

# Supplemental Appendix: “Equal Pay for *Similar Work*”

Diego Gentile Passaro, Fuhito Kojima, and Bobak Pakzad-Hurson

Appendix A presents proofs omitted from the main text. Appendix B presents microfoundations and discussions regarding our theoretical analyses. 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 A1 shows that any outcome satisfying these two enumerated conditions is a core outcome, Lemma A2 shows that there are uncountably many non-equivalent outcomes satisfying these two conditions (implying that there are uncountably many core outcomes), and Lemma A3 shows that there are no core outcomes that fail either of the two enumerated conditions.<sup>31</sup>

**LEMMA A1.** *Any outcome  $O$  satisfying the two enumerated conditions in the Proof of Proposition 1 is a core outcome.*

*Proof of Lemma A1.* 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

---

<sup>31</sup>Kojima and Pakzad-Hurson (2025) study Bertrand competition under a wage monotonicity constraint. Their analysis and proof differ due to the discrepancy in setting.

$\tilde{O}_j$  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.  $\square$

**LEMMA A2.** *There exist a continuum of (non-equivalent) core outcomes satisfying the two enumerated conditions in the Proof of Proposition 1.*

*Proof of Lemma A2.* 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  $v > v^*$ .  $\square$

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 A3.** *There exist no core outcomes which do not satisfy the two enumerated conditions in the Proof of Proposition 1.*

*Proof of Lemma A3.* 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  $\beta$ .

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:

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

$\tilde{O}_j$  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:

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

$\tilde{O}_j$  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:

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

$\tilde{O}_j$  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:

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

$\tilde{O}_{-j}$  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  $v$ :

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

$\tilde{O}_j$  blocks  $O_j$  for the following reasons: Condition 1 of the definition of block is satisfied for all  $v \in V'$  since  $w_{-j}^g(v) \leq w_j^g(v) < \tilde{w}_j^g(v)$  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  $\varepsilon'$ , 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:

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

$\tilde{O}_j$  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.  $\square$

## 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

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

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  $g \in \{A,B\}$ :

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

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  $w_i^g(v) \leq v$  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 A3, 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  $\tilde{O}_j$ :

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

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 A3.

**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  $g \in \{A, B\}$ :

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

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,

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

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  $\epsilon' > 0$ :

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

For sufficiently small  $\epsilon'$ ,  $\pi_1^{\tilde{O}_1} > \pi_{1,A} + \pi_{2,A} \geq \pi_1$ , 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

$$(A1) \quad \pi_{1,A} + \pi_{2,A} > \pi_{1,B} + \pi_{2,B} + \epsilon.$$

By the Portmanteu Theorem,<sup>32</sup> for any  $\tilde{w}_2 \in [0,1]$  it follows that

$$\beta \int_0^{\tilde{w}_2} v f'_A(v) dv \rightarrow \beta \int_0^{\tilde{w}_2} v f_A(v) dv,$$

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

$$(A2) \quad \beta \int_0^{\tilde{w}_2} v f'_A(v) dv \in \left( \beta \int_0^{\tilde{w}_2} v f_A(v) dv - \frac{\epsilon}{8}, \beta \int_0^{\tilde{w}_2} v f_A(v) dv + \frac{\epsilon}{8} \right).$$

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

---

<sup>32</sup>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  $(F_A, F_B)$ ,

$$\beta \int_0^{w_2-\gamma} v f'_A(v) dv + \int_0^{w_2-\gamma} v f'_B(v) dv \rightarrow \beta \int_0^{w_2-\gamma} v f_A(v) dv + \int_0^{w_2-\gamma} v f_B(v) dv$$

and

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

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

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

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

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

and, with an analogous argument, that

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

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  $w'_2 \in (w_2 - \gamma, w_2 + \gamma)$ . Moreover, by monotonicity of the expression  $\beta \int_0^{w'_2} v f'_A(v) dv$  in  $w'_2$ , it follows that  $\beta \int_0^{w'_2} v f'_A(v) dv \in \left( \beta \int_0^{w_2-\gamma} v f'_A(v) dv, \beta \int_0^{w_2+\gamma} v f'_A(v) dv \right)$  for any distribution  $(F'_A, F'_B)$  in the neighborhood. Take a sufficiently small  $\gamma > 0$  such that

$$(A3) \quad \beta \int_0^{w_2-\gamma} v f_A(v) dv > \beta \int_0^{w_2} v f_A(v) dv - \frac{\epsilon}{8} \quad \text{and} \quad \beta \int_0^{w_2+\gamma} v f_A(v) dv < \beta \int_0^{w_2} v f_A(v) dv + \frac{\epsilon}{8}.$$

