Equal Pay for *Similar* Work*

Diego Gentile Passaro Fuhito Kojima Bobak Pakzad-Hurson †

December, 2024

Abstract

Equal pay laws increasingly require that workers with different group identities doing "similar" work are paid equal wages within firm. We study such "equal pay for similar work" (EPSW) policies theoretically and test our models' predictions empirically using evidence from a 2009 gender-based Chilean EPSW. Under EPSW, firms segregate their workforce by gender. When there are more men than women in a labor market, EPSW increases the gender wage gap.

1 Introduction

> "No employee with status within one or more protected class or classes shall be paid a wage at a rate less than the rate at which an employee without status within the same protected class or classes in the same establishment is paid for... **similar work** [emphasis added]"
>
> -New York Labor Code, Section 194

Firms have some degree of wage-setting power in many labor markets (see, e.g. Card, 2022; Manning, 2005). Because of this, they may pay workers different relative salaries in ways that are repugnant to society at large. In particular, wage gaps between groups of workers, often men

*We are especially grateful to Koji Yokote for numerous helpful comments and for directing us to mathematical results used for our analysis. We thank Bo Cowgill, Zoe Cullen, Mayara Felix, Michihiro Kandori, Erik Madsen, Akihiko Matsui, Teddy Mekonnen, Ellen Muir, Yusuke Narita, Chris Neilson, Yoko Okuyama, Debraj Ray, Jesse Shapiro, Neil Thakral, Rakesh Vohra, and Seth Zimmerman for their helpful comments and conversations, as well as to seminar audiences at UPenn, NYU, Yale, Brown, USC, UCSD, Princeton, Wash U, Northwestern, and conference audiences at NBER Market Design, SITE (Market Design, Gender), EC, SOLE, EAAMO, Virtual Market Design Seminar, MAD, and Tokyo Gender Economics Conference. We acknowledge research assistance from Atsumu Akazawa, Nanami Aoi, Kento Hashimoto, Asuka Hirano, Nadeen Kablawi, Masanori Kobayashi, Shinji Koiso, Kevin Li, Anya Marchenko, Sota Minowa, Daiji Nagara, Leo Nonaka, Ryosuke Sato, and Kenji Utagawa. Fuhito Kojima is supported by the JSPS KAKENHI Grant-In-Aid 21H04979 and JST ERATO Grant Number JPMJER2301. Bobak Pakzad-Hurson acknowledges support from the James M. and Cathleen D. Stone Inequality Initiative.

†Gentile Passaro: Amazon, diegogp@amazon.com. Kojima: Department of Economics, University of Tokyo, fuhitokojima1979@gmail.com. Pakzad-Hurson: Department of Economics, Brown University, bph@brown.edu.

and women, are frequent rallying points for governmental action. A popular form of legislation seeks to prohibit firms from paying disparate wages to different workers, guided by the principle of "equal pay for equal work" (EPEW). In the United States, 49 states had EPEW laws in effect in 2015, requiring each firm to pay equal wages to all of its workers doing equal work.

However, EPEW may be difficult to enforce; "equal pay" is straightforward to define, but it is likely that no two workers do exactly "equal work" within a firm. Firms can avoid the intent of these laws by pointing out differences between workers or making other maneuvers such as job title proliferation to marginally heterogenize their workforce (Baron and Bielby, 1986; Goldin, 1990).[1] To combat this enforceability issue, many EPEW laws have been updated to include a measure of coarseness–they require a firm to set "equal pay for similar work" (EPSW).[2] California, among the first states in the US that moved from EPEW to EPSW, described a "key difference" of EPSW as "mak[ing] it more difficult for employers to justify unequal pay."[3] The judicial system has taken a similar view of the "coarseness" of EPSW; the highest court in the US to have adjudicated an alleged violation of EPSW was the 9^{th} Circuit Court of Appeals, whose majority opinion in *Freyd v. University of Oregon* states "'[similar] work' is a more inclusive standard than equal work," and interprets "similar" work in the university faculty setting to be defined by the department-by-rank cell (e.g. all full professors in the economics department of a particular university are "similar" to one another). As of January 2023, more of the US workforce is under the jurisdiction of a state EPSW law than a state EPEW law.[4] The equal pay provision in EPSW is frequently group based in that it binds only across groups of workers, prohibiting for example, that a man is paid more than a "similar" woman (and vice versa).

Despite the rapid growth of EPSW, little is known about its effects on labor market outcomes. Since EPSW is more constraining on firms than EPEW, EPSW may lead to a larger direct effect on wages. But how will firms adjust their employment policies to adapt in equilibrium? How will potential employment changes affect the goal of ensuring fairer pay?

We theoretically and empirically study the labor market effects of EPSW. Our findings

[1] For example, a manufacturer told the *Washington Post* (1964) that his firm would "downgrade some job classifications for women and reassign higher-level, higher-paying duties to men" in response to EPEW. We are grateful to Martha Bailey for bringing this article to our attention.

[2] Some jurisdictions refer to such policies as "Equal Pay for Comparable Work" or "Equal Pay for Substantially Similar Work." Given the similarity in both the stated goals and remaining wording of these policies to EPSW, we view them all synonymously.

[3] See `https://www.dir.ca.gov/dlse/california_equal_pay_act.htm`.

[4] The percent of the US workforce that is under the jurisdiction of EPSW and EPEW, respectively, are 45.9 and 45.6. These figures are calculated from 1) finding all states covered by each of these policies (see `https://www.dol.gov/agencies/wb/equal-pay-protections`), and 2) the share of the US workforce employed in each state (see `https://www.statista.com/statistics/223669/state-employment-in-the-us/`).

suggest that the equilibrium effects of EPSW overwhelm the direct effects, leading to increased occupational segregation and a shift in the wage gap in favor of the majority group of workers in a labor market. Therefore, these policies may counterintuitively exacerbate the problem they were intended to solve: for example, as in our empirical setting, if male labor force participation is higher than female labor force participation and men receive higher average wages than women, then implementing a gender-based EPSW will *increase* the gender wage gap, exacerbating the problem it was designed to solve.

We develop a theoretical framework to elucidate key economic forces at play. Homogeneous firms compete for a continuum of heterogeneous workers who all perform "similar" work in the eyes of the law. Each worker belongs to one of two groups, A or B (e.g. men or women), and is endowed with a "productivity" drawn from a distribution that potentially differs by group identity. Each firm has a constant-returns-to-scale production function.

Several remarks are in order regarding our modeling decisions. First, we are agnostic about the underlying fundamentals of worker productivity, and therefore the differences in distributions between groups; we interpret productivity as a measure of firm willingness to pay for an individual worker net of any unmodeled, discriminatory factors. For example, in consumer-facing industries where consumers have discriminatory preferences in favor of a particular group of workers (Bar and Zussman, 2017; Holzer and Ihlanfeldt, 1998; Kelley et al., 2023; Kline et al., 2022), we may expect that group's productivity distribution to first-order stochastically dominate the other group's distribution. Similarly, taste-based bias by (managers of) firms can be incorporated. Second, even though our model assumes complete information in which each worker's productivity is common knowledge, a worker's productivity can instead be interpreted as the expected productivity at the time the hiring decision is made. Therefore, we can similarly include statistical discrimination into our framework. Third, our model assumes that all workers perform "similar" work in the eyes of the law.[5] Thus, our theoretical predictions should be viewed as applying "within job" in a particular labor market, and should not be used to predict differential effects on, for example, custodians and lawyers working within the same firm.

Our model analysis reveals important effects of EPSW on worker sorting across firms. We begin by describing the results of a simple, static model containing two firms and no frictions for intuition. Without EPSW, each worker can be hired by either firm in equilibrium,

[5]We highlight that our modeling choices preclude an alternative where two workers are deemed "similar" if their productivities are within some distance. This is because (as we have discussed in this paragraph) we are agnostic about what "productivity" represents and, in particular, may include subjective factors not tied to an objective factor like output. Presumably, any EPSW uses objective measures to define similarity, and thus assuming that "similarity" is a function of the distance of worker "productivities" in our model is in conflict with our agnosticism about our interpretation of productivity.

regardless of group identity or productivity. Similarly to the classic Bertrand model, firms compete fiercely for each worker, and as a result, the average gap in pay across groups A and B is equal to the difference in average productivity between the groups. Thus, any discriminatory factors affecting firms' willingness to pay are exactly reflected in the wage gap.

With EPSW, we show that firms segregate the workforce, with one firm hiring all $A-$group workers and the other hiring all $B-$group workers. To understand why this is the case, note that because each firm hires from only one group of workers, no firm is exposed to the constraint of equal pay in equilibrium. By contrast, EPSW makes poaching workers from its competitor costly: EPSW requires equal pay to any two workers from different groups and, by transitivity, this implies that equal wages must be paid to *all* workers it hires. Thus, EPSW serves as the enforcement mechanism for segregation, similarly to location choices in Hotelling's competition model.

We then show that EPSW moves the wage gap in favor of the majority group of workers. This result follows from an equal profit condition between firms that must be satisfied in equilibrium. More specifically, if there are more $A-$group workers than $B-$group workers, the firm that hires these workers must receive smaller average profit from each worker than the other firm receives from the average $B-$group worker, so $A-$group workers' average wage is relatively higher. Notably, the directional effect of EPSW on the wage gap follows simply because the majority group has a larger population and, in particular, this conclusion holds regardless of the distributions of productivities of the two groups. Moreover, we also show that firm profit and the magnitude of increase in the wage gap co-move, implying that firms would benefit from selecting equilibria with larger wage gaps.

One might claim that our model stacks the deck against EPSW because the outcome without EPSW is already "fair" in the sense that workers from both groups are paid their productivity. We disagree. Recall that our model does not make strong assumptions on discriminatory factors, so "productivity" could incorporate firms' bias. Given this observation, our model does *not* take a stance on whether or not the outcome without EPSW is fair. What we do show, by contrast, is that EPSW is relatively more advantageous to the majority group. And this conclusion is robust in that it does not depend on whether the outcome without EPSW is fair or not. Of course, if the labor market is not fair in that the majority group is favored without EPSW for discriminatory reasons, then our result implies that EPSW makes the labor market even less fair.

Our basic model is stylized in several respects. To study whether our main predictions hold under more realistic assumptions, we also study a dynamic model of search and bargaining.

The model features an arbitrary number of firms, search frictions (firms do not bargain simultaneously with all workers), and reallocation frictions (firms retain workers from previous periods, and so segregating for one group entails firing existing workers from the other group). We show that our search model produces qualitatively similar results to those in the static model. Specifically, firms segregate eventually, if not necessarily immediately, after the introduction of EPSW. A worker's bargaining power is endogenously determined by the number of firms willing to hire her, therefore, the wage gap moves in favor of the group for which more firms segregate. Under an additional regularity condition that workers search "evenly" across firms, more firms segregate for the majority group, and thus the wage gap moves in favor of the majority group. These results are dynamic analogues of our main theoretical results in the static model. Moreover, we also find that firms with small reallocation frictions in the sense of having a nearly segregated pool of existing employees segregate early, while those with large frictions in the sense of having a less segregated pool of existing employees may opt to stay desegregated for longer.

Our theoretical framework yields clear empirical tests. We predict that EPSW leads firms to *fully* segregate, and not merely that it increases the share of the majority group of workers at each firm. Moreover, we predict EPSW leads to an increase in the wage gap toward majority-group workers in a particular labor market, therefore, EPSW's directional impact on wage inequality is heterogeneous depending on the labor market. These specific, heterogeneous predictions serve to differentiate our framework from others. For instance, consider an alternative hypothesis that EPSW reduces demand for labor from the historically-discriminated against group g. We would expect a decrease in the share of firms' workforce from group g, but not necessarily all-or-nothing segregation. Moreover, we would expect a relative decrease in the wages of $g-$group workers, regardless of whether $g-$group workers are in the majority of a particular labor market.

We test our theoretical predictions by studying the effect of the enactment of EPSW in Chile in 2009. This EPSW was the first equal pay law in Chile, and it constrained how a firm could pay its workers across gender: no firm is permitted to pay a woman less than it pays a man for similar work, and vice versa. The law subjects firms in violation to substantial fines, and through a public-records request, we find direct evidence of policy enforcement. Importantly, EPSW binds only for firms above a particular size threshold. This allows for a straightforward event-study (difference-in-differences) design to estimate the causal effects of EPSW, wherein we compare differences in outcomes of firms above ("treated") and below ("control") the threshold following EPSW. Following Bennedsen et al. (2022); Böheim and Gust (2021); Duchini et al. (2022); Gulyas et al. (2023), we restrict our sample to firms

just above and below this size threshold to limit size-based wage dynamics. That is, our identifying assumption is that parallel trends hold for similarly-sized firms.

Using matched employer-employee data from 2005–2013 we identify the following effects of EPSW consistent with our model predictions. First, EPSW increases gender segregation across firms. The share of firms with workers of only one gender increases by 4.4 percentage points, off a baseline of 34.3% of gender-segregated firms prior to EPSW enactment. We show that EPSW leads to a "missing mass" of firms that are nearly-but-not-fully gender segregated, and moreover, that the share of "missing," nearly segregated firms is of a similar magnitude as the increased share of fully segregated firms. Moreover, firms that are nearly-but-not-fully segregated pre policy are the most likely to fully segregate post policy. These patterns are consistent with the predictions from our search model, and suggest that firms on the margin of full segregation can more easily transition to full segregation (e.g. by firing the only worker of the "wrong" gender).

Second, we show that EPSW moves the gender wage gap in favor of the majority group of workers in a local labor market. In local labor markets—defined by firm industry and county—where a majority of workers are men, we find that EPSW *increases* (in favor of men) the gender wage gap by 4.3 percentage points, while in local labor markets where a majority of workers are women, we find that EPSW *decreases* (in favor of women) the gender wage gap by 6.2 percentage points. These findings exactly match our prediction that EPSW benefits the majority group of workers in a labor market. Because men in Chile dominate the overall labor market (5/6 of all workers are employed in majority male local labor markets), the overall effect of EPSW is to increase the gender wage gap (in favor of men) by 2.7 percentage points.

1.1 Related Literature

While we are the first, we believe, to analyze the novel equilibrium effects of EPSW, there are rich theoretical and empirical literatures related to EPEW.

Theoretical studies of EPEW have typically focused on its unintended effects. This focus can be traced back to Milton Friedman who once famously said, "What you are doing, not intentionally, but by misunderstanding, when you try to get equal pay for equal work law... is reducing to zero the cost imposed on people who are discriminating for irrelevant reasons."[6] More recent work studies EPEW in Salop's classic location model; the first such paper is Bhaskar et al. (2002) and is succeeded by Berson (2016); Kaas (2009); Lagerlöf (2020);

[6]See `https://www.aei.org/carpe-diem/milton-friedman-makes-the-case-against-equal-pay-for-equal-work-laws/`.

Lanning (2014). These papers must contend with the very motivation that led to EPSW: what is "equal work"? Doing so results in at least two difficulties. First, the authors interpret "equal work" literally and assume workers are equally productive, while in reality there may be very few workers whose productivities are exactly equal. Second, their analyses predict that EPEW can either increase or decrease differences in outcomes across groups of workers, often within the same paper. The lack of clear policy-relevant predictions is reflected in the empirical literature on EPEW, which we discuss shortly. By contrast, we find that EPSW has clear, if unintended, effects: our theoretical analysis unambiguously predicts both job segregation and a change in the wage gap that favors the majority group of workers in the labor market, and our empirical analysis of Chilean data confirms both predictions.

The empirical literature investigating equal-labor-rights legislation primarily considers US policies in the 1960s and 1970s. As with the theoretical literature we detail above, this empirical literature draws mixed conclusions about whether such legislation improves the employment rate or wages of protected classes of workers (see Altonji and Blank (1999); Bailey et al. (2022); Blau and Kahn (1992); Donohue III and Heckman (1991); Hyland et al. (2020); Neumark and Stock (2006) for detailed discussions).[7] One difficulty in much of the literature is assessing the impacts of individual policies, as many related labor policies were enacted in quick succession.[8] Donohue III and Heckman (1991) argue that it is difficult to attribute observed effects to any one of the contemporaneous policies, as there may be complementarities between them. Our empirical setting of Chile is notable as no existing equal pay laws were on the books at the time EPSW was enacted in 2009, and no significant related policies were enacted in quick succession.

Methodologically, our paper is more related to the literature on "best-price" guarantees, which commit firms to rebating past consumers if prices fall in the future. These policies have the direct effect of equalizing payments across heterogeneous buyers, but have the unintended equilibrium effect of raising firm market power (Butz, 1990; Cooper and Fries, 1991; Scott Morton, 1997a,b). In our paper, EPSW plays the role of a best-wage guarantee, but due to competition between firms, it importantly binds off the equilibrium path; firms segregate in equilibrium because a deviating firm faces a competitive disadvantage in hiring. As a result, while firms in our model have an ex-ante identical willingness to pay for each

[7]The findings of Bailey et al. (2022) suggest that the conclusions from this literature may be sensitive to the econometric methods used.

[8]The Equal Pay Act of 1963 requires equal pay for men and women for equal work, while Title VII of the Civil Rights Act of 1964 prohibits discrimination in hiring, layoffs, and promotions. There were also other federal equal pay policies—Executive order 11246 in 1965 banning discrimination in hiring by federal contractors against minority candidates, and an extension to include women in 1967; the Equal Employment Opportunity Act in 1972 to increase enforcement; and many individual state policies.

particular worker, the costs of hiring workers from the "wrong" group are differentiated in equilibrium, driving the unintended wage effects of EPSW. This force is similar to "artificial" costs that heterogenize ex-ante identical products in consumer markets, which can lead to local market power for firms (Klemperer, 1987).

Therefore, a key force in our model is firms' equilibrium behavior to segregate their workforce by group identity. Indeed, we show empirically that the Chilean EPSW leads to an increase in gender segregation across firms. One may suspect that such segregation is less likely to occur in other localities that enact EPSW.[9] Speaking to this point, however, group-based segregation across firms has been noted in the US (Blau, 1977; Goldin, 1990; Hellerstein and Neumark, 2008; Neumark et al., 1996), and recent research (Ferguson and Koning, 2018) argues that this segregation has increased over time. Therefore, it seems plausible that EPSW could further affect segregation in a wide variety of labor markets.

2 Static Model

2.1 Model Description

There are two firms, 1 and 2, and a continuum of workers. Each worker is endowed with a type $e=(g,v)\in\{A,B\}\times[0,1]$, where $g\in\{A,B\}$ is the worker's group identity (say, men and women) and $v\in[0,1]$ is the worker's productivity. There is a $\beta\geq 1$ measure of $A-$group workers and 1 measure of $B-$group workers. F_A and F_B are cumulative distribution functions governing the productivities of workers in groups A and B, respectively. F_A and F_B are absolutely continuous and thus admit density functions f_A and f_B, respectively. For each $g\in\{A,B\}$, we assume that $0<\underline{f}_g\leq\bar{f}_g<+\infty$, where $\underline{f}_g=\inf\{f_g(v)|v\in[0,1]\}$ and $\bar{f}_g=\sup\{f_g(v)|v\in[0,1]\}$. Note that the distribution of $A-$group workers may be different from that of $B-$group workers, allowing us to model situations in which firms discriminate against one of the groups of workers (i.e. the output of $B-$group workers are drawn from the same distribution as $A-$group workers, but the firms' willingness to pay for them is lower because firms have a taste-based preference for $A-$group workers). For each $g\in\{A,B\}$ we define $\mathbb{E}_g(v):=\int_0^1 v f_g(v)dv$.

A *(labor) market* is a tuple (β,F_A,F_B). For expositional ease, we study this environment

[9] For example, recent research shows that gender-based occupational segregation may be especially likely when the local language has gendered nouns, as firms can target their hiring to workers of specific genders (Card et al., 2024; Kuhn and Shen, 2023; Kuhn et al., 2020). This may explain the high baseline level of gender segregation in Chile, where Spanish is the official language. Gendered nouns and targeted hiring may also play a role in the ability of Chilean firms to further segregate once EPSW is enacted. Our search model includes this feature, by allowing workers to direct their search based on the segregation status of firms.

via a cooperative game (all of our model predictions are unchanged if we instead consider a non-cooperative game, see Remark 1). An outcome specifies, for each worker, the firm she works for (or the outside option of staying unemployed) and the wage received (if employed). Each worker only cares about her wage. A firm generates infinitesimal profit $v-w$ if it hires a worker of productivity v and pays her wage w. The firm does not have any capacity constraint (i.e. the firm can hire any measure of workers), and its payoff is the integral of profit generated from workers it employs.

Formally, an outcome for firm i is $O_i := \{(f_g^i(v), w_i^g(v))\}_{v\in[0,1], g=A,B}$, where:

1. $f_g^i(v) \in [0, f_g(v)]$ is the density of workers of type $e=(g,v)$ hired by firm i,
2. $w_i^g(v) \in [0,\infty)$ is the wage firm i pays to each worker of type $e=(g,v)$ it hires. If $f_g^i(v)=0$, then we fix $w_i^g(v)=0$.

An outcome is a tuple $O := (O_1, O_2)$ where O_i is the outcome for firm i such that $f_g^1(v) + f_g^2(v) \leq f_g(v)$ for each v and g. That is, the (overall) outcome specifies the outcome for both firms such that the total hiring does not exceed the supply of workers (a feasibility requirement). We assume that f_g^i and w_i^g are measurable functions for each i and g.

Under an outcome for i, $O_i := \{(f_g^i(v), w_i^g(v))\}_{v\in[0,1], g=A,B}$, firm i receives profit

$$\pi_i^{O_i} := \beta \int_0^1 [v - w_i^A(v)] f_A^i(v) dv + \int_0^1 [v - w_i^B(v)] f_B^i(v) dv.$$

Given an outcome $O=(O_1,O_2)$ and firm i, we denote $\pi_i^O := \pi_i^{O_i}$, and denote by AW_g^O the average wages for group $g \in \{A,B\}$, i.e.[10]

$$AW_g^O := \int_0^1 w_1^g(v) f_g^1(v) dv + \int_0^1 w_2^g(v) f_g^2(v) dv.$$

We refer to $AW_A^O - AW_B^O$ as the *wage gap* in outcome O. Similarly, for $(AW_A^O, AW_B^O) \neq (0,0)$, we refer to AW_A^O / AW_B^O as the *wage ratio (or log wage gap)* in outcome O (assuming that $AW_B^O \neq 0$, and if $AW_A^O > AW_B^O = 0$ then we define the wage ratio in outcome O to be $+\infty$).[11]

[10] Note that each unemployed worker from group g contributes a wage of 0 to the calculation of the average wage for group g.

[11] The wage ratio can be interpreted as the log wage gap because the latter, i.e. the gap in the log wages between the two groups, $\ln(AW_A^O) - \ln(AW_B^O)$, is a monotonic transformation of the former by the natural logarithm; $\ln(AW_A^O / AW_B^O) = \ln(AW_A^O) - \ln(AW_B^O)$.

We view two outcomes for firm i, $O_i := \{(f_g^i(v), w_i^g(v))\}_{v\in[0,1], g=A,B}$ and $\tilde{O}_i := \{(\tilde{f}_g^i(v), \tilde{w}_i^g(v))\}_{v\in[0,1], g=A,B}$ as equivalent if, for each $g \in \{A,B\}$, $f_g^i(v) = \tilde{f}_g^i(v)$ and $w_i^g(v) = \tilde{w}_i^g(v)$ for almost all v. We view two outcomes O and $\tilde{O}$ as equivalent if either 1) O_i is equivalent to $\tilde{O}_i$ for every $i \in \{1,2\}$, or 2) O_1 is equivalent to $\tilde{O}_2$ and O_2 is equivalent to $\tilde{O}_1$. The first condition captures the usual notion that the outcomes are regarded as equivalent if both the employment patterns and wages are identical between them except for a measure-zero set. The second condition captures the case in which the employment patterns and wages are identical almost everywhere once the names of the firms are relabeled—recall that firms are homogeneous in the present model.

An outcome is said to be a *core outcome* (without EPSW) if there is no firm and an alternative wage schedule for a subset of workers such that they are made better off being matched to each other, that is, both the firm and each worker in the hired subset obtain a higher payoff than in the present outcome. Formally, we say that an outcome $O := \{(f_g^i(v), w_i^g(v))\}_{v\in[0,1], i=1,2, g=A,B}$ is *blocked* by firm j via an alternative outcome (for j) $\tilde{O}_j := \{(\tilde{f}_g^j(v), \tilde{w}_j^g(v))\}_{v\in[0,1], g=A,B}$ if $\pi_j^{\tilde{O}_j} > \pi_j^{O_j}$ and, for each $g \in \{A,B\}$ and almost all $v \in [0,1]$, one of the following conditions hold. Note that, because we define $\tilde{O}_j$ to be an outcome, it must satisfy all restrictions imposed on an outcome in addition to those listed below:

1. $\tilde{w}_j^g(v) \geq w_j^g(v)$ and $\tilde{w}_j^g(v) > w_{-j}^g(v)$,
2. $\tilde{w}_j^g(v) \geq w_j^g(v)$ and $\tilde{f}_g^j(v) + f_g^{-j}(v) \leq f_g(v)$,
3. $\tilde{w}_j^g(v) > w_{-j}^g(v)$ and $\tilde{f}_g^j(v) + f_g^j(v) \leq f_g(v)$, or
4. $\tilde{f}_g^j(v) + f_g^j(v) + f_g^{-j}(v) \leq f_g(v)$.

These cases enumerate all possibilities for the formation of a blocking coalition. Condition 1 states a "no wage cuts" requirement; if firm j weakly raises the wages of all workers involved, and strictly raises wages for workers employed by the other firm $-j$, then these workers are all willing to join the blocking coalition. Condition 2 considers the case in which firm j does not need to poach workers from firm $-j$ to construct the blocking outcome, so the only constraint on wages is that existing workers' wages are not reduced. Condition 3 considers the case in which firm j does not need to keep any existing workers to construct the blocking outcome, so the only restriction on wages is that the wage paid to poached workers is higher than those paid by $-j$ to the same workers. Condition 4 considers the case in which firm j can hire from unemployed workers to construct the blocking outcome, so there is no additional restriction on the wages of these workers.

We say that an outcome O is a *core outcome* if there exists no outcome that blocks it.

REMARK **1.** Implicit in the above machinery is the assumption that each firm pays the same wage to all workers of the same type that it employs. We also note that our model is cooperative in nature. Core is the standard solution concept in models of matching with transfers, e.g. Shapley and Shubik (1971) and Kelso and Crawford (1982). Indeed, our setting without EPSW can be regarded as a specific case of the latter paper with two homogeneous firms with additive surplus functions (except that we have a continuum of workers, a feature we regard as merely a technical difference).

An alternative approach would be to set up a non-cooperative game and analyze its equilibria. In Appendix B, we present a non-cooperative game wherein the firms simultaneously make wage offers to workers, and each worker accepts at most one of the offers she receives. The pure-strategy subgame perfect Nash equilibrium outcomes of this game exactly coincide with the set of core outcomes of the cooperative game we describe above under the assumption that firms must pay the same wage to all workers of the same type. We choose to present the cooperative framework in the main text because its exposition is simpler, and the equivalence mentioned here provides a noncooperative foundation for doing so. We also note that our results extend (in both the cooperative and non-cooperative models) straightforwardly to the case in which firms can pay workers of the same type different wages under a maintained assumption of measurability of the wage schedule.

REMARK **2.** The definition of block implies two restrictions any core outcome must satisfy. First, Condition 3 of the definition of block immediately implies the following **Equal Profit Condition**:

> In any core outcome O, $\pi_1^O = \pi_2^O$.

