

Supplemental Appendix

**The Gender Pay Gap:
Micro Sources and Macro Consequences**

By IACOPO MORCHIO AND CHRISTIAN MOSER[†]

[†]Morchio: University of Bristol (iacopo.morchio@bristol.ac.uk); Moser: Columbia University, NBER, CEPR, CESifo, and IZA (c.moser@columbia.edu).

A. DATA DESCRIPTION APPENDIX

1. Comparison of Actual versus Potential Experience

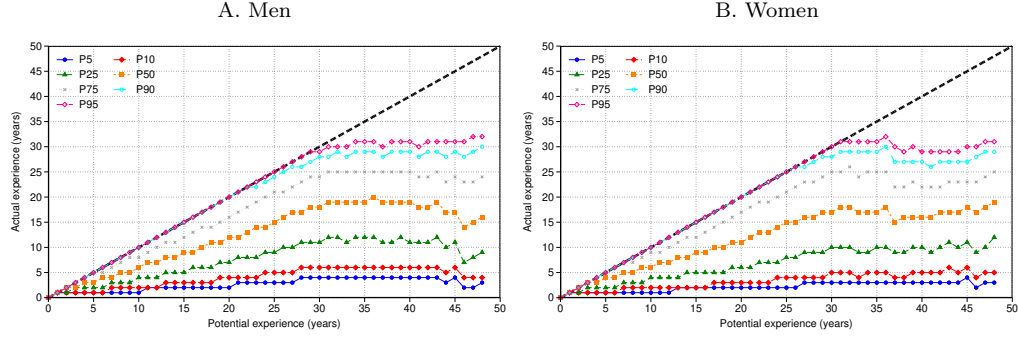


FIGURE A.1. PERCENTILES OF ACTUAL EXPERIENCE CONDITIONAL ON POTENTIAL EXPERIENCE

Note: Figure shows percentiles of actual against potential experience separately for men (Panel A) and women (Panel B). Actual experience is constructed from panel data for 1985–2014, while potential experience = age – years of education + 6. Solid line represents the 45-degree line, for which actual experience equals potential experience.
Source: RAIS, 2007–2014.

2. Additional Summary Statistics

TABLE A.1—SUMMARY STATISTICS FOR THE RAW DATA

	2007			2014			Pooled 2007–2014		
	Overall	Men	Women	Overall	Men	Women	Overall	Men	Women
Share Nonwhite	0.338	0.369	0.288	0.390	0.420	0.347	0.365	0.397	0.318
Share primary school	0.107	0.143	0.051	0.067	0.093	0.031	0.085	0.115	0.039
Share middle school	0.221	0.268	0.146	0.180	0.223	0.119	0.199	0.245	0.130
Share high school	0.496	0.470	0.538	0.589	0.569	0.617	0.548	0.524	0.584
Share college	0.176	0.119	0.265	0.164	0.114	0.233	0.168	0.115	0.246
Mean years of education	10.8	10.1	11.9	11.2	10.7	12.1	11.0	10.4	12.0
(std. dev.)	(3.4)	(3.5)	(3.0)	(3.0)	(3.1)	(2.6)	(3.2)	(3.3)	(2.8)
Mean years of age	32.8	32.8	32.8	33.5	33.5	33.5	33.0	33.1	33.0
(std. dev.)	(9.4)	(9.4)	(9.4)	(9.5)	(9.5)	(9.4)	(9.4)	(9.4)	(9.4)
Mean employer size	2,566	1,401	4,421	2,436	1,694	3,479	2,406	1,518	3,742
(std. dev.)	(18,390)	(12,175)	(25,204)	(16,366)	(12,398)	(20,651)	(16,264)	(11,559)	(21,410)
Mean gender-employer size	1,600	613	3,171	1,503	896	2,357	1,495	761	2,599
(std. dev.)	(11,892)	(3,325)	(18,577)	(9,927)	(4,978)	(14,178)	(10,194)	(4,173)	(15,231)
Mean employer age	23.3	21.6	25.8	23.4	22.0	25.5	23.1	21.6	25.4
(std. dev.)	(21.3)	(19.9)	(23.1)	(22.3)	(20.9)	(23.8)	(21.8)	(20.4)	(23.5)
Mean months employed during year	9.6	9.5	9.7	9.7	9.6	9.8	9.6	9.5	9.7
(std. dev.)	(3.3)	(3.3)	(3.3)	(3.2)	(3.1)	(3.2)	(3.2)	(3.2)	(3.3)
Mean contractual work hours	41.8	42.6	40.4	41.8	42.6	40.8	41.9	42.6	40.7
(std. dev.)	(5.7)	(4.5)	(7.0)	(5.4)	(4.3)	(6.4)	(5.4)	(4.3)	(6.6)
Mean years of tenure	3.4	3.1	3.9	3.2	3.0	3.4	3.2	2.9	3.5
(std. dev.)	(5.1)	(4.8)	(5.6)	(5.0)	(4.8)	(5.2)	(5.0)	(4.7)	(5.4)
Mean log real monthly earnings	6.984	7.032	6.908	7.222	7.283	7.136	7.112	7.166	7.032
(std. dev.)	(0.680)	(0.689)	(0.659)	(0.631)	(0.640)	(0.606)	(0.656)	(0.666)	(0.633)
Number of worker-years	46,347,012	28,474,352	17,872,660	62,571,376	36,561,824	26,009,552	449,390,272	269,897,824	179,492,464
Number of unique workers	38,412,504	23,362,020	15,050,485	50,872,600	29,455,352	21,417,268	77,397,648	44,436,752	32,960,898
Number of unique employers	2,775,579	1,710,549	1,065,030	3,677,552	2,131,329	1,546,223	6,044,593	3,404,144	2,640,449
Share female	0.386			0.416			0.399		
Mean log gender earnings gap	0.124			0.147			0.134		

Note: Table shows summary statistics for the raw data for 2007, 2014, and the pooled years 2007–2014.

Source: RAIS, 2007–2014.