Moreover, we apply (A2) to  $\tilde{w}_2 = w_2 - \gamma$  and  $\tilde{w}_2 = w_2 + \gamma$  to obtain

$$(A4) \quad \begin{aligned} \beta \int_0^{w_2 - \gamma} v f'_A(v) dv &\in \left( \beta \int_0^{w_2 - \gamma} v f_A(v) dv - \frac{\epsilon}{8}, \beta \int_0^{w_2 - \gamma} v f_A(v) dv + \frac{\epsilon}{8} \right), \\ \beta \int_0^{w_2 + \gamma} v f'_A(v) dv &\in \left( \beta \int_0^{w_2 + \gamma} v f_A(v) dv - \frac{\epsilon}{8}, \beta \int_0^{w_2 + \gamma} v f_A(v) dv + \frac{\epsilon}{8} \right). \end{aligned}$$

Thus, by (A3) and (A4) 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 (A1) 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  $\gamma_1 > 0$  there exist distributions  $F'_A, F'_B$  satisfying our regularity conditions such that  $|F'_A(v) - F_A(v)| \leq \gamma_1$  and  $|F'_B(v) - F_B(v)| \leq \gamma_1$  for all  $v \in [0, 1]$  and for which there does not exist a non-segregation core outcome.<sup>33</sup> 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  $\gamma \in (0, \min\{f_A, f_B, \gamma_1\})$  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:

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

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

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

<sup>33</sup>Note that our pointwise notion of ‘‘closeness’’ corresponds to that of the weak-\* topology on the set of distributions satisfying our regularity conditions.

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

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

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

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.  $\square$

### Proof of Proposition 3

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

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

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

$$(A5) \quad \pi_1^{O(\gamma, \gamma')} = \beta \int_{\gamma}^1 (v - \gamma) f_A(v) dv = \int_0^{\gamma'} v f_B(v) dv = \pi_2^{O(\gamma, \gamma')}.$$

To see that the above wage functions permit a core outcome  $O(\gamma, \gamma')$ , first note that there is no firm  $i$  and outcome  $\tilde{O}_i$  with group-based segregation that blocks  $O(\gamma, \gamma')$ . 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  $\pi_i^{\tilde{O}_i} \leq \pi_1^{O(\gamma, \gamma')} = \pi_2^{O(\gamma, \gamma')}$ . 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  $w \leq \gamma$ . 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  $w_1(v) = \min\{v, \gamma\} \geq w$  in outcome  $O(\gamma, \gamma')$ . 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  $w > \gamma$ . 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  $O(\gamma, \gamma')$ . 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  $O(\gamma, \gamma')$  with  $\gamma \geq \gamma'$  and  $\pi_1^{O(\gamma, \gamma')} = \pi_2^{O(\gamma, \gamma')}$  is a core outcome. To see that there is a continuum of pairs  $\gamma, \gamma'$  that satisfy these conditions, note that if  $\gamma = \gamma' = 1$ , then the right-hand side of (A5) 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  $\gamma$ , the left-hand side is strictly smaller than the right-hand side for any  $\gamma$  that is sufficiently close to 1. Now, noting continuity of the right-hand side in  $\gamma'$  and the fact that it is equal to zero for  $\gamma' = 0$ , by intermediate value theorem, it follows that there exists  $\gamma'$  such that (A5) holds with equality. This concludes the proof of part 1.  $\square$

**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

$$(A6) \quad \begin{aligned} TS_A - \beta AW_A^O &= TS_B - AW_B^O \\ \iff TS_A - TS_B + (1 - \beta) AW_A^O &= AW_A^O - AW_B^O. \end{aligned}$$

*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 (A6) 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.  $\square$

*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 (A6),  $AW_A^O - AW_B^O \leq AW_A^{O'} - AW_B^{O'}$ .  $\square$

**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.  $\square$

**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  $\infty$ . Therefore, it remains only to show that the wage ratio is strictly increasing in  $\pi^O$  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.  $\square$

## 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 stage 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  $\{-\underline{t}, \underline{t}+1, \dots, -1, 0, 1, 2, \dots\}$ .
- 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  $r$ , (3), and  $d$ , respectively.
- All players in  $\mathcal{G}'_r$  discount future payoffs according to the same geometric rate  $\delta$ .