This is because otherwise the firm earning strictly lower profit could fire all of its existing workers and hire all of the workers employed by the other firm with an arbitrarily small wage increase. Second, the definition of the core implies the following **Individual Rationality Condition** for firms:

> In any core outcome, there is no set $V \subset [0,1]$ of positive Lebesgue measure, a group $g \in \{A,B\}$, and a firm $i \in \{1,2\}$ such that $w_i^g(v) > v$ for all $v \in V$.

Intuitively this is because, if there were, firm i could simply fire all of the workers in question (i.e. set $\tilde{f}_g^i(v) = 0$ for all $v \in V$) and increase its profit. A formal argument for the case without EPSW is given in the proof of Proposition 1, and an essentially identical argument extends this observation to the case with EPSW as well.

2.2 Results

In this section, we present theoretical results from our model. Throughout, we fix an arbitrary labor market (β,F_A,F_B) and present results within this labor market.

2.2.1 Core without EPSW

PROPOSITION **1.** *Without EPSW, there exist a continuum of (non-equivalent) core outcomes. In any core outcome, almost every worker is employed and earns a wage equal to her productivity (formally, for all $i\in\{1,2\}$, all $g\in\{A,B\}$, and almost all $v\in[0,1]$: $f_g^1(v)+f_g^2(v)=f_g(v)$, and $w_i^g(v)=v$ if $f_g^i(v)>0$).*

Proposition 1 establishes that, while there are multiple core outcomes, they all feature full employment and result in wages to each worker equal to their productivity. Our proof demonstrates that this result obtains via classic "Bertrand" reasoning: in any outcome satisfying individual rationality that fails the conditions of Proposition 1 for (a positive measure set of) workers of type (g,v) one firm can outcompete the other and hire (g,v) workers in a way that increases its profit.

By Proposition 1, the wage gap and wage ratio in any core outcome O without EPSW are, respectively,

$$AW_A^O-AW_B^O=\int_0^1 vf_A(v)dv-\int_0^1 vf_B(v)dv=\mathbb{E}_A(v)-\mathbb{E}_B(v) \quad \text{and} \quad AW_A^O/AW_B^O=\mathbb{E}_A(v)/\mathbb{E}_B(v).$$

2.2.2 Core with EPSW

Now we study core outcomes under EPSW. Informally, this restriction requires that each firm pays the same wages to almost all workers it hires if it hires a positive measure of workers from both groups. Formally, we modify the definition of outcome $O_i=\{(f_g^i(v),w_i^g(v))\}_{v\in[0,1],g=A,B}$ for all $i\in\{1,2\}$ to include the following restriction:

> There exist no sets $V_g\subset[0,1]$ and $V_{-g}\subset[0,1]$ with positive Lebesgue measure such that:
>
> 1. $f_g^i(v)>0$ for all $v\in V_g$ and $f_{-g}^i(v)>0$ for all $v\in V_{-g}$, and
> 2. $\inf_{v\in V_{-g}} w_i^{-g}(v)>\sup_{v\in V_g} w_i^g(v)$.

The preceding restriction prevents a firm from employing sets of workers from both groups with positive measure (point 1) such that all workers in one set receive strictly higher pay than all workers in the other (point 2).[12] Given its symmetry across groups, EPSW implies, by transitivity, that if a firm hires a positive measure of workers from both groups, it must pay almost all workers the same wages.

The next result shows that generically firms must fully segregate by group in any core outcome under EPSW.[13] Our proof shows that any other market configuration–both firms segregate for the same group, both desegregate, or only one firm desegregates–generically cannot be supported in a core outcome. The intuition for this result can be obtained by analogy to Hotelling's model: a firm can "cater to" a group g by segregating towards g, which allows this firm to fiercely compete for these workers as it is not bound by EPSW. Both firms segregating toward some group g leads them to "Bertrand compete" away all profit, while workers of the other group could instead be employed at low wages to form a block. If only one firm desegregates, it must pay a common wage to all workers, and can thus be outcompeted for workers of the group the other firm segregates toward to form a block. If both firms desegregate, we show that the only core outcome that can potentially exist involves exactly one firm paying a wage of zero to all its workers, but that such an outcome cannot be in the core for "almost all" distribution functions F_A and F_B.

PROPOSITION **2.** *Generically, in any core outcome under EPSW, firms completely segregate. Specifically, one firm hires almost all $A-$group workers, and the other hires almost all $B-$group workers (formally, for some $i \in \{1,2\}$, $f_A^i(v) = f_A(v)$ for almost all $v \in [0,1]$ and $f_B^{-i}(v) = f_B(v)$ for almost all $v \in [0,1]$).*

Following the previous result, we assume throughout that any core outcome O under EPSW exhibits full segregation by group: without loss of generality let firm 1 hire all $A-$group workers ($f_A^1(v) = f_A(v)$ for all v) and let firm 2 hire all $B-$group workers ($f_B^2(v) = f_B(v)$ for all v). Moreover, letting $w_1(\cdot)$ specify firm 1's wages to $A-$group workers and letting $w_2(\cdot)$

[12] Note that because the above restriction must hold for every set V_g and V_{-g} of positive measure, we could equivalently state point 2 using the essential infemum and essential supremum of the wages, respectively.

[13] We consider the space of distributions F_A and F_B describing the distributions of productivities of $A-$ and $B-$group workers, respectively, that admit respective density functions f_A and f_B with respective lower bounds $\underline{f}_A, \underline{f}_B > 0$. We endow this space with the weak-* topology and consider genericity with respect to the product topology over the product set of distributions, where we say that a property holds generically if the property holds in an open and dense subset. Note that the conclusion of this result holds only generically. An example of a non-generic case in which the conclusion fails features $\beta = 1$ and $F_A(v) = F_B(v) = v$ for all $v \in [0,1]$. For this parameterization, it is straightforward to verify that there exists a core outcome where firm 1 hires all workers from both groups with $v \in [0, \frac{1}{2}]$ at wage zero while firm 2 hires all other workers at wage $\frac{1}{2}$.

specify firm 2's wages to $B-$group workers, Individual Rationality (see Remark 2) implies that it suffices to consider $w_i(\cdot):[0,1]\to[0,1]$ for each $i\in\{1,2\}$ in any core outcome. Note that we can therefore represent the wage gap and the wage ratio in a core outcome O under EPSW as, respectively,[14]

$$AW_A^O - AW_B^O = \int_0^1 w_1(v)f_A(v)dv - \int_0^1 w_2(v)f_B(v)dv, \quad \text{and} \quad AW_A^O/AW_B^O = \int_0^1 w_1(v)f_A(v)dv \Big/ \int_0^1 w_2(v)f_B(v)dv.$$

The existence of a core outcome holds generally under EPSW; consider the zero-profit outcome in which firm 1 hires all $A-$group workers and firm 2 hires all $B-$group workers, and all workers are paid wages equal to their productivities. Such an outcome is a core outcome because no firm is bound by the equal pay constraint of EPSW due to group-level segregation and because adding any new workers requires poaching workers from the competing firm, and such poaching would violate individual rationality. This outcome is similar to core outcomes without EPSW, except that the workforce is now necessarily segregated. However, the next proposition shows that there are also other core outcomes under EPSW; Indeed, there always exist a continuum of (non-equivalent) core outcomes. The proposition also shows that EPSW relatively benefits the majority group of workers, and both firms prefer core outcomes in which wage inequality is maximized.

PROPOSITION **3.** *Suppose there is an EPSW restriction.*

1. *There exist a continuum of non-equivalent core outcomes.*
2. *Let $\beta>1$. There exists one core outcome (and its equivalent outcomes) that yields the same wage gap as the (essentially unique) core outcome without EPSW. In all other core outcomes under EPSW, the wage gap is strictly larger.*
3. *Let $\beta>1$. Consider any two core outcomes. The wage gap is larger in the first outcome if and only if firm profit is higher in the first outcome.*
4. *Let $\beta\mathbb{E}_A[v]>\mathbb{E}_B[v]$. Then the statements of Parts 2 and 3 hold (for any $\beta\geq 1$) when replacing "wage gap" with "wage ratio."*

As demonstrated by Part 2 of Proposition 3, EPSW moves the wage gap in favor of the majority group (and does so strictly except for one core outcome among a continuum).

[14]The given expression for the wage ratio holds under the assumption that $w_2(v)$ is not equal to zero almost everywhere. Otherwise, by construction, the wage ratio is defined as ∞.

Moreover, Part 3 shows that larger wage gaps are associated with higher firm profits under EPSW. An implication of Part 3 is that firms prefer core outcomes that result in larger wage gaps, suggesting that a core outcome with a larger wage gap may be more likely to occur if firms can coordinate to select an outcome from the core. Part 4 replicates the findings of Parts 2 and 3 when inequality is measured via the wage ratio instead of the wage gap, under the assumption that $\beta\mathbb{E}_A[v] > \mathbb{E}_B[v]$.

When interpreting this result, we emphasize the comparative paucity of assumptions made about the relative productivity distributions of the two groups.[15] Parts 2 and 3 make no restrictions on the relative distributions, meaning that, regardless of differences in the distribution of output or in discrimination against either group, our model predicts EPSW moves the wage gap in favor of the majority group. The condition that $\beta\mathbb{E}_A[v] > \mathbb{E}_B[v]$, which is sufficient to extend the same conclusion to the wage ratio, says that the majority group is not "much" less productive (or is not "much" more discriminated against) than the minority group. While this is clearly a restriction, evidence from our empirical setting suggests that it is true: across male-majority labor markets, the pre-EPSW log wage gap is larger than one, which implies by Proposition 1 that $\beta\mathbb{E}_A[v] > \mathbb{E}_B[v]$. Across female-majority labor markets the pre-EPSW log wage gap is 0.993 while β is 1.32, implying $\beta\mathbb{E}_A[v] > \mathbb{E}_B[v]$. We provide details of our empirical approach in Section 4.

3 Search Model

In this section, we present a search model and use it to study the evolution of wages and segregation upon the enactment of EPSW. In order to consider the incentives facing agents over time, we consider a non-cooperative dynamic game. A continuum of workers enters a labor market in each period, and each worker searches at, and bargains with, firms in set I where $|I| = n \geq 1$. Each worker has a type (g,v), where $g \in \{A,B\}$ is her group identity and $v \in [0,1]$ is her productivity.

Before proceeding, we discuss several important features of our model which match our empirical context. First, we study a dynamic setting in which EPSW is a "surprise," meaning that agents initially optimize their strategies without consideration of EPSW and then unex-

[15] Contrast our distributional assumptions with those of Bharahwaj et al. (2024), who consider the measurement of wage inequality and discrimination across groups. Their Theorem 1 assumes the average productivities of the two groups are the same, while their Theorem 2 makes a weaker assumption that the average productivity of the higher-average-wage group is not strictly below that of the lower-average-wage group. The latter assumption is similar to our assumption that $\beta\mathbb{E}_A[v] > \mathbb{E}_B[v]$.

pectedly must contend with it. Second, not all firms are bound by EPSW when it is enacted; such "unconstrained" firms represent those below a binding size threshold, or those which hire predominantly fixed-term labor, for which EPSW in Chile does not bind. Third, workers' choice sets are endogenously determined by firm segregation decisions, as firms segregated for a group g will be unwilling to hire workers of group g' as doing so would expose the firm to the bite of the policy. Moreover, recent research (Card et al., 2024; Kuhn and Shen, 2023; Kuhn et al., 2020) suggests that a firm in a labor market where the written language has gendered nouns–as does Spanish, the de facto official language of Chile–are able to significantly affect the set of workers that apply for employment through use of masculine or feminine nouns in job postings (e.g. "mecánico" and "mecánica" are respectively the masculine and feminine Spanish translations of "mechanic"). We capture both of these features by allowing workers to observe firms' segregation choices before engaging in directed search. Finally, the workers entering in each period have productivities distributed according to F_A and F_B (satisfying the same regularity conditions as in Section 2), where informally, the density $f_g(v)$ denotes the proportion of $g-$group workers with productivity v. We proceed in the sequel to study the economic forces underlying our agents with this informal understanding of the meaning of our distribution functions. In Appendix B we show that this interpretation can be made rigorous in the sense that we construct measures of worker types that mimic the underlying distributions F_A and F_B.

3.1 Model Description

3.1.1 Basic Setup

There is a countably infinite number of time periods, $\{-\underline{t},-\underline{t}+1,...,-1,0,1,2,...\}$, with $\underline{t}$ being a positive integer. EPSW is introduced between the end of period -1 and the beginning of period 0. We refer to the set of periods $\{-\underline{t},-\underline{t}+1,...,-1\}$ as the "pre EPSW" periods and the set of periods $\{0,1,2,...\}$ as the "post EPSW" periods, respectively.

We study a game of perfect information, where each player observes all prior actions taken by all players. At the beginning of period $-\underline{t}$, no worker is employed by any firm. Each period $t\in\{-\underline{t},-\underline{t}+1,...,-1,0,1,2,...\}$ unfolds as follows:

1. A measure $\beta\geq 1$ of $A-$group workers and a unit measure of $B-$group workers are born. Worker productivities are distributed according to distributions F_A and F_B, respectively.

2. Newly born workers and firms are matched and bargain over wages. Details will be described in subsequent sections.

3. All newly born workers who do not come to a bargaining agreement with any firm exit the market. All other workers exit the market, each with independent probability $d \in (0,1)$.

Each worker of type (g,v) who is employed by a firm at a going wage of w receives a per-period payoff of w, and the employing firm receives an infinitesimal per-period payoff of $v-w$ from each such worker if it employs a measurable set of workers and pays a measurable wage schedule, and otherwise the firm receives a per-period payoff of -1. Unemployed workers receive per-period payoffs of zero. All agents geometrically discount future payoffs according to discount factor $\delta \in [0,1)$. Note that firms face no aggregate uncertainty due to the continuum of workers.

3.1.2 Bargaining in the Pre-EPSW Periods

Fix an exogenous order over unemployed workers.[16] Sequentially, each of these workers bargains with firms according to the following protocol, where we denote the relevant worker's type by $(g,v) \in \{A,B\} \times [0,1]$:

1. The worker decides which of the firms to bargain with.

2. The worker and firm engage in an alternating-offer bilateral bargaining game, with the firm being the first proposer. If the proposer's offer is rejected, then there is probability $1-\delta' \in (0,1)$ that the bargaining breaks down, and with the complementary probability bargaining continues with the other player becoming the proposer. Additionally, if the worker's and the firm's strategies are such that the proposer's offer is rejected in all future rounds of bargaining, then we also say that the bargaining breaks down without an agreement being reached.

3. If bargaining breaks down with a particular firm, then the worker selects another eligible firm with which to bargain (the worker is not eligible to match with a firm she has previously bargained with). If there are no eligible firms remaining, the worker permanently remains unmatched and obtains a payoff of zero.

[16] Formally, let $\mathcal{W}^t$ denote the set of newly born workers at period t. Within each period t, there are uncountably many "moments" of time comprising the set $[0,1]$. Let there be a bijection $\varphi : [0,1] \to \mathcal{W}^t$ such that at each moment $y \in [0,1]$ of period t, the worker $\varphi(y) \in \mathcal{W}^t$ engages in search and bargaining according to the specified protocol. This construction and the specific timing of bargaining is not crucial for our main results. See Remark 4 for an alternative static bargaining protocol.

3.1.3 Bargaining in the Post-EPSW Periods

When EPSW is introduced between the end of period -1 and the beginning of period 0, there are $n_U \geq 1$ exogenously fixed firms that are "unconstrained" by the policy, and the remaining $n_C = n - n_U$ "constrained" firms are initially desegregated.[17] In each period $t \in \{0,1,...\}$, the actions available to unconstrained firms do not change, while for constrained firms, the following step is added immediately prior to Step 1 of the timing of the game in Section 3.1.1.

0:

(a) Following an exogenously given order, each currently desegregated firm sequentially chooses from the following options: segregate for group A, segregate for group B, or desegregate at any endogenously selected wage $w \in [0,1]$.[18] The decision to segregate is irreversible, i.e. once the firm has segregated for a group g, then the firm continues to be segregated for group g in all subsequent periods. The (de)segregation decision potentially involves job separation and wage adjustments:

- If a firm is segregated for group g, then the firm is immediately unmatched from all previously employed workers with group identity $g' \neq g$.
- If a firm is desegregated at wage w, then all previously employed workers with productivity $v < w$ and all previously employed workers earning wage $w' > w$ are immediately unmatched from the firm. The wages of all remaining workers are immediately increased to w.

(b) Each worker who is separated from their previous employer at this stage exits the market, and with probability $\rho \in [0,1]$ is replaced by a new unemployed worker of the same type.[19]

[17] As mentioned earlier, some firms are unconstrained by EPSW in our empirical setting. All of our results also hold under an alternative assumption that there exists a firm that is already segregated for each group $g \in \{A,B\}$ at period 0. In our empirical setting, 34.3% of firms are already segregated at the time of announcement of EPSW.

[18] The assumption that the firms' segregation decisions are sequential is made for tractability. Specifically, this assumption gives each firm i perfect information about the (de)segregation decisions of others when i makes its (de)segregation decision in any period $t > 0$. In the proof of Proposition 4, we show how the presence of perfect information leads to the existence of a pure-strategy subgame perfect Nash equilibrium, following the well-known existence result of Harris (1985).

[19] $\rho \in \{0,1\}$ corresponds, respectively, to markets without and with replacement, which are commonly considered settings in the existing literature (for examples of search models with replacement, see Burdett and Coles (1999), Bloch and Ryder (2000), Chade (2006), and Cullen and Pakzad-Hurson (2023)). We allow for the possibility of "partial" replacement, $\rho \in (0,1)$, for generality.

Moreover, compared to the case without EPSW, the bargaining protocol is restricted as follows:

- A worker with group identity g is not eligible to be matched with any firm that is segregated for group $g' \neq g$. Similarly, a worker with productivity v is not eligible to be matched with any desegregated firm whose wage offer is strictly higher than v.[20]
- Any firm that desegregates at some wage w in the current period must always offer workers wage w in the current period and will accept a worker's offer in the current period if and only if it is w.

3.2 Preliminary Analysis

As mentioned earlier, we assume that the introduction of EPSW is a surprise to firms and workers. Specifically, we assume that the equilibrium play of the game for periods $\{-\underline{t},...,-1\}$ follows a strategy profile that is a subgame perfect Nash equilibrium in a game in which there is no EPSW for all periods $\{-\underline{t},...,-1,0,1,...\}$. ESPW is announced and introduced between the end of period -1 and the beginning of period 0, and the game proceeds under EPSW thereafter. Thus we assume that, for any history at the beginning of period 0, the continuation equilibrium play for periods $\{0,1,...\}$ must be a subgame perfect Nash equilibrium of a subgame beginning at that history.

To analyze equilibrium strategies in our game, consider a period t and consider a newly born worker of some type $(g,v) \in \{A,B\} \times [0,1]$. Let D_t^v be the set of all desegregated firms that set a wage $w_i(t) \leq v$ in period t. Throughout, we adopt the expositional convention that all firms are unconstrained in pre-EPSW periods $t<0$. Therefore, $D_t^v = \emptyset$ for all v if $t<0$. Let $T_i^v \in \mathbb{N} \cup \{\infty\}$ be the number of periods until $i \in D_t^v$ fires the relevant worker, i.e. period $T_i^v + t$ is defined as the first period in which either i segregates for group $g' \neq g$ or i desegregates at some wage $w_i(T_i^v + t) > v$ or some wage $w_i(T_i^v + t) < w_i(\tau)$ for some $\tau \in \{t, t+1, ..., T_i^v + t - 1\}$. Therefore, T_i^v is an equilibrium object for workers born in periods $t \geq 0$, and $T_i^v = \infty$ for all

[20] Here, we are directly prohibiting a worker from being matched with firms that are segregated for the opposite group. This assumption is made only for convenience. If we assume that there is a sufficiently large penalty for firms hiring from both groups while paying different wages, then no such matching occurs in equilibrium. In a similar spirit, no worker is matched to a desegregated firm whose wage is higher than her productivity in equilibrium.

workers born in periods $t<0$.[21] Define

$$V^v := \sum_{\tau=1}^{\infty} \delta^{\tau-1}(1-d)^{\tau-1} v, \quad \text{and} \quad W_i^v := \sum_{\tau=1}^{T_i} \delta^{\tau-1}(1-d)^{\tau-1} w_i(t+\tau-1). \tag{1}$$

V^v and W_i^v represent, respectively, the total expected surplus generated by the worker if she is never fired and her total expected earnings if she accepts a job at desegregated firm i. Let $W^v = \max\{W_i^v | i \in D_t^v\}$ if $D_t^v \neq \emptyset$ and $W^v = 0$ otherwise. Note that $V^v \geq W^v$ by definition, i.e. W^v serves as the worker's "outside option" that she has access to by working at a desegregated firm. Implicit behind (1) is a "Markovian" assumption on firm strategies. Specifically, we focus on subgame perfect equilibria such that a firm's decision to segregate for some group g or to desegregate at some wage w at any time t does not depend on the specific "bargaining path" that any single worker takes before reaching an agreement with that firm.

In general, a wide range of bargaining outcomes can be consistent with subgame perfect Nash equilibrium, but we view some of them as being unrealistic. First, our model contains a continuum of infinitesimally small workers. This implies that firms can "throw their weight around" in potentially unrealistic ways because the surplus obtained through bargaining with any single worker is always zero. Suppose, for instance, a firm's strategy were to propose a wage of 0 and accept only a wage of 0 when bargaining with any worker, and a worker's strategy were to always offer and accept a wage of 0. If each worker made a strictly positive impact on firm profit, it would be farfetched for an equilibrium to contain such a strategy: if the worker were to propose a sufficiently small positive wage, the firm would rather accept that wage offer and obtain positive surplus from this worker at the present round rather than risking a breakdown of the bargaining with that worker. However, our continuum model makes the firm willing to forgo the infinitesimal surplus from bargaining with any one worker, so the above behavior does not formally violate sequential rationality. Second, some equilibria may not feature stationarity. Specifically, strategies employed by players may depend on the specific bargaining outcome struck between a ("small") worker and a single firm in the past.

To avoid such nuisances, we impose a refinement of subgame perfect Nash equilibria as follows. Let x denote the equilibrium offer that a worker makes to the firm as the worker's own share, and y denote the equilibrium offer that a firm proposes to the worker as the worker's share. Denote the total surplus between the firm and the worker by V^v, namely

[21] $T_i^v = \infty$ for all firms i and all workers born in periods $t<0$ due to the surprise feature of EPSW, that is, agents do not anticipate that any worker will ever be fired (workers still exit with probability d due to exogenous separation in each period).

as defined in (1), and let $\bar{V}^v$ denote the worker's continuation value if the bargaining breaks down with the present firm. In a bargaining model between a firm and a non-negligible worker, sequential rationality would require that x and y satisfy

$$y=\delta' x+(1-\delta')\bar{V}^v, \qquad\qquad V^v-x=\delta'(V^v-y), \tag{2}$$

and the firm and the worker accept the offers if and only if the offers are at least x and y specified by the above equations, respectively. We require the same: any equilibrium strategy profile is such that the bargaining between any firm and worker follows (2) and the acceptance rules as described above both on and off path. Solving these equations, we obtain

$$x=\frac{1}{1+\delta'}V^v+\frac{\delta'}{1+\delta'}\bar{V}^v, \qquad\qquad y=\frac{\delta'}{1+\delta'}V^v+\frac{1}{1+\delta'}\bar{V}^v.$$

Now, let there be m firms that are either segregated for the worker's group g or unconstrained, and let W^v be as defined in (1). Then, consider the strategy of the worker such that the worker bargains with the m segregated or unconstrained firms and then bargains with a desegregated firm with wage W^v if $D_t^v\neq\emptyset$. Writing V_j^v to be the expected payoff for the worker immediately before beginning bargaining with the j'th segregated firm, we have

$$V_j^v=\frac{1}{(1+\delta')^{m-j+1}}W^v+\left(1-\frac{1}{(1+\delta')^{m-j+1}}\right)V^v, \tag{3}$$

and thus, in particular,

$$V_1^v=\frac{1}{(1+\delta')^{m}}W^v+\left(1-\frac{1}{(1+\delta')^{m}}\right)V^v. \tag{4}$$

Note that the payoff V_1^v in (4) is weakly increasing in m, and strictly so if and only if $W^v<V^v$. There is a unique per-period wage $w\in[0,v]$ that corresponds with expected worker lifetime earnings specified in (4).[22]

Note that the above ordering, namely to bargain with the m segregated and unconstrained firms and then bargain with a desegregated firm with wage W^v if $D_t^v\neq\emptyset$, is an optimal ordering

[22]To see this, note that $V_1\in[0,\sum_{\tau=1}^{\infty}\delta^{\tau-1}(1-d)^{\tau-1}v]$ because $0\leq V_1\leq V^v=\sum_{\tau=1}^{\infty}\delta^{\tau-1}(1-d)^{\tau-1}v$, where the first inequality follows because both terms in (4) are weakly positive, the second inequality follows because $W^v\leq V^v$, and the equality follows from (1). Therefore, the worker's expected lifetime earnings, as a function of the per-period wage w, is $\sum_{\tau=1}^{\infty}\delta^{\tau-1}(1-d)^{\tau-1}w$, which equals 0 for $w=0$, equals $V=\sum_{\tau=1}^{\infty}\delta^{\tau-1}(1-d)^{\tau-1}v$ for $w=v$ and is strictly increasing and continuous in w over the interval $[0,v]$. Therefore, there is a unique per-period wage $w\in[0,v]$ that corresponds with expected worker lifetime earnings specified in (4).

for the worker. To see this, let i be the first desegregated firm in the worker's bargaining order (if such a firm exists) with the property that, conditional on failing to reach a bargaining agreement with all firms ordered before it, the worker and firm reach an agreement. Let $m' \leq m$ be the number of desegregated or unconstrained firms ordered before i, and let $W' \leq W^v$ be the wage set by firm i if such firm i exists, and let $m' = m$ and $W' = 0$ if no such firm i exists. Then, by the same argument as before, on path the worker reaches an agreement with the first firm (potentially excepting desegregated firms other than i) and obtains expected surplus

$$\frac{1}{(1+\delta')^{m'}} W' + \left(1 - \frac{1}{(1+\delta')^{m'}}\right) V^v, \tag{5}$$

where we note that the set of desegregated firms ordered before i does not affect this surplus because they will not reach an agreement with the worker. Because $m' \leq m$ and $W' \leq W^v$ by definition of m and W^v, this expression is no larger than (4), showing the optimality of the indicated order.

We focus on equilibria in which worker search is affected by firms' (de)segregation choices, and not by other aspects of the history of play. In particular, we exclude strategies in which workers coordinate such that they jointly avoid searching first at some firm i, as this prevents firm i from hiring workers regardless of its segregation decision. To formally define our refinement, for $t \geq 0$, let I_t^g be the set of unconstrained firms and firms segregated toward group g by the hiring stage of period t, and for $t < 0$ let $I_t^g = I$. We assume there is a *search intensity* $r = (r_i^g)_{i \in I, g \in \{A,B\}} \in (0,1)^{2n}$ such that at any time t and any history of play, for any firm $i \in I_t^g$, and for each productivity v, $\frac{r_i^g}{\sum_{i' \in I_t^g} r_{i'}^g}$ share of the entering workforce of type (g,v) first elect to bargain with firm i.