TABLE A.2—SUMMARY STATISTICS FOR SELECTED SAMPLE

	2007			2014			Pooled 2007–2014		
	Overall	Men	Women	Overall	Men	Women	Overall	Men	Women
Share Nonwhite	0.351	0.381	0.301	0.398	0.429	0.350	0.378	0.409	0.329
Share primary school	0.106	0.138	0.053	0.072	0.097	0.034	0.088	0.117	0.043
Share middle school	0.219	0.263	0.144	0.185	0.226	0.122	0.203	0.245	0.134
Share high school	0.488	0.472	0.515	0.562	0.553	0.577	0.531	0.516	0.554
Share college	0.187	0.126	0.289	0.180	0.124	0.267	0.179	0.122	0.269
Mean years of education	10.9	10.2	12.0	11.3	10.7	12.2	11.1	10.4	12.1
(std. dev.)	(3.4)	(3.5)	(3.1)	(3.1)	(3.2)	(2.8)	(3.3)	(3.3)	(2.9)
Mean years of age	33.0	32.9	33.3	34.6	34.5	34.8	33.6	33.5	33.8
(std. dev.)	(9.2)	(9.2)	(9.3)	(9.3)	(9.3)	(9.3)	(9.4)	(9.4)	(9.4)
Mean employer size	2,784	1,528	4,875	2,961	2,065	4,332	2,775	1,757	4,402
(std. dev.)	(17,415)	(11,394)	(24,187)	(16,618)	(12,451)	(21,405)	(16,298)	(11,448)	(21,827)
Mean gender-employer size	1,744	698	3,485	1,852	1,154	2,920	1,743	926	3,048
(std. dev.)	(11,287)	(3,138)	(17,839)	(10,211)	(5,419)	(14,727)	(10,314)	(4,413)	(15,573)
Mean employer age	27.9	25.4	31.9	29.7	27.3	33.5	28.0	25.7	31.7
(std. dev.)	(22.1)	(20.8)	(23.7)	(23.4)	(22.2)	(24.8)	(22.8)	(21.5)	(24.3)
Mean months employed during year	9.9	9.7	10.1	10.2	10.0	10.3	9.9	9.8	10.0
(std. dev.)	(3.2)	(3.3)	(3.2)	(3.0)	(3.0)	(2.9)	(3.2)	(3.2)	(3.1)
Mean contractual work hours	41.6	42.6	40.0	41.6	42.5	40.3	41.7	42.6	40.3
(std. dev.)	(5.4)	(4.1)	(6.8)	(5.1)	(3.9)	(6.3)	(5.1)	(3.9)	(6.4)
Mean years of tenure	4.0	3.5	4.8	4.2	3.9	4.6	3.9	3.6	4.5
(std. dev.)	(5.6)	(5.2)	(6.2)	(5.7)	(5.4)	(6.1)	(5.6)	(5.2)	(6.1)
Mean log real monthly earnings	7.072	7.119	6.995	7.348	7.407	7.259	7.206	7.259	7.122
(std. dev.)	(0.709)	(0.713)	(0.696)	(0.674)	(0.675)	(0.662)	(0.693)	(0.697)	(0.678)
Number of worker-years	27,609,252	17,248,300	10,360,952	35,134,128	21,238,274	13,895,856	271,401,248	166,964,944	104,436,320
Number of unique workers	27,609,252	17,248,300	10,360,952	35,134,128	21,238,274	13,895,856	56,868,844	33,975,636	22,893,206
Number of unique employers	425,186	295,142	130,044	492,174	334,898	157,276	640,625	419,717	220,908
Share female	0.375			0.396			0.385		
Mean log gender earnings gap	0.124			0.149			0.137		

Note: Table shows summary statistics for the selected sample for 2007, 2014, and the pooled years 2007–2014.

Source: RAIS, 2007–2014.

TABLE A.3—SUMMARY STATISTICS FOR THE CONNECTED SET

	2007			2014			Pooled 2007–2014		
	Overall	Men	Women	Overall	Men	Women	Overall	Men	Women
Share Nonwhite	0.349	0.380	0.299	0.397	0.429	0.349	0.378	0.409	0.327
Share primary school	0.105	0.137	0.052	0.072	0.096	0.034	0.088	0.116	0.042
Share middle school	0.217	0.262	0.142	0.185	0.226	0.122	0.202	0.245	0.133
Share high school	0.487	0.473	0.512	0.562	0.553	0.576	0.531	0.516	0.554
Share college	0.190	0.128	0.294	0.181	0.124	0.268	0.179	0.122	0.272
Mean years of education	10.9	10.2	12.1	11.3	10.7	12.2	11.1	10.4	12.1
(std. dev.)	(3.4)	(3.5)	(3.1)	(3.1)	(3.2)	(2.8)	(3.3)	(3.3)	(2.9)
Mean years of age	33.1	32.9	33.4	34.6	34.5	34.8	33.6	33.5	33.8
(std. dev.)	(9.3)	(9.2)	(9.3)	(9.3)	(9.3)	(9.3)	(9.4)	(9.4)	(9.4)
Mean employer size	2.881	1.573	5.083	2.987	2.075	4.394	2.815	1.774	4.496
(std. dev.)	(17.751)	(11.591)	(24.734)	(16.688)	(12.482)	(21.552)	(16.418)	(11.510)	(22.059)
Mean gender-employer size	1.805	719	3.635	1.868	1.160	2.961	1.768	936	3.114
(std. dev.)	(11.506)	(3.191)	(18.243)	(10.255)	(5.432)	(14.828)	(10.390)	(4.436)	(15.739)
Mean employer age	28.3	25.8	32.6	29.8	27.3	33.6	28.1	25.7	32.0
(std. dev.)	(22.2)	(20.8)	(23.7)	(23.5)	(22.2)	(24.8)	(22.8)	(21.5)	(24.3)
Mean months employed during year	9.9	9.8	10.1	10.1	10.0	10.3	9.9	9.8	10.0
(std. dev.)	(3.2)	(3.2)	(3.2)	(3.0)	(3.0)	(2.9)	(3.2)	(3.2)	(3.1)
Mean contractual work hours	41.6	42.5	39.9	41.6	42.5	40.3	41.7	42.6	40.3
(std. dev.)	(5.5)	(4.1)	(6.9)	(5.1)	(3.9)	(6.3)	(5.1)	(3.9)	(6.4)
Mean years of tenure	4.0	3.6	4.8	4.2	3.9	4.6	3.9	3.6	4.5
(std. dev.)	(5.7)	(5.2)	(6.3)	(5.7)	(5.4)	(6.1)	(5.6)	(5.2)	(6.1)
Mean log real monthly earnings	7.081	7.125	7.007	7.351	7.408	7.263	7.211	7.262	7.129
(std. dev.)	(0.712)	(0.715)	(0.699)	(0.674)	(0.675)	(0.662)	(0.693)	(0.697)	(0.679)
Number of worker-years	26,540,816	16,652,198	9,888,619	34,829,068	21,129,852	13,699,215	267,312,672	165,144,176	102,168,504
Number of unique workers	26,540,816	16,652,198	9,888,619	34,829,068	21,129,852	13,699,215	56,297,216	33,761,352	22,535,864
Number of unique employers	396,307	278,473	117,834	478,031	328,691	149,340	607,106	402,880	204,226
Share female	0.373			0.393			0.382		
Mean log gender earnings gap	0.118			0.146			0.133		

Note: Table shows summary statistics for the connected set for 2007, 2014, and the pooled years 2007–2014.

Source: RAIS, 2007–2014.

TABLE A.4—COMPARISON OF SUMMARY STATISTICS (SELECTION VS. ALL)