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  $r$ . 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 worker in  $\mathcal{G}$  has an incentive to deviate her search behavior, as the selected bargaining path has been shown to be optimal for each worker. 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.  $\square$

**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  $\pi_i = \frac{r_i^A}{\sum_{i' \in I} r_{i'}^A} \Delta^n \beta \mathbb{E}_A[v] > 0$  be the per-period profit firm  $i$  receives from hiring  $A$ -group workers at wages determined by (3) born in the current period when all other constrained firms are also segregated for  $A$ .<sup>34</sup> 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

$$(A7) \quad \sum_{t=\underline{T}}^{\infty} \delta^{t-\underline{T}} (1-d)^{t-\underline{T}} \pi_i > \delta^{T_i^1-\underline{T}} \hat{V} + \sum_{t=\underline{T}}^{T_i^1-1} \delta^{t-\underline{T}} (1-d)^{t-\underline{T}} \varepsilon.$$

Because  $\pi_i > \varepsilon$  and  $\hat{V}$  is finite, such  $T_i^1$  exists. The left-hand side of (A7) is a lower

---

<sup>34</sup> Noting that  $W^v$  as denoted in (3) is equal to zero because all firms are assumed to be segregated for group  $A$ , the (infinitesimal) discounted expected lifetime profit that a worker of type  $(v, g)$  accrues to firm  $i$  is  $\Delta^n V^v$ , and therefore the (unique, as shown in the discussion following (3)) per-period wage for such a worker is  $w = (1 - \Delta^n)v$  because  $\sum_{\tau=1}^{\infty} \delta^{\tau-1} (1-d)^{\tau-1} w = (1 - \Delta^n) \sum_{\tau=1}^{\infty} \delta^{\tau-1} (1-d)^{\tau-1} v = (1 - \Delta^n) V^v$ , 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 (3). Therefore,  $\pi_i = \frac{r_i^A}{\sum_{i' \in I} r_{i'}^A} \Delta^n \beta \mathbb{E}_A[v] > 0$ .

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, (A7) 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.  $\square$

**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 (3) 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 (3) implies that the surplus generated to a worker of type  $(g, v)$  who is born at any  $t \geq T^1$  is given by

$$(A8) \quad (1 - \Delta^{n_g + n_U})V,$$

and the worker’s wage at each period (see Footnote 34) is

$$(A9) \quad (1 - \Delta^{n_g + n_U})v.$$

Thus, the average wage of workers in group  $g$  is

$$(A10) \quad (1 - \Delta^{n_g + n_U})\mathbb{E}_g[v] + \epsilon_{g,t},$$

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  $t \rightarrow \infty$  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

$$(A11) \quad \frac{1 - \Delta^{n_A + n_U}}{1 - \Delta^{n_B + n_U}} \cdot \frac{\mathbb{E}_A[v]}{\mathbb{E}_B[v]} > \frac{\mathbb{E}_A[v]}{\mathbb{E}_B[v]},$$

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  $t' < 0$ .  $\square$

### 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 (5) are decreasing in variables  $x_A$  and  $x_B$ , respectively, (5) implies that

$$(A12) \quad \frac{\Delta^{n_A + n_U - 1}}{n_A + n_U - 1} \cdot \beta \mathbb{E}_A[v] < \frac{\Delta^{x_B + n_U + 1}}{n_B + n_U + 1} \cdot \mathbb{E}_B[v].$$

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

$$(A13) \quad \frac{\Delta^{n_A + n_U}}{n_A + n_U} \cdot \beta \mathbb{E}_A[v] < \frac{\Delta^{x_B + n_U + 1}}{n_B + n_U + 1} \cdot \mathbb{E}_B[v].$$

Let

$$(A14) \quad \Pi := \sum_{t=0}^{\infty} [\delta^t (1-d)^t] \left[ \sum_{t'=-t}^{-1} (1-d)^{t'+t} \beta \mathbb{E}_A[v] \right].$$

We note that  $\Pi$  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,  $\Pi$  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  $\Pi$  converges to zero as  $d \rightarrow 1$ , (A13) implies that there

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

$$(A15) \quad \Pi + \sum_{t'=t_0}^{\infty} \sum_{t=t'}^{\infty} \left[ \delta^{t-t_0} (1-d)^{t-t'} \frac{\Delta^{n_A+n_U}}{n_A+n_U} \cdot \beta \mathbb{E}_A[v] \right] < \sum_{t'=t_0}^{\infty} \sum_{t=t'}^{\infty} \left[ \delta^{t-t_0} (1-d)^{t-t'} \frac{\Delta^{x_B+n_U+1}}{n_B+n_U+1} \cdot \mathbb{E}_B[v] \right].$$

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 (A15) 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 (A15) with  $t_0 = T$ .<sup>35</sup> Therefore, (A15) implies that  $i$  is strictly better off by deviating, a contradiction to the equilibrium assumption.  $\square$

**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.  $\square$

## B Microfoundations and Theoretical Discussions

In this appendix, we present several microfoundations and additional discussions pertaining to our static model in Section I and to our search model in Section II.

