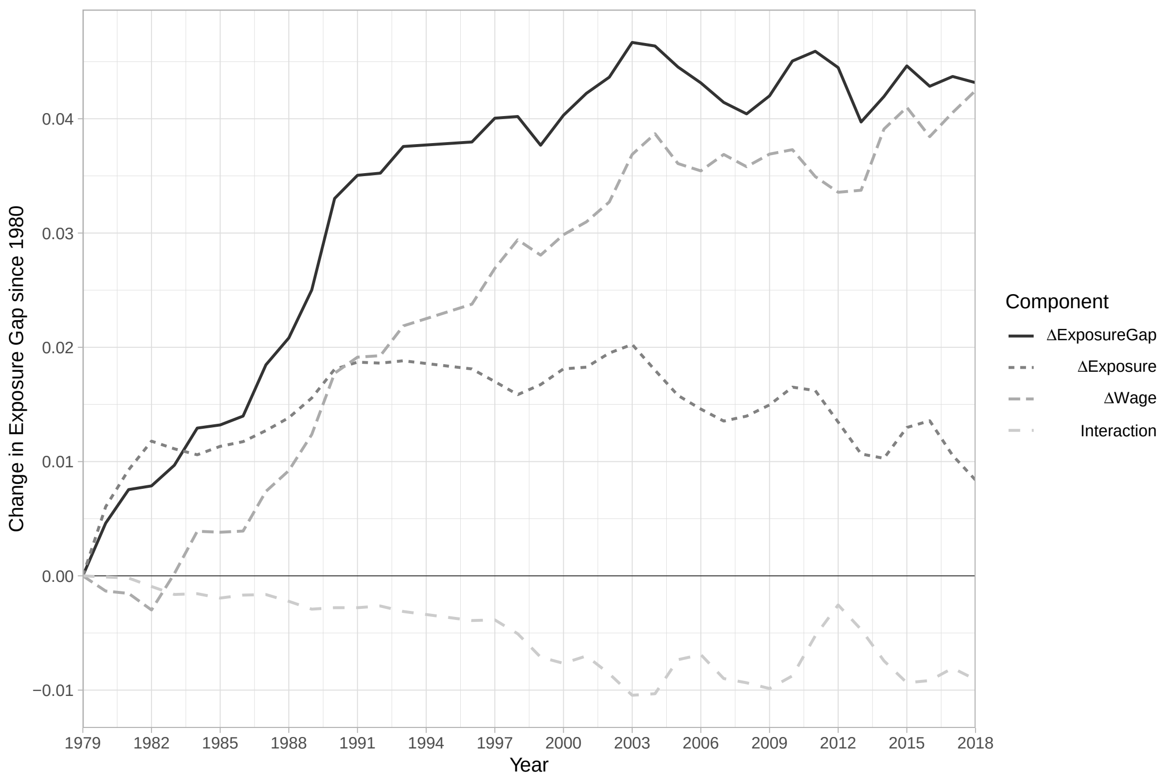


Figure E.1: Decomposing Changes in the Exposure Gap

Notes: The figure uses data from Outgoing Rotation Group (ORG) Supplements of the Current Population Survey. The figure displays each component of the decomposition given in equation (E.35). Industry categories are aggregated into 45 industry groups (details are provided in Appendix C).

Figure E.2: Decomposing the Exposure Component

Notes: The figure uses data from Outgoing Rotation Group (ORG) Supplements of the Current Population Survey. The figure displays each component of the decomposition given in equation (E.36). Industry categories are aggregated into 45 industry groups (details are provided in Appendix C).

actual decline in the INTERACTION component. That is, exposure changes continued to favour industries that were initially more female, but this was offset by two forces: the jobs that were becoming increasingly female (i) ceased being those that were relatively high paying, and (ii) started becoming those that were 'declining' (in the sense of experiencing the largest declines in in relative wages).