	Pooled 2007–2014, selection set			Pooled 2007–2014, all			Ratio: selection set to all		
	Overall	Men	Women	Overall	Men	Women	Overall	Men	Women
Share Nonwhite	0.378	0.409	0.329	0.365	0.397	0.318	103.6	103.0	103.5
Share primary school	0.088	0.117	0.043	0.085	0.115	0.039	103.5	101.7	110.3
Share middle school	0.203	0.245	0.134	0.199	0.245	0.130	102.0	100.0	103.1
Share high school	0.531	0.516	0.554	0.548	0.524	0.584	96.9	98.5	94.9
Share college	0.179	0.122	0.269	0.168	0.115	0.246	106.5	106.1	109.3
Mean years of education	11.1	10.4	12.1	11.0	10.4	12.0	100.9	100.0	100.8
(std. dev.)	(3.3)	(3.3)	(2.9)	(3.2)	(3.3)	(2.8)	103.1	100.0	103.6
Mean years of age	33.6	33.5	33.8	33.0	33.1	33.0	101.8	101.2	102.4
(std. dev.)	(9.4)	(9.4)	(9.4)	(9.4)	(9.4)	(9.4)	100.0	100.0	100.0
Mean employer size	2,775	1,757	4,402	2,406	1,518	3,742	115.3	115.7	117.6
(std. dev.)	(16,298)	(11,448)	(21,827)	(16,264)	(11,559)	(21,410)	100.2	99.0	101.9
Mean gender-employer size	1,743	926	3,048	1,495	761	2,599	116.6	121.7	117.3
(std. dev.)	(10,314)	(4,413)	(15,573)	(10,194)	(4,173)	(15,231)	101.2	105.8	102.2
Mean employer age	28.0	25.7	31.7	23.1	21.6	25.4	121.2	119.0	124.8
(std. dev.)	(22.8)	(21.5)	(24.3)	(21.8)	(20.4)	(23.5)	104.6	105.4	103.4
Mean months employed during year	9.9	9.8	10.0	9.6	9.5	9.7	103.1	103.2	103.1
(std. dev.)	(3.2)	(3.2)	(3.1)	(3.2)	(3.2)	(3.3)	100.0	100.0	93.9
Mean contractual work hours	41.7	42.6	40.3	41.9	42.6	40.7	99.5	100.0	99.0
(std. dev.)	(5.1)	(3.9)	(6.4)	(5.4)	(4.3)	(6.6)	94.4	90.7	97.0
Mean years of tenure	3.9	3.6	4.5	3.2	2.9	3.5	121.9	124.1	128.6
(std. dev.)	(5.6)	(5.2)	(6.1)	(5.0)	(4.7)	(5.4)	112.0	110.6	113.0
Mean log real monthly earnings	7,206	7,259	7,122	7,112	7,166	7,032	101.3	101.3	101.3
(std. dev.)	(0.693)	(0.697)	(0.678)	(0.656)	(0.666)	(0.633)	105.6	104.7	107.1
Number of worker-years	271,401,248	166,964,944	104,436,320	449,390,272	269,897,824	179,492,464	60.4	61.9	58.2
Number of unique workers	56,868,844	33,975,636	22,893,206	77,397,648	44,436,752	32,960,898	73.5	76.5	69.5
Number of unique employers	640,625	419,717	220,908	6,044,593	3,404,144	2,640,449	10.6	12.3	8.4
Share female	0.385			0.399			96.5		
Mean log gender earnings gap	0.137			0.134			102.2		

Note: Table shows summary statistics for the selected sample, all observations, and their ratio for the pooled years 2007–2014.

Source: RAIS, 2007–2014.

TABLE A.5—COMPARISON OF SUMMARY STATISTICS (CONNECTED VS. SELECTION)

	Pooled 2007–2014, connected set			Pooled 2007–2014, selection set			Ratio: connected set to selection set		
	Overall	Men	Women	Overall	Men	Women	Overall	Men	Women
Share Nonwhite	0.378	0.409	0.327	0.378	0.409	0.329	100.0	100.0	99.4
Share primary school	0.088	0.116	0.042	0.088	0.117	0.043	100.0	99.1	97.7
Share middle school	0.202	0.245	0.133	0.203	0.245	0.134	99.5	100.0	99.3
Share high school	0.531	0.516	0.554	0.531	0.516	0.554	100.0	100.0	100.0
Share college	0.179	0.122	0.272	0.179	0.122	0.269	100.0	100.0	101.1
Mean years of education	11.1	10.4	12.1	11.1	10.4	12.1	100.0	100.0	100.0
(std. dev.)	(3.3)	(3.3)	(2.9)	(3.3)	(3.3)	(2.9)	100.0	100.0	100.0
Mean years of age	33.6	33.5	33.8	33.6	33.5	33.8	100.0	100.0	100.0
(std. dev.)	(9.4)	(9.4)	(9.4)	(9.4)	(9.4)	(9.4)	100.0	100.0	100.0
Mean employer size	2,815	1,774	4,496	2,775	1,757	4,402	101.4	101.0	102.1
(std. dev.)	(16,418)	(11,510)	(22,059)	(16,298)	(11,448)	(21,827)	100.7	100.5	101.1
Mean gender-employer size	1,768	936	3,114	1,743	926	3,048	101.4	101.1	102.2
(std. dev.)	(10,390)	(4,436)	(15,739)	(10,314)	(4,413)	(15,573)	100.7	100.5	101.1
Mean employer age	28.1	25.7	32.0	28.0	25.7	31.7	100.4	100.0	100.9
(std. dev.)	(22.8)	(21.5)	(24.3)	(22.8)	(21.5)	(24.3)	100.0	100.0	100.0
Mean months employed during year	9.9	9.8	10.0	9.9	9.8	10.0	100.0	100.0	100.0
(std. dev.)	(3.2)	(3.2)	(3.1)	(3.2)	(3.2)	(3.1)	100.0	100.0	100.0
Mean contractual work hours	41.7	42.6	40.3	41.7	42.6	40.3	100.0	100.0	100.0
(std. dev.)	(5.1)	(3.9)	(6.4)	(5.1)	(3.9)	(6.4)	100.0	100.0	100.0
Mean years of tenure	3.9	3.6	4.5	3.9	3.6	4.5	100.0	100.0	100.0
(std. dev.)	(5.6)	(5.2)	(6.1)	(5.6)	(5.2)	(6.1)	100.0	100.0	100.0
Mean log real monthly earnings	7.211	7.262	7.129	7.206	7.259	7.122	100.1	100.0	100.1
(std. dev.)	(0.693)	(0.697)	(0.679)	(0.693)	(0.697)	(0.678)	100.0	100.0	100.1
Number of worker-years	267,312,672	165,144,176	102,168,504	271,401,248	166,964,944	104,436,320	98.5	98.9	97.8
Number of unique workers	56,297,216	33,761,352	22,535,864	56,868,844	33,975,636	22,893,206	99.0	99.4	98.4
Number of unique employers	607,106	402,880	204,226	640,625	419,717	220,908	94.8	96.0	92.4
Share female	0.382			0.385			99.2		
Mean log gender earnings gap	0.133			0.137			97.1		

Note: Table shows summary statistics for the connected set, all observations, and their ratio for the pooled years 2007–2014.

Source: RAIS, 2007–2014.

B. EMPIRICAL GENDER PAY GAPS AND EMPLOYER HETEROGENEITY APPENDIX

1. Detailed AKM Estimation Results

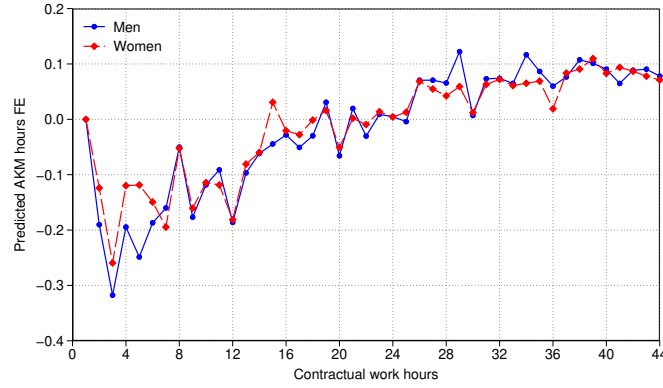


FIGURE B.1. PREDICTED AKM CONTRACTUAL WORK HOURS FES, BY GENDER

Note: Figure shows predicted AKM contractual work hours FEs separately for men and women based on estimating earnings equation (1). The omitted category is 1 hour, for which the FE value is normalized to 0.
Source: RAIS, 2007–2014.