In the following, we focus on pure-strategy subgame perfect Nash equilibria with the above restrictions on bargaining and search intensity (and the aforementioned "surprise" feature of EPSW), and refer to them simply as *equilibria.*

REMARK **3.** Our model does not explicitly place any constraints on the measure of workers each firm can hire in each period. One relevant consideration may be whether segregation and wage setting decisions are artificially driven by "unlimited" hiring capacities. We do not believe this is the case.

First, note that the unconstrained firms mechanically hire at least a fixed fraction of the workforce, as determined by the vector r of search intensities. Letting I^U represent the set of unconstrained firms, no firm $i \notin I^U$ ever hires more than $\frac{r_i^g}{\sum_{i' \in I^U \cup \{i\}} r_{i'}^g}$ fraction of workers of

group g in any time period t (i.e. even if i is the only firm segregated toward group g). Fixing r_i^g for each $i \notin I^U$ and each $g \in \{A,B\}$, increasing $r_{i'}^g$ for each $i' \in I^U$ and each $g \in \{A,B\}$ mechanically lowers the maximum hiring rate of any firm $i \notin I^U$. Of note is that increasing search intensities as described above decreases the measure of workers hired by i in a given period, regardless of the segregation patterns of the remaining firms.

Second, consider a more standard capacity interpretation: hiring is constrained in equilibrium if and only if "too many" workers seek positions at a firm. Specifically, let $\kappa = (\kappa_i^g)_{i \in I, g \in \{A,B\}} \in (0,\infty)^{2n}$ and consider equilibria in which each firm i is to hire no more than κ_i^g measure of $g-$group workers in each period on path. We assume that unconstrained firms jointly have excess capacity: $\sum_{i \in U} \kappa_i^A \geq \beta$ and $\sum_{i \in I^U} \kappa_i^B \geq 1$. To accommodate hiring limits, the search intensity at any time t (i.e. the measure of workers of each (g,v) type who initially search for each firm $i \in I_t^g$) is proportionally determined by weights r_i^g such that no firm is initially searched for by more than κ_i^g $g-$group workers.[23] Our results are qualitatively similar given such a modification to our model.

REMARK **4.** (3) presents the equilibrium split of total surplus V^v accruing to each of the firm and the worker given a worker "outside option" of W^v (which she earns at the top-paying de-segregated firm) in the equilibrium of our alternating-offer bargaining protocol. One question may be, given the finding of Binmore et al. (1988), whether the alternating-offer solution presented has an analogue as the solution to Nash bargaining (or a variation thereof).

While we note that Nash presents a cooperative solution to the bargaining game and our model necessarily studies bargaining in the context of a non-cooperative game, it is possible to nest Nash-like bargaining into our model a la Horn and Wolinsky (1988). Specifically, this "Nash-in-Nash" solution takes the following structure: the "available" surplus $V^v - \bar{V}^v$ between a segregated or unconstrained firm and worker is fractionally split, assuming that the outcome between the worker and each subsequent firm she potentially bargains with follows a fractional split, where $\bar{V}^v$ is the "disagreement point" that the worker receives if bargaining breaks down with the present firm, i.e. it is the surplus she accrues if bargaining were to break down with the current firm but with no subsequent firms. Agent strategies, which take as fixed the outcome of any bilateral bargaining as described above, must satisfy

[23] Formally, for each firm $i \in I_t^A$, the fraction of $A-$group workers that first search at firm i is $\hat{r}_i^A(x^A) := \min\{\kappa_i^A/\beta, r_i^A x^A\}$ where x^A solves $\sum_{i \in I_t^A} \hat{r}_i^A(x^A) = 1$ and for each firm $i \in I_t^B$, the fraction of $B-$group workers that first search at firm i is $\hat{r}_i^B(x^B) := \min\{\kappa_i^B, r_i^B x^B\}$ where x^B solves $\sum_{i \in I_t^B} \hat{r}_i^B(x^B) = 1$. Notice that, for each $g \in \{A,B\}$, there is a unique solution for x^g: the function $\sum_{i \in I_t^g} \hat{r}_i^g(x^g)$ equals 0 at $x^g = 0$, and is continuous and strictly increasing in x^g until the capacity constraints of all firms bind, i.e. the minimum in the definition of $\hat{r}_i^g(\cdot)$ is achieved by the first element for each group g and firm i.

subgame perfection in the standard way.

Let $\Delta \in (1/2,1)$ denote the share of the available surplus that accrues to a segregated or unconstrained firm when bargaining with a worker. Let $\bar{\Delta}(\delta') := \frac{1}{1+\delta'}$ for any $\delta' \in (0,1)$. Note that $\bar{\Delta}(\cdot)$ is continuous and strictly decreasing, the limit of the function as $\delta' \to 0$ is 1, and the limit of the function as $\delta' \to 1$ is $1/2$. These observations together imply that there is a unique value of δ' such that $\bar{\Delta}(\delta') = \Delta$. By (3), each bargaining outcome in the "Nash-in-Nash" model corresponds to the outcome in equilibrium of our non-cooperative model for appropriately chosen Δ and δ'.

There is nothing in the "Nash-in-Nash" procedure that suggests a surplus split $\Delta \in (1/2,1)$ is more viable than a split $\Delta \in (0,1/2)$. Any split $\Delta \in (0,1/2)$ can be obtained by setting the first mover in any bargaining situation to be the worker instead of the firm, and appropriately selecting δ'.

3.3 Results

In this section, we present the main results arising from our search model. Our first such result, Proposition 4, makes three points. First, for any $r \in (0,1)^{2n}$, the set of equilibria with search intensity r is non-empty and in each equilibrium every newly born worker is hired.

Second, all firms eventually segregate in any equilibrium. Moreover, there is a uniform bound $T \geq 0$ such that in any equilibrium, all firms have segregated by time T. For any $t \geq 0$, let n_g^t, $g \in \{D,A,B\}$ represent the number of firms that are desegregated, segregated toward A, and segregated toward B, respectively, at the end of time period t. Therefore, $n_D^{t+1} = n_D^t (= 0)$, $n_A^{t+1} = n_A^t$, and $n_B^{t+1} = n_B^t$ for any $t \geq T$ in equilibrium. We will denote those constant values by n_D,n_A, and n_B, respectively. Note that because every newly born worker is hired, there is full employment for workers born after period T when all firms have segregated. Thus, the only possible source of unemployment is worker displacement in periods $t \in \{0,...,T\}$ as firms make (de)segregation choices. These displaced workers will be "rehired" at rate ρ, and therefore, full employment is achieved for all t if $\rho = 1$.

Third, the long-run wage ratio moves in favor of the group toward which more firms segregate in equilibrium. For any t, we define the wage ratio (or log wage gap) at time t to be the average wage of $A-$group workers over the average wage of $B-$group workers who are employed at the end of time t. In the pre-EPSW periods, in equilibrium, every worker is hired by the first firm she bargains with, and therefore both groups' wages follow (4) with $m = n$ and $W^v = 0$ for all $v \in [0,1]$, which implies:

OBSERVATION **1.** *The wage ratio is* $\mathbb{E}_A[v]/\mathbb{E}_B[v]$ *in any period* $t<0$.

We say that the wage ratio is *more in favor of* group A (B) in period t compared to period t' if the wage ratio is higher (lower) in period t than in period t'. Part 2 of Proposition 4 shows that the the key determinant of the wage ratio in the long run–that is, the wage ratio between any pre-EPSW period and any sufficiently large period since the introduction of EPSW–is the relationship between n_A and n_B: if $n_A > n_B$ then the wage ratio is more in favor of group A in the long run compared to pre-EPSW, and vice versa, and if $n_A = n_B$, then EPSW has no impact on the long-run wage ratio.

PROPOSITION **4.** *For any* $r \in (0,1)^{2n}$*, the set of equilibria with search intensity* r *is nonempty. In any equilibrium, every newly born worker is hired. Moreover, there exists* $T \geq 0$ *such that in any equilibrium:*

1. $n_D^t = 0$ *for all* $t \geq T$*, and*
2. *If* $n_A > n_B$*, then the wage ratio is more in favor of group* A *in any period* $t \geq T$ *when compared to the wage ratio in any period* $t' < 0$*. If* $n_B > n_A$*, then the wage ratio is more in favor of group* B *in any period* $t \geq T$ *when compared to the wage ratio in any period* $t' < 0$*. If* $n_A = n_B$ *then the wage ratio in period* t *converges to the wage ratio in any period* $t' < 0$ *as* $t \to \infty$*.*

Part 2 of Proposition 4 reveals that the number of firms segregating for each group is the key determinant of EPSW's effect on the long-run wage ratio. While it may seem intuitive for $n_A \geq n_B$ to hold if there are more $A-$group workers than $B-$group workers, this condition does not necessarily hold in equilibrium. Three forces can counteract this intuition. First, the minority group may be more productive than the majority: if $\beta \mathbb{E}_A[v] < \mathbb{E}_B[v]$ then the majority group is "less valuable" to firms than the minority group, potentially incentivizing more firms to segregate for group B. Second, worker search intensity may be skewed: if too many $A-$group workers search for some firm i (i.e. r_j^A is small compared to β for all $j \neq i$) while $B-$group workers search more equitably across firms, any firm $j \neq i$ that segregates in favor of group A hires only a small fraction of workers if firm i also segregates toward group A, or is unconstrained. Thus, skewed search intensity could incentivize more firms to segregate for group B. Third, firms may face high reallocation frictions: if d is low, then firms retain a large fraction of their workforce across periods implying that they have "more to lose" in competition with other firms. Therefore, a firm may segregate toward group B in equilibrium.

Therefore, a policy-relevant consideration is under what conditions $n_A > n_B$ holds. Indeed, we note a near-ubiquitous empirical pattern in Section 4.4— more than 94% of labor markets feature more firms segregating toward the gender that makes up a majority of the workforce post-EPSW—suggesting that there are realistic cases under which equilibria feature more segregation toward the majority group. Proposition 5 below considers equilibria with *equitable search* defined as an equilibrium with search intensity $r_i^g = 1$ for every $i \in I$ and $g \in \{A,B\}$, and makes two points. First, given small reallocation frictions, i.e. large d, a near equal profit condition arises whereby all firms receive similar profit for sufficiently large t, approximating the equal profit condition that arises in any equilibrium of our static game. To define what we mean by "near equal profit" let $x_A, x_B \in [0, n_C]$ be real numbers such that

$$\frac{1}{x_A + n_U} \cdot \frac{1}{(1+\delta')^{x_A+n_U}} \cdot \beta \mathbb{E}_A[v] = \frac{1}{x_B + n_U} \cdot \frac{1}{(1+\delta')^{x_B+n_U}} \cdot \mathbb{E}_B[v], \tag{6}$$

$$x_A + x_B = n_C \tag{7}$$

if such numbers exist. If the left-hand side of (6) is larger than the right-hand side for every $x_A \in [0, n_C]$, then let $x_A = n_C$. If the left-hand side of (6) is smaller than the right-hand side for every $x_A \in [0, n_C]$, then let $x_B = n_C$. If $n_A = x_A$, then (4) implies that] all firms will earn equal profit from workers born in periods $t \geq T$, where T is the time period after which all firms have segregated in equilibrium. If x_A is not an integer, then it is impossible to have exactly $n_A = x_A$ but if n_A is "close to" x_A, then all firms will earn approximately equal profit from workers born in periods $t \geq T$. Proposition 5 shows that for sufficiently large d (i.e. for any sufficiently large values of d that depend on other parameters but not on the selection of a specific equilibrium) in any equilibrium with equitable search, n_A is "close to" x_A (Part 1).[24] If $\beta \mathbb{E}_A[v] > \mathbb{E}_B[v]$ also holds, then $A-$group workers are collectively more valuable than $B-$group workers, implying by (6) that $x_A > x_B$, and we show that $n_A \geq n_B$ must hold (Part 2).

PROPOSITION **5.** *There exists $d^* < 1$ such that for all $d > d^*$, and in any equilibrium with equitable search:*

1. *$n_A \in [x_A - 1, x_A + 1]$ and $n_B \in [x_B - 1, x_B + 1]$.*
2. *Let $\beta \mathbb{E}_A[v] > \mathbb{E}_B[v]$. Then $n_A \geq n_B$.*

REMARK **5.** As previously discussed in Section 2, we have empirical evidence to support that $\beta \mathbb{E}_A[v] > \mathbb{E}_B[v]$ holds. Without the equitable search assumption, the conclusion of Part

[24] We note that x_A is not a function of d.

2 of Proposition 5 holds under more stringent conditions. Specifically, for any given search intensity r, an essentially identical argument as that presented in the Proof of Proposition 5 shows that $n_A \geq n_B$ (and also $n_A > n_B$) holds for any sufficiently large $\beta \geq 1$.

Another empirically motivated question is to what extent firms that are "nearly" segregated upon EPSW enactment will quickly switch to segregation. In our model, the fraction of the employed workforce at each firm at $t = 0$ is driven by search intensities. For distinct $g,g' \in \{A,B\}$, define $R_g^i := r_i^g / r_i^{g'}$ to be the ratio of search intensities. Then a firm will be "nearly" segregated toward group g' at the beginning of period 0 if R_g^i is sufficiently low. The following remark presents conditions under which "nearly" segregated firms elect to segregate sooner than all other firms. When the discount rate is sufficiently low, firms become myopic, meaning that the cost of remaining desegregated–separating with some existing workers in both groups, increasing the pay of almost all other employed workers, and being unable to hire new workers–exceeds the costs of segregating–separating with a small measure of employed workers from the "wrong" group g. The proof is straightforward, and therefore omitted.

PROPOSITION **6.** *There exist $\delta^* \in (0,1)$ and $R^* \in (0,\infty)$ such that if the discount factor is $\delta < \delta^*$ then in any equilibrium with search intensity r, any unconstrained firm i with $R_g^i < R^*$ for some $g \in \{A,B\}$ will segregate toward group $g' \neq g$ at $t = 0$.*

4 Empirical Analysis

4.1 Institutional Background

Chile is an OECD country with nearly 20 million inhabitants. The Chilean labor market is relatively concentrated in the formal employment sector; the informal labor market share around the time of the policy was 25%, the lowest in Latin America (Gasparini and Tornarolli, 2007). Only 10% of the (formal) workforce is unionized, and only union members are covered by collective bargaining agreements, implying that Chilean firms plausibly have a high degree of wage-setting power. Workers can be fired without cause and without notice at the cost of one month's wages.

The gender wage gap in Chile is similar to that in the United States. Chilean female workers earn 18–23% less than their male counterparts (Perticará and Bueno, 2009). Female labor force participation was roughly 30% in 2009, which is lower than in many OECD and other Latin American countries (Verick, 2014).

4.2 EPSW implementation

In June 2009, Law 20.348 was signed as an amendment to Chile's labor code with the General Secretary of the Chilean Senate declaring, "The main objective of the initiative is to establish the right to equal remuneration between men and women for the provision of services of similar value." We refer throughout to June 2009 as the time of announcement. The law took effect in November 2009, which we refer to as the time of enactment. An important part of the discussion and debate surrounding the law was providing a definition of "similar work." The law specifies that a firm cannot pay a man and a woman different wages for "arbitrary reasons"; pay differences across genders are allowable only if workers fall into different coarse categories based on skills, qualifications, suitability, responsibility, or productivity. Firms that do not comply are subject to sizable monetary fines per offense, as we discuss below. The law also establishes a 10% discount for any other labor fines a firm is subject to if it pays men and women the same wages for "similar jobs and responsibilities." We therefore classify the law for our purposes as (a gender-based) EPSW.[25]

The law has different consequences for firms of different sizes, based on the number of a firm's workers with long-term employment contracts.[26] Firms with 10 or more long-term workers are required to explicitly have a grievance procedure for gender-based pay discrimination. Workers in firms above this threshold who allege the firm has violated Law 20.348 must receive a sufficient response from the firm within 30 days. If no such response is received, the worker can file a complaint at the Labor Inspection Office or can directly raise the issue with a labor court. Financial penalties also differ per infraction by firm size. Firms with 10–49 long-term workers found to violate the law are subject to a fine of 69-1,384 USD per worker-month of violation, while firms with fewer than 10 long-term workers are not subject to a financial penalty. [27]

[25]Guideline 1187/018 published in April 2010 by the Directorate of Labor clarifies that 1) the law does not bind within gender group, and 2) that a firm paying even a single man more than a single woman (or vice versa) despite both performing similar work is in violation of the law.

[26]The vast majority of workers in Chile have either long-term contracts (no end date is specified ex ante) or fixed-term contracts (an end date is specified ex ante, although such contracts are automatically transitioned into long-term contracts if the worker continues to be employed beyond the contract end date). EPSW protections for workers of different contract types within the same firm are identical. Nearly half of firm-months in our dataset contain no workers with fixed-term contracts.

[27]Cruz and Rau (2022) further discuss how the law imposed mandatory transparency guidelines on worker roles in the firm, and additional fines for violations, for firms with at least 200 long-term workers. They show that the disclosure policy reduced the gender wage gap through a bargaining channel. Our analysis avoids firms treated by this additional policy, due to the potential confounding equilibrium effects of transparency policies (for further discussion on potential equilibrium effects, see Cullen and Pakzad-Hurson, 2023).

Initial evidence suggests that the law was both widely known to workers, and enforced. In a 2013 governmental survey,[28] 11% of respondents stated that they know someone who had complained using the law. Through a public-records request, we found that 9,577 complaints were filed by workers alleging violations of Law 20.348, 9,723 inspections were carried out by the government, and that 489 individual firms were punished. The average fine amount was 1,167 USD per violation (each worker-month of unfair pay is a separate violation of the law).

4.3 Data

We study the effects of EPSW using matched worker-firm administrative data from the Chilean unemployment insurance system from January 2005 to December 2013. We observe a random sample of firms, stratified by size, totaling roughly 4% of all firms. In our data, an observation is a worker-firm-month. For each observation, the data include worker pay[29] and demographic information including gender, education level, contract status, age, and marital status; additionally, we observe the firm's geographic location and industry code. We observe the entire employment history of each worker ever employed at a sampled firm. We discuss further details of our dataset in Appendix C.4.

Firms in our sample are typically small, with a median of 7 concurrent workers. However, there are outliers. Following Bennedsen et al. (2022); Böheim and Gust (2021); Duchini et al. (2022); Gulyas et al. (2023) we only consider firms of similar sizes at the time of policy announcement in order to limit size-based wage dynamics. In our main specifications, we consider firms with at least 6 and no more than 13 total workers at announcement, which account for 40% of all active firms.

In Table 1, we present descriptive statistics for (column I) the set of firms prior to the size restriction, (column II) the set of firms following the size restriction, (column III) the set of firms with that will form our control group, and (column IV) the set of firms following the size restriction that are active over our entire period of analysis.

4.4 Identification Strategy: Mapping Theory into Empirics

Based on our theoretical findings, we investigate the effect of EPSW on gender segregation within firms and on the differential pay of men and women. We make several conservative

[28] See `https://www.evaluaciondelaley.cl/wp-content/uploads/2019/07/ley_20348_igualdad_remuneraciones.pdf`.

[29] Worker monthly pay at each firm is top coded in our data. The threshold for top coding varies over time; in June 2009, the top-code threshold was roughly $3,550. The share of observations in our data that are top coded in June 2009 is 0.07%.

Table 1: Descriptive Statistics

	(I) All firms	(II) Estimation Sample	(III) Estimation Sample, Control	(IV) Estimation Sample, Balanced
Panel (a): Employees				
Average age in June 2009	32.418	32.180	32.186	32.123
Share with tertiary education	0.124	0.124	0.120	0.128
Share male	0.693	0.734	0.736	0.732
Share married	0.341	0.337	0.336	0.336
Share with residence in Santiago Region	0.369	0.389	0.380	0.391
Share in female majority industry-county in June 2009	0.174	0.154	0.155	0.155
Number of workers	955,602	473,991	435,485	311,046
Panel (b): Firms				
Average size in June 2009	12.207	8.685	8.303	8.754
Share in Santiago Region	0.484	0.553	0.539	0.567
Share in Agriculture	0.101	0.091	0.095	0.100
Share in Manufacturing	0.088	0.132	0.126	0.165
Share in female majority industry-county in June 2009	0.137	0.149	0.153	0.162
Number of firms	25,436	6,551	5,791	3,418

Notes: This table displays summary statistics for the different samples used in the paper. The unit of our panels is the worker-firm-month. In Panel (a), we display figures about the workers present in our data. In Panel (b), we display figures about the firms in our data. In column I, we display figures corresponding to all firms. In column II, we display figures for the data in column I restricting to (workers of) firms that employed between 6 and 13 workers in June 2009. In column III, we display figures for the data in column I restricting to (workers of) firms that employed between 6 and 9 workers in June 2009. In column IV, we display figures for the data in column II after further restricting to (workers of) firms that are in our data in every month between January 2005 and December 2013. Column II is the sample we use in the majority of our analysis.

choices, which we describe presently, in an attempt to provide a plausible, causal lower bound on the effect of EPSW on our outcomes of interest.

To obtain the causal effect of EPSW on our outcomes of interest, we consider an event-study analysis based on firm size *at the time of policy announcement*; firms are considered "treated" if they were subject to EPSW at announcement (i.e. the firm in question employed at least 10 long-term workers in June 2009) and "control" otherwise.[30] This yields an "intent-to-treat" framework as we do not adjust our designation of firm treatment status based on firm choices (e.g. to fire workers to escape the bite of EPSW).

There are two standard assumptions for our approach to yield causal estimates of EPSW.

[30] Following our analysis in Section 3, the effects of EPSW on a firm are predicted to be smaller in magnitude for untreated firms; they are unconstrained by EPSW, obviating the benefits of segregation, and thus also avoiding changes in wages for workers who were hired prior to EPSW. We discuss this point further in Appendix C.1.

First, we assume the policy is not anticipated prior to its announcement. Since we use the time of policy announcement to designate treatment status, "no anticipation" is plausible in our context. However, the choice to use the time of announcement to determine treatment status serves to attenuate our estimates because some firms in our "treatment" group will not be bound by EPSW and some firms in our "control" group will be in future periods. Therefore, this decision likely contributes to our analysis understating the effect of EPSW on increased segregation rates and wage gaps. We present and discuss supportive evidence for our designation of treatment status in Appendix C.1. Appendix C.1 also presents "placebo" tests which do not find statistically or economically meaningful effects at alternative firm size thresholds, supporting that the observed effects of EPSW around the size threshold specified in the law are plausibly causal. Second, we assume that outcomes of treated versus control units follow parallel trends in the absence of EPSW. As discussed above, treatment status is plausibly as good as random, and we show balance between treatment and control firms across several observable dimensions in Table 1 which makes the parallel trends assumption plausible in our context. In our analysis below, we discuss a partial validation of this assumption by showing a lack of pre-trends between our control and treatment groups.

Our theoretical analyses predict an increase in segregation for firms that are treated by EPSW. Importantly, our theoretical predictions hold within sets of "similar" workers, which leaves open the possibility that a firm could be segregated within, but not across, roles. For example, a firm could be segregated for the purposes of EPSW if all of its custodians are men and all of its lawyers are women. We do not observe which workers perform "similar" work within a firm. Therefore, we define our segregation measure as complete firm-level gender segregation, i.e. only workers of one gender are employed by the firm. This is intuitively the strongest measure of segregation, implying the gender segregation of every set of similar workers within the firm, but it potentially misses firms that are gender segregated by, but not across, roles. Therefore, the causal effect of EPSW on segregation that we present is plausibly a lower bound. In Appendix C.2 we discuss how wage effects in non-fully-segregated firms are consistent with the presence of segregation by role.

We also predict EPSW causes a divergence in wages across groups, depending on features of the local labor market (LLM) in question. An important question is, therefore, how we define LLMs. In our dataset, there are 21 industry codes and 16 geographic regions, which are further disaggregated into 321 counties. We define the LLM a firm operates in by the firm's geographic county and industry code pair. We augment our upcoming analysis by allowing wages to trend differently for workers in different LLMs, capturing aggregate changes in

productivity for workers in different sectors of the economy.[31]

Proposition 3 finds that EPSW increases the wage gap in favor of the majority group, and that it increases the log wage gap in favor of the "more productive" group (recall that group A is collectively more productive if $\beta\mathbb{E}_A[v]>\mathbb{E}_B[v]$). We identify the majority gender in a LLM at the time of policy announcement, and we find evidence that the condition $\beta\mathbb{E}_A[v]>\mathbb{E}_B[v]$ holds across LLMs. Specifically, and as we discuss at the end of Section 2, our upcoming regression analysis predicts the (residual) log wage gap pre-policy—which equals $\mathbb{E}_A[v]/\mathbb{E}_B[v]$, see Observation 1—is greater than one across male-majority LLMs, and is 0.993 across female-majority LLMs. The desired condition is satisfied in male-majority LLMs, as $\beta>1$ by virtue of there being more male than female workers, and it is also satisfied in female-majority LLMs where $\beta=1.32$. In our search model, Proposition 4 finds that EPSW increases the log wage gap in favor of the group for which more firms have segregated. After the introduction of EPSW, 94% of LLM-by-month units feature more firms segregating toward the majority group in the LLM. Taken together, these figures imply that the conditions of our two theoretical models are met and are consistent with one another: they predict the wage gap moves in favor of the majority group of workers within a LLM after the onset of EPSW.

4.5 Effect of EPSW on Segregation

To study the effect of EPSW on gender segregation within the firm, we consider a panel in which an observation is a firm-month. We let j index a firm and let t index a month. We construct a full segregation indicator, $full_{jt}$, that equals 1 in time t if all workers employed by firm j are of the same gender at time t, and 0 otherwise. We also construct an indicator for whether the firm employs at least 10 long-term workers in June 2009 (policy announcement date), which we call $above10_j$. Finally, we construct a post-treatment indicator $post_t$ that is 1 if the time t of the observation is from June 2009 or later, and zero otherwise. We estimate difference-in-difference models of the following form:

$$full_{jt}=\alpha_j+\alpha_{k(j)t}+\beta^{seg}(above10_j\times post_t)+X_{jt}\Lambda+\epsilon_{jt} \tag{8}$$

where X_{jt} is a vector of controls indicating the share of workers (strictly) younger than the median age in the industry-region, the share of workers with tertiary education, the share of workers that have long-term contracts, and the total wages paid by the firm. α_j is a fixed

[31] As we discuss in the following pages and in Appendix C.3, our findings are qualitatively similar under alternative time trends, corresponding to more aggregated designations of labor markets.

effect for firm j, and $\alpha_{k(j)t}$ are time-varying fixed effects. These time-varying fixed effects compare the firm's segregation outcome in a given month to the average level of segregation among firms within the same "comparison group": in our baseline specification, time-varying fixed effects are at the level of 1) the firm's year of exit from the data panel by the firm's industry by the firm's region, and separately, 2) the firm's year of exit from the panel by the firm's county.[32] We provide results under alternative time trends (i.e. different "comparison groups") in Appendix C.2, and in Appendix C.3 we discuss alternative specifications and data samples. Our coefficient of interest is β^{seg}, and we interpret it as the effect of the policy on the share of gender-segregated firms.

To understand more about the dynamic effects of EPSW, we consider a difference-in-differences model period by period, where we define each period to be a half year (January-June, July-December). We label the period immediately prior to policy announcement as period 0 and omit it as the reference period. Therefore, the set of non-reference periods included in our analysis is $\mathcal{T} := \{-7,...,-1,1,...,10\}$. By construction, each period (indexed by τ) is composed of six months (indexed by t).

$$full_{jt} = \alpha_j + \alpha_{k(j)t} + \sum_{\tau \in \mathcal{T}} \beta_\tau^{seg} D_{jt} + X_{jt}\Lambda + \epsilon_{jt} \tag{9}$$

where D_{jt} is an indicator that equals 1 in time period t if firm j employs at least 10 long-term workers at policy announcement, and zero otherwise. β_τ^{seg} is the average difference in segregation between treated and control firms in period τ (relative to period 0).