To unpack the ΔEXPOSURE, Figure E.2 shows its decomposition into male and female parts. We see that the rise of the ΔEXPOSURE component until the mid-1990s was due mostly to men shifting out of jobs that were initially good. It appears that women saw a modest shift toward better jobs, but by the mid-1980s women too had started to shift out of jobs that were initially good. The period between the mid-1980s and the mid-1990s thus saw initially good jobs become more female because men were moving out of such jobs faster than were women. By the mid-1990s men and women were shifting out of initially good jobs at around the same pace, leaving the overall ΔEXPOSURE component relatively stable. This decomposition provides evidence that is consistent with the rapid narrowing of the within-industry pay gap in the 1980s being more about men losing out than women gaining. Indeed, by the mid-1990s when the narrowing slowed, men and women had both experienced absolute declines in the ΔEXPOSURE component of exposure but men had lost almost three times as much.

To summarize, the period until the mid-1990s was characterized by initially male jobs with falling relative wages to initially female jobs, and to a lesser extent because initially good jobs became more female (owing to men shifting out of such jobs to a greater extent than women). Since the mid-1990s relative wages continue to shift in favor of initially female jobs, but this is partially offset by relative wages also shifting toward jobs that had become less female.

References

Akerlof, G. A. and R. E. Kranton (2000). Economics and identity. *The Quarterly Journal of Economics 115*(3), 715–753.

Autor, D. H. and D. Dorn (2013, aug). The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market. *American Economic Review 103*(5), 1553–1597.

Bar, M., S. Kim, and O. Leukhina (2015, October). Gender Wage Gap Accounting: The Role of Selection Bias. *Demography 52*(5), 1729–1750.

Bayer, P., S. L. Ross, and G. Topa (2008, December). Place of work and place of residence: Informal hiring networks and labor market outcomes. *Journal of Political Economy 116*(6), 1150–1196.

Beaudry, P., D. A. Green, and B. Sand (2012). Does industrial composition matter for wages? a test of search and bargaining theory. *Econometrica 80*(3), 1063–1104.

Brashears, M. E. (2008). Gender and homophily: Differences in male and female association in blau space. *Social Science Research 37*(2), 400 – 415.

Chandrasekhar, A., V. Chernozhukov, F. Molinari, and P. Schrimpf (2012, December). Inference for best linear approximations to set identified functions. CeMMAP working papers CWP43/12, Centre for Microdata Methods and Practice, Institute for Fiscal Studies.

Doms, M. and E. Lewis (2006). Labor supply and personal computer adoption. Technical Report 2006-18, Federal Reserve Bank of Philadelphia.

Dorn, D. (2009, September). Essays on inequality, spatial interaction, and the demand for skills. Technical Report 3613, Dissertation University of St. Gallen.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020, August). Bartik instruments: What, when, why, and how. *American Economic Review 110*(8), 2586–2624.

Herrmann, M. and C. Machado (2012). Patterns of selection in labor market participation. In *11th IZA/SOLE Transatlantic Meeting of Labor Economists*.

Lemieux, T. (2006). Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill? *The American Economic Review 96*(3), pp. 461–498.

Machado, C. (2012). Selection, Heterogeneity and the Gender Wage Gap. Institute for the Study of Labor (IZA) Working Paper 7005, Institute for the Study of Labor (IZA).

Mulligan, C. B. and Y. Rubinstein (2008). Selection, investment, and women's relative wages over time. *The Quarterly Journal of Economics 123*(3), 1061–1110.

Munshi, K. (2003, May). Networks in the modern economy: Mexican migrants in the u.s. labor market. *The Quarterly Journal of Economics 118*(2), 549–599.

Oreopoulos, P., T. von Wachter, and A. Heisz (2012, January). The short- and long-term career effects of graduating in a recession. *American Economic Journal: Applied Economics 4*(1), 1–29.

Park, J. H. (1994). Estimation of Sheepskin Effects and Returns to Schooling Using he Old and the New CPS Measures of Educational Attainment. Working Paper 717, Princeton University, Industrial Relations Section.

Pissarides, C. (2000). *Equilibrium Unemployment Theory* (Second ed.). Cambridge, MA: MIT press.

Ruggles, S., T. J. Alexander, K. Genadek, R. Goeken, and M. Schroeder, Matthew B. Sobek (2010). Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]. Technical report, University of Minnesota.

Shapiro, C. and J. E. Stiglitz (1984). Equilibrium unemployment as a worker discipline device. *The American Economic Review 74*(3), 433–444.

Sorkin, I. (2017, May). The role of firms in gender earnings inequality: Evidence from the united states. *American Economic Review 107*(5), 384–87.