<sup>35</sup>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  $B$ .

## B.A Non-cooperative Game Formulation

We describe a non-cooperative game played by firms and workers which is an analogue to our model in Section I. We show that the pure-strategy subgame perfect Nash equilibrium outcomes of this game are equivalent to the core of the static game analyzed in the main text. We then discuss how some of the mathematical machinery we produce in this appendix is applicable to the search model presented in Section II.

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  $(g,v,\beta) \in \{A,B\} \times [0,1]^2$ , where  $g$  is the group that the worker belongs to,  $v$  is her productivity, and  $\beta$  is an index. For each  $g$ , we assume that there is measure  $\mu_g$  that is given as a product measure of  $\mu_g^p$  and  $\mu_g^w$ . More specifically, let  $\mu_g^p$  be the Lebesgue measure on Lebesgue  $\sigma$ -algebra  $\mathcal{B}^p$  on  $[0,1]$ , representing the measure of productivity.<sup>36</sup> Let  $\mu_g^w$  be a measure on a  $\sigma$ -algebra  $\mathcal{B}^w$  on  $[0,1]$ , representing the measure of workers. We assume that  $\mu_g^w$  is non-atomic. The density function associated with  $\mu_g$  is given by  $f_g(v,\beta) = f_g^p(v) \times f_g^w(\beta)$ , where  $f_g^p(v)$  is associated with measure  $\mu_g^p$  and represents the density of  $g$ -group workers with productivity  $v$  while  $f_g^w(\beta)$  is associated with measure  $\mu_g^w$  and represents the density of workers whose indices are  $\beta$ .

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  $\mu_g$  denote the product measure induced from  $\mu_g^p$  and  $\mu_g^w$  with the corresponding  $\sigma$ -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  $Y_v \in \mathcal{B}^w$ . We say that  $h: [0,1] \rightarrow \mathbb{R}$  is  $\mathcal{B}^p$ -measurable if  $\{v \in [0,1] : h(v) \leq a\} \in \mathcal{B}^p$  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 B1.** *Let  $h: [0,1] \rightarrow \mathbb{R}_{\geq 0}$  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).

---

<sup>36</sup>Formally, we define a Borel measure  $\tilde{\mu}_g^p$  such that  $\tilde{\mu}_g^p([0,x]) = F_g(x)$  for all  $x \in [0,1]$ , which exists and is unique (Royden and Fitzpatrick, 2010, Proposition 25, Section 20.3). Let  $\mu_g^p$  be the unique measure defined on the Lebesgue measurable sets and coincides with  $\tilde{\mu}_g^p$  on Borel measurable sets: such  $\mu_g^p$  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’).

LEMMA **B1**. *Suppose that  $h: [0,1] \rightarrow \mathbb{R}_{\geq 0}$  is  $\mathcal{B}^p$ -measurable. Then, there exists a sequence of  $\mathcal{B}^p$ -measurable simple functions  $(h^n)_{n=1}^\infty$  such that*

$$(B1) \quad 0 \leq h^n(v) \leq h^{n+1}(v) \leq h(v) \text{ for all } v \in [0,1] \text{ and } n=1,2,\dots,$$

$$(B2) \quad h^n(v) \rightarrow h(v) \text{ (as } n \rightarrow \infty) \text{ for all } v \in [0,1].$$

For  $A \subseteq [0,1]$ , let  $\chi_A: [0,1] \rightarrow \{0,1\}$  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 B1, there exists a sequence of  $\mathcal{B}^p$ -measurable simple functions  $(h^n)_{n=1}^\infty$  that satisfies (B1) and (B2). For each  $n=1,2,\dots$ , 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  $\mathcal{B}^p$ -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.<sup>37</sup>

LEMMA **B2**. *Suppose that  $\mu_g^w$  is a non-atomic measure. Let  $c = \mu_g^w([0,1])$ . Then, there exists a function  $\phi: [0,c] \rightarrow \mathcal{B}^w$  such that*

$$(B3) \quad \mu_g^w(\phi(t)) = t \text{ for all } t \in [0,c], \text{ and}$$

$$(B4) \quad \phi(t) \subseteq \phi(t') \text{ for all } t, t' \in [0,c] \text{ with } t \leq t'.$$

By Lemma B2, there exists a function  $\phi$  that satisfies (B3) and (B4). For each  $n=1,2,\dots$ , and  $r=1,2,\dots,k(n)$ , by (B1) 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  $n=1,2,\dots$ , and  $r=1,2,\dots,k(n)$ , and we define  $Y^* = \bigcup_{n=1}^\infty \bigcup_{r=1}^{k(n)} (A^{n,r} \times \bar{A}^{n,r})$ . It holds that  $A^{n,r} \in \mathcal{B}^p$  (because  $h^n$  is  $\mathcal{B}^p$ -measurable) and  $\bar{A}^{n,r} = \phi(a^{n,r}) \in \mathcal{B}^w$ . Since  $\mu_h$  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  $Y^* \in \mathcal{B}$ .