Our identifying assumption is that parallel trends hold between treated and control firms. That is, $\mathbb{E}[\epsilon_{jt} \cdot D_{jt}] = 0$ for all t. This strategy builds in a partial falsification test, in that we expect coefficient estimates of β_τ^{seg} to be zero for all $\tau < 1$.

Panel (a) of Table 2 presents estimates on the effect of EPSW on segregation. Column I presents our baseline results from (8). We find a 4.41 percentage point increase in segregation following EPSW, from a baseline of 34.3% of firms that were fully segregated at EPSW announcement. Columns II-VI present results on segregation from alternative empirical specifications and alternative sample selections. Across all specifications, our point estimate of β^{seg} is positive, indicating that EPSW led to an increase in segregation; in all but one specification, this increase is statistically significant at conventional levels. While we cannot reject the null hypothesis that the coefficient of interest in the "doughnut hole" sample is

[32] We include an interaction with the firm's exit year from the panel to control for potential cohort effects leading firms facing the similar labor market conditions to shut down.

zero at conventional levels, the p-value for the test of this null equals 0.177; also, under the assumption that the coefficient of interest across our baseline and doughnut hole samples are independently distributed, we fail to reject the null hypothesis that the effect of EPSW on the doughnut hole sample is equal to the effect on our baseline sample at conventional significance levels. Additionally, we present the mean of the pre-treatment estimates (i.e. for $\tau < 1$) from (9) to support our parallel trends assumption. Across specifications, these estimates are small and statistically insignificant.

Panel (a) of Figure 1 displays the estimated coefficients of interest from (9). Prior to EPSW, the coefficient of interest is statistically indistinguishable from 0 in all periods $\tau < 1$. In $\tau = 1$ segregation rises by 1.75 percentage points (SE = 0.0079) in the treated group compared to the control group and rises to 4.96 percentage points (SE=0.0185) by $\tau = 10$.

To better understand how EPSW affects firm incentives to segregate, we consider its effect on the share of firms that are mostly-but-not-fully segregated. Our model makes strong predictions that firms bound by EPSW will have incentives to fully segregate in equilibrium, but not that there are particular incentives to partially segregate. Indeed, Proposition 6 suggests that treated firms are unlikely to be mostly-but-not-fully segregated post EPSW. That is, firms that would have otherwise had only a small number of workers of the "wrong" gender that prevent full segregation may be particularly likely to end their relationship with these workers to achieve full segregation. Therefore, the economic forces present in our model likely indicate a decrease in the share of treated firms that are almost-but-not-fully segregated after EPSW.

We reanalyze our firm-based analysis with a different dependent variable: "near" segregation. Specifically, we define a firm j to be nearly segregated at time t if the share of workers in the majority gender of its workforce is in the interval [0.8,1). Note that j is classified as "nearly" segregated at time t only if it is not fully segregated at time t. We select 0.8 as the lower end of the interval for the definition of this outcome variable due to our size restrictions; firms in our sample typically are nearly segregated only in time periods in which they employ either 1 or 2 workers of the non-majority gender; however, our findings are robust to other selections of the lower end of the range. We re-estimate (8) and (9) with near segregation as the outcome variable and present these results in Panel (b) of Table 2 and Panel (b) of Figure 1, respectively. We refer to the associated coefficients of interest as $\beta^{nearseg}$ and $\beta_{\tau}^{nearseg}$, respectively.

The findings presented in Panel (b) of Table 2 support our hypothesis on the effects of EPSW on near segregation. Each column presents a specification corresponding to the same column in Panel (a). Our baseline specification in column I reveals that EPSW lowers the share of nearly segregated firms 4.02 percentage points following EPSW. Across all columns,

Table 2: Effect of EPSW on Segregation

	(I) Baseline	(II) No firm FEs	(III) No controls	(IV) Doughnut hole	(V) Firm growth	(VI) Balanced sample
Panel (a): Dep. var. "Full Segregation"						
($\hat{\beta}^{seg}$) Post × Treated	0.0441***	0.0438***	0.0385***	0.0233	0.0394***	0.0349**
	(0.0133)	(0.0136)	(0.0134)	(0.0171)	(0.0140)	(0.0161)
Mean Pre-Treatment	−0.0014	−0.0040	0.0030	−0.0055	−0.0024	−0.0110
	(0.0118)	(0.0113)	(0.0118)	(0.0158)	(0.0127)	(0.0130)
Panel (b): Dep. var. "Near Segregation"						
($\hat{\beta}^{nearseg}$) Post × Treated	−0.0402***	−0.0371***	−0.0350**	−0.0300*	−0.0344**	−0.0344**
	(0.0139)	(0.0139)	(0.0140)	(0.0176)	(0.0146)	(0.0165)
Mean Pre-Treatment	0.0047	0.0048	0.0003	0.0189	0.0051	0.0106
	(0.0132)	(0.0123)	(0.0130)	(0.0175)	(0.0138)	(0.0152)
Number of Firms	5312	5312	5312	4660	4578	3311
Number of Observations	515361	515363	515361	452638	457649	357588
Fixed effects						
Firm	Yes	No	Yes	Yes	Yes	Yes
Month×exit year×county	Yes	Yes	Yes	Yes	Yes	Yes
Month×exit year×region×industry	Yes	Yes	Yes	Yes	Yes	Yes
Firm-month level controls	Yes	Yes	No	Yes	Yes	Yes

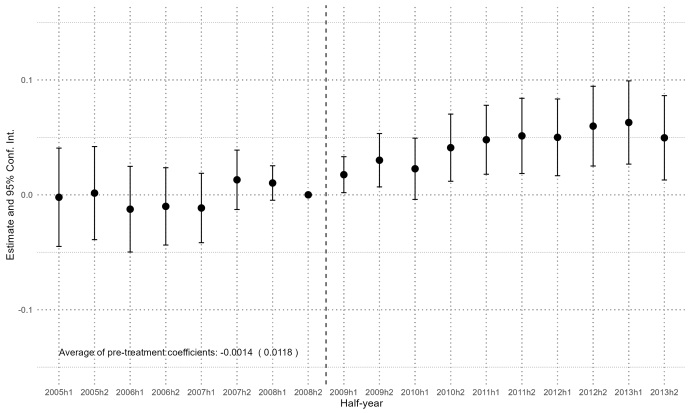
Notes: The unit of analysis is the firm-month. In panel (a) the dependent variable is a binary variable that equals 1 if and only if all workers at the firm in question are of a single gender in a given month. In panel (b) the dependent variable is a binary variable that equals 1 if and only if the share of the majority gender of workers at firm j at time t is an element of [.8,1). Across panels, column I presents estimated coefficients $\hat{\beta}^{seg}$ and $\hat{\beta}^{nearseg}$, respectively, for our baseline difference-in-differences regression presented in (8). Firm-month controls included are: the share of workers younger than the median age in the industry-region, the share of workers with tertiary education, the share of workers that have long-term contracts, and the total wages paid by the firm. The mean pre-treatment effect is the mean of $\hat{\beta}^{seg}_{\tau}$ ($\hat{\beta}^{nearseg}_{\tau}$) for $\tau<1$ calculated from (9). Columns II-VI present the analogous information for alternative sample selections and empirical specifications as described in Appendix C.3: column II removes firm fixed effects, column III removes the vector of controls, column IV drops firms with 9 or 10 workers at announcement from our baseline sample, column V interacts time-varying fixed effects with quartiles of firms growth in the 6 months prior to policy announcement, and column VI drops firms that are not present in every month of our panel. Throughout, standard errors in parentheses are two-way clustered at the firm and month levels. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

these estimates are similar in magnitude (and opposite in sign) to the increase in the share of fully segregated, treated firms following EPSW (see Panel (a)). All estimates are statistically significant at conventional levels. This "missing mass" of firms with near-but-not-full segregation suggests that firms that face relatively lower costs of fully segregating are the most likely to do so within our window of analysis. Panel (b) of Figure 1 shows the dynamic

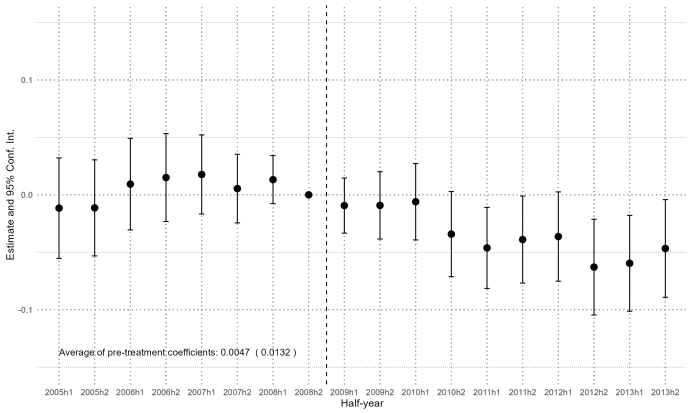
path of near segregation following EPSW. The pre-treatment coefficients are not statistically different from 0 in any period $\tau < 1$.

Figure 1: Dynamic Path of EPSW's Effect on Segregation

(a): Full Segregation



(b): Near Segregation



Notes: This figure displays the estimated coefficients $\hat{\beta}_\tau$ for the difference-in-differences regression described in (9). In Panel (a) the outcome variable is "full" segregation. In Panel (b) the outcome variable is "near" segregation, i.e. the dependent variable equals 1 if and only if the share of the majority gender of workers at firm j at time t is an element of [.8,1). In both panels, bars depict 95% confidence intervals.

Proposition 6 predicts that firms that are initially nearly segregated are the most likely to segregate within our window of analysis. While near-segregation at announcement is not randomly assigned, we can test to see whether this predicted association is present in our data. To do so, we add an additional independent variable (and its relevant interactions) to (8) which captures whether a firm is nearly segregated at announcement. Specifically, we construct an indicator, NSA_j, that equals 1 if the share of workers in the majority gender of firm j's workforce *at announcement* is in the interval [0.8,1), and zero otherwise. We estimate difference models of the following form:

$$\begin{aligned} full_{jt} = {} & \alpha_j + \alpha_{k(j)t} + \gamma_1 (NSA_j \times post_t) \\ & + \beta^{seg} (above10_j \times post_t) + \beta^{NSA} (above10_j \times post_t \times NSA_j) + X_{jt}\Lambda + \epsilon_{jt} \end{aligned} \tag{10}$$

Our coefficients of interest are β^{seg} and β^{NSA}. We interpret β^{seg} as the association between the policy and the share of gender-segregated firms among those that are not nearly segregated at announcement, and $\beta^{seg} + \beta^{NSA}$ as the association between the policy and the share of gender-segregated firms among those that are nearly segregated at announcement.

Table 3 presents estimates on the association between EPSW and segregation by near segregation status at announcement. Column I presents our baseline results from (10). We find a 2.78 percentage point increase in segregation following EPSW for firms that are not nearly segregated at announcement, and a 8.71 percentage point increase in segregation for firms that are nearly segregated at announcement. Columns II-VI present results from alternative empirical specifications and alternative sample selections. The same pattern holds across all specifications: the estimated coefficient for firms that are nearly segregated at announcement is roughly three times larger than the estimated coefficient for firms that are not nearly segregated at announcement, suggesting that firms that are nearly segregated prior to EPSW are the most likely to segregate post EPSW.

Table 3: Effect of EPSW on Segregation, by Nearly Segregated at Announcement

	(I) Baseline	(II) No firm FEs	(III) No controls	(IV) Doughnut hole	(V) Firm growth	(VI) Balanced sample
($\hat{\beta}^{seg}$) Post × Treated	0.0278*	0.0330**	0.0212	0.0186	0.0268*	0.0203
	(0.0144)	(0.0148)	(0.0145)	(0.0181)	(0.0152)	(0.0175)
($\hat{\beta}^{seg}$+$\hat{\beta}^{NSA}$) Effect for Firms	0.0871***	0.0810***	0.0836***	0.0418	0.0711**	0.0782**
Nearly Segregated at Announcement	(0.0267)	(0.0269)	(0.0271)	(0.0355)	(0.0278)	(0.0320)
Number of Firms	5312	5312	5312	4660	4578	3311
Number of Observations	515361	515363	515361	452638	457649	357588
Fixed effects						
Firm	Yes	No	Yes	Yes	Yes	Yes
Month×exit year×county	Yes	Yes	Yes	Yes	Yes	Yes
Month×exit year×region×industry	Yes	Yes	Yes	Yes	Yes	Yes
Firm-month level controls	Yes	Yes	No	Yes	Yes	Yes

Notes: The unit of analysis is the firm-month. The dependent variable is a binary variable that equals 1 if and only if all workers at the firm in question are of a single gender in a given month. Column I presents estimated coefficients of interest for our baseline difference regression presented in (10). Firm-month controls included are: the share of workers younger than the median age in the industry-region, the share of workers with tertiary education, the share of workers that have long-term contracts, and the total wages paid by the firm. Columns II-VI present the analogous information for alternative sample selections and empirical specifications as described in Appendix C.3: column II removes firm fixed effects, column III removes the vector of controls, column IV drops firms with 9 or 10 workers at announcement from our baseline sample, column V interacts time-varying fixed effects with quartiles of firms growth in the 6 months prior to policy announcement, and column VI drops firms that are not present in every month of our panel. Throughout, standard errors in parentheses are two-way clustered at the firm and month levels. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

4.6 Effect of EPSW on the Gender Wage Gap

Guided by our theory, we are interested in studying the effect of EPSW on the wage gap between male and female workers. Our theory predicts a relative benefit to men in male majority LLMs and a relative benefit to women in female majority LLMs. We define a LLM as being female majority if the share of female workers across all firms in a particular industry-county pair is greater than 0.5 in June 2009, and otherwise, we define it as a male majority LLM.

To estimate the effects of EPSW on the gender wage gap across LLMs, we consider a panel in which an observation is a worker-firm-month. In order to remove potential confounds from workers who are simultaneously enrolled in formal schooling and those beyond Chile's official retirement age, we only consider workers aged 22-65. We let i index a worker, j index a firm, and t index a month. Let w_{ijt} be the earnings of worker i at firm j in month t. We construct an indicator, $male_i$, that equals 1 if the worker is a male, and 0 otherwise. We construct indicator $femalemaj_j$ that equals 1 if firm j is in a female majority LLM, and 0 otherwise. We estimate difference models of the following form:

$$\begin{aligned}\ln w_{ijt} = {} & \alpha_i + \alpha_j + \alpha_{k(ij)t} + \gamma_1(above10_j \times post_t) + \psi_1(above10_j \times post_t \times femalemaj_j) \\ & + \gamma_2(male_i \times post_t) + \psi_2(male_i \times post_t \times femalemaj_j) \\ & + \gamma_3(above10_j \times male_i) + \psi_3(above10_j \times male_i \times femalemaj_j) \\ & + \beta^{Mgap}(above10_j \times male_i \times post_t) + \beta^{Fgap}(above10_j \times male_i \times post_t \times femalemaj_j) \\ & + X_{ijt}\Lambda + \epsilon_{ijt} \qquad (11)\end{aligned}$$

where α_i is a fixed effect for worker i and X_{ijt} is a vector of firm-month level controls and worker-firm-month level controls. The firm-month level controls are the share of workers younger than the median age at the industry-region, the share of workers with tertiary education, and the share of workers that have long-term contracts. The worker-firm-month level controls are the number of months at the firm and an indicator for whether the worker's earnings reach the top-coding threshold. $\alpha_{k(ij)t}$ are time fixed effects for workers in set $k(ij)$, where $k(ij)$ is a comparison group of workers to i employed at a comparison group of firms to j. Worker comparison groups are defined by equivalence across three binary "human capital" dimensions at time t at firm j: an indicator for the highest level of educational attainment (high school, tertiary, or neither), an indicator for long-term versus fixed-term contract, and an indicator for the worker's age in decades (i.e. 20-29 years old, 30-39 years old, etc). Firm comparison groups are defined as in (8). Because workers can switch from one LLM to

another, the inclusion of these fixed effects controls for composition changes in the distribution of observable characteristics within a LLM caused by EPSW. We provide results controlling for alternative time trends at more aggregated levels in Appendix C.2; our estimates are qualitatively and quantitatively similar which suggests that worker substitution across LLMs is not driving our findings. In Appendix C.3 we discuss alternative specifications and data samples.

Our coefficients of interest are β^{Mgap} and β^{Fgap}. We interpret β^{Mgap} as the effect of the policy on the (percentage) wage gap between male and female workers in male majority LLMs. We interpret $\beta^{Fgap}+\beta^{Mgap}$ as the effect of the policy on the (percentage) wage gap between male and female workers in female LLMs.

To understand more about the dynamic effects of EPSW, we consider a difference model period by period, where we define each period to be a half year (January-June, July-December). We label the period immediately prior to policy announcement as period 0 and omit it as the reference period. Therefore, the set of non-reference periods included in our analysis is $\mathcal{T}=\{-7,...,-1,1,...,10\}$. By construction, each period (indexed by τ) is composed of six months (indexed by t). Let $period_\tau$ be an indicator that equals 1 in period τ, and zero otherwise.

$$\begin{aligned}\ln w_{ijt} = {} & \alpha_i+\alpha_j+\alpha_{k(ij)t}+\sum_{\tau\in\mathcal{T}}\gamma_{1\tau}(above10_j\times period_\tau)+\sum_{\tau\in\mathcal{T}}\psi_{1\tau}(above10_j\times femalemaj_j\times period_\tau)\\ & +\sum_{\tau\in\mathcal{T}}\gamma_{2\tau}(male_i\times period_\tau)+\sum_{\tau\in\mathcal{T}}\psi_{2\tau}(male_i\times femalemaj_j\times period_\tau)\\ & +\gamma_3(above10_j\times male_i)+\psi_3(above10_j\times male_i\times femalemaj_j)\\ & +\sum_{\tau\in\mathcal{T}}\beta_\tau^{Mgap}D_{ijt}^M+\sum_{\tau\in\mathcal{T}}\beta_\tau^{Fgap}D_{ijt}^F\\ & +X_{ijt}\Lambda+\epsilon_{ijt} \qquad (12)\end{aligned}$$

where D_{ijt}^M is an indicator that equals 1 in time t if firm j employs at least 10 long-term workers at the time of policy announcement, j's LLM is coded as male majority, and i is male, and zero otherwise. β_τ^{Mgap} is the average difference in log wages between men and women in treated versus control firms in period τ (relative to period 0) in male majority LLMs. Similarly, D_{ijt}^F is an indicator that equals 1 in time t if firm j employs at least 10 long-term workers at the time of policy announcement, j's LLM is coded as female majority, and i is male, and zero otherwise. β_τ^{Fgap} is the average difference in log wages between men and women in treated versus control firms in period τ (relative to period 0) in female majority LLMs *relative to male majority LLMs*. Therefore, $\beta_\tau^{Mgap}+\beta_\tau^{Fgap}$ is the average difference in log wages between men and women in

treated versus control firms in period τ (relative to period 0) in female majority LLMs.

Our identifying assumption is that parallel trends hold between treated and control firms, that is, $\mathbb{E}[\epsilon_{ijt} \cdot D^M_{ijt}] = 0$ and $\mathbb{E}[\epsilon_{ijt} \cdot (D^M_{ijt} + D^F_{ijt})] = 0$ for all t (Olden and Møen, 2022). This strategy builds in a partial falsification test, in that we expect coefficient estimates of β_τ^{Mgap}, and $\beta_\tau^{Mgap} + \beta_\tau^{Fgap}$ to be zero for $\tau < 1$.

Table 4 presents our estimates on the effect of EPSW on the gender wage gap. Column I presents our baseline results from (11). We find that EPSW increases the gender wage gap (in favor of men) by 4.27 percentage points in male majority LLMs, but decreases the gender wage gap (in favor of women) by 6.24 percentage points in female majority LLMs. For reference, men typically outearn women prior to EPSW: the wage gap (controlling for human capital factors, as in (11)) is 8.63% (SE=0.0058) in male-majority LLMs and 0.71% (SE=0.0118) in female-majority LLMs. Columns II-VII present results on the gender wage gap from alternative empirical specifications and alternative sample selections. We discuss these specifications further in Appendix C.3. Across all specifications, our findings of an increase in the gender wage gap in male majority LLMs, and a decrease in the gender wage gap in female majority LLMs, are statistically significant at conventional levels. Additionally, we present the mean of the pre-treatment estimates (i.e. for $\tau < 1$) from (12) to support our parallel trends assumption. Across specifications, these estimates are small and statistically insignificant.

Moreover, our empirical findings match our theoretical predictions that the change in the wage gap caused by EPSW is primarily driven by a reduction in average wages of the minority group of workers in an LLM. Specifically, women's average wages in male-majority labor markets fall post EPSW by 3.3 percentage points (SE= 0.011) in treated versus control firms, and men's average wages in female-majority labor markets fall post EPSW by 4.5 percentage points (SE=0.018) in treated versus control firms.[33] There are no statistically significant changes in the average wage of the majority group of workers within a LLM.

Figure 2 displays the estimated coefficients of interest from (12). Panel (a) presents estimates for male majority LLMs. The coefficient of interest is statistically indistinguishable from 0 in all periods $\tau < 1$. In $\tau = 1$ the wage gap rises (in favor of men) by 1.38 percentage points (SE = 0.0123) in the treated group compared to the control group, and rises by 3.76 percentage points (SE = 0.0187) by $\tau = 10$. Panel (b) presents estimates for female majority LLMs. The coefficient of interest is statistically indistinguishable from 0 in all periods $\tau < 1$. In $\tau = 1$, the wage gap falls (in favor of women) by 2.34 percentage points (SE=0.0210) in the treated group compared to the control group, and falls by 10.30 percentage points (SE = 0.0313) by $\tau = 10$.

[33] These values are estimates, respectively, of γ_1 and $\gamma_1 + \psi_1 + \beta^{Mgap} + \beta^{Fgap}$ in (11).

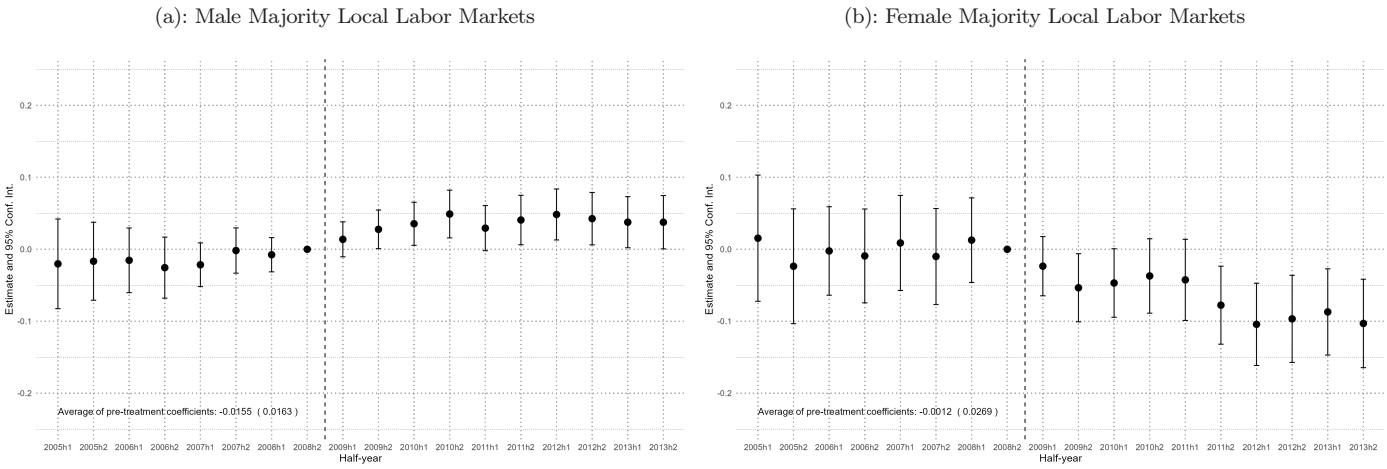
Table 4: Effect of EPSW on Gender Wage Gap, by Majority Worker Group

	(I) Baseline	(II) No firm FEs	(III) No controls	(IV) Doughnut hole	(V) Firm growth	(VI) Balanced sample
($\hat{\beta}^{Mgap}$) Treated × Male × Post	0.0427***	0.0364***	0.0432***	0.0409***	0.0439***	0.0407***
	(0.0116)	(0.0114)	(0.0116)	(0.0158)	(0.0117)	(0.0144)
($\hat{\beta}^{Mgap}+\hat{\beta}^{Fgap}$) Effect in Female	−0.0624***	−0.0475**	−0.0575***	−0.0840***	−0.0622***	−0.0511**
Majority LLM	(0.0197)	(0.0206)	(0.0196)	(0.0251)	(0.0230)	(0.0234)
Mean Pre-Treatment	−0.0155	−0.0134	−0.0136	−0.0082	−0.0159	−0.0101
(Male Majority LLM)	(0.0163)	(0.0155)	(0.0166)	(0.0229)	(0.0158)	(0.0201)
Mean Pre-Treatment	−0.0012	−0.0167	−0.0072	0.0343	0.0102	0.0335
(Female Majority LLM)	(0.0269)	(0.0332)	(0.0268)	(0.0319)	(0.0324)	(0.0284)
Number of Firms	6424	6424	6424	5692	6153	3351
Number of Observations	5551100	5551100	5551100	4800898	5108499	3563150
Fixed effects						
Firm	Yes	No	Yes	Yes	Yes	Yes
Worker	Yes	Yes	Yes	Yes	Yes	Yes
Month×exit year×county×hum. cap.	Yes	Yes	Yes	Yes	Yes	Yes
Month×exit year×region×industry×hum. cap.	Yes	Yes	Yes	Yes	Yes	Yes
Firm-month level controls	Yes	Yes	No	Yes	Yes	Yes
Worker-firm-month level controls	Yes	Yes	No	Yes	Yes	Yes

Notes: The unit of analysis is the worker-firm-month and the dependent variable is the natural logarithm of the worker's wage at the firm in a given month. Column I presents estimated coefficients $\hat{\beta}^{Mgap}$ and $\hat{\beta}^{Fgap}$ for our baseline regression specification presented in (11). "Human capital" fixed effects are comprised of an indicator for the highest level of educational attainment (high school, tertiary, or neither), an indicator for long-term versus fixed-term contract, and an indicator for the worker's age in decades (i.e. 20-29 years old, 30-39 years old, etc). Firm-month controls included are: the share of workers younger than the median age in the industry-region, the share of workers with tertiary education, and the share of workers that have long-term contracts. Worker-firm-month level controls included are the number of months at the firm, and an indicator for reaching the earnings truncation threshold. The mean pre-treatment effects are the mean of $\hat{\beta}_{\tau}^{Mgap}$ and $\hat{\beta}_{\tau}^{Mgap}+\hat{\beta}_{\tau}^{Fgap}$, respectively, for $\tau<1$ calculated from (12). Columns II-VI present the analogous information for alternative sample selections and empirical specifications as described in Appendix C.3: column II removes firm fixed effects, column III removes the vector of controls, column IV drops firms with 9 or 10 workers at announcement from our baseline sample, column V interacts time-varying fixed effects with quartiles of firms growth in the 6 months prior to policy announcement, and column VI drops firms that are not present in every month of our panel. Throughout, standard errors in parentheses are two-way clustered at the firm and month levels. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

One potentially interesting policy-related question is the overall effect of EPSW on the gender wage gap across all LLMs, given that the majority of workers are employed in male-majority LLMs. We find that EPSW increases the overall gender wage gap (in favor of men) by 2.74 percentage points (SE=0.0097).