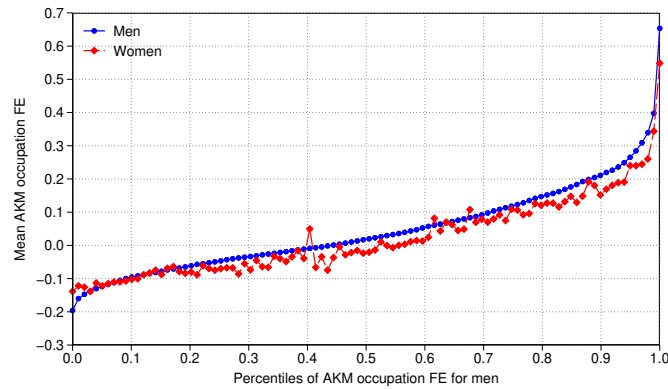


FIGURE B.2. PREDICTED AKM OCCUPATION FES, BY GENDER

Note: This figure shows the predicted AKM occupation FEs separately for men and women based on estimating the earnings equation (1). The fixed effects of both genders are sorted by the mean FEs of men's FE quantiles.
Source: RAIS, 2007–2014.

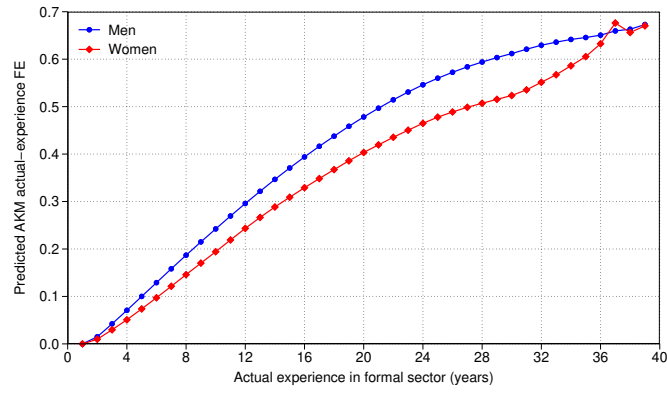


FIGURE B.3. PREDICTED AKM ACTUAL-EXPERIENCE FEs, BY GENDER

Note: This figure shows predicted AKM actual-experience FEs separately for men and women based on estimating the earnings equation (1).
Source: RAIS, 2007–2014.

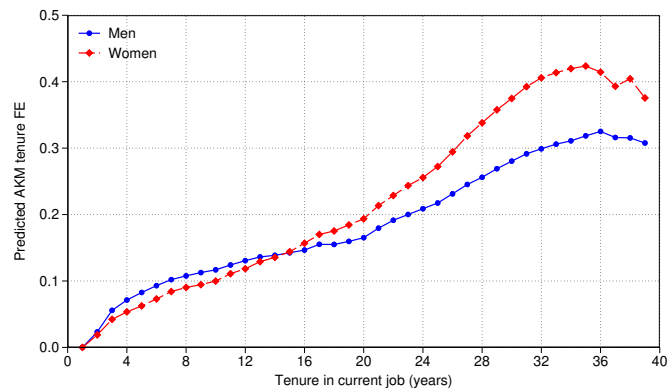


FIGURE B.4. PREDICTED AKM TENURE FEs, BY GENDER

Note: This figure shows predicted AKM tenure FEs separately for men and women based on estimating earnings equation (1).
Source: RAIS, 2007–2014.

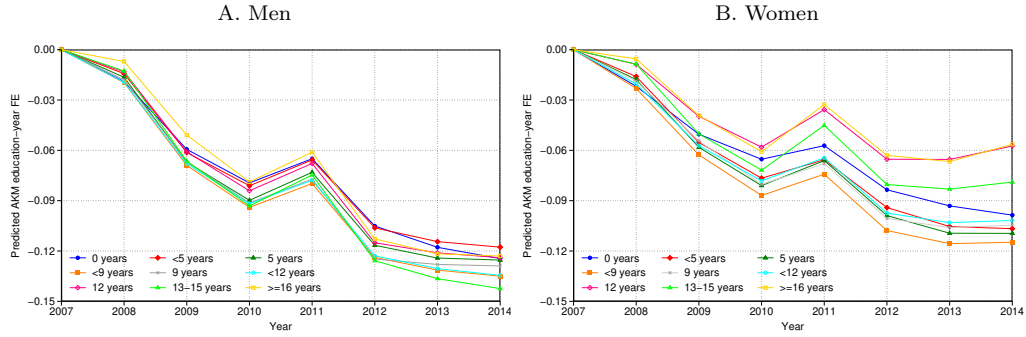


FIGURE B.5. PREDICTED AKM EDUCATION-YEAR FES, BY GENDER

Note: Figure shows predicted AKM education-year FEs separately for men and women based on estimating earnings equation (1). Note that the declining pattern for both genders and all education categories is due to measuring earnings in multiples of the prevailing minimum wage, which increased over this period—see Engbom and Moser (2022a) for details.
Source: RAIS, 2007–2014.

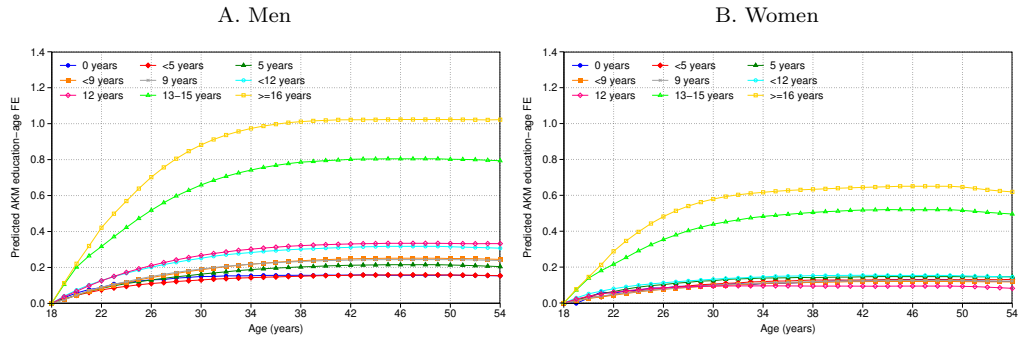


FIGURE B.6. PREDICTED AKM EDUCATION-AGE FES, BY GENDER

Note: Figure shows predicted AKM education-age FEs separately for men and women based on estimating earnings equation (1). Age-pay profiles for all education groups are constrained to be constant from age 45 to age 49 and unconstrained otherwise.
Source: RAIS, 2007–2014.

2. Further Details on Between vs. Within-Employer Pay Differences

Here, we present two alternative Kitagawa-Oaxaca-Blinder decompositions of the gender pay gap, only the first of which is presented in the main text. The overall gender gap in employer FEs can be written as

$$\begin{aligned}
 \text{(B.1)} \quad & \mathbb{E} [\psi_{MJ(i,t)} | G(i) = M] - \mathbb{E} [\psi_{FJ(i,t)} | G(i) = F] \\
 \text{(B.2)} \quad & = \underbrace{(\mathbb{E} [\psi_{MJ(i,t)} | G(i) = M] - \mathbb{E} [\psi_{MJ(i,t)} | G(i) = F])}_{\text{between-employer gap}} + \underbrace{\mathbb{E} [\psi_{MJ(i,t)} - \psi_{FJ(i,t)} | G(i) = F]}_{\text{within-employer gap}} \\
 \text{(B.3)} \quad & = \underbrace{(\mathbb{E} [\psi_{FJ(i,t)} | G(i) = M] - \mathbb{E} [\psi_{FJ(i,t)} | G(i) = F])}_{\text{between-employer gap}} + \underbrace{\mathbb{E} [\psi_{MJ(i,t)} - \psi_{FJ(i,t)} | G(i) = M]}_{\text{within-employer gap}}.
 \end{aligned}$$

Table B.1 shows the results from the two alternative decompositions of the gender log pay gap corresponding to equations (B.2) and (B.3).