<sup>37</sup>See Sierpiński (1922) for the original result, and see [https://en.wikipedia.org/wiki/Atom\\_\(measure\\_theory\)](https://en.wikipedia.org/wiki/Atom_(measure_theory)) for the modified result we state.

Fix an arbitrary  $v \in [0, 1]$ . Recall that, for any  $n = 1, 2, \dots$ , 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

$$(B5) \quad v \in A^{n,r(n)} \text{ for all } n = 1, 2, \dots$$

By (B1), we have  $a^{n,r(n)} = h^n(v) \leq h^{n+1}(v) = a^{n+1,r(n+1)}$  for all  $n = 1, 2, \dots$ . Together with (B4),

$$(B6) \quad \bar{A}^{n,r(n)} = \phi(a^{n,r(n)}) \subseteq \phi(a^{n+1,r(n+1)}) = \bar{A}^{n+1,r(n+1)} \text{ for all } n = 1, 2, \dots$$

We obtain

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

where the first equality follows from (B5), the second equality follows from (B6) and the monotone convergence theorem of a measure, the third equality follows from (B3), and the fourth equality follows (B5), and the fifth equality follows from (B2).  $\square$

Now we proceed to describe our non-cooperative game. In the first period, each firm  $i$  simultaneously announces measurable sets of workers  $\xi_i^A$  and  $\xi_i^B$  of groups  $A$  and  $B$ , respectively, to which it makes job offers, as well as a measurable function  $w_i^g$  on  $\xi_i^g$  for each  $g \in \{A, B\}$  where  $w_i^g(v, \mathfrak{B})$  is the wage that the firm makes to worker  $(g, v, \mathfrak{B})$ . We assume that  $w_i^g(v, \mathfrak{B}) = w_i^g(v', \mathfrak{B}')$  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  $\xi_i^A \cup \xi_i^B$  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  $\tilde{\xi}_i^A$  and  $\tilde{\xi}_i^B$  be the sets of workers from groups  $A$  and  $B$ , respectively, who accepted firm  $i$ 's offer. If  $\tilde{\xi}_i^A$  and  $\tilde{\xi}_i^B$  are both measurable, then the firm  $i$  obtains a payoff of

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

If at least one of  $\tilde{\xi}_i^A$  and  $\tilde{\xi}_i^B$  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  $\tilde{\xi}_i^g \subseteq g \times [0,1]^2$  and  $\tilde{\xi}_i^{g'} \subseteq g' \times [0,1]^2$  hired by it such that  $g \neq g'$  and  $w_i^g(v, \beta) \neq w_i^{g'}(v', \beta')$  for all  $(g, v, \beta) \in \tilde{\xi}_i^g$  and  $(g', v', \beta') \in \tilde{\xi}_i^{g'}$ .

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 B2.** *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 B1. 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 construct 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  $\beta$ . To show a profitable deviation, we consider exhaustive cases as in Lemma A3.

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 by firm  $j$  in which it continues making the same wage offers to all workers with productivities  $v \notin V$  and makes no offers to any workers with  $v \in V$ . This will be a profitable deviation, because the firm now fails to hire workers in  $V$  at wages strictly greater than their productivities, while retaining the rest of its workforce at the same wages. Therefore, the original strategy profile is not a subgame perfect Nash equilibrium.

Other cases, as enumerated in the proof of Lemma A3, can be treated in a similar manner.  $\square$

Finally, we note that an analogous analysis to the one offered here also establishes that the search model in Section II using density functions has a measure-theoretic foundation. Specifically, each worker is associated with a type  $(g, v, \beta, t)$  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.

## B.B Capacity constraints

Our search 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  $\kappa_i^g$  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  $\kappa_i^g$   $g$ -group workers.<sup>38</sup> Our results are qualitatively similar given such a modification to our model.

## B.C Alternating offer microfoundation

Our search model considers that the total surplus accruing to each of the firm and a given worker is split in an exogenous manner given the worker’s “outside option,” leading to a “Nash-in-Nash” bargaining protocol. One question may be, given the finding of Binmore et al. (1988), whether any equilibrium presented above has an analogue as the equilibrium of a game with alternating-offer bargaining. In this appendix, we show that this is indeed true.