Figure 2: Dynamic Path of EPSW's Effect on Gender Wage Gap, by Majority Worker Group



Notes: Panel (a) displays estimated coefficients $\hat{\beta}_{\tau}^{Mgap}$ for the regression described in (12). Panel (b) displays estimated coefficients $\hat{\beta}_{\tau}^{Mgap}+\hat{\beta}_{\tau}^{Fgap}$ for the regression described in (12). Bars depict 95% confidence intervals.

5 Conclusion

We find that the equilibrium effects of Equal Pay for Similar Work (EPSW) policies dominate their direct effects. Therefore, imposing these policies may lead to unintended outcomes. Our modeling demonstrates that EPSW targeted specifically to equalize pay across protected classes of workers leads to firms segregating their workforce in equilibrium to avoid the bite of the policy. Although discriminatory forces may lead to pay gaps across groups without EPSW, segregation caused by EPSW results in the minority group of workers in a labor market receiving even lower relative wages. Our empirical evaluation of Chile's 2009 EPSW—which prohibited unequal pay for similar work across genders—supports these predictions. The policy caused gender segregation within firm to rise, and the gender wage gap to rise. Importantly, the rise in the wage gap only occurred in male-majority local labor markets; in local labor markets with majority female workers, the wage gap closed. Both of these findings are as predicted by our theoretical analysis.

Many important questions remain in understanding the equilibrium impacts of EPSW, and equity-related labor market policies in general. One difficulty is understanding the role of complementary policies; a benefit of studying the Chilean labor market is the relative dearth of alternative anti-discrimination policies prior to and contemporaneously with the enactment of EPSW. For this reason, further theoretical study of labor-market policies may

be particularly fruitful. One particular avenue for further research is understanding firm incentives–especially those of large firms, for which segregation may be especially costly–to heterogenize on-the-job responsibilities and duties to evade EPSW.

References

Joseph G. Altonji and Rebecca M. Blank. Race and gender in the labor market. In *Handbook of Labor Economics*, volume 3, pages 3143–3259. Elsevier, 1999.

Robert B. Ash and Catherine A. Doléans-Dade. *Probability & Measure Theory, Second Edition*. Academic Press, 2000.

Martha Bailey, Thomas Helgerman, and Bryan Stuart. How the 1963 Equal Pay Act and 1964 Civil Rights Act Shaped the U.S. Gender Gap. *Mimeo*, 2022.

Revital Bar and Asaf Zussman. Customer Discrimination: Evidence from Israel. *Journal of Labor Economics*, 35(4):1031–1059, 2017.

James N. Baron and William T. Bielby. The proliferation of job titles in organizations. *Administrative Science Quarterly*, 31(4):561–586, 1986.

Morten Bennedsen, Elena Simintzi, Margarita Tsoutsoura, and Daniel Wolfenzon. Do Firms Respond to Gender Pay Gap Transparency? *Journal of Finance*, 77(4):2051–2091, 2022.

Clémence Berson. Local labor markets and taste-based discrimination. *IZA Journal of Labor Economics*, 5(5), 2016.

Prashant Bharahwaj, Rahul Deb, and Ludovic Renou. Statistical discrimination and the distribution of wages. *Mimeo*, 2024.

V. Bhaskar, Alan Manning, and Ted To. Oligopsony and Monopsonistic Competition in Labor Markets. *Journal of Economic Perspectives*, 16(2):155–174, 2002.

Ken Binmore, Ariel Rubinstein, and Asher Wolinsky. The Nash bargaining solution in economic modelling. *RAND Journal of Economics*, 17(2):176–109, 1988.

Francine D. Blau. *Equal pay in the office*. Lexington Books, 1977.

Francine D. Blau and Lawrence M. Kahn. Race and Gender Pay Differentials. *NBER Working Paper 4120*, 1992.

Francis Bloch and Harl Ryder. Two-Sided Search, Marriages, and Matchmakers. *International Economic Review*, 41(1):93–115, 2000.

René Böheim and Sarah Gust. The Austrian Pay Transparency Law and the Gender Wage Gap. *IZA Discussion Paper 14206*, 2021.

Kenneth Burdett and Melvyn G. Coles. Long-Term Partnership Formation: Marriage and Employment. *The Economic Journal*, 109(456):F307–F334, 1999.

David A. Butz. Durable-Good Monopoly and Best-Price Provisions. *American Economic Review*, 80(5):1062–1076, 1990.

Sebastian Calonico, Matias D. Cattaneo, and Rocío Titiunik. Optimal Data-Driven Regression Discontinuity Plots. *Journal of Econometrics*, 110(512):1753–1769, 2015.

David Card. Who set *Your* wage? *American Economic Review*, 112(4):1075–1090, 2022.

David Card, Fabrizio Colella, and Rafael Lalive. Gender Preferences in Job Vacancies and Workplace Gender Diversity. *Review of Economic Studies*, 2024.

Hector Chade. Matching with Noise and the Acceptance Curse. *Journal of Economic Theory*, 129(1):81–113, 2006.

Thomas E. Cooper and Timothy L. Fries. The most-favored-nation pricing policy and negotiated prices. *International Journal of Industrial Organization*, 9:209–223, 1991.

Gabriel Cruz and Tomás Rau. The effects of equal pay laws on firm pay premiums: Evidence from chile. *Labour Economics*, 75, 2022.

Zoe Cullen and Bobak Pakzad-Hurson. Equilibrium Effects of Pay Transparency. *Econometrica*, 91(4):765–802, 2023.

John J. Donohue III and James Heckman. Continuous versus episodic change: the impact of civil rights policy on the economic status of blacks. *Journal of Economic Literature*, 29:1603–1643, 1991.

Emma Duchini, Stefania Simion, Arthur Turrell, and Jack Blundell. Pay Transparency and Gender Equality. *mimeo*, 2022.

John-Paul Ferguson and Rembrand Koning. Firm Turnover and the Return of Racial Establishment Segregation. *American Sociological Review*, 83(3):445–474, 2018.

Leonardo Gasparini and Leopoldo Tornarolli. Labor Informality in Latin America and the Caribbean: Patterns and Trends from Household Survey Microdata. *CEDLAS, Working Papers*, 0046, 2007.

Claudia Goldin. *Understanding the Gender Gap: An Economic History of American Women*. Oxford University Press, 1990.

Andreas Gulyas, Sebastian Seritz, and Sourav Sinha. Does Pay Transparency Affect the Gender Wage Gap? Evidence from Austria. *American Economic Journal: Economic Policy*, 15(2):236–255, 2023.

Christopher Harris. Existence and Characterization of Perfect Equilibrium in Games of Perfect Information. *Econometrica*, 53(3):613–628, 1985.

Judith K. Hellerstein and David Neumark. Workplace Segregation in the United States: Race, Ethnicity, and Skill. *The Review of Economics and Statistics*, 90(3):459–477, 2008.

Harry J. Holzer and Keith R. Ihlanfeldt. Customer Discrimination and Employment Outcomes for Minority Workers. *The Quarterly Journal of Economics*, 113(3):835–867, 1998.

Henrick Horn and Asher Wolinsky. Bilateral Monopolies and Incentives for Merger. *RAND Journal of Economics*, 19(3):408–419, 1988.

Marie Hyland, Simeon Djankov, and Pinelopi Koujianou Goldberg. Gendered Laws and Women in the Workforce. *American Economic Review: Insights*, 2(4):475–490, 2020.

Leo Kaas. Does Equal Pay Legislation Reduce Labour Market Inequality? *The Scandinavian Journal of Economics*, 111(1):51–71, 2009.

Erin Kelley, Gregory Lane, Matthew Pecenco, and Edward Rubin. Customer Discrimination in the Workplace: Evidence from Online Sales. *mimeo*, 2023.

Alexander S Kelso and Vincent P Crawford. Job matching, coalition formation, and gross substitutes. *Econometrica: Journal of the Econometric Society*, pages 1483–1504, 1982.

Paul Klemperer. Markets with consumer switching costs. *The Quarterly Journal of Economics*, 102(2):375–394, 1987.

Patrick Kline, Evan K. Rose, and Christopher R. Walters. Systemic discrimination among large US employers. *The Quarterly Journal of Economics*, 137(4):1963–2036, 2022.

Peter Kuhn and Kailing Shen. What Happens When Employers Can No Longer Discriminate in Job Ads? *American Economic Review*, 113(4):1013–1048, 2023.

Peter Kuhn, Kailing Shen, and Shuo Zhang. Gender-targeted job ads in the recruitment process: Facts from a Chinese job board. *Journal of Development Economics*, 147, 2020.

Johan Lagerlöf. Strategic gains from discrimination. *European Economic Review*, 122, 2020.

Jonathan Lanning. A search model with endogenous job destruction and discrimination: Why equal wage policies may not eliminate wage disparity. *Labour Economics*, 26:51–71, 2014.

Alan Manning. *Monopsony in Motion: Imperfect Competition in Labor Markets*. Princeton University Press, 2005.

Justin McCrary. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714, 2008.

David Neumark and Wendy A. Stock. The Labor Market Effects of Sex and Race Discrimination Laws. *Economic Inquiry*, 44(3):385–419, 2006.

David Neumark, Roy J. Bank, and Kyle D. Van Nort. Sex Discrimination in Restaurant Hiring: An Audit Study. *The Quarterly Journal of Economics*, 111(3):915–941, 1996.

Andreas Olden and Jarle Møen. The triple difference estimator. *The Econometrics Journal*, 25(3):531–553, 2022.

Marcela Perticará and Ivonne Bueno. A new approach to gender wage gaps in Chile. *CEPAL Review*, 99, 2009.

H.L. Royden and P.M. Fitzpatrick. *Real Analysis (Fourth Edition)*. Prentice Hall, 2010.

Fiona Scott Morton. The Strategic Response by Pharmaceutical Firms to the Medicaid Most-Favored-Customer Rules. *The RAND Journal of Economics*, 28(2):269–290, 1997a.

Fiona Scott Morton. The Interaction between a Most-Favored-Customer Clause and Price Dispersion: An Empirical Examination of the Medicaid Rebate Rules of 1990. *Journal of Economics & Management Strategy*, 6(1):151–174, 1997b.

Lloyd S Shapley and Martin Shubik. The assignment game i: The core. *International Journal of game theory*, 1(1):111–130, 1971.

Waclaw Sierpiński. Sur les fonctions d'ensemble additives et continues. *Fundamenta Mathematicae*, 3:240–246, 1922.

Nancy L. Stokey and Robert E. Lucas, Jr. *Recursive methods in economic dynamics*. Harvard University Press, 1989.

Sher Verick. Female labor force participation in developing countries. *IZA World of Labor*, 87, 2014.

Washington Post and Times Herald. Pay is up: Will jobs go down? page F14, June 14 1964.

Appendix: For Online Publication

Appendix A presents proofs omitted from the main text. Appendix B presents a non-cooperative formulation of the cooperative game studied in the main text. Appendix C presents additional empirical results and descriptions.

A Proofs

Proof of Proposition 1

Consider any outcome $O=\{(f_g^i(v),w_i^g(v))\}_{v\in[0,1],i=1,2,g=A,B}$ in which the following hold for almost all $v\in[0,1]$ and all $g\in\{A,B\}$:

1. $f_g^1(v)+f_g^2(v)=f_g(v)$, and
2. for all $i\in\{1,2\}$, $w_i^g(v)=v$ if $f_g^i(v)>0$.

We establish the desired result through three lemmas regarding these two enumerated conditions. Lemma 1 shows that any outcome satisfying these two enumerated conditions is a core outcome, Lemma 2 shows that there are uncountably many non-equivalent outcomes satisfying these two conditions (implying that there are uncountably many core outcomes), and Lemma 3 shows that there are no core outcomes that fail either of the two enumerated conditions.

LEMMA **1.** *Any outcome O satisfying the two enumerated conditions in the Proof of Proposition 1 is a core outcome.*

Proof of Lemma 1. Suppose not for the sake of contradiction. Then there are a firm j and a distinct outcome (for firm j) $\tilde{O}_j:=\{(\tilde{f}_g^j(v),\tilde{w}_j^g(v))\}_{v\in[0,1],g=A,B}$ that blocks O. In order for $\tilde{O}_j$ to block O it must be that $\pi_j^{\tilde{O}_j}>\pi_j^{O_j}$. However, it is also the case that

$$\begin{aligned}\pi_j^{O_j}&=\beta\int_0^1[v-w_j^A(v)]f_A^j(v)dv+\int_0^1[v-w_i^B(v)]f_B^j(v)dv\\&=0\\&\geq\beta\int_0^1[v-\tilde{w}_j^A(v)]\tilde{f}_A^j(v)dv+\int_0^1[v-\tilde{w}_i^B(v)]\tilde{f}_B^j(v)dv\\&=\pi_j^{\tilde{O}_j}\end{aligned}$$

The second equality follows because, by the construction of O, either $f_g^j(v)=0$ or $w_j^g(v)=v$ for almost all v and all g, therefore, the integrand is almost always equal to zero. The inequality follows because of the following exhaustive cases for almost all v and all g, corresponding, respectively, to Conditions 1-4 of the definition of block:

- Suppose $\tilde{w}_j^g(v) \geq w_j^g(v)$ and $\tilde{w}_j^g(v) > w_{-j}^g(v)$, then it must be that $\tilde{w}_j^g(v) \geq v$ since $\max\{w_j^g(v), w_{-j}^g(v)\}=v$, which makes the integrand weakly negative,

- Suppose $\tilde{w}_j^g(v) \geq w_j^g(v)$ and $\tilde{f}_g^j(v)+f_g^{-j}(v) \leq f_g(v)$. If $\tilde{f}_g^j(v)=0$ then the integrand is weakly negative. If $\tilde{f}_g^j(v)>0$ then it must be that $f_g^{-j}(v)<f_g(v)$, and by the construction of O that $f_g^j(v)+f_g^{-j}(v)=f_g(v)$, it must be that $f_g^j(v)>0$. Therefore, it must be that $w_j^g(v)=v$, and the requirement that $\tilde{w}_j^g(v) \geq w_j^g(v)$ makes the integrand weakly negative.

- Suppose $\tilde{w}_j^g(v) > w_{-j}^g(v)$ and $\tilde{f}_g^j(v)+f_g^j(v) \leq f_g(v)$. If $\tilde{f}_g^j(v)=0$ then the integrand is weakly negative. If $\tilde{f}_g^j(v)>0$ then it must be that $f_g^j(v)<f_g(v)$, and by the construction of O that $f_g^j(v)+f_g^{-j}(v)=f_g(v)$, it must be that $f_g^{-j}(v)>0$. Therefore, it must be that $w_{-j}^g(v)=v$, and the requirement that $\tilde{w}_j^g(v) > w_{-j}^g(v)$ makes the integrand strictly negative.

- Suppose $\tilde{f}_g^j(v)+f_g^j(v)+f_g^{-j}(v) \leq f_g(v)$ then it must be that $\tilde{f}_g^j(v)=0$ since by the construction of O it is the case that $f_g^j(v)+f_g^{-j}(v)=f_g(v)$. Therefore, the integrand is weakly negative.

These cases reveal a contradiction with the premise that $\pi_j^{\tilde{O}_j} > \pi_j^{O_j}$. Therefore, O is a core outcome. □

LEMMA **2.** *There exist a continuum of (non-equivalent) core outcomes satisfying the two enumerated conditions in the Proof of Proposition 1.*

Proof of Lemma 2. For any $v^* \in [0,1]$ there is an outcome O such that $f_g^1(v)=f_g(v)$ and $w_1^g(v)=v$ for all g and $v \leq v^*$, and $f_g^2(v)=f_g(v)$ and $w_2^g(v)=v$ for all g and $v>v^*$. □

It remains to show that any core outcome must be such that all workers are hired and all workers receive a wage equal to their productivity.

LEMMA **3.** *There exist no core outcomes which do not satisfy the two enumerated conditions in the Proof of Proposition 1.*

Proof of Lemma 3. Suppose for contradiction that there is a core outcome $O=\{(f_g^i(v),w_i^g(v))\}_{v\in[0,1],i=1,2,g=A,B}$ such that there exists a set with positive measure V where the following fails for some $g\in\{A,B\}$ and all $v\in V$:

1. $f_g^1(v)+f_g^2(v)=f_g(v)$, and
2. for all $i\in\{1,2\}$, $w_i^g(v)=v$ if $f_g^i(v)>0$.

Throughout, we assume that the above conditions are violated for $g=B$, and we construct a blocking outcome $\tilde{O}_j=\{(\tilde{f}_g^j(v),\tilde{w}_j^g(v))\}_{v\in[0,1],g=A,B}$, where $(\tilde{f}_A^j(v),\tilde{w}_j^A(v))=(f_A^j(v),w_j^A(v))$ (or $\tilde{O}_{-j}=\{(\tilde{f}_g^{-j}(v),\tilde{w}_{-j}^g(v))\}_{v\in[0,1],g=A,B}$ where $(\tilde{f}_A^{-j}(v),\tilde{w}_{-j}^A(v))=(f_A^{-j}(v),w_{-j}^A(v))$), i.e. we do not change the outcome for $A-$group workers for either firm. The argument in which the violations occur for $g=A$ is analogous, where terms related to the firm's profit must be multiplied by β.

We show a contradiction by considering six exhaustive cases: By countable additivity of measure, the set of productivities that fails one of the above two enumerated points has positive measure if and only if at least one of the sets in the following six cases has a positive measure.

First, suppose there exist a firm j and a subset of productivities $V\subset[0,1]$ with positive measure such that $w_j^g(v)>v$ for all $v\in V$. We proceed by showing that firm j can fire workers of type (g,v), $v\in V$ and increase its profit. Let $\tilde{O}_j$ be defined as follows:

$$\tilde{f}_g^j(v):=\begin{cases} f_g^j(v) & \text{if } v\notin V,\\ 0 & \text{if } v\in V.\end{cases}\qquad \tilde{w}_j^g(v):=\begin{cases} w_j^g(v) & \text{if } v\notin V,\\ 0 & \text{if } v\in V.\end{cases}$$

$\tilde{O}_j$ blocks O as j's profit increases and Condition 4 of the definition of block is satisfied for all $v\in V$ (i.e. the workers in V are fired) and for all $v\in[0,1]\setminus V$ Condition 2 of the definition of block is satisfied (i.e. there is no change in the hiring or wages of workers with $v\in[0,1]\setminus V$).

Therefore, we proceed with the assumption that for each firm j, $w_j^g(v)\leq v$ for almost all v.

Second, suppose there exist a firm j and a subset of productivities V with positive measure such that $f_g^1(v)+f_g^2(v)<f_g(v)$ and $w_j^g(v)<v$ for all $v\in V$. We proceed by showing that firm j can hire unemployed workers of type (g,v), $v\in V$ at the going wage and increase its profit. Let $\tilde{O}_j$ be defined as follows:

$$\tilde{f}_g^j(v):=\begin{cases} f_g(v)-f_g^{-j}(v) & \text{if } v\in V,\\ f_g^j(v) & \text{if } v\notin V.\end{cases}\qquad \tilde{w}_j^g(v):=w_j^g(v) \text{ for all } v.$$

$\tilde{O}_j$ blocks O as firm j's profit increases as some previously unemployed workers are hired at a wage strictly less than their productivity while all existing workers at j continue to be employed at the same wage as before, and Condition 2 of the definition of a block is satisfied for all $v \in [0,1]$ (i.e. no worker receives a wage cut and no workers are poached from firm $-j$).

Third, suppose there exists a firm j and a subset of productivities V with positive measure such that $f_g^1(v)+f_g^2(v)<f_g(v)$ and $w_j^g(v)=v$ for all $v \in V$. We proceed by showing that firm j can fire its current workers of type (g,v), $v \in V$ and hire the unemployed workers of the same type at a lower wage to increase its profit. Let $\tilde{O}_j$ be defined as follows:

$$\tilde{f}_g^j(v) := \begin{cases} f_g(v) - f_g^1(v) - f_g^2(v) & \text{if } v \in V, \\ f_g^j(v) & \text{otherwise.} \end{cases} \qquad \tilde{w}_j^g(v) := \begin{cases} 0 & \text{if } v \in V, \\ w_g^j(v) & \text{otherwise.} \end{cases}$$

$\tilde{O}_j$ blocks O. To see that $\tilde{O}_j$ blocks O, note that Condition 4 of the definition of block is satisfied for all $v \in V$ by construction, and Condition 2 of the definition of block is satisfied for all $v \in [0,1] \setminus V$ (i.e. there are no changes in employment or wages for workers with $v \in [0,1] \setminus V$). Therefore, it remains only to show that $\tilde{O}_j$ yields firm j a strictly higher profit than does O_j, and this fact follows because j earns no profit in O_j from worker types (g,v), $v \in V$ because $w_j^g(v)=v$, and because j earns strictly positive profit in $\tilde{O}_j$ from worker types (g,v), $v \in V$ because $\tilde{f}_g^j(v) = f_g(v) - f_g^1(v) - f_g^2(v) > 0$ (where the inequality follows from the ongoing assumption of this case) and $\tilde{w}_j^g(v) = 0 < v$ for all $v \in V \setminus \{0\}$.

The previous two cases exhaust the possibility of a core outcome in which $f_g^1(v)+f_g^2(v)<f_g(v)$ for any g and a subset of productivities with positive measure. Therefore, we proceed with the assumption that $f_g^1(v)+f_g^2(v)=f_g(v)$ for almost all v.

Fourth, suppose that there exist j and a set V of productivities with positive measure such that $w_j^g(v)<v$ and $f_g^j(v)=f_g(v)$ for all $v \in V$. We proceed by showing that firm $-j$ can poach workers of type (g,v), $v \in V$ from firm j at marginally higher wages and increase its profit. Let $\tilde{O}_{-j}$ be defined as follows:

$$\tilde{f}_g^{-j}(v) := \begin{cases} f_g(v) & \text{if } v \in V, \\ f_g^{-j}(v) & \text{otherwise.} \end{cases} \qquad \tilde{w}_{-j}^g(v) := \begin{cases} (w_j^g(v)+v)/2 & \text{if } v \in V, \\ w_{-j}^g(v) & \text{otherwise.} \end{cases}$$

$\tilde{O}_{-j}$ blocks O. To see that $\tilde{O}_{-j}$ blocks O, note that Condition 1 of the definition of block is satisfied for all $v \in V$ by construction, and Condition 2 of the definition of block is satisfied for all $v \in [0,1] \setminus V$ (i.e. there are no changes in employment or wages for workers with $v \in [0,1] \setminus V$).

Therefore, it remains only to show that $\tilde{O}_{-j}$ yields firm $-j$ a strictly higher profit than does O_{-j}, and this fact follows because $-j$ earns no profit in O_{-j} from worker types (g,v), $v\in V$ because $f_g^{-j}(v)=0$, and because $-j$ earns strictly positive profit in $\tilde{O}_{-j}$ from worker types (g,v), $v\in V$ which follows from $w_{-j}^g(v)=(w_j^g(v)+v)/2<v$ and $f_g^{-j}(v)>0$ for all $v\in V\setminus\{0\}$.

Fifth, suppose that there exist j and $V\subset[0,1]$ with positive measure such that $0\leq w_{-j}^g(v)\leq w_j^g(v)<v$ and $f_g^j(v)\in(0,f_g(v))$ for all $v\in V$. Then, there exists $\varepsilon>0$ and $V'\subset V$ with positive measure such that $0\leq w_{-j}^g(v)\leq w_j^g(v)<v-\varepsilon$ and $f_g^j(v)\in(0,f_g(v)-\varepsilon)$ for all $v\in V'$. We proceed by showing that firm j can retain workers of type (g,v), $v\in V'$, and poach all workers of the same type from firm $-j$ by marginally increasing its wages and increase its profit. For a constant $\varepsilon'>0$, consider $\tilde{O}_j$ where for all v:

$$\tilde{f}_g^j(v):=\begin{cases} f_g(v) & \text{if } v\in V',\\ f_g^j(v) & \text{otherwise.}\end{cases}\qquad \tilde{w}_j^g(v):=\begin{cases} 0 & \text{if } \tilde{f}_g^j(v)=0,\\ w_j^g(v)+\varepsilon' & \text{otherwise.}\end{cases}$$

$\tilde{O}_j$ blocks O_j for the following reasons: Condition 1 of the definition of block is satisfied for all $v\in V'$ since $w_{-j}^g(v)\leq w_j^g(\bar{v})<\tilde{w}_j^g(v)$ for all $v\in V'$ by construction, and Condition 2 of the definition of the block is satisfied for all $v\notin V'$. To see that firm j's profit increases, first note that firm j benefits from hiring additional workers from V', which results in an additional profit of at least $(\varepsilon-\varepsilon')\varepsilon\mu(V')$. Meanwhile, the firm may lose from paying more for existing workers, but the associated loss is bounded from above by $\varepsilon'\beta$. Therefore, for any sufficiently small ε', firm j's profit increases, as desired.

Cases 4 and 5 exhaust the possibility of a core outcome in which there exists a set V of positive Lebesgue measure such that $\max\{w_1^g(v),w_2^g(v)\}<v$ for almost all $v\in V$. Therefore, we proceed with the assumption that for almost any $v\in[0,1]$ there exists a firm j such that $w_j^g(v)=v$.

Sixth, suppose there exist a set V of positive Lebesgue measure and a firm j such that $0\leq w_{-j}^g(v)<w_j^g(v)=v$ and $f_g^{-j}(v)\in(0,f_g(v))$ for all $v\in V$. We proceed by showing that firm j can fire all workers of type (g,v), $v\in V$ and poach all workers of the same type from firm $-j$ by paying higher wages than $-j$ does and increase its profit. Let $\tilde{O}_j$ be defined as follows:

$$\tilde{f}_g^j(v):=\begin{cases} f_g^{-j}(v) & \text{if } v\in V,\\ f_g^j(v) & \text{otherwise.}\end{cases}\qquad \tilde{w}_j^g(v):=\begin{cases} (w_{-j}^g(v)+v)/2 & \text{if } v\in V,\\ w_g^j(v) & \text{otherwise.}\end{cases}$$

$\tilde{O}_j$ blocks O. To see that $\tilde{O}_j$ blocks O, note that Condition 3 of the definition of block is satisfied for all $v\in V$ by construction, and Condition 2 of the definition of block is satisfied for

all $v \in [0,1] \setminus V$ (i.e. there are no changes in employment or wages for workers with $v \in [0,1] \setminus V$). Therefore, it remains only to show that $\tilde{O}_j$ yields firm j a strictly higher profit than does O_j, and this fact follows because j earns no profit in O_j from worker types (g,v), $v \in V$, because $w_j^g(v) = v$, and because j earns strictly positive profit in $\tilde{O}_j$ from worker types (g,v), $v \in V$ because $\tilde{f}_g^j(v) = f_g^{-j}(v) > 0$ (where last inequality follows from the ongoing assumption of this case) and $\tilde{w}_j^g(v) < v$ for all $v \in V \setminus \{0\}$.

As these six cases are exhaustive and none of them admits a core outcome, we have completed the argument that any core outcome must be such that all workers are hired and all workers receive a wage equal to their productivity. □

Proof of Proposition 2

We show the desired conclusions via multiple steps. First, we show that if firm (without loss of generality) 1 hires almost no $B-$group workers, then it hires almost all $A-$group workers and firm 2 hires almost all $B-$group workers. Therefore, the only possible core outcomes that do not feature segregation involve each firm hiring from both groups at a common wage. Second, we show that if there exists a core outcome featuring common wages, it must be that the two firms pay different wages. Third, we show that generically there exists no core outcome in which each firm pays common wages and hires from both groups. Throughout, it suffices to assume $w_i^g(v) \leq v$ for all $v \in [0,1]$, all $i \in \{1,2\}$, and all $g \in \{A,B\}$ by the Individual Rationality Condition.