TABLE B.1—ALTERNATIVE KITAGAWA-OAXACA-BLINDER DECOMPOSITIONS OF THE GENDER LOG PAY GAP

	Gender log pay gap	Between-employer gap		Within-employer gap	
		Level	Share (%)	Level	Share (%)
Decomposition 1	0.112	0.089	79.5	0.023	20.5
Decomposition 2	0.112	0.073	64.9	0.039	35.1

Note: This table shows results from the Kitagawa-Oaxaca-Blinder decomposition of the overall gender log pay gap into a between-employer gap (i.e., pay-policy component) and a within-employer gap (i.e., sorting component). Decomposition 1 corresponds to equation (B.2) and uses men’s employer FEs for computing the between-employer component. Decomposition 2 corresponds to equation (B.3) and uses women’s employer FEs for computing the between-employer component.

Source: RAIS, 2007–2014.

To illustrate these two decompositions, Figure B.7 shows the distributions of gender-specific employer FEs underlying the individual terms in equations (B.2) and (B.3).

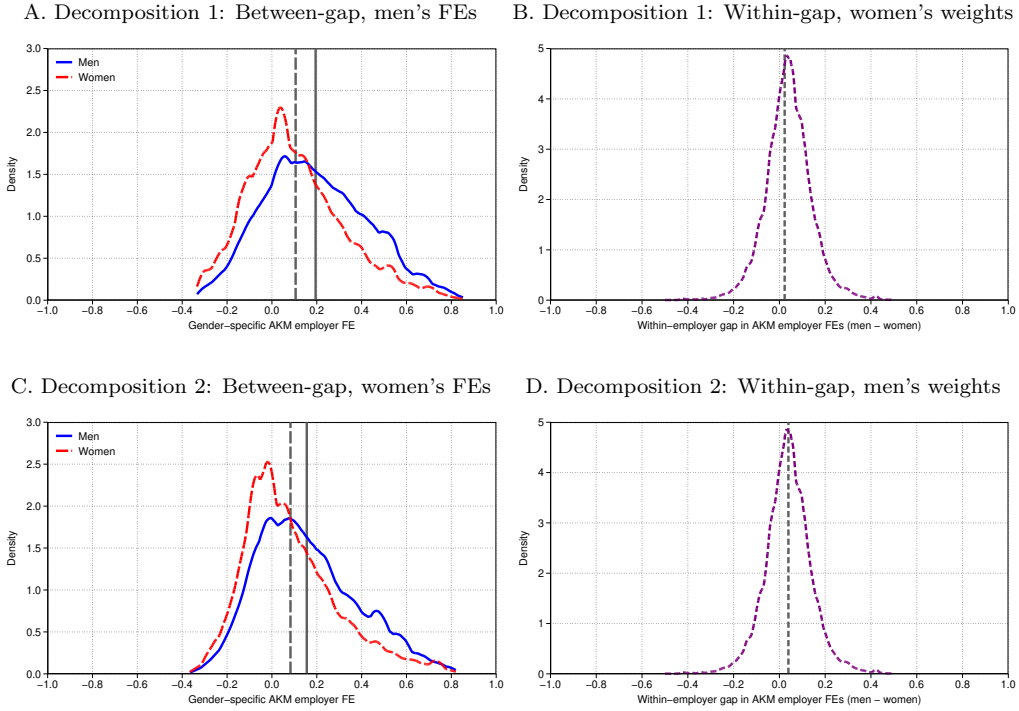


FIGURE B.7. COMPONENTS OF KITAGAWA-OAXACA-BLINDER DECOMPOSITIONS

Note: This figure shows the pay distributions underlying the Kitagawa-Oaxaca-Blinder decompositions—specifically, the between-gap using men's FEs (Panel A) and using women's FEs (Panel C), as well as the within-gap using women's weights (Panel B) and using men's weights (Panel D). Decomposition 1 (Panels A and B) and decomposition 2 (Panels C and D) correspond to equations (B.2) and (B.3), respectively. The dashed vertical line shows the mean of the distribution.
Source: RAIS, 2007–2014.

C. EQUILIBRIUM MODEL OF EMPLOYER PAY, AMENITIES, AND SIZE APPENDIX

1. Definition of a Stationary Equilibrium

We are now ready to define a *stationary equilibrium* for this economy.

DEFINITION 1: A stationary search equilibrium is a set of worker value functions $\{S_{gz}, W_{gz}\}_{gz}$ and policy functions $\{\phi_{gz}\}_{gz}$; firm value function Π and policy functions $\{w_{gz}, a_{gz}, v_{gz}\}_{gz}$; utility offer distributions $\{F_{gz}(x)\}_{gz}$; measures of nonemployed workers $\{u_{gz}\}_{gz}$, aggregate job searchers $\{U_{gz}\}_{gz}$, aggregate vacancies $\{V_{gz}\}_{gz}$, and labor market tightnesses $\{\theta_{gz}\}_{gz}$; job offer arrival rates $\{\lambda_{gz}^U, \lambda_{gz}^E, \lambda_{gz}^G\}_{gz}$; and firm sizes $\{l_{gz}\}_{gz}$ such that, for all (g, z) :

- given $F_{gz}(x)$ and $\{\lambda_{gz}^U, \lambda_{gz}^E, \lambda_{gz}^G\}$, workers' value functions S_{gz} and W_{gz} satisfy equations (3) and (4);
- nonemployed workers' job acceptance policy follows a threshold rule ϕ_{gz} given by equation (5), and employed workers with flow utility x accept voluntarily any job x' such that $x' > x$;
- firm sizes $l_{gz}(\cdot)$ solve equation (16);
- given $l_{gz}(\cdot)$, firms' value function $\Pi_{gz}(\cdot)$ satisfies equation (11);
- firms choose $\{w_{gz}, a_{gz}, v_{gz}\}$ to maximize their objective given by equation (11);
- measures of nonemployed workers are given by equation (6), aggregate job searchers U_{gz} and aggregate vacancies V_{gz} are given by equation (12), and labor market tightness θ_{gz} is given by equation (13);
- given θ_{gz} , the job offer arrival rates $\{\lambda_{gz}^U, \lambda_{gz}^E, \lambda_{gz}^G\}$ satisfy equation (14);
- given $F_{gz}(x)$, $\{\lambda_{gz}^U, \lambda_{gz}^E\}_{gz}$, λ_{gz}^G , and V_{gz} , firm sizes satisfy equation (16); and
- the offer distribution satisfies $F_{gz}(x) = \int_j v_{gz}(j) \mathbf{1}[x_{gz}(j) \leq x] d\Gamma(j) / V_{gz}$.

2. Proof of Lemma 1 (Optimal Amenities)

RESTATEMENT OF LEMMA 1 (OPTIMAL AMENITIES).

A firm's optimal amenity policy function $a_{gz}^*(\cdot)$ can be written as $a_z^*(\tilde{c}_g^{a,0}) = (\tilde{c}_g^{a,0})^{1/(1-\eta^a)} z$. The optimal amenity cost function $c_{gz}^{a,*}(\cdot)$ can be written as $c_z^{a,*}(\tilde{c}_g^{a,0}) = (\tilde{c}_g^{a,0})^{1/(1-\eta^a)} z / \eta^a$.