Formally, we alter Stage 2 of the timing of the game presented in Section II.A.1 in the following manner. Fix an exogenous order over unemployed workers.<sup>39</sup> 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.

---

<sup>38</sup>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  $i$ .

<sup>39</sup>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  $\varphi: [0, 1] \rightarrow \mathcal{W}^t$  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.

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.

In post EPSW periods  $t \geq 0$ , we assume that any firm  $i$  that is desegregated at some wage  $w$  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  $w$ .

All other aspects of the game are as before.

In general, a wide range of bargaining outcomes can be consistent with subgame perfect Nash equilibrium in this alternating-offer game, 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 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 standard alternating offer bargaining between two “non-negligible” parties, sequential rationality requires that  $x$  and  $y$  satisfy

$$(B7) \quad y = \delta'x + (1 - \delta')\bar{V}^v, \quad V^v - x = \delta'(V^v - y),$$

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 (B7) and the acceptance rules as described above both on and off path. Solving these equations, we obtain

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

By the same logic as in our base game, it is optimal for the worker to first bargain with all segregated or unconstrained firms, and then bargain with a desegregated firm offering surplus  $W^v$  if  $D_t^v \neq \emptyset$ . Therefore, applying the above acceptance and offer rules along the optimal bargaining path yields worker surplus of

$$(B8) \quad V_1^v = \frac{1}{(1 + \delta')^m}W^v + \left(1 - \frac{1}{(1 + \delta')^m}\right)V^v.$$

Comparing (B8) to (3) reveals an obvious similarity. 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 in our base game. 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  $\delta' \rightarrow 0$  is 1, and the limit of the function as  $\delta' \rightarrow 1$  is 1/2. These observations together imply that there is a unique value of  $\delta'$  for which  $\bar{\Delta}(\delta') = \Delta$ , that is, there is a unique value of the bargaining friction  $\delta'$  in our alternating-offer game that yields the same equilibrium outcomes as does the exogenous split  $\Delta \in (1/2, 1)$  in our “Nash-in-Nash” game.

We note that the previous paragraph establishes the equivalence of the set of equilibrium outcomes of the two games for  $\Delta \in (1/2, 1)$ . 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  $\Delta \in (0, 1/2)$ . By a similar logic, any split  $\Delta \in (0, 1/2)$  can be obtained in our alternating-offer game by setting the first mover in any bargaining situation to be the worker instead of the firm, and appropriately selecting  $\delta'$ .

REMARK B1. For any  $\Delta \in (0,1) \setminus \{1/2\}$ , there is a unique value of  $\delta' \in (0,1)$  such that the set of equilibria (with our imposed regularity conditions) of our “Nash-in-Nash” bargaining game with exogenous split  $\Delta$  corresponds to the set of equilibria of the alternating-offer game with negotiation breakdown probability  $\delta'$ , and the firm as first proposer if  $\Delta > 1/2$  and the worker as first proposer if  $\Delta < 1/2$ .

## C Empirical appendix

### C.A 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.<sup>40</sup>

---

<sup>40</sup>We believe that earlier anticipation of the policy is unlikely because EPSW was introduced to the Chilean senate only in May 2009. Moreover, in unreported results, we show that our findings are qualitatively

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.<sup>41</sup> Our approach, while conservative, avoids such ad hoc assumptions.

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 II 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.