Step 1: We show that in any core outcome O if $f_B^1(v) = 0$ for almost all v, then it must be that $f_A^1(v) = f_A(v)$ for almost all v and $f_B^2(v) = f_B(v)$ for almost all v. We show this in several exhaustive cases.

First, we will show that if $f_B^1(v) = 0$ for almost all v, then it must be that there exists a set $V \subset [0,1]$ of positive Lebesgue measure such that $f_A^1(v) > 0$ for all $v \in V$. Suppose not toward a contradiction so that $f_B^1(v) = 0$ for almost all v and $f_A^1(v) = 0$ for almost all v. Then, $\pi_1^O = 0$. By the Equal Profit Condition it must be that $\pi_2^O = 0$. There are two possibilities to consider. First, it may be that there exist a group $g \in \{A,B\}$ and a set V with positive Lebesgue measure such that $f_g^2(v) < f_g(v)$ for all $v \in V$. Without loss of generality let $g = A$.

Then firm 1 can block outcome O via $\tilde{O}_1$ where

$$\tilde{f}_A^1(v) := f_A(v) - f_A^2(v) - f_A^1(v) \text{ for all } v \qquad \tilde{w}_1^A(v) := 0 \text{ for all } v,$$
$$\tilde{f}_B^1(v) := 0 \text{ for all } v \qquad \tilde{w}_1^B(v) := 0 \text{ for all } v.$$

which yields a positive profit, contradiction. Second, it may be that $f_A^2(v) = f_A(v)$ and $f_B^2(v) = f_B(v)$ for almost all v. By EPSW, it must be that firm 2 pays a common wage to almost all workers: there exists $w_2 \geq 0$ such that $w_2 = w_2^A(v) = w_2^B(v)$ for almost all v. Moreover, by the ongoing assumption that $w_i^g(v) \leq v$ for all $v \in [0,1]$, all $i \in \{1,2\}$, and all $g \in \{A,B\}$, it must be the case that $w_2 = 0$. But then for any $w^* \in (0,1)$ firm 1 can block O via $\tilde{O}_1$ where for each $g \in \{A,B\}$:

$$\tilde{f}_g^1(v) := \begin{cases} f_g(v) - f_g^1(v) & \text{if } v \geq w^*, \\ 0 & \text{otherwise.} \end{cases} \qquad \tilde{w}_1^g(v) := \begin{cases} 0 & \text{if } \tilde{f}_g^1(v) = 0, \\ w^* & \text{otherwise.} \end{cases}$$

which yields positive profit, contradiction.

Second, we will prove that if $f_B^1(v) = 0$ for almost all v and there exists a positive Lebesgue measure set V such that $f_A^1(v) > 0$ for all $v \in V$, then there does not exist any pair of sets V_A and V_B with positive Lebesgue measure such that $f_A^2(v) > 0$ for all $v \in V_A$ and $f_B^2(v) > 0$ for all $v \in V_B$. To show this, assume for contradiction that there exist sets V_A and V_B with positive Lebesgue measure such that $f_A^2(v) > 0$ for all $v \in V_A$ and $f_B^2(v) > 0$ for all $v \in V_B$. Then due to the EPSW, it must be that firm 2 pays a common wage w_2 to almost all workers it hires, i.e. $w_2^A(v) = w_2^B(v) = w_2$ for almost all v such that $f_A^2(v) > 0$ and $f_B^2(v) > 0$, respectively. By our ongoing assumption that $w_2^g(v) \leq v$ for all $v \in [0,1]$ and all $g \in \{A,B\}$, it must be the case that $f_A^2(v) = 0$ for all $v \leq w_2$. This implies that firm 2 is earning a positive profit from the $A-$group workers it hires as $w_2 < v$ for almost all $v \in V_A$. By an argument similar to those made in the proof of Case 4 of Lemma 3, firm 1 can block O by "poaching" some subset of these workers and increase its profits.

Third, we will establish that if $f_B^1(v) = 0$ for almost all v and there exists a set V with positive Lebesgue measure such that $f_A^1(v) > 0$ for all $v \in V$ then $f_A^2(v) = 0$ for almost all v. To show this, by the previous case, it suffices to consider for contradiction that $f_B^2(v) = 0$ for almost all v and that there exists a positive Lebesgue measure set V_A such that $f_A^2(v) > 0$ for all $v \in V_A$. The conclusion of Proposition 1 applies, and the two firms "Bertrand" compete away profits from $A-$group workers such that they both earn zero profits, i.e. $w_1^A(v) = v$

for almost all v such that $f_A^1(v)>0$ and $w_2^A(v)=v$ for almost all v such that $f_A^2(v)>0$. But then an arbitrary firm j can block via outcome $\tilde{O}_j$:

$$\tilde{f}_A^j(v):=0 \text{ for all } v \qquad \tilde{w}_j^A(v):=0 \text{ for all } v$$
$$\tilde{f}_B^j(v):=f_B(v)-f_B^{-j}(v) \text{ for all } v \qquad \tilde{w}_j^B(v):=0 \text{ for all } v$$

The ongoing assumption that $f_B^1(v)=0$ for almost all v and $f_B^2(v)=0$ for almost all v demonstrates that $\tilde{O}_j$ indeed blocks O.

Fourth, we will show that if $f_B^1(v)=0$ for almost all v, there exists a set V with positive Lebesgue measure such that $f_A^1(v)>0$ for all $v\in V$, and $f_A^2(v)=0$ for almost all v then it must be that $f_A^1(v)=f_A(v)$ for almost all v. This claim is shown via the second and third cases in the proof of Lemma 3.

Step 2: Step 1 implies the only remaining possible core outcomes in which firms do not completely segregate involves each firm hiring both groups of workers, i.e. there exist sets V_A^1,V_A^2,V_B^1,V_B^2 with positive Lebesgue measure such that $f_A^1(v)>0$ for all $v\in V_A^1$, $f_A^2(v)>0$ for all $v\in V_A^2$, $f_B^1(v)>0$ for all $v\in V_B^1$, and $f_B^2(v)>0$ for all $v\in V_B^2$. Suppose for contradiction that there exists such a non-segregation core outcome O. We have argued that in any such core outcome, each firm $i\in\{1,2\}$ must pay a common wage w_i to almost every worker it hires. Without loss of generality, let $w_1\leq w_2$. We first claim that, if $w_1<w_2$:

1. For almost all $v<w_1$, all workers of productivity v are unemployed: for almost all $v<w_1$, $f_A^1(v)=f_B^1(v)=f_A^2(v)=f_B^2(v)=0$,
2. for almost all $v\in(w_2,1]$, all workers of productivity v are hired by firm 2, i.e. $f_A^2(v)=f_A(v)$ and $f_B^2(v)=f_B(v)$, and
3. for almost all $v\in[w_1,w_2]$, all workers of productivity v are hired by firm 1, i.e. $f_A^1(v)=f_A(v)$ and $f_B^1(v)=f_B(v)$.

Point 1 follows from our previous individual rationality argument that no firm hires a positive measure of workers at wage higher than productivity in any core outcome. Point 2 is demonstrated with the following argument. Individual rationality implies firm 2 will hire almost no workers with productivity $v\in[w_1,w_2]$. If there exists a set $V\subset[w_2,1]$ with positive measure such that $f_A^2(v)<f_A(v)$ or $f_B^2(v)<f_B(v)$ for all $v\in V$, then firm 2 can

block outcome O via $\tilde{O}_2$ such that for all $g \in \{A,B\}$:

$$\tilde{f}_g^2(v) := \begin{cases} f_g(v) & \text{if } v \in (w_2,1], \\ 0 & \text{otherwise.} \end{cases} \qquad \tilde{w}_2^g(v) := \begin{cases} 0 & \text{if } \tilde{f}_g^2(v) = 0, \\ w_2 & \text{otherwise.} \end{cases}$$

which clearly blocks O as firm 2's profit increases (from hiring additional workers at a positive marginal profit) and Condition 1 of the definition of block is satisfied for all $v \in (w_2,1]$. Point 3 follows from a similar argument as the proof of point 2.

We now claim that in any core outcome O it must be that $w_1 < w_2$ (recall that we have assumed $w_1 \leq w_2$). To see this, suppose for contradiction that $w_1 = w_2$. If $w_1 = w_2 < 1$, then at least one firm $i \in \{1,2\}$ receives total profit $\pi_i^O < \int_{w_1}^1 (v - w_1)[\beta f_A(v) + f_B(v)]dv$. But because F_A and F_B are atomless, for sufficiently small $\epsilon > 0$, firm i would receive $\int_{w_1+\epsilon}^1 (v - w_1 - \epsilon)[\beta f_A(v) + f_B(v)]dv > \pi_i^O$ by instead setting wage $w' = w_1 + \epsilon$ and hiring all workers with productivity strictly greater than $w_1 + \epsilon$. This is a contradiction to the assumption that O is a core outcome. If $w_1 = w_2 = 1$, then $\pi_1^O = \pi_2^O = 0$ and so either firm could block O by setting any wage $w' \in (0,1)$ and hiring all workers with productivities $v \in [w',1)$, which yields profit $\int_{w'}^1 (v - w')[\beta f_A(v) + f_B(v)]dv > 0$. This contradicts the assumption that O is a core outcome. As these two cases are exhaustive, it must be the case that $w_1 < w_2$ in any core outcome in which $w_1 \leq w_2$.

Step 3: We now show that generically, any outcome O in which $f_A^1(v) = f_B^1(v) = f_A^2(v) = f_B^2(v) = 0$ for almost all $v < w_1$, $f_A^1(v) = f_A(v)$ and $f_B^1(v) = f_B(v)$ for almost all $v \in [w_1, w_2]$, and $f_A^2(v) = f_A(v)$ and $f_B^2(v) = f_B(v)$ for almost all $v > w_2$ is not a core outcome. We consider two exhaustive cases.

First, suppose that $0 < w_1 < w_2$. Let $\pi_{i,g}$ represent the profit firm $i \in \{1,2\}$ receives from all workers in group $g \in \{A,B\}$ in outcome O, or more formally,

$$\pi_{i,A} := \beta \int_0^1 (v - w_i(v)) f_A^i(v) dv, \qquad \pi_{i,B} := \int_0^1 (v - w_i(v)) f_B^i(v) dv.$$

It must be the case that $\pi_{1,A} + \pi_{2,A} \geq \pi_{1,B} + \pi_{2,B}$, or vice versa. Without loss of generality, assume $\pi_{1,A} + \pi_{2,A} \geq \pi_{1,B} + \pi_{2,B}$. We claim that firm 1 can block outcome O via outcome

$\tilde{O}_1$ where for some small $\epsilon'>0$:

$$\tilde{f}_A^1(v):=f_A(v) \qquad \text{for all } v\in[0,1] \qquad \tilde{w}_1^A(v):=\begin{cases}0 & \text{if } v<w_1,\\ w_1 & \text{if } v\in[w_1,w_2],\\ w_2+\epsilon' & \text{if } v>w_2.\end{cases}$$

$$\tilde{f}_B^1(v):=0 \qquad \text{for all } v\in[0,1] \qquad \tilde{w}_1^B(v):=0 \qquad \text{for all } v\in[0,1]$$

For sufficiently small ϵ', $\pi_1^{\tilde{O}_1} > \pi_{1,A}+\pi_{2,A} \geq \pi_1$, where the last inequality comes from the assumption that $\pi_{1,A}+\pi_{2,A}\geq\pi_{1,B}+\pi_{2,B}$ and the Equal Profit Condition $\pi_1=\pi_2=\frac{1}{2}(\pi_{1,A}+\pi_{2,A}+\pi_{1,B}+\pi_{2,B})$. For almost all $A-$group workers, Condition 1 of the definition of block is satisfied, and for almost all $B-$group workers, Condition 4 of the definition of block is satisfied. Therefore, $\tilde{O}_1$ blocks O.

Second, suppose that $0=w_1<w_2$. The argument is analogous to case 1 if $\pi_{1,A}+\pi_{2,A}>\pi_{1,B}+\pi_{2,B}$ or $\pi_{1,A}+\pi_{2,A}<\pi_{1,B}+\pi_{2,B}$. Below, we argue that the condition $\pi_{1,A}+\pi_{2,A}=\pi_{1,B}+\pi_{2,B}$ is non-generic on the space of distributions:

First, we will show the openness of the set of distributions F_A,F_B such that $\pi_{1,A}+\pi_{2,A}>\pi_{1,B}+\pi_{2,B}$ or $\pi_{1,A}+\pi_{2,A}<\pi_{1,B}+\pi_{2,B}$. To do so, assume that the former inequality holds at F_A,F_B (the other case is analogous), where the core outcome wages of firms 1 and 2 are 0 and w_2, respectively. Let $\epsilon>0$ be small enough that

$$\pi_{1,A}+\pi_{2,A}>\pi_{1,B}+\pi_{2,B}+\epsilon. \tag{13}$$

By the Portmanteu Theorem,[34] for any $\tilde{w}_2\in[0,1]$ it follows that

$$\beta\int_0^{\tilde{w}_2} vf'_A(v)dv \to \beta\int_0^{\tilde{w}_2} vf_A(v)dv,$$

as (F'_A,F'_B) with densities f'_A,f'_B converges in weak* topology to (F_A,F_B). Thus, there is a neighborhood of (F_A,F_B) such that, for any (F'_A,F'_B) in that neighborhood, we have

$$\beta\int_0^{\tilde{w}_2} vf'_A(v)dv\in\left(\beta\int_0^{\tilde{w}_2} vf_A(v)dv-\frac{\epsilon}{8},\beta\int_0^{\tilde{w}_2} vf_A(v)dv+\frac{\epsilon}{8}\right). \tag{14}$$

Next, again by Portmanteu Theorem, it follows that for any $\delta<w_2$ (where w_2 is the wage

[34]See Theorem 2.8.1 (c) of Ash and Doléans-Dade (2000).

paid by firm 2 when the distributions of worker productivities are given by (F_A,F_B)),

$$\beta\int_0^{w_2-\delta} vf'_A(v)dv+\int_0^{w_2-\delta} vf'_B(v)dv\to\beta\int_0^{w_2-\delta} vf_A(v)dv+\int_0^{w_2-\delta} vf_B(v)dv$$

and

$$\beta\int_{w_2-\delta}^1 (v-(w_2-\delta))f'_A(v)dv+\int_{w_2-\delta}^1 (v-(w_2-\delta))f'_B(v)dv\to\beta\int_{w_2-\delta}^1 (v-(w_2-\delta))f_A(v)dv+\int_{w_2-\delta}^1 (v-(w_2-\delta))f_B(v)dv,$$

as (F'_A,F'_B) converges in weak* topology to (F_A,F_B). Because we know that

$$\beta\int_0^{w_2-\delta} vf_A(v)dv+\int_0^{w_2-\delta} vf_B(v)dv<\beta\int_{w_2-\delta}^1 (v-(w_2-\delta))f_A(v)dv+\int_{w_2-\delta}^1 (v-(w_2-\delta))f_B(v)dv$$

from the construction of outcome O at (F_A,F_B), it follows that there exists a neighborhood of (F_A,F_B) such that, for any pair of distributions (F'_A,F'_B) in that neighborhood, we have

$$\beta\int_0^{w_2-\delta} vf'_A(v)dv+\int_0^{w_2-\delta} vf'_B(v)dv<\beta\int_{w_2-\delta}^1 (v-(w_2-\delta))f'_A(v)dv+\int_{w_2-\delta}^1 (v-(w_2-\delta))f'_B(v)dv,$$

and, with an analogous argument, that

$$\beta\int_0^{w_2+\delta} vf'_A(v)dv+\int_0^{w_2+\delta} vf'_B(v)dv>\beta\int_{w_2+\delta}^1 (v-(w_2+\delta))f'_A(v)dv+\int_{w_2+\delta}^1 (v-(w_2+\delta))f'_B(v)dv.$$

These inequalities imply that if firm 2 pays w'_2 to all workers it employs and firm 1 pays $w'1=0$ to all workers it employs, both given distributions (F'_A,F'_B), such that the Equal Profit Condition is satisfied, then $w'_2\in(w_2-\delta,w_2+\delta)$. Moreover, by monotonicity of the expression $\beta\int_0^{w'_2} vf'_A(v)dv$ in w'_2, it follows that $\beta\int_0^{w'_2} vf'_A(v)dv\in\left(\beta\int_0^{w_2-\delta} vf'_A(v)dv,\beta\int_0^{w_2+\delta} vf'_A(v)dv\right)$ for any distribution (F'_A,F'_B) in the neighborhood. Take a sufficiently small $\delta>0$ such that

$$\beta\int_0^{w_2-\delta} vf_A(v)dv>\beta\int_0^{w_2} vf_A(v)dv-\frac{\epsilon}{8} \text{ and } \beta\int_0^{w_2+\delta} vf_A(v)dv<\beta\int_0^{w_2} vf_A(v)dv+\frac{\epsilon}{8}. \quad (15)$$

Moreover, we apply (14) to $\tilde{w}_2 = w_2 - \delta$ and $\tilde{w}_2 = w_2 + \delta$ to obtain

$$\beta \int_0^{w_2-\delta} v f'_A(v)dv \in \left(\beta \int_0^{w_2-\delta} v f_A(v)dv - \frac{\epsilon}{8}, \beta \int_0^{w_2-\delta} v f_A(v)dv + \frac{\epsilon}{8}\right),$$
$$\beta \int_0^{w_2+\delta} v f'_A(v)dv \in \left(\beta \int_0^{w_2+\delta} v f_A(v)dv - \frac{\epsilon}{8}, \beta \int_0^{w_2+\delta} v f_A(v)dv + \frac{\epsilon}{8}\right). \tag{16}$$

Thus, by (15) and (16) it follows that $\pi'_{1,A} \in (\pi_{1,A} - \frac{\epsilon}{4}, \pi_{1,A} + \frac{\epsilon}{4})$. Similarly we can establish that $\pi'_{i,g}$ for $i \in \{1,2\}, g \in \{A,B\}$, each satisfies $\pi'_{i,g} \in (\pi_{i,g} - \frac{\epsilon}{4}, \pi_{i,g} + \frac{\epsilon}{4})$. This and (13) imply $\pi'_{1,A} + \pi'_{2,A} > \pi'_{1,B} + \pi'_{2,B}$, completing the proof of openness.

Second, we will show the denseness of the set of distributions F_A, F_B such that $\pi_{1,A} + \pi_{2,A} > \pi_{1,B} + \pi_{2,B}$ or $\pi_{1,A} + \pi_{2,A} < \pi_{1,B} + \pi_{2,B}$. To do so, suppose for a given F_A, F_B satisfying our regularity conditions there is a core outcome that does not feature segregation, i.e. $\pi_{1,A} + \pi_{2,A} = \pi_{1,B} + \pi_{2,B}$. We show that for any $\delta_1 > 0$ there exist distributions F'_A, F'_B satisfying our regularity conditions such that $|F'_A(v) - F_A(v)| \leq \delta_1$ and $|F'_B(v) - F_B(v)| \leq \delta_1$ for all $v \in [0,1]$ and for which there does not exist a non-segregation core outcome.[35] We show this by modifying the distributions such that the same wages by firms 1 and 2, 0 and w_2, respectively, yield equal profit, but show that aggregate profit derived from $A-$group workers no longer equals that derived from $B-$group workers.

Take any $\delta \in (0, \min\{\underline{f}_A, \underline{f}_B, \delta_1\})$ and any $w^* \in (0, w_2)$ such that $w_2 - w^* < 1 - w_2$. Define F'_A and F'_B via their respective densities:

$$f'_A(v) := \begin{cases} f_A(v) & \text{if } v < w^*, \\ f_A(v) - \frac{\delta}{\beta} & \text{if } v \in [w^*, w_2), \\ f_A(v) + \frac{\delta}{\beta} & \text{if } v \in [w_2, w_2 + (w_2 - w^*)), \\ f_A(v) & \text{otherwise} \end{cases}, \qquad f'_B(v) := \begin{cases} f_B(v) & \text{if } v < w^*, \\ f_B(v) + \delta & \text{if } v \in [w^*, w_2), \\ f_B(v) - \delta & \text{if } v \in [w_2, w_2 + (w_2 - w^*)), \\ f_B(v) & \text{otherwise} \end{cases}.$$

Noting there exists a β measure of $A-$group workers, the construction of F'_A and F'_B imply

$$\pi_{2,A} + \pi_{2,B} = \beta \int_{w_2}^1 [v - w_2] f_A(v)dv + \int_{w_2}^1 [v - w_2] f_B(v)dv = \beta \int_{w_2}^1 [v - w_2] f'_A(v)dv + \int_{w_2}^1 [v - w_2] f'_B(v)dv = \pi'_{2,A} + \pi'_{2,B},$$

[35] Note that our pointwise notion of "closeness" corresponds to that of the weak-* topology on the set of distributions satisfying our regularity conditions.

$$\pi_{1,A}+\pi_{1,B}=\beta\int_0^{w_2} vf_A(v)dv+\int_0^{w_2} vf_B(v)dv=\beta\int_0^{w_2} vf'_A(v)dv+\int_0^{w_2} vf'_B(v)dv=\pi'_{1,A}+\pi'_{1,B},$$

where $\pi'_{i,g}$ represents firm i's profit from $g-$group workers it hires given distribution F'_g. The above two equations imply that $\pi'_i:=\pi'_{i,A}+\pi'_{i,B}=\pi_i^{O_i}$ for $i\in\{1,2\}$. It is also the case that

$$\pi_{1,A}+\pi_{2,A}-(\pi'_{1,A}+\pi'_{2,A})=\beta\int_{w_2}^{w_2+(w_2-w^*)}[v-w_2]\frac{\delta}{\beta}dv+\beta\int_{w^*}^{w_2} v\frac{\delta}{\beta}dv=\delta(w_2-w^*)w_2>0,$$

where the final inequality comes from the assumptions that $w_2>w^*>0$. Similarly it can be shown that

$$\pi_{1,B}+\pi_{2,B}-(\pi'_{1,B}+\pi'_{2,B})<0.$$

The two preceding equations imply that $\pi'_{1,A}+\pi'_{2,A}<\pi_{1,A}+\pi_{2,A}$, and $\pi_{1,B}+\pi_{2,B}<\pi'_{1,B}+\pi'_{2,B}$. The initial condition that $\pi_{1,A}+\pi_{2,A}=\pi_{1,B}+\pi_{2,B}$ further implies that $\pi'_{1,A}+\pi'_{2,A}<\pi'_{1,B}+\pi'_{2,B}$ which completes the claim. □

Proof of Proposition 3

Proof of Part 1: We first constructively show that there exist a continuum of core outcomes $O(\delta,\delta')$ with group-based segregation, parameterized by the following class of wage functions for $\delta,\delta'\in[0,1]$:

$$w_1(v):=\begin{cases} v & \text{if } v<\delta,\\ \delta & \text{otherwise.}\end{cases}\qquad w_2(v):=\begin{cases} 0 & \text{if } v<\delta',\\ v & \text{otherwise.}\end{cases}$$

with firms 1 and 2 hiring all $A-$ and $B-$group workers, respectively. We show that these wage functions permit a core outcome if $\delta\geq\delta'$ and

$$\pi_1^{O(\delta,\delta')}=\beta\int_\delta^1(v-\delta)f_A(v)dv=\int_0^{\delta'} vf_B(v)dv=\pi_2^{O(\delta,\delta')}. \tag{17}$$

To see that the above wage functions permit a core outcome $O(\delta,\delta')$, first note that there is no firm i and outcome $\tilde{O}_i$ with group-based segregation that blocks $O(\delta,\delta')$. Suppose without loss of generality that in outcome $\tilde{O}_i$, firm i hires a positive measure of $A-$group workers only

with associated wage function $\tilde{w}_i(\cdot)$. By the definition of block, for almost every v, $\tilde{f}_A^i(v)>0$ only if $\tilde{w}_i(v)\geq w_1(v)$, implying that $\pi_i^{\tilde{O}_i}\leq\pi_1^{O(\delta,\delta')}=\pi_2^{O(\delta,\delta')}$. Therefore, $\tilde{O}_i$ is not a block.

Thus, consider a firm i and a potential blocking outcome $\tilde{O}_i$ in which i employs positive measures of both $A-$ and $B-$group workers. Under EPSW, such a firm pays a common wage w to almost all workers. There are two exhaustive cases:

- Consider $w\leq\delta$. By the definition of block, firm i does not benefit from hiring $A-$group workers, as almost all $A-$group workers with $v\geq w$ are paid $w_1(v)=\min\{v,\delta\}\geq w$ in outcome $O(\delta,\delta')$. Therefore, $\pi_i^{\tilde{O}_i}$ is upper bounded by the case in which firm i hires only $B-$group workers. It has been shown in the preceding paragraph, however, that there exists no such block.

- Consider $w>\delta$. By the definition of block, firm i does not benefit from hiring $B-$group workers, as almost all $B-$group workers with $v\geq w$ are paid $w_2(v)=v$ in outcome $O(\delta,\delta')$. Therefore, $\pi_i^{\tilde{O}_i}$ is upper bounded by the case in which firm i hires only $A-$group workers. It has been shown in the preceding paragraph, however, that there exists no such block.