PROOF:

Based on the insight that workers care only about the flow utility of a job, we can rewrite the problem of a firm in equation (11) as one of choosing in each market a flow utility $x = w + \beta_g a$ and vacancies v that solve the following problem:

$$(C.1) \quad \max_{x,v} \left\{ [(1 - \tau_g)pz - c_{gz}^x(x)] l_{gz}(x, v) - c_{gz}^v(v) \right\}, \quad \forall (g, z),$$

where $c_{gz}^x(x)$ is the solution to the following cost-minimization subproblem in each market (g, z) :

$$(C.2) \quad c_{gz}^x(x) = \min_{w,a} \{w + c_{gz}^a(a)\} \quad \text{s.t.} \quad w + \beta_g a = x.$$

Once written in this way, it is evident that an interior solution to the firm's cost-minimization problem in equation (C.2) is characterized by the following set of first-order conditions (FOCs):

$$(C.3) \quad \frac{\partial c_{gz}^a(a^*)}{\partial a} = \beta_g,$$

$$(C.4) \quad w^* = x - \beta_g a^*.$$

Equation (C.3) uniquely pins down a firm's optimal amenity choice a_{gz}^* for every market (g, z) .

Using the functional form of amenity costs in equation (8), we can rearrange equation (C.3) to get the following expression for the optimal amenity level a_{gz}^* as a function of the preference-adjusted amenity cost shifter $\tilde{c}_g^{a,0}$:

$$(C.5) \quad a_z^* (\tilde{c}_g^{a,0}) = (\tilde{c}_g^{a,0})^{\frac{1}{1-\eta^a}} z.$$

Equation (C.5) shows that a firm's optimal amenity provision is a function of only the preference-adjusted amenity cost shifter $\tilde{c}_g^{a,0}$ and z . Specifically, the optimal amenity policy is linear in worker ability z . Since $\eta^a > 1$, optimal amenities are also decreasing in the preference-adjusted amenity cost shifter $\tilde{c}_g^{a,0}$. Optimal amenities are invariant to all other parameters. Consequently, a firm with preference-adjusted amenity cost shifter $\tilde{c}_g^{a,0}$ delivers an amenity valuation of $\beta_g(j)a_z^*(\tilde{c}_g^{a,0})$ to workers of type (g, z) . The optimal wage is then chosen to deliver the remainder of flow utility x according to equation (C.4).⁵⁵

We obtain an explicit expression for the preference-adjusted amenity cost shifter

⁵⁵Taking into account possible corner solutions, the optimal wage-amenity combination is

$$a_z^{**}(x, \tilde{c}_g^{a,0}) = \begin{cases} x/\beta_g & \text{if } x < \bar{x}(\tilde{c}_g^{a,0}) \\ a_z^*(\tilde{c}_g^{a,0}) & \text{if } x \geq \bar{x}(\tilde{c}_g^{a,0}) \end{cases}, \quad w_{gz}^{**}(x, \tilde{c}_g^{a,0}) = \begin{cases} 0 & \text{if } x < \bar{x}(\tilde{c}_g^{a,0}) \\ x - \beta_g a_z^{**}(\tilde{c}_g^{a,0}, x) & \text{if } x \geq \bar{x}(\tilde{c}_g^{a,0}) \end{cases},$$

where $\bar{x}(\tilde{c}_g^{a,0})$ solves $\partial c_{gz}^a(\bar{x}(\tilde{c}_g^{a,0}))/\partial a = \beta_g$. Note, however, that in such corner solutions, the optimal wage is $w^{**} = 0$, which is empirically not relevant. Going forward, we focus on the interior solution.

$\tilde{c}_g^{a,0}$ corresponding to a given equilibrium amenity level a_{gz}^* by rearranging equation (C.5) as

$$(C.6) \quad \tilde{c}_g^{a,0} = \left(\frac{a_{gz}^*}{z} \right)^{1-\eta^a}.$$

By plugging equation (C.6) into the amenity cost function from equation (8) and using the expression for optimal amenities in equation (C.5), we can rewrite the equilibrium cost of delivering an amenity valuation $\beta_g a^*$ to the worker as

$$(C.7) \quad c_{gz}^a(a_{gz}^*) = c_g^{a,0} \frac{(a_{gz}^*/z)^{\eta^a}}{\eta^a} z$$

$$(C.8) \quad = \beta_g \left(\frac{a_{gz}^*}{z} \right)^{1-\eta^a} \frac{(a_{gz}^*/z)^{\eta^a}}{\eta^a} z$$

$$(C.9) \quad = \beta_g \frac{a_{gz}^*}{\eta^a}$$

$$(C.10) \quad = \beta_g \frac{\left(\tilde{c}_g^{a,0} \right)^{\frac{1}{1-\eta^a}} z}{\eta^a}.$$

3. Proof of Lemma 2 (Optimal Vacancies)

RESTATEMENT OF LEMMA 2 (OPTIMAL VACANCIES).

Keeping fixed all other parameters, a firm's optimal vacancy policy $v_{gz}^*(\cdot)$ is strictly increasing in composite productivity $\tilde{p}_{gz}$ and, thus, strictly increasing in productivity p , strictly decreasing in the gender wedge τ for women, and strictly decreasing in the preference-adjusted amenity cost shifter $\tilde{c}_g^{a,0}$.

PROOF:

We first reformulate the firm's problem. The expected profits per worker contacted by a firm are

$$(C.11) \quad P_{gz}(\tilde{p}_{gz}, x) = h_{gz}(x) J_{gz}(\tilde{p}_{gz}, x),$$

where $h_{gz}(x)$ is the acceptance probability and $J_{gz}(\tilde{p}_{gz}, x)$ is the value of employing a worker to a firm with composite productivity $\tilde{p}_{gz}$ providing flow utility x . Under the assumption that firms maximize long-run profits, the value of employing a worker is simply

$$(C.12) \quad J_{gz}(\tilde{p}_{gz}, x) = \frac{\tilde{p}_{gz} - x}{\delta_{gz} + \lambda_{gz}^E (1 - F_{gz}(x)) + \lambda_{gz}^G}$$

$$(C.13) \quad = \frac{(\tilde{p}_{gz} - x) / (\delta_{gz} + \lambda_{gz}^G)}{1 + \kappa_{gz}^E (1 - F_{gz}(x))},$$

where $\kappa_{gz}^e = \lambda_{gz}^e / (\delta_{gz} + \lambda_{gz}^G)$. The acceptance probability for a firm offering x is

(C.14)

$$h_{gz}(x) = \frac{u_{gz} + s_{gz}^E (1 - u_{gz}) G_{gz}(x) + s_{gz}^G}{u_{gz} + s_{gz}^E (1 - u_{gz}) + s_{gz}^G}$$

(C.15)

$$= \frac{\delta_{gz} + s_{gz}^E (\lambda_{gz}^U + \lambda_{gz}^G) G_{gz}(x) + s_{gz}^G (\delta_{gz} + \lambda_{gz}^U + \lambda_{gz}^G)}{\delta_{gz} + s_{gz}^E (\lambda_{gz}^U + \lambda_{gz}^G) + s_{gz}^G (\delta_{gz} + \lambda_{gz}^U + \lambda_{gz}^G)}$$

(C.16)

$$= \frac{1 + s_{gz}^E \kappa_{gz}^U G_{gz}(x) + s_{gz}^G (1 + \kappa_{gz}^U)}{1 + s_{gz}^E \kappa_{gz}^U + s_{gz}^G (1 + \kappa_{gz}^U)}$$

(C.17)

$$= \frac{1 + s_{gz}^E \kappa_{gz}^U \left[\frac{F_{gz}(x)}{1 + \kappa_{gz}^E [1 - F_{gz}(x)]} \right] + s_{gz}^G (1 + \kappa_{gz}^U)}{1 + s_{gz}^E \kappa_{gz}^U + s_{gz}^G (1 + \kappa_{gz}^U)}$$