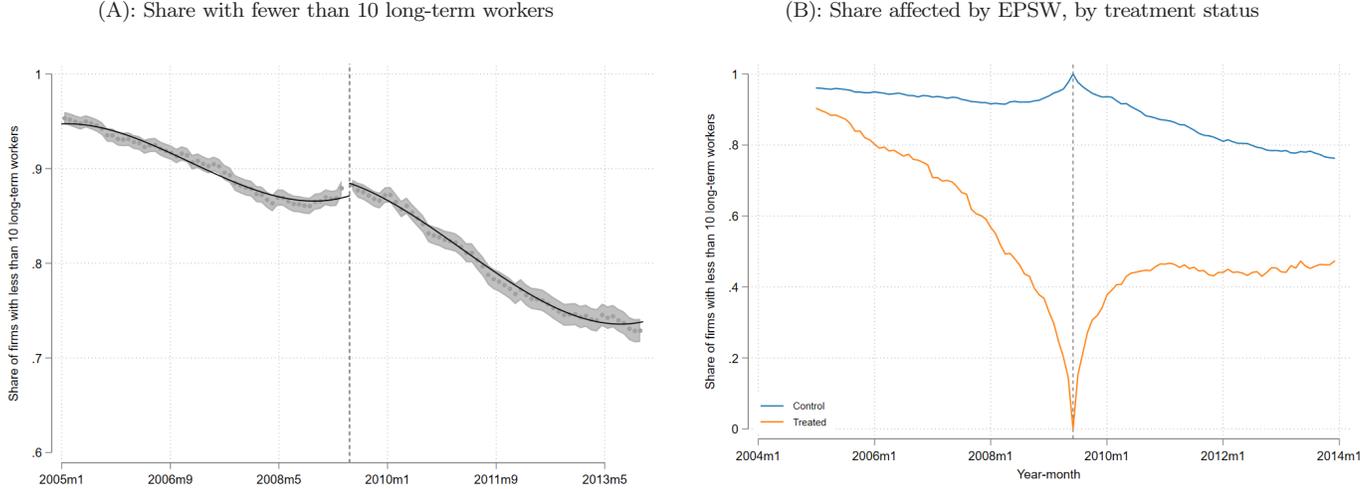
Fifth, one may worry that there is an alternative reason for the effects that we observe. We are unaware of any contemporaneous labor-market policy that differentially affects firms

---

and quantitatively similar if we change our analysis to designate policy time to be April 2009.

<sup>41</sup>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  $t-10, \dots, t-1$ . Consider another firm  $i$  that employs strictly fewer than 10 long-term workers in times  $t-10, \dots, t-1$ , but at least 10 long-term workers in time period  $t$ . It is not obvious, given potential wage rigidities, which of these two firms is more affected by EPSW in period  $t$ .

Figure 3: Firm size and treatment status over time



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.

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 (7) and (10) 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. Under most of the alternative cutoffs, the effects are insignificant and/or the point estimates are reversed in sign from our baseline estimates,<sup>42</sup> suggesting that

<sup>42</sup>The sole exception is in column IV of 6. The effect in male-majority LLMs is close to zero and statistically insignificant, while the effect in female-majority LLMs is statistically significant at the 5% level. Given that there is no significant effect on segregation in this sample (see Table 5) and no significant effect on wages in male-majority LLMs, it is likely that the effect in female-majority LLMs is spurious.

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 $\times$ Treated	0.0441 (0.0133)	-0.0067 (0.0219)	0.0089 (0.0232)	-0.0167 (0.0271)	0.0121 (0.0292)
Number of Firms	5312	1811	1496	1265	1097
Number of Observations	515361	181298	150478	127427	110409
<i>Fixed effects</i>					
Firm	Yes	Yes	Yes	Yes	Yes
Month $\times$ exit year $\times$ county	Yes	Yes	Yes	Yes	Yes
Month $\times$ exit year $\times$ region $\times$ 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 (7). 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.

### C.B 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 III, 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 (7). 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 (10). The specifications in columns I–IV differ in the

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

	(I)	(II)	(III)	(IV)	(V)
	Baseline - 6 to 13	14 to 21	16 to 23	18 to 25	20 to 27
$(\hat{\beta}^{Mgap})$ Treated $\times$ Male $\times$ Post	0.0427 (0.0116)	0.0209 0.0175	0.0039 0.0169	-0.0022 0.0199	-0.0439 0.0181
$(\hat{\beta}^{Mgap} + \hat{\beta}^{Fgap})$ Effect in Female Majority LLM	-0.0624 (0.0197)	-0.0541 0.0372	0.0053 0.0297	-0.1054 0.0353	-0.0561 0.0423
Number of Firms	6424	2286	1916	1648	1436
Number of Observations	5551100	3892193	3626862	3433948	3252092
<i>Fixed effects</i>					
Firm	Yes	Yes	Yes	Yes	Yes
Worker	Yes	Yes	Yes	Yes	Yes
Month $\times$ exit year $\times$ county $\times$ hum. cap.	Yes	Yes	Yes	Yes	Yes
Month $\times$ exit year $\times$ region $\times$ industry $\times$ hum. cap.	Yes	Yes	Yes	Yes	Yes
Firm-month level controls	Yes	Yes	Yes	Yes	Yes
Worker-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 (10). 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.