The preceding arguments demonstrate that any outcome $O(\delta,\delta')$ with $\delta\geq\delta'$ and $\pi_1^{O(\delta,\delta')}=\pi_2^{O(\delta,\delta')}$ is a core outcome. To see that there is a continuum of pairs δ,δ' that satisfy these conditions, note that if $\delta=\delta'=1$, then the right-hand side of (17) is strictly positive while the left-hand side is zero, and hence the former is strictly larger than the latter. Because of continuity of the left-hand side in δ, the left-hand side is strictly smaller than the right-hand side for any δ that is sufficiently close to 1. Now, noting continuity of the right-hand side in δ' and the fact that it is equal to zero for $\delta'=0$, by intermediate value theorem, it follows that there exists δ' such that (17) holds with equality. This concludes the proof of part 1. □

Proof of Part 3: Recall that AW_A^O and AW_B^O denote the average wages for $A-$ and $B-$group workers, respectively. In addition, denote $TS_A:=\beta\mathbb{E}_A(v)$ and $TS_B:=\mathbb{E}_B(v)$ to be the total surpluses created by $A-$ and $B-$group workers (together with the firms hiring them), respectively. Then, note that $\pi_1^O=TS_A-\beta AW_A^O$ and $\pi_2^O=TS_B-AW_B^O$. The Equal Profit Condition tells us

$$\begin{aligned} &TS_A-\beta AW_A^O=TS_B-AW_B^O \\ \Longleftrightarrow &TS_A-TS_B+(1-\beta)AW_A^O=AW_A^O-AW_B^O. \end{aligned} \tag{18}$$

Proof of the "only if" part: Suppose that $AW_A^O - AW_B^O \leq AW_A^{O'} - AW_B^{O'}$ for core outcomes O and O'. Then (18) implies that $TS_A - TS_B + (1-\beta)AW_A^O \leq TS_A - TS_B + (1-\beta)AW_A^{O'}$. Note that $TS_A - TS_B$ is a constant. Since $\beta > 1$, this implies that $AW_A^O \geq AW_A^{O'}$. Since firm 1 hires almost all $A-$group workers in both O and O', this implies that $\pi_1^O \leq \pi_1^{O'}$. Then $\pi_2^O = \pi_1^O \leq \pi_1^{O'} = \pi_2^{O'}$, where the equalities follow from the Equal Profit Condition. □

Proof of the "if" part: Suppose $\pi_1^O \leq \pi_1^{O'}$ for two core outcomes O and O' (which implies $\pi_2^O = \pi_1^O \leq \pi_1^{O'} = \pi_2^{O'}$ by the Equal Profit Condition). Since firm 1 hires almost all $A-$group workers in both O and O', this implies that $AW_A^O \geq AW_A^{O'}$. Since $\beta > 1$, this implies that $TS_A - TS_B + (1-\beta)AW_A^O \leq TS_A - TS_B + (1-\beta)AW_A^{O'}$. Thus, by (18), $AW_A^O - AW_B^O \leq AW_A^{O'} - AW_B^{O'}$. □

Proof of Part 2: Consider the core outcome in which $w_1(v) = w_2(v) = v$ for all $v \in [0,1]$ and all $A-$ and $B-$group workers are hired by firms 1 and 2, respectively. The wage gap in this core outcome is the same as in any core outcome without EPSW, and both firms' profits are zero. It is straightforward to see that in any non-equivalent core outcome firm profits are strictly positive. Thus, applying the conclusion of Part 3 completes the claim. □

Proof of Part 4: We first show that if $\beta\mathbb{E}_A[v] > \mathbb{E}_B[v]$ then the analogue of Part 3 holds. Let $\pi^O := \pi_1^O = \pi_2^O$ be the profit of each firm in core outcome O, where the equality follows from the Equal Profit Condition. Then, $AW_A^O = (\beta\mathbb{E}_A[v] - \pi^O)/\beta$ and $AW_B^O = \mathbb{E}_B[v] - \pi^O$ because every worker is hired in any core outcome. So, the wage ratio is $AW_A^O/AW_B^O = (\beta\mathbb{E}_A[v] - \pi^O)/[\beta(\mathbb{E}_B[v] - \pi^O)]$ if $AW_B^O \neq 0$. By individual rationality, it must be the case that $\pi^O \geq 0$ and by the Equal Profit Condition and the assumption that $\beta\mathbb{E}_A[v] > \mathbb{E}_B[v]$, it must be the case that $\pi^O \leq \mathbb{E}_B[v]$. Note that if $\pi^O = 0$ (which is equivalent to $AW_B^O = \mathbb{E}_B[v]$) then $AW_A^O/AW_B^O = \mathbb{E}_A[v]/\mathbb{E}_B[v]$, and if $\pi^O = \mathbb{E}_B[v]$ (which is equivalent to $AW_B^O = 0$) then by definition the wage ratio is ∞. Therefore, it remains only to show that the wage ratio is strictly increasing in π^O over the interval $(0, \mathbb{E}_B[v])$. To see that this is true, consider two core outcomes O and O' such that $\pi^O > \pi^{O'}$. $AW_A^O/AW_B^O > AW_A^{O'}/AW_B^{O'}$ if and only if $(\beta\mathbb{E}_A[v] - \pi^O)/(\mathbb{E}_B[v] - \pi^O) > (\beta\mathbb{E}_A[v] - \pi^{O'})/(\mathbb{E}_B[v] - \pi^{O'})$. Cross multiplying reveals this latter condition is equivalent to $(\mathbb{E}_B[v] - \pi^{O'})(\beta\mathbb{E}_A[v] - \pi^O) > (\beta\mathbb{E}_A[v] - \pi^{O'})(\mathbb{E}_B[v] - \pi^O)$, which is in turn equivalent to $\beta\mathbb{E}_A[v] > \mathbb{E}_B[v]$, which is true by assumption, thus demonstrating

the desired claim.

To see that the analogue of Part 2 holds, consider the core outcome O in which $w_1^O(v)=w_2^O(v)=v$ for all $v\in[0,1]$ and all $A-$ and $B-$group workers are hired by firms 1 and 2, respectively. Then $\pi^O=0$, and we have established above that the wage ratio in this core outcome is the same as in any core outcome without EPSW. Now, it is straightforward to see that in any non-equivalent core outcome firm profits are strictly positive. Therefore, the fact that the wage ratio is strictly increasing in the profit in any core outcome completes the claim. □

Proof of Proposition 4

Proof of Equilibrium Existence: Fix any search intensity r satisfying the previously imposed regularity conditions. In order to show the existence of an equilibrium comporting with r, we consider another game which "automates" the search and bargaining stages of our original game and is payoff equivalent for firms to our original game. Let $\mathcal{G}$ denote the original game, where we suppress model primitives. Let $\mathcal{G}'_r$ be a game of perfect information defined as follows:

- The set of players in $\mathcal{G}'_r$ is the set of firms in $\mathcal{G}$, i.e. workers are not players in $\mathcal{G}'_r$. The same set of firms are "unconstrained" in both games.
- The set of time periods in $\mathcal{G}'_r$ is $\{-\underline{t},\underline{t}+1,...,-1,0,1,2,...\}$.
- Firms have no actions available to them in $\mathcal{G}'_r$ in periods $t<0$. In each period $t\geq 0$, unconstrained firms have no actions available to them, and each of the other firms can select from the following options: segregate for group A, segregate for group B, or desegregate at any endogenously selected wage $w\in[0,1]$. As in $\mathcal{G}$, the decision to segregate is irreversible, i.e. once a firm has segregated for a group g, then that firm continues to be segregated for group g in all subsequent periods. Moreover, the segregation decision is made sequentially, following the same order as in $\mathcal{G}$.
- Given any history of play up to and including some arbitrary time period t in $\mathcal{G}'_r$, period-t payoffs for each firm equal the period-t payoff each firm receives in $\mathcal{G}$ given the same history of (de)segregation decisions and assuming search, bargaining, and exogenous departure rates follow according to r, (4), and d, respectively.
- All players in $\mathcal{G}'_r$ discount future payoffs according to the same geometric rate δ.

Let s' be a subgame perfect equilibrium of $\mathcal{G}'_r$. We claim that there is a corresponding equilibrium s in $\mathcal{G}$ in which the firms make the same (de)segregation decision following a

common history of (de)segregation decisions in s and s', and, both on and off path, workers search according to intensity r and bargaining offers/acceptances follow (4) for all players. Clearly such a strategy profile s satisfies our regularity conditions, and so it remains only to show that there is no decision node at any $t \geq 0$ at which any player can improve her payoff by deviating from s. To see this, note that by construction no firm in $\mathcal{G}$ has an incentive to deviate at any (de)segregation decision node, because no firm in $\mathcal{G}'_r$ has any incentive to deviate at any decision node by the equilibrium hypothesis. Also, no firm nor worker in $\mathcal{G}$ has an incentive to deviate at any other decision node due to the construction of bargaining offers resulting in (4), and worker search intensity r. By the one-shot deviation principle, s then satisfies subgame perfection in $\mathcal{G}$, as desired.

It therefore remains only to show that there exists a pure-strategy subgame perfect equilibrium s' in game $\mathcal{G}'_r$, which follows from Harris (1985): $\mathcal{G}'_r$ is a game of perfect information and clearly satisfies all the regularity conditions required in Harris (1985), and therefore admits at least one pure-strategy subgame perfect equilibrium. □

Proof of Part 1: To show Part 1, suppose for contradiction that, for any $t \geq 0$, there exists a pure-strategy equilibrium in which there exists a firm i that is desegregated at time t. Let $\pi_i = \frac{r_i^A}{\sum_{i' \in I} r_{i'}^A} \frac{\beta}{(1+\delta')^n} \mathbb{E}_A[v] > 0$ be the per-period profit firm i receives from hiring A-group workers at wages determined by (4) born in the current period when all other constrained firms are also segregated for A.[36] Also, let $\hat{V}$ be any finite upper bound on the present discounted value for a firm when it hires all workers at wage zero in each period, starting at the present period indefinitely into the future. Let $\underline{T}$ be a period such that firm i's profit from existing workers at period $\underline{T}$ is $\varepsilon \in (0, \pi_i)$; note that because desegregated firms hire no new workers on path and because workers exit the market independently with probability $d > 0$ in each period, such a period $\underline{T}$ must exist. Let $T_i^1 > \underline{T}$ be such that

$$\sum_{t=\underline{T}}^{\infty} \delta^{t-\underline{T}} (1-d)^{t-\underline{T}} \pi_i > \delta^{T_i^1 - \underline{T}} \hat{V} + \sum_{t=\underline{T}}^{T_i^1 - 1} \delta^{t-\underline{T}} (1-d)^{t-\underline{T}} \varepsilon. \tag{19}$$

[36] Noting that W^v as denoted in (4) is equal to zero because all firms are assumed to be segregated for A, the (infinitesimal) discounted expected lifetime profit that a worker of type (v,g) accrues to firm i is $\frac{1}{(1+\delta')^n} V^v$, and therefore the (unique, as shown in the discussion following (4)) per-period wage for such a worker is $w = \left(1 - \frac{1}{(1+\delta')^n}\right) v$ because $\sum_{\tau=1}^{\infty} \delta^{\tau-1} (1-d)^{\tau-1} w = \left(1 - \frac{1}{(1+\delta')^n}\right) \sum_{\tau=1}^{\infty} \delta^{\tau-1} (1-d)^{\tau-1} v = \left(1 - \frac{1}{(1+\delta')^n}\right) V^v$, where the first equality follows by substituting in the specified value for w and the second inequality follows from (1). Note that the final term is the residual surplus accruing to the worker from (4). Therefore, $\pi_i = \frac{r_i^A}{\sum_{i' \in I} r_{i'}^A} \frac{\beta}{(1+\delta')^n} \mathbb{E}_A[v] > 0$.

Because $\pi_i > \varepsilon$ and $\hat{V}$ is finite, such T_i^1 exists. The left-hand side of (19) is a lower bound of the continuation value at period $\underline{T}$ of segregating for A at period $\underline{T}$, while the right-hand side is an upper bound of the continuation value at period $\underline{T}$ from playing any strategy in which the firm desegregates at least until period T_i^1. Thus, (19) implies firm i has a profitable deviation, leading to a contradiction with the hypothesis that there exists an equilibrium in which i remains desegregated after time T_i^1. Noting that there are finitely many firms, our argument further establishes the existence of a uniform bound T^1 such that all firms are segregated by the end of time T^1 in any equilibrium, as desired. □

Proof of Part 2 and that every newly born worker is hired in equilibrium:

We first show that every newly born worker is hired in equilibrium. To see this, note that (4) implies that every newly born worker immediately reaches a bargaining agreement with the first firm that she bargains with in any time t.

Next, we show the claim in Part 2 holds. Let T^1 be the time period identified in the proof of Part 1 such that all firms are segregated by the end of time T^1 in any equilibrium. For any time $t \geq T^1$, the "outside option" $W^v = 0$ for all newly born workers with any productivity v because there are no desegregated firms. Then (4) implies that the surplus generated to a worker of type (g,v) who is born at any $t \geq T^1$ is given by

$$\left(1 - \frac{1}{(1+\delta')^{n_g+n_U}}\right)V, \tag{20}$$

and the worker's wage at each period (see Footnote 36) is

$$\left(1 - \frac{1}{(1+\delta')^{n_g+n_U}}\right)v. \tag{21}$$

Thus, the average wage of workers in group g is

$$\left(1 - \frac{1}{(1+\delta')^{n_g+n_U}}\right)\mathbb{E}_g[v] + \epsilon_{g,t}, \tag{22}$$

where $\epsilon_{g,t}$ denotes the contribution to the average wage of the workers of group g who are hired at earlier periods and have not exited the market by period t. Note that, because of exogenous exit of workers, $\epsilon_{g,t}$ converges to zero as $t \to \infty$ for each group g. Consider the case in which $n_A > n_B$ (the other cases of interest follow a nearly identical argument). The

wage ratio converges to

$$\frac{1-\frac{1}{(1+\delta')^{n_A+n_U}}}{1-\frac{1}{(1+\delta')^{n_B+n_U}}}\cdot\frac{\mathbb{E}_A[v]}{\mathbb{E}_B[v]}>\frac{\mathbb{E}_A[v]}{\mathbb{E}_B[v]}, \tag{23}$$

where the inequality follows because $n_A > n_B$. Noting that this inequality is strict and that the right-hand side is equal to the wage ratio before EPSW (see Observation 1), this inequality implies that there is some $T \geq T^1$ such that the wage ratio is strictly more in favor of group A for any $t \geq T$ compared to any period $t' < 0$. □

Proof of Proposition 5

Proof of Part 1: Suppose for contradiction that $n_A > x_A + 1$ (the case with $n_A < x_A - 1$ follows a symmetric argument, and is therefore omitted). Then, n_A and n_B satisfy $n_A - 1 > x_A$ and $n_B + 1 < x_B$. Noting that the left- and right-hand sides of (6) are decreasing in variables x_A and x_B, respectively, (6) implies that

$$\frac{1}{n_A+n_U-1}\cdot\frac{1}{(1+\delta')^{n_A+n_U-1}}\cdot\beta\mathbb{E}_A[v]<\frac{1}{n_B+n_U+1}\cdot\frac{1}{(1+\delta')^{x_B+n_U+1}}\cdot\mathbb{E}_B[v]. \tag{24}$$

Considering the left-hand side of the previous equation and noting $n_A > n_A - 1$, it follows that

$$\frac{1}{n_A+n_U}\cdot\frac{1}{(1+\delta')^{n_A+n_U}}\cdot\beta\mathbb{E}_A[v]<\frac{1}{n_B+n_U+1}\cdot\frac{1}{(1+\delta')^{x_B+n_U+1}}\cdot\mathbb{E}_B[v]. \tag{25}$$

Let

$$\Pi := \sum_{t=0}^{\infty}\left[\delta^t(1-d)^t\right]\left[\sum_{t'=-\underline{t}}^{-1}(1-d)^{t'+\underline{t}}\beta\mathbb{E}_A[v]\right]. \tag{26}$$

We note that Π is the discounted sum of any firm's payoff (evaluated at $t=0$) from hiring and keeping all $A-$group workers who are born before EPSW at wage zero. Therefore, Π gives an upperbound on the discounted sum of any firm's payoffs (evaluated at any period post EPSW) from hiring and keeping all $A-$group workers who are born and matched to the firm before EPSW. Noting that Π converges to zero as $d \to 1$, (25) implies that there

exists $d^* \in (0,1)$ such that, for any $d > d^*$ and $t_0 \in \mathbb{N}$,

$$\Pi + \sum_{t'=t_0}^{\infty} \sum_{t=t'}^{\infty} \left[\delta^{t-t_0} (1-d)^{t-t'} \frac{1}{n_A + n_U} \cdot \frac{1}{(1+\delta')^{n_A+n_U}} \cdot \beta \mathbb{E}_A[v] \right] < \sum_{t'=t_0}^{\infty} \sum_{t=t'}^{\infty} \left[\delta^{t-t_0} (1-d)^{t-t'} \frac{1}{n_B + n_U + 1} \cdot \frac{1}{(1+\delta')^{x_B+n_U+1}} \cdot \mathbb{E}_B[v] \right]. \tag{27}$$

Now, let T be the final time period during which any firm segregates for group A on equilibrium path–note that since the equilibrium is in pure strategies, such a period T is well defined. Then, the payoff (evaluated at T) of any firm i that segregates for A at time T is at most as high as the left-hand side of (27) with $t_0 = T$. Meanwhile, if firm i deviates and segregates for group B at time T, then i's continuation payoff (evaluated at T) following this deviation is at least as high as the right-hand side of (27) with $t_0 = T$.[37] Therefore, (27) implies that i is strictly better off by deviating, a contradiction to the equilibrium assumption. □

Proof of Part 2: Suppose $n_A^t < n_B^t$ for large t for contradiction. Assume, without loss of generality, t is large enough that every firm has segregated by then on the equilibrium path. Consider the last period at which a firm segregates for B. For this firm, the payoff from newly hired workers in each period is strictly higher if it deviates to segregating for A rather than segregating for B, because both the firm's surplus from each worker and the volume of workers reaching a bargaining agreement with it are larger if it is segregated for A than when it is segregated for B. Because of the assumption that $d \in (0,1)$ is sufficiently large, this means that this firm has an incentive to deviate to segregating for A in this period, contradicting the equilibrium hypothesis. □

B Non-cooperative Game Formulation

In this appendix, we describe a non-cooperative game played by firms and workers. We show that the pure-strategy subgame perfect Nash equilibrium outcomes of this game are equivalent to the core of the game analyzed in the main text.

The set of players are composed of two firms 1,2 and a continuum of workers. The set of workers are given by $\{A,B\} \times [0,1]^2$. A worker is identified by a tuple $(g,v,ß) \in \{A,B\} \times [0,1]^2$, where g is the group that the worker belongs to, v is her productivity, and ß is an index.

[37] To see this, note that $n_A - 1$ firms other than i segregate for A by the end of time t_0. So, in any period following i's deviation, there are at most $n_B + 1$ firms segregating for B.

For each g, we assume that there is measure μ_g that is given as a product measure of μ_g^{p} and μ_g^{w}. More specifically, let μ_g^{p} be the Lebesgue measure on Lebesgue σ-algebra $\mathcal{B}^{\mathrm{p}}$ on [0,1], representing the measure of productivity.[38] Let μ_g^{w} be a measure on a σ-algebra $\mathcal{B}^{\mathrm{w}}$ on [0,1], representing the measure of workers. We assume that μ_g^{w} is non-atomic. The density function associated with μ_g is given by $f_g(v,ß) = f_g^{\mathrm{p}}(v) \times f_g^{\mathrm{w}}(ß)$, where $f_g^{\mathrm{p}}(v)$ is associated with measure μ_g^{p} and represents the density of $g-$group workers with productivity v while $f_g^{\mathrm{w}}(ß)$ is associated with measure μ_g^{w} and represents the density of workers whose indices are ß.

We first provide a "microfoundation" to our approach to work with distribution functions over productivities in our game. Let $X = [0,1] \times [0,1]$. For $Y \subseteq X$ and $v \in [0,1]$, let $Y_v = \{t \in [0,1] : (v,t) \in Y\}$. Let μ_g denote the product measure induced from μ_g^{p} and μ_g^{w} with the corresponding σ-algebra $\mathcal{B}$. By the product measure theorem, there exists a unique product measure induced from the two measures. Theorem 7.14 of Stokey and Lucas (1989) implies that for any $Y \in \mathcal{B}$ and $v \in [0,1]$, it holds that $Y_v \in \mathcal{B}^{\mathrm{w}}$. We say that $h : [0,1] \to \mathbb{R}$ is $\mathcal{B}^{\mathrm{p}}$-measurable if $\{v \in [0,1] : h(v) \leq a\} \in \mathcal{B}^{\mathrm{p}}$ for all $a \in \mathbb{R}$.

The following proposition shows how we can interpret $\mu_g^w(\cdot)$ as the "density" of $g-$group workers with productivity v.

PROPOSITION **7.** *Let $h : [0,1] \to \mathbb{R}_{\geq 0}$ be a $\mathcal{B}^p$-measurable function such that $h(v) \leq c$ for all $v \in [0,1]$. Then, there exists a subset $Y^* \subseteq X$ such that $Y^* \in \mathcal{B}$ and $\mu_g^w(Y_v^*) = h(v)$ for all $v \in [0,1]$.*

Proof. We say that a function mapping [0,1] to $\mathbb{R}$ is *simple* if its range is a finite set. The following useful result is stated as Theorem 7.5 of Stokey and Lucas (1989, p. 180).

LEMMA **4.** *Suppose that $h : [0,1] \to \mathbb{R}_{\geq 0}$ is $\mathcal{B}^p$-measurable. Then, there exists a sequence of $\mathcal{B}^p$-measurable simple functions $(h^n)_{n=1}^{\infty}$ such that*

$$0 \leq h^n(v) \leq h^{n+1}(v) \leq h(v) \text{ for all } v \in [0,1] \text{ and } n = 1,2,..., \tag{28}$$

$$h^n(v) \to h(v) \text{ (as } n \to \infty) \text{ for all } v \in [0,1]. \tag{29}$$

[38] Formally, we define a Borel measure $\tilde{\mu}_g^{\mathrm{p}}$ such that $\tilde{\mu}_g^{\mathrm{p}}([0,x]) = F_g(x)$ for all $x \in [0,1]$, which exists and is unique (Royden and Fitzpatrick, 2010, Proposition 25, Section 20.3). Let μ_g^{p} be the unique measure defined on the Lebesgue measurable sets and coincides with $\tilde{\mu}_g^{\mathrm{p}}$ on Borel measurable sets: such μ_g^{p} exists and is unique due to the Caratheodory Extension Theorem and the Hahn Extension Theorem (see Stokey and Lucas, 1989, Theorems 7.3 and 7.2').

For $A \subseteq [0,1]$, let $\chi_A : [0,1] \to \{0,1\}$ denote the *indicator function* for A, i.e.

$$\chi_A(v) = \begin{cases} 1 & \text{if } v \in A, \\ 0 & \text{otherwise.} \end{cases}$$

By Lemma 4, there exists a sequence of $\mathcal{B}^{\mathrm{p}}$-measurable simple functions $(h^n)_{n=1}^{\infty}$ that satisfies (28) and (29). For each $n = 1,2,...$, since h^n is a simple function, its range consists of a finite number of reals; let $k(n)$ denote the number. Then, there exists a sequence of reals $(a^{n,r})_{r=1}^{k(n)}$ and a sequence of mutually disjoint $\mathcal{B}^{\mathrm{p}}$-measurable sets $(A^{n,r})_{r=1}^{k(n)}$ such that

$$h^n = \sum_{r=1}^{k(n)} a^{n,r} \cdot \chi_{A^{n,r}}.$$

We next state the following mathematical result, which is a known modification of Sierpiński's theorem on non-atomic measures.[39]

LEMMA **5.** *Suppose that μ_g^w is a non-atomic measure. Let $c = \mu_g^w([0,1])$. Then, there exists a function $\phi : [0,c] \to \mathcal{B}^w$ such that*

$$\mu_g^w(\phi(t)) = t \text{ for all } t \in [0,c], \text{ and} \tag{30}$$

$$\phi(t) \subseteq \phi(t') \text{ for all } t,t' \in [0,c] \text{ with } t \leq t'. \tag{31}$$

By Lemma 5, there exists a function ϕ that satisfies (30) and (31). For each $n = 1,2,...$, and $r = 1,2,...,k(n)$, by (28) and $h(v) \leq c$ for all $v \in [0,1]$, it holds that $a^{n,r} \leq c$. We define $\bar{A}^{n,r} = \phi(a^{n,r})$ for all $n = 1,2,...$, and $r = 1,2,...,k(n)$, and we define $Y^* = \cup_{n=1}^{\infty} \cup_{r=1}^{k(n)} (A^{n,r} \times \bar{A}^{n,r})$. It holds that $A^{n,r} \in \mathcal{B}^{\mathrm{p}}$ (because h^n is $\mathcal{B}^{\mathrm{p}}$-measurable) and $\bar{A}^{n,r} = \phi(a^{n,r}) \in \mathcal{B}^{\mathrm{w}}$. Since μ_h is a product measure, $A^{n,r} \times \bar{A}^{n,r} \in \mathcal{B}$. Since Y^* is obtained by taking union of measurable sets countably many times, we have $Y^* \in \mathcal{B}$.

Fix an arbitrary $v \in [0,1]$. Recall that, for any $n = 1,2,...$, the subsets $(A^{n,r})_{r=1}^{k(n)}$ are mutually disjoint. Thus, there exists a unique sequence $(A^{n,r(n)})_{n=1}^{\infty}$ such that

$$v \in A^{n,r(n)} \text{ for all } n = 1,2,.... \tag{32}$$

[39] See Sierpiński (1922) for the original result, and see `https://en.wikipedia.org/wiki/Atom\_(measure\_theory)` for the modified result we state.

By (28), we have $a^{n,r(n)}=h^n(v)\leq h^{n+1}(v)=a^{n+1,r(n+1)}$ for all $n=1,2,....$ Together with (31),

$$\bar{A}^{n,r(n)}=\phi(a^{n,r(n)})\subseteq\phi(a^{n+1,r(n+1)})=\bar{A}^{n+1,r(n+1)} \text{ for all } n=1,2,.... \tag{33}$$

We obtain

$$\mu_g^{\mathrm{w}}(Y_v^*)=\mu_g^{\mathrm{w}}(\cup_{n=1}^{\infty}\bar{A}^{n,r(n)})=\lim_{n\to\infty}\mu_g^{\mathrm{w}}(\bar{A}^{n,r(n)})=\lim_{n\to\infty}a^{n,r(n)}=\lim_{n\to\infty}h^n(v)=h(v),$$

where the first equality follows from (32), the second equality follows from (33) and the monotone convergence theorem of a measure, the third equality follows from (30), and the fourth equality follows (32), and the fifth equality follows from (29). □

Now we proceed to describe our non-cooperative game. In the first period, each firm i simultaneously announces measurable sets of workers $\S_i^A$ and $\S_i^B$ of groups A and B, respectively, to which it makes job offers, as well as a measurable function w_i^g on $\S_i^g$ for each $g\in\{A,B\}$ where $w_i^g(v,ß)$ is the wage that the firm makes to worker $(g,v,ß)$. We assume that $w_i^g(v,ß)=w_i^g(v',ß')$ if $v=v'$. Each worker observes the identity of the firm that made an offer to her (if any) and the associated wage offered to her and chooses to accept one of the offers or stay unassigned and receive a wage of zero. Then, each firm i is matched to the workers from $\S_i^A\cup\S_i^B$ who accepted its offers and pays wages to the hired workers according to the offer it made to them.

We assume that each worker is only interested in monetary transfers, so the worker's payoff is equal to the wage paid to her if she accepts an offer from a firm and zero otherwise. Next, we describe the firm's payoff. First consider the case where there is no EPSW. Let $\tilde{\S}_i^A$ and $\tilde{\S}_i^B$ be the sets of workers from groups A and B, respectively, who accepted firm i's offer. If $\tilde{\S}_i^A$ and $\tilde{\S}_i^B$ are both measurable, then the firm i obtains a payoff of

$$\beta\int_{\tilde{\S}_i^A}[v-w_i^A(v,ß)]d\mu_A+\int_{\tilde{\S}_i^B}[v-w_i^B(v,ß)]d\mu_B.$$

If at least one of $\tilde{\S}_i^A$ and $\tilde{\S}_i^B$ is nonmeasurable, then the firm's payoff is -1.

When there is a EPSW, we modify the firm's payoff such that it receives a payoff of -1 if there are sets of workers with positive measure $\tilde{\S}_i^g\subseteq g\times[0,1]^2$ and $\tilde{\S}_i^{g'}\subseteq g'\times[0,1]^2$ hired by it such that $g\neq g'$ and $w_i^g(v,ß)\neq w_i^{g'}(v',ß')$ for all $(g,v,ß)\in\tilde{\S}_i^g$ and $(g',v',ß')\in\tilde{\S}_i^{g'}$.

Given an action profile, the associated *outcome* is defined as $O:=(O_1,O_2)$, $O_i=\{(f_g^i(v),w_i^g(v))\}_{v\in[0,1],g=A,B}$ for $i=1,2$, where $f_g^i(v)$ is the density of workers of group g with productivity v who are hired by firm i in equilibrium, and $w_i^g(v)$ be the wage paid to those

workers if $f_g^i(v)>0$ and zero otherwise.

Our solution concept is pure-strategy subgame perfect Nash equilibrium.