(C.18)

$$= \frac{1 + \kappa_{gz}^E [1 - F_{gz}(x)] + s_{gz}^E \kappa_{gz}^U F_{gz}(x) + s_{gz}^G (1 + \kappa_{gz}^U) [1 + \kappa_{gz}^E [1 - F_{gz}(x)]]}{[1 + s_{gz}^E \kappa_{gz}^U + s_{gz}^G (1 + \kappa_{gz}^U)] [1 + \kappa_{gz}^E [1 - F_{gz}(x)]]},$$

where $\kappa_{gz}^U = (\lambda_{gz}^U + \lambda_{gz}^G) / \delta_{gz}$ and u_{gz} is substituted with its expression in equation (6). Combining expressions, the expected profits per contacted worker are

(C.19)

$$\begin{aligned} z(\tilde{p}_{gz}, x) &= h(x) J(\tilde{p}, x) \\ (C.20) \quad &= \frac{\{1 + \kappa_{gz}^E [1 - F_{gz}(x)] + s_{gz}^E \kappa_{gz}^U F_{gz}(x) + s_{gz}^G (1 + \kappa_{gz}^U) [1 + \kappa_{gz}^E [1 - F_{gz}(x)]]\} (\tilde{p}_{gz} - x)}{[1 + s_{gz}^E \kappa_{gz}^U + s_{gz}^G (1 + \kappa_{gz}^U)] [1 + \kappa_{gz}^E (1 - F_{gz}(x))]^2 (\delta_{gz} + \lambda_{gz}^G)}. \end{aligned}$$

Then, the firm's problem becomes

$$(C.21) \quad \max_{x, v} \{P_{gz}(\tilde{p}_{gz}, x) v q_{gz} - c_{gz}^v(v)\},$$

where q_{gz} is defined as in equation (14). Therefore, the optimal flow-utility and vacancy policy functions satisfy

$$\begin{aligned} x_{gz}^*(\tilde{p}_{gz}, \cdot) &= \arg \max_x P_{gz}(\tilde{p}_{gz}, x) \\ (C.22) \quad \frac{\partial c_{gz}^v(v^*(\tilde{p}_{gz}, \cdot))}{\partial v} &= q_{gz} \max_x P_{gz}(\tilde{p}_{gz}, x). \end{aligned}$$

Since the vacancy cost function $c^v(\cdot)$ is convex, and $z(\tilde{p}_{gz}, x)$ in equation (C.20) is strictly increasing in $\tilde{p}_{gz}$, then it follows from an application of the envelope

theorem to equation (C.22) that $v^*(\tilde{p}_{gz}, \cdot)$ is strictly increasing in $\tilde{p}_{gz}$. Therefore, $v_{gz}^*(\cdot)$ is strictly increasing in productivity p , strictly decreasing (constant) in z_a for women (men). Finally, optimal vacancies are also strictly decreasing in the preference-adjusted amenity cost shifter $\tilde{c}_g^{a,0}$ due to the result in Lemma 1.

4. Proof of Lemma 3 (Optimal Flow Utility and Wages)

RESTATEMENT OF LEMMA 3 (OPTIMAL FLOW UTILITY AND WAGES).

Keeping fixed all other parameters, a firm's optimal flow utility offer $x_{gz}^*(\cdot)$ is strictly increasing in composite productivity $\tilde{p}_{gz}$ and, thus, strictly increasing in productivity p for all worker types, strictly decreasing in the gender wedge τ for women, and strictly decreasing in the preference-adjusted amenity cost shifter $\tilde{c}_g^{a,0}$. A firm's optimal wage offer $w_{gz}^*(\cdot)$ is strictly increasing in productivity p for all worker types and strictly decreasing in the gender wedge τ for women.

PROOF:

We proceed in two steps.

Step 1.

In the first step, we prove monotonicity of x_{gz}^* in components of $\tilde{p}_{gz}$. Lemma 1 implies that at the optimum, amenities can be equivalently considered exogenous. Thus, we rewrite the FOCs as functions of exogenous parameters, the endogenous offer distribution, and x_{gz} :

$$(C.23) \quad [\partial v_{gz}] : \quad c_{gz}^{v,0} \frac{\partial \tilde{c}^v(v_{gz})}{\partial v_{gz}} = T_{gz}(\tilde{p}_{gz} - x_{gz}) \left(\frac{1}{\delta_{gz} + \lambda_{gz}^G + \lambda_{gz}^E(1 - F_{gz}(x_{gz}))} \right)^2,$$

$$(C.24) \quad [\partial x_{gz}] : \quad 1 = (\tilde{p}_{gz} - x_{gz}) \frac{2\lambda_{gz}^E f_{gz}(x_{gz})}{\delta_{gz} + \lambda_{gz}^G + \lambda_{gz}^E(1 - F_{gz}(x_{gz}))},$$

where $T_{gz} = \mu_{gz}[(u_{gz} + s_{gz}^G)\lambda_{gz}^U(\delta_{gz} + \lambda_{gz}^G + \lambda_{gz}^E)]/V_{gz}$. Now consider equation (C.23); because the term on the right-hand side is always positive for $\tilde{p}_{gz} > \phi_{gz}$, it follows that optimal vacancies $v_{gz}^*(\tilde{p}_{gz}, c_{gz}^{v,0})$ are always strictly positive.

We now show that the derivative of wages with respect to $\tilde{p}_{gz}$ is always positive. Define $h_{gz}(\tilde{p}_{gz}) = F_{gz}(x_{gz}^*(\tilde{p}_{gz}))$. Thus

$$(C.25) \quad h_{gz}(\tilde{p}_{gz}) = \frac{\int_{\tilde{p}' \geq \phi_{gz}}^{\tilde{p}_{gz}} v_{gz}^*(\tilde{p}') \gamma_{gz}(\tilde{p}')}{V_{gz}} d\tilde{p}'$$

$$(C.26) \quad h'_{gz}(\tilde{p}_{gz}) = f_{gz}(x_{gz}^*(\tilde{p}_{gz})) x_{gz}^{*'}(\tilde{p}_{gz})$$

$$(C.27) \quad f_{gz}(x_{gz}^*(\tilde{p}_{gz})) = h'_{gz}(\tilde{p}_{gz})/x_{gz}^{*'}(\tilde{p}_{gz}),$$

where $v_{gz}^*(\tilde{p}_{gz})$ are optimal vacancies conditional on $\tilde{p}_{gz}$, $\gamma_{gz}(\tilde{p}_{gz})$ is the marginal density of composite productivity $\tilde{p}_{gz}$, and $\partial x_{gz}^*(\tilde{p}_{gz})/\partial \tilde{p}_{gz} = x_{gz}^{*'}(\tilde{p}_{gz})$ is the derivative of equilibrium flow utility with respect to $\tilde{p}_{gz}$. Thus, we can rewrite

$h'_{gz}(\tilde{p}_{gz}) = v_{gz}^*(\tilde{p}_{gz})/V_{gz}\gamma(\tilde{p}_{gz})$ by differentiating equation (C.27) using Leibniz's integral rule.

Using these identities, we can write

$$f_{gz}(x_{gz}^*(\tilde{p}_{gz})) = v_{gz}^*(\tilde{p}_{gz})/V_{gz}\gamma_{gz}(\tilde{p}_{gz})\partial\tilde{p}_{gz}/\partial x_{gz}^*(\tilde{p}_{gz}).$$

Thus, we can rewrite equation (C.24) as

$$(C.28) \quad \frac{\partial x_{gz}^*(\tilde{p}_{gz})}{\partial \tilde{p}_{gz}} = (\tilde{p}_{gz} - x_{gz}^*) \frac{2\lambda_{gz}^E}{\delta_{gz} + \lambda_{gz}^G + \lambda_{gz}^E(1 - h_{gz}(\tilde{p}_{gz}))} \frac{v_{gz}^*(\tilde{p}_{gz})}{V_{gz}} \gamma_{gz}(\tilde{p}_{gz}).$$

Because the right-hand side of this expression is positive for $\tilde{p}_{gz} > \phi_{gz}$, we have $\partial x_{gz}^*(\tilde{p}_{gz})/\partial \tilde{p}_{gz} > 0$, thus proving that equilibrium flow utility is increasing in $\tilde{p}_{gz}$.