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— $\alpha_i$  and  $\alpha_j$ , respectively—with a worker $\times$ firm fixed effect  $\alpha_{ij}$ . 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

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

	(I)	(II)	(III)	(IV)	(V)
$(\hat{\beta}^{seg})$ Post $\times$ Treated	0.0401 (0.0118)	0.0398 (0.0124)	0.0459 (0.0124)	0.0389 (0.0141)	0.0441 (0.0133)
Mean Pre-Treatment	−0.0032 (0.0104)	−0.0065 (0.0108)	−0.0025 (0.0111)	0.0037 (0.0129)	−0.0014 (0.0118)
Number of Firms	6542	6326	5612	4599	5312
Number of Observations	602426	584252	536079	449306	515361
<i>Fixed effects</i>					
Firm	Yes	Yes	Yes	Yes	Yes
Month $\times$ exit year	Yes	No	No	No	No
Month $\times$ exit year $\times$ industry	No	Yes	No	No	No
Month $\times$ exit year $\times$ county	No	No	Yes	No	Yes
Month $\times$ exit year $\times$ county $\times$ industry	No	No	No	Yes	No
Month $\times$ exit year $\times$ region $\times$ 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 (7). 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  $\hat{\beta}_\tau^{seg}$  for  $\tau < 0$  calculated from (8). Throughout, standard errors in parentheses are two-way clustered at the firm and month levels.

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 II).

### C.C Alternative empirical specifications

Section III 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 “ $\alpha_j$ ” firm fixed effects and add “ $\delta_{above10_j}$ ” to each of the specifications presented in (7), (8), (10), and (11). 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 (7) and (8), and we remove the vector of worker-firm-month level controls  $X_{ijt}$  from (10) and (11). Because this specification removes observable characteristics as covariates, 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 (7), (8), (10), and (11) 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 (7) and (8), and  $\alpha_{k(ij)t}$  in (10) and (11)) 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 $\times$ Male $\times$ Post	0.0349 (0.0123)	0.0403 (0.0123)	0.0356 (0.0118)	0.0467 (0.0121)	0.0407 (0.0119)	0.0427 (0.0116)
$(\hat{\beta}^{Mgap} + \hat{\beta}^{Fgap})$ Effect in Female Majority LLM	-0.0564 (0.0204)	-0.0583 (0.0201)	-0.0499 (0.0201)	-0.0747 (0.0225)	-0.0506 (0.0184)	-0.0624 (0.0197)
Mean Pre-Treatment (Male Majority LLM)	-0.0159 (0.0154)	-0.0214 (0.0143)	-0.0114 (0.0141)	-0.0232 (0.0145)	-0.0125 (0.0144)	-0.0136 (0.0143)
Mean Pre-Treatment (Female Majority LLM)	-0.0195 (0.0210)	-0.0145 (0.0218)	-0.0183 (0.0223)	0.0025 (0.0264)	-0.0013 (0.0259)	-0.0010 (0.0235)
Number of Firms	6436	6436	6433	6424	6421	6424
Number of Observations	6045315	5885503	5677415	5181609	5500022	5551100
<i>Fixed effects</i>						
Firm	Yes	Yes	Yes	Yes	No	Yes
Worker	Yes	Yes	Yes	Yes	No	Yes
Worker $\times$ firm	No	No	No	No	Yes	No
Month $\times$ exit year $\times$ hum. cap.	Yes	No	No	No	No	No
Month $\times$ exit year $\times$ industry $\times$ hum. cap.	No	Yes	No	No	No	No
Month $\times$ exit year $\times$ county $\times$ hum. cap.	No	No	Yes	No	Yes	Yes
Month $\times$ exit year $\times$ county $\times$ industry $\times$ hum. cap.	No	No	No	Yes	No	No
Month $\times$ exit year $\times$ region $\times$ industry $\times$ 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 (10). 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  $k(i,j)$  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  $\tau < 0$  calculated from (11). Throughout, standard errors in parentheses are two-way clustered at the firm and month levels.

## C.D 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.

## References

- Robert B. Ash and Catherine A. Doléans-Dade. *Probability & Measure Theory, Second Edition*. Academic Press, 2000.
- Ken Binmore, Ariel Rubinstein, and Asher Wolinsky. The Nash bargaining solution in economic modelling. *RAND Journal of Economics*, 17(2):176–109, 1988.
- Sebastian Calonico, Matias D. Cattaneo, and Rocío Titiunik. Optimal Data-Driven Regression Discontinuity Plots. *Journal of Econometrics*, 110(512):1753–1769, 2015.

- Christopher Harris. Existence and Characterization of Perfect Equilibrium in Games of Perfect Information. *Econometrica*, 53(3):613–628, 1985.
- Fuhito Kojima and Bobak Pakzad-Hurson. Bertrand Menu Competition. *arXiv Working Paper 2504.16842*, 2025.
- Justin McCrary. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698–714, 2008.
- H.L. Royden and P.M. Fitzpatrick. *Real Analysis (Fourth Edition)*. Prentice Hall, 2010.
- 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.