PROPOSITION **8.** *The set of pure-strategy subgame perfect Nash equilibrium outcomes of the noncooperative game without EPSW and with EPSW, respectively, coincides with the core of the cooperative game without EPSW and with EPSW, respectively.*

Proof. We first prove that for any core outcome of the cooperative game $O:=\{(f_g^i(v),w_i^g(v))\}_{i\in\{1,2\},v\in[0,1],g=A,B}$, there exists a subgame perfect Nash equilibrium in the noncooperative game whose outcome is O. To show this, note by Proposition 1 that, for almost all i,g and v, we have $f_g^1(v)+f_g^2(v)=f_g(v)$ and $w_i^g(v)=v$ if $f_g^i(v)>0$ in any outcome in the core.

For each g, let $Y_g^*\subseteq\{g\}\times[0,1]^2$ be a measurable set of workers containing, for each v, $f_g^1(v)$ share of workers–note that such sets exist thanks to Proposition 7. Consider the following strategy profile in the non-cooperative game:

1. Both firms make job offers to all workers $(g,v,t)\in\{g\}\times[0,1]^2$ where the wage offer for each worker is equal to her productivity v.

2. All workers accept an offer from the firm whose wage offer is the highest and weakly positive, if any. If they receive the same wage offers from both firms, then the workers in set $Y_A^*\cup Y_B^*$ accept an offer of firm 1 while all other workers accept an offer from firm 2.

It is straightforward to verify that the above strategy profile results in outcome O. To see that this is a subgame perfect Nash equilibrium, first observe that by definition of the workers' strategies, each worker is clearly maximizing their payoffs against all possible strategies of other players. For the firms, there is no other strategy that results in a strictly higher payoff than the payoff of zero under the prescribed strategy profile because all workers are offered a wage that is equal to their own productivity from both firms under the prescribed strategy profile. These observations show that the specified strategy profile constitutes a subgame perfect Nash equilibrium, as desired.

We next prove that any pure-strategy subgame perfect Nash equilibrium outcome of the noncooperative game is in the core of the cooperative game. To show this, assume that the outcome $O:=\{(f_g^i(v),w_i^g(v))\}_{v\in[0,1],i=1,2,g=A,B}$ is an outcome associated with a strategy profile of the noncooperative game and is not in the core of the cooperative game. Then we constuct an outcome $\tilde{O}_j=\{(\tilde{f}_g^j(v),\tilde{w}_j^g(v))\}_{v\in[0,1],g=A,B}$ such that the associated strategy profile of the noncooperative game yields a profitable deviation for at least one of the firms.

We take $(\tilde{f}_A^j(v),\tilde{w}_j^A(v)) = (f_A^j(v),w_j^A(v))$ (or $\tilde{O}_{-j} = \{(\tilde{f}_g^{-j}(v),\tilde{w}_{-j}^g(v))\}_{v\in[0,1],g=A,B}$ where $(\tilde{f}_A^{-j}(v),\tilde{w}_{-j}^A(v))=(f_A^{-j}(v),w_{-j}^A(v))$), i.e. we do not change the outcome for $A-$group workers for either firm. The argument in which the deviation involves workers in group $g=A$ is analogous, where terms related to the firm's profit must be multiplied by β. To show a profitable deviation, we consider exhaustive cases as in Lemma 3.

First, suppose there exist a firm j and a subset of productivities $V\subset[0,1]$ with positive measure such that $w_j^g(v)>v$ for all $v\in V$. Consider a deviation of firm j such that it makes an offer to the same set of workers as in the present strategy profile while offering wage of $\min\{w_j^g(v),v\}$ to workers with productivity v (while not changing offers to workers in the other group–this feature is kept for other cases and thus we omit mentioning this for other cases). This will be a profitable deviation, because the firm now fails to hire any workers at wages strictly greater than their productivities. Therefore, the original strategy profile is not a subgame perfect Nash equilibrium.

Other cases, as enumerated in the proof of Lemma 3, can be treated in a similar manner. □

Finally, we note that an analogous analysis to the one offered here also establishes that the search model in Section 3 using density functions has a measure-theoretic foundation. Specifically, each worker is associated with a type $(g,v,ß,t)$ where t is the period at which the worker enters the market. The worker distribution is given as a measure on a sigma algebra defined over the set of worker types, and the wage function is measurable with respect to the same sigma algebra.

C Empirical appendix

C.1 Designation of treatment status

We define a treated firm as a firm that employs at least 10 long-term workers at the time of EPSW announcement. In this section, we discuss several reasons lending validity to this choice.

First, as firm size is endogenous, a potential concern is that manipulations in size at the time of EPSW announcement could affect our analysis. As discussed by McCrary (2008), a discontinuity in the share of firms with fewer than 10 long-term workers at the time of announcement suggests that firms may strategically alter their workforce quickly to avoid EPSW, which would mean our designation of treatment would not be as good as random. Panel (a) of Figure 3 investigates this concern by plotting the share of firms in our sample

with strictly fewer than 10 long-term workers across time, and overlays a separate best-fit polynomial for the time period before versus after policy announcement. As can be seen visually, there is a small increase in the share of firms with fewer than 10 long-term workers around the time of policy announcement. However, there is no statistically significant discontinuity; the p-value for the test of the null hypothesis that the share of firms with fewer than 10 long-term workers is equal in May and June 2009 is 0.43.

Second, would an alternative time, such as policy enactment instead of policy announcement, be more appropriate to denote the "post" period? We believe not. As seen in panel (a) of Figure 3, there is one notable time interval over which the share of firms with strictly fewer than 10 long-term workers increases, and this is centered around policy announcement. This indicates EPSW announcement likely led to anticipatory firm responses. Note that firm size responses are consistent with firms attempting to avoid the bite of the policy, suggesting policy announcement was salient to firms (but as discussed in the previous paragraph, the lack of discontinuity around the announcement date allows us to proceed with our difference-in-difference analysis). No such change is discernible around the time of policy enactment.[40]

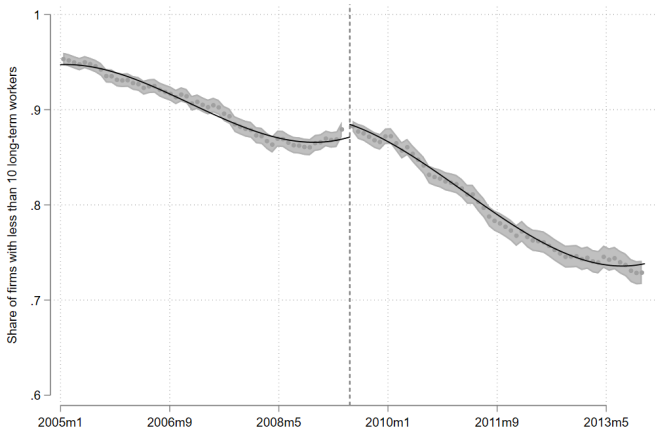
Third, a concern may be that our definition of treatment using time of announcement may not affect the probability that a firm is bound by EPSW in subsequent months. Panel (b) of Figure 3 replicates Panel (a) but generates separate series by firm treatment status. While there is a mechanical mean reversion due to our definition of treatment (i.e. control firms all have fewer than 10 long-term workers at EPSW announcement, and treated firms all have at least 10 long-term workers at EPSW announcement), our treatment variable is positively correlated with a firm being affected by EPSW in future time periods. In all months after announcement, strictly fewer than 25% of firms in our control group are bound by EPSW while strictly more than 50% of firms in our treatment group are bound by EPSW. Panel (b) of Figure 3 also suggests that the magnitudes of our estimates are likely conservative, because some control firms are bound by EPSW and some treated firms are not bound by EPSW in every time period after announcement. An alternative empirical approach would be to instrument for these shares across the two groups. However, such an analysis would require additional assumptions on how a firm's specific history of long-term workers translates into policy bite.[41] Our approach, while conservative, avoids such ad hoc assumptions.

[40] We believe that earlier anticipation of the policy is unlikely because EPSW was introduced to the Chilean senate only in May 2009. Moreovoer, in unreported results, we show that our findings are qualitatively and quantitatively similar if we change our analysis to designate policy time to be April 2009.

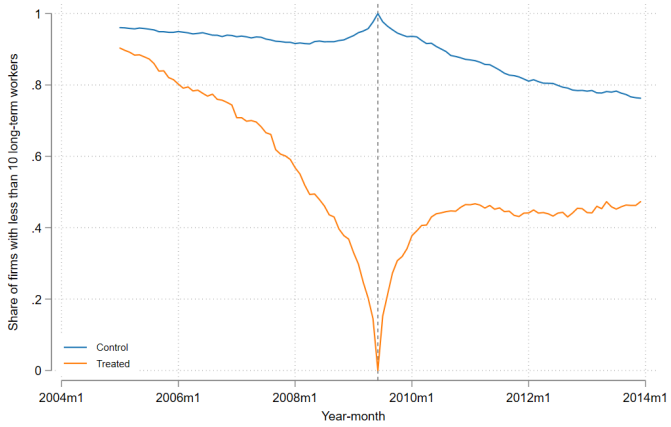
[41] For example, consider a firm j that employs strictly fewer than 10 long-term workers at some time t, but at least 10 long-term workers for time periods $t-10$,...,$t-1$. Consider another firm i that employs strictly fewer than 10 long-term workers in times $t-10$,...,$t-1$, but at least 10 long-term workers in time period t. It

Figure 3: Firm size and treatment status over time

(a): Share with fewer than 10 long-term workers



(b): Share affected by EPSW, by treatment status



Notes: Both panels display size information about firms in our baseline estimation sample. Panel (a) presents binned means (across time, where each bin corresponds to a month-year) for the share of firms with fewer than 10 long-term workers. The black curve is a 3rd-degree polynomial, fitted on each side of the cutoff. The shaded gray region represents 95% confidence bands for the local means, computed as in Calonico et al. (2015). Plots were implemented using rdplot from the rdrobust STATA package (See `https://rdpackages.github.io/rdrobust/`). Panel (b) displays the monthly share of firms that employ fewer than 10 long-term workers by treatment status, where treatment is defined as having at least 10 long-term workers in June 2009.

Fourth, is the comparison between control and treatment firms in our empirical specifications plausibly revealing a lower bound on the causal effect of EPSW? We believe it is. Our model in Section 3 finds that EPSW causes treated firms to segregate but not control firms. That is, there are no equilibrium spillovers from treated to control firms (or vice versa) on the outcome of segregation. Our model also finds that EPSW causes treated firms to make larger contributions to shifting the wage gap in favor of the majority group, as control firms retain more of their pre-EPSW workforce at existing wages. That is, the equilibrium effects serve only to attenuate the presented empirical results on wage inequality, as EPSW affects the wages set by both treated and control firms in the same direction, but it affects treated firms to a larger degree. In other words, our presented estimates on the effect of EPSW on the wage gap are smaller in magnitude than the difference between wage gaps between a particular market in which EPSW was never introduced and that same market under the counterfactual in which all firms were treated by EPSW.

Fifth, one may worry that there is an alternative reason for the effects that we observe.

is not obvious, given potential wage rigidities, which of these two firms is more affected by EPSW in period t.

We are unaware of any contemporaneous labor-market policy that differentially affects firms in our treatment and control groups. However, it is possible that other labor-market forces could be differentially affecting firms with more long-term workers. To assuage this concern, we run several "placebo" tests. Specifically, we rerun our analyses in (8) and (11) but instead suppose the treatment cutoff for number of long-term workers at a firm varies from 10. If the true cause for the effects we observe in our baseline specification is the policy itself, we should expect little to no effects around these placebo cutoffs for firms with the same actual treatment status according to the real policy. For each cutoff c we consider, we construct a sample of firms that have between $[c-4,c+3]$ workers at the time of EPSW announcement, and firms are considered treated if and only if they employ at least c long-term workers at announcement. All other details are as in our baseline specification. Results are presented below in Tables 5 and 6. The alternative cutoffs we consider do not lead to statistically significant effects, and point estimates are often close to zero or reversed in sign from our baseline estimates, suggesting that the effects we observe in our analysis are not due to other factors that differentially affect firms with different numbers of long-term workers.

Table 5: Effect of EPSW on Segregation–Placebos

	(I)	(II)	(III)	(IV)	(V)
	Baseline - 6 to 13	14 to 21	16 to 23	18 to 25	20 to 27
($\hat{\beta}^{seg}$) Post × Treated	0.0441***	−0.0067	0.0089	−0.0167	0.0121
	(0.0133)	(0.0219)	(0.0232)	(0.0271)	(0.0292)
Number of Firms	5312	1811	1496	1265	1097
Fixed effects					
Firm	Yes	Yes	Yes	Yes	Yes
Month×exit year×county	Yes	Yes	Yes	Yes	Yes
Month×exit year×region×industry	Yes	Yes	Yes	Yes	Yes
Firm-month level controls	Yes	Yes	Yes	Yes	Yes

Notes: In this table we display estimated coefficients for the difference-in-differences regression described in (8). Column I presents our baseline specification for our balanced sample with firms that employed between 6 to 13 workers at June 2009, and corresponds to column I of Table 2. Subsequent columns are identical except they consider placebo cutoffs in the number of long-term workers $c \in \{18,20,22,24\}$. For each cutoff c we consider, we construct a sample of firms that have between $[c-4,c+3]$ workers at the time of EPSW announcement. For each cutoff c, firms are considered treated if they employ at least c long-term workers at EPSW announcement, and are considered control otherwise. Throughout, standard errors in parentheses are two-way clustered at the firm and month levels. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 6: Effect of EPSW on Gender Wage Gap, by Majority Worker Group–Placebos

	(I) Baseline - 6 to 13	(II) 14 to 21	(III) 16 to 23	(IV) 18 to 25	(V) 20 to 27
($\hat{\beta}^{Mgap}$) Treated × Male × Post	0.0427***	0.0252	−0.0148	0.0033	−0.0302
	(0.0116)	(0.0234)	(0.0252)	(0.0281)	(0.0310)
($\hat{\beta}^{Mgap}+\hat{\beta}^{Fgap}$) Effect in Female	−0.0624***	−0.0195	−0.0251	−0.0046	−0.0194
Majority LLM	(0.0197)	(0.0507)	(0.0667)	(0.0929)	(0.0870)
Number of Firms	6424	2286	1916	1648	1436
Number of Observations	5551100	4229034	3942092	3728834	3518566
Fixed effects					
Firm	Yes	Yes	Yes	Yes	Yes
Worker	Yes	Yes	Yes	Yes	Yes
Month×exit year×county×hum. cap.	Yes	Yes	Yes	Yes	Yes
Month×exit year×region×industry×hum. cap.	Yes	Yes	Yes	Yes	Yes
Firm-month level controls	Yes	Yes	Yes	Yes	Yes
Worker-month level controls	Yes	Yes	Yes	Yes	Yes

Notes: In this table we display estimated coefficients for the difference-in-differences regression described in (11). Column I presents our baseline specification for our sample with firms that employed between 6 to 13 workers at June 2009, and corresponds to column I of Table 4. Subsequent columns are identical except they consider placebo cutoffs in the number of long-term workers $c \in \{18,20,22,24\}$. For each cutoff c we consider, we construct a sample of firms that have between $[c-4,c+3]$ workers at the time of EPSW announcement. For each cutoff c, firms are considered treated if they employ at least c long-term workers at EPSW announcement, and are considered control otherwise. Throughout, standard errors in parentheses are two-way clustered at the firm and month levels. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

C.2 Alternative firm comparison groups

In this section, we re-estimate our baseline specifications with different fixed effects. Primarily, we consider time-varying fixed effects that alter those in our segregation and wage gap analyses by altering the firm comparison groups. As discussed in Section 4, these time-varying fixed effects control for composition changes across different "sectors" of the labor market, as defined by the firm comparison group. Therefore, if our results are quantitatively similar across various specification of "sector," we interpret this as evidence that composition changes across sectors of the labor market are not driving the findings we present in the main body.

Table 7 presents estimates from (8). The specifications in columns I–IV differ in the firm comparison groups $k(j)$, and column V recreates the results of our baseline specification presented initially in Table 2, column I. Across specifications, we find a 3.9–4.6 percentage point increase in segregation due to EPSW.

Table 8 presents estimates from (11). The specifications in columns I–IV differ in the

Table 7: Effect of EPSW on Segregation–Alternative Time Trends

	(I)	(II)	(III)	(IV)	(V)
($\hat{\beta}^{seg}$) Post × Treated	0.0401***	0.0398***	0.0459***	0.0389***	0.0441***
	(0.0118)	(0.0124)	(0.0124)	(0.0141)	(0.0133)
Mean Pre-Treatment	−0.0032	−0.0065	−0.0025	0.0037	−0.0014
	(0.0104)	(0.0108)	(0.0111)	(0.0129)	(0.0118)
Number of Firms	6542	6326	5612	4599	5312
Number of Observations	602426	584252	536079	449306	515361
Fixed effects					
Firm	Yes	Yes	Yes	Yes	Yes
Month×exit year	Yes	No	No	No	No
Month×exit year×industry	No	Yes	No	No	No
Month×exit year×county	No	No	Yes	No	Yes
Month×exit year×county×industry	No	No	No	Yes	No
Month×exit year×region×industry	No	No	No	No	Yes
Firm-month level controls	Yes	Yes	Yes	Yes	Yes

Notes: This table displays estimated coefficient $\hat{\beta}^{seg}$ for the difference-in-differences regression described in (8). The unit of the panel is the firm-month and the dependent variable is a binary variable that indicates whether all workers at the firm in question are of a single gender in a given month. Each column presents time-varying fixed effects corresponding to a different comparison group of firms $k(j)$ for each firm j. In column I, all firms are included in the comparison group. In column II firms with the same industry code are included in the comparison group. In column III firms in the same geographic county are included in the comparison group. In column IV firms in both the same county and industry code are included in the comparison group. Column V is our baseline specification (see column I of Table 2). The mean pre-treatment effect is the mean of $\hat{\beta}^{seg}_{\tau}$ for $\tau < 1$ calculated from (9). Throughout, standard errors in parentheses are two-way clustered at the firm and month levels. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

firms included in comparison group $k(ij)$, and column VI recreates the results of our baseline specification presented initially in Table 4, column I. Across specifications, we find a 3.5-4.7 percentage point increase in the wage gap due to EPSW in male-majority LLMs, and a 5.0-7.5 percentage point decrease in the wage gap due to EPSW in female-majority LLMs. Column V presents our baseline specification, but replaces the separate worker and firm fixed effects–α_i and α_j, respectively–with a worker×firm fixed effect α_{ij}. By the construction of our triple difference specification, a worker-firm pair contributes to the estimates of the main coefficients of interest only if the pair are matched both before and after June 2009 (and only in the time periods in which the pair is matched). In other words, workers who change jobs due to EPSW are not driving our presented estimates. It is moreover the case that workers of firms that are fully segregated either before or after the policy do not contribute to the main coefficient

estimates. Our findings resemble those in our baseline, but are smaller in magnitude; there is a smaller absolute change in the wage gap in both male- and female-majority LLMs. Recalling that EPSW is predicted to increase gender segregation *by "role" but not across "roles,"* within a firm, these findings are consistent with further gender segregation within roles among firms that employ workers of both genders post EPSW (because firms that are desegregated by role within firm can contribute to the wage gap in the "opposite" direction, see Section 3).

C.3 Alternative empirical specifications

Section 4 of the paper discusses alternative empirical specifications and samples as robustness checks. Specifically, results from each of these robustness specifications is presented in a separate column in each of Tables 2 and 4. In this section, we describe these robustness specifications.

Removing firm fixed effects The results from this specification are presented in column II ("No firm FEs") of the aforementioned tables. The sample of firms and workers used is the same as in our main analysis, but we remove "α_j" firm fixed effects and add "$\delta above10_j$" to each of the specifications presented in (8), (9), (11), and (12). This specification removes the restriction of identifying effects only through within-firm variation, and also allows for between-firm comparisons. In our firm-month panel, corresponding to our segregation results, the presence of firm fixed effects accounts for the possibility that firms sort differentially into segregation based on unobservables. Because such sorting does not contribute to the treatment effects in our baseline specification with firm fixed effects, the comparison to the specification without firm fixed effects provides evidence on the relative importance of sorting. In the worker-firm-month panel, corresponding to our wage gap results, the presence of firm fixed effects additionally accounts for the possibility that *workers* sort differentially into firms based on firm unobservables. Therefore, the comparison to the model without firm fixed effects provides evidence on the relative importance of worker sorting.

Removing controls The results from this specification are presented in column III ("No controls") of the aforementioned tables. The sample of firms and workers used is the same as in our main analysis, but we remove the vector of firm-month level controls X_{jt} from (8) and (9), and we remove the vector of worker-firm-month level controls X_{ijt} from (11) and (12). Because this specification removes observable characteristics as covariariates, stability of the coefficient estimates of interest between this and our baseline specifications indicates that

plausibly-related characteristics are not driving our results. Moreover, it suggests that accounting for other, similar observable characteristics that could be present in some hypothetical data set would not greatly affect our results.

Doughnut hole Recall that our baseline sample includes firms with between 6 and 13 workers at announcement. Column IV ("Doughnut hole") in the aforementioned tables displays results from a sample which excludes all firms with either 9 or 10 workers at announcement from our baseline sample, resulting in a total of 5,799 firms. This sample is to account for the mechanical increase in likelihood that the excluded firms' treatment status does not match whether they are bound by EPSW at any given point in time. Our results reanalyze the specifications in (8), (9), (11), and (12) with this alternative sample.

Firm growth The results from this specification are presented in column V ("Firm growth") of the aforementioned tables. We additionally interact each of the time varying fixed effects (i.e. $\alpha_{k(j)t}$ in (8) and (9), and $\alpha_{k(ij)t}$ in (11) and (12)) with a firm-level indicator that tracks the absolute growth of each firm in terms of its workforce in the 6 months prior to EPSW announcement. Specifically, for each firm, we subtract the number of employed workers in December 2008 from the number of employed workers in June 2009. There are 285 firms that are not in existence in December 2008 but are in our baseline sample. We exclude these firms, leading to a total of 6,266 firms in the current sample. We then define an indicator variable for each quartile of the growth distribution. Therefore, this specification identifies the effects of EPSW on labor market outcomes of interest only comparing (workers of) firms to those of other firms that are on a similar growth path, and not to those of firms that are on different growth paths and may therefore respond differently to EPSW.

Balanced panel of firms The results from this specification are presented in column VI ("balanced sample") of the aforementioned tables. The sample of firms and workers is defined as all firms that are present in the market in all time periods between January 2005 and December 2013. That is, it includes only the firms present in our main analysis that neither enter nor exit our panel. This leads to a total of 3,418 firms. This sample therefore studies the impact of EPSW on segregation and wage gaps for firms that neither enter nor exit our panel "in the middle."

Table 8: Effect of EPSW on Gender Wage Gap, by Majority Worker Group
Alternative Fixed Effects

	(I)	(II)	(III)	(IV)	(V)	(VI)
($\hat{\beta}^{Mgap}$) Treated × Male × Post	0.0349***	0.0403***	0.0356***	0.0467***	0.0407***	0.0427***
	(0.0123)	(0.0123)	(0.0118)	(0.0121)	(0.0119)	(0.0116)
($\hat{\beta}^{Mgap}+\hat{\beta}^{Fgap}$) Effect in Female	−0.0564***	−0.0583***	−0.0499**	−0.0747***	−0.0506***	−0.0624***
Majority LLM	(0.0204)	(0.0201)	(0.0201)	(0.0225)	(0.0184)	(0.0197)
Mean Pre-Treatment	−0.0182	−0.0244	−0.0130	−0.0266	−0.0143	−0.0155
(Male Majority LLM)	(0.0176)	(0.0164)	(0.0161)	(0.0165)	(0.0165)	(0.0163)
Mean Pre-Treatment	−0.0223	−0.0165	−0.0209	0.0028	−0.0015	−0.0012
(Female Majority LLM)	(0.0240)	(0.0249)	(0.0255)	(0.0301)	(0.0296)	(0.0269)
Number of Firms	6436	6436	6433	6424	6421	6424
Number of Observations	6045315	5885503	5677415	5181609	5500022	5551100
Fixed effects						
Firm	Yes	Yes	Yes	Yes	No	Yes
Worker	Yes	Yes	Yes	Yes	No	Yes
Worker×firm	No	No	No	No	Yes	No
Month×exit year×hum. cap.	Yes	No	No	No	No	No
Month×exit year×industry×hum. cap.	No	Yes	No	No	No	No
Month×exit year×county×hum. cap.	No	No	Yes	No	Yes	Yes
Month×exit year×county×industry×hum. cap.	No	No	No	Yes	No	No
Month×exit year×region×industry×hum. cap.	No	No	No	No	Yes	Yes
Firm-month level controls	Yes	Yes	Yes	Yes	Yes	Yes
Worker-firm-month level controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table displays estimated coefficients for the regression described in (11). In particular, we present estimates of $\hat{\beta}^{Mgap}$ and $\hat{\beta}^{Fgap}$. The unit of the panel is the worker-firm-month and the dependent variable is the natural logarithm of monthly earnings. Each column presents time-varying fixed effects corresponding to a different comparison group of firms $k(ij)$ for each worker i and each firm j. All columns include worker "human capital" comparison groups defined by equivalence across three binary dimensions at time t at firm j: an indicator for tertiary education, an indicator for long-term versus fixed-term contract, and an indicator for being above median age in the particular industry-region in which firm j operates. Columns differentially include firm comparison groups. In column I, all firms are included in the comparison group. In column II firms with the same industry code are included in the comparison group. In column III firms in the same geographic county are included in the comparison group. In column IV firms in both the same county and industry code are included in the comparison group. Columns V and VI have time-varying fixed effects at the same level as in our baseline specification in column I of Table 4; Column V removes firm and worker fixed effect and replaces them with a fixed effect for each worker-firm pair, while column VI is precisely our baseline specification. The mean pre-treatment effects are the mean of $\hat{\beta}_{\tau}^{Mgap}$ and $\hat{\beta}_{\tau}^{Mgap}+\hat{\beta}_{\tau}^{Fgap}$, respectively, for $\tau<1$ calculated from (12). Throughout, standard errors in parentheses are two-way clustered at the firm and month levels. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

C.4 Details on data sample

We use administrative data from the Chilean unemployment insurance system from January 2005 to December 2013. In our data, an observation is a worker-firm-month. We observe two stratified (by firm size) random samples—a 1% sample and a 3% sample—of firms, and for each sampled firm we observe the entire monthly working history of every worker that was ever employed by the sampled firm, regardless of whether the worker remains at the sampled firm or not. Therefore, we observe some workers during time periods they are employed by other, non-sampled firms. We do not directly observe which firms are sampled. For that reason, our dataset includes "incidental firms" for which we do not necessarily observe the entire workforce at any given moment in time. This happens, for example, if a worker from a sampled firm switched to a non-sampled firm.

Because we do not observe the entire workforce for these incidental firms, we do not observe the size (number of workers) for these firms. This naturally leads to a potential concern with our size-based empirical strategy. To address this potential concern, we attempt to filter out the incidental firms. Our filtering approach is built on the notion that, due to the sampling procedure, we anticipate incidental firms to have high variance in the number (and presence) of workers across time periods.

Our filtering procedure does the following. 1) For each of the 1% and 3% samples, we compute for each firm the first and the last month it is observed. We drop firms that employ no workers for some months in between these two dates in any of the two samples, i.e. if the firm has "holes" in its employment history. 2) For firms in our sample, we compute the monthly average number of workers in each firm in our data and we drop firms that have an average of fewer than 4 in any of the two samples. 3) For each of the firms present in both the 1% and 3% samples, we compute the average number of workers of each firm across time, and we drop firms that have different average numbers of workers in the two samples. Descriptive statistics for the set of firms and workers left after the filtering are presented in column I of Table 1.