Since $\tilde{p}_{gz}$ is strictly increasing in productivity p , strictly decreasing (constant) in the gender wedge τ_g for women (men), and strictly decreasing in the preference-adjusted amenity cost shifter $\tilde{c}_g^{a,0}$ due to Lemma 1, it follows that optimal flow utility is strictly increasing in p , strictly decreasing (constant) in τ_g for women (men), and strictly decreasing in $\tilde{c}_g^{a,0}$.

Step 2.

In the second step, we prove the monotonicity of w_{gz} in components of $\tilde{p}_{gz}$. The characterization of $w_{gz} = x_{gz} - \beta_g a_{gz}$ follows from combining Lemmas 1 and 2 with Step 1 above.

5. Alternative Modeling Assumption on Amenity Production

We here present an alternative formulation of firms' amenity production technology. While in the baseline formulation of the model, firms produce gender-specific amenity values for each worker, an alternative formulation has firms produce a vector of amenities with gender-specific utility weights for each worker. We establish conditions for *observational equivalence* and *counterfactual equivalence* between the baseline model and the alternative model. For ease of exposition, we work with the piece-rate version of our model in which wages and amenity values scale with worker ability z .

Firms post an amenity vector $\vec{a} = (a_1, a_2, \dots, a_N) \in \mathbb{R}^{N \times 1}$ subject to cost $c^a(\vec{a})$. Workers derive gender-specific utility from amenities given a preference vector $\vec{\beta}_g = (\beta_{g,1}, \beta_{g,2}, \dots, \beta_{g,N}) \in \mathbb{R}^{N \times 1}$. A worker of gender g at a firm with amenity vector $\vec{a}$ enjoys amenity utility $\vec{a}'\vec{\beta}_g$.

If $N = 1$, then $\vec{a} = a \in \mathbb{R}$ and $\vec{\beta}_g = \beta_g \in \mathbb{R}$. In this case, men and women agree on the ranking of employers in terms of their amenity values as long as $\text{sign}(\beta_M) = \text{sign}(\beta_F)$. Otherwise, if $\text{sign}(\beta_M) \neq \text{sign}(\beta_F)$, then men and women have opposite rankings of employers in terms of their amenity values. This formulation is too restrictive to match the data, which motivates the following two assumptions.

ASSUMPTION C.1: *Amenities are at least twofold:*

$$(C.29) \quad N \geq 2.$$

ASSUMPTION C.2: *The gender-specific preference vectors $\vec{\beta}_M^a$ for men and $\vec{\beta}_F^a$ for women are linearly independent:*

$$(C.30) \quad \nexists c \in \mathbb{R} \quad \text{s.t.} \quad \vec{\beta}_M = c\vec{\beta}_F.$$

Under these assumptions, we obtain the following result, which helps us rationalize the data:

LEMMA C.1 (Existence of amenity vector): *Suppose Assumptions C.1 and C.2 hold. Then for any duplet of gender-specific utilities (U_M^a, U_F^a) at a given employer, there exists an amenity vector $\vec{a} = (a_1, a_2, \dots, a_N)$ such that*

$$(C.31) \quad \begin{bmatrix} U_M^a \\ U_F^a \end{bmatrix}_{2 \times 1} = \begin{bmatrix} \vec{\beta}_M' \\ \vec{\beta}_F' \end{bmatrix}_{2 \times N} \vec{a}_{N \times 1}.$$

PROOF:

This is simply a system of two linear equations in $N \geq 2$ unknowns. By linear independence of $\vec{\beta}_M$ and $\vec{\beta}_F$ due to Assumption C.2, the matrix that premultiplies $\vec{a}$ in equation (C.31) has full rank. Therefore, the system admits at least one solution.

If $N = 2$, then Lemma C.1 admits a unique amenity vector $\vec{a} = (a_1, a_2)$ that rationalizes any amenity-utility duplet (U_M^a, U_F^a) . If $N > 2$, then there exist multiple amenity vectors $\vec{a}$ that rationalize the same duplet of gender-specific utilities (U_M^a, U_F^a) , among which a profit-maximizing firm will pick the cost-minimizing one.

In addition, we make the following assumption as a natural extension of that in the baseline model:

ASSUMPTION C.3: *Firms provide firm-wide amenities $\vec{a}$, and the cost of amenity provision is given by*

$$(C.32) \quad c^a(\vec{a}, l_M, l_F) = \int_{j=0}^{l_M+l_F} \sum_{i=1}^N c_i^{a,0} \tilde{c}^a(a_i) dj,$$

where j indexes workers, i indexes amenities, $c_i^{a,0}$ is an amenity-specific cost shifter that differs across firms, and $\tilde{c}^a(a_i)$ is increasing convex such that $\tilde{c}^a(0) = 0$ and $\partial \tilde{c}^a / \partial a_i(0) = 0$.

Next, we show that there exist (unique) values of productivity p , the gender wedge τ_g ; there also exist amenity cost shifters $c_i^{a,0}$ for $i \in \{1, \dots, N\}$ that rationalize a

given vector of amenities $\vec{a}$ along with wages w_M and w_F and vacancies v_M and v_F as firms' equilibrium choices.

LEMMA C.2 (Observational equivalence): *Suppose Assumptions C.1, C.2, and C.3 hold. Then, for a given firm-level amenity vector $\vec{a}$, there exists a firm-specific amenity cost function $c^a(\vec{a})$ such that $\vec{a}$ solves*

$$(C.33) \quad (\vec{a}, w_M, w_F, v_M, v_F) = \arg \max_{\vec{a}, \tilde{w}_M, \tilde{w}_F, \tilde{v}_M, \tilde{v}_F} \left\{ \sum_{g=M,F} \left[(1 - \tau_g)p - \tilde{w}_g - c^a(\vec{a}) \right] l_g(\vec{a}, \tilde{w}_g, \tilde{v}_g) - \sum_{g=M,F} c^v(\tilde{v}_g) \right\}$$

for some levels of firm productivity p and gender wedge τ_g .

PROOF:

The system of FOCs associated with firm optimality is the following:

$$(C.34) \quad [\partial a_i]: \quad c_i^{a,0} \frac{\partial \tilde{c}^a(a_i)}{\partial a_i} = \frac{[p - w_M - c^a(\vec{a})] \frac{\partial l_M(\vec{a}, w_M, v_M)}{\partial a_i} + [(1 - \tau)p - w_F - c^a(\vec{a})] \frac{\partial l_F(\vec{a}, w_F, v_F)}{\partial a_i}}{l_M(\vec{a}, w_g, v_g) + l_F(\vec{a}, w_F, v_F)}$$

$$(C.35) \quad [\partial w_M]: \quad 1 = [p - w_M - c^a(\vec{a})] \frac{\frac{\partial l_M(\vec{a}, w_M, v_M)}{\partial w_M}}{l_M(\vec{a}, w_M, v_M)}$$

$$(C.36) \quad [\partial w_F]: \quad 1 = [(1 - \tau)p - w_F - c^a(\vec{a})] \frac{\frac{\partial l_F(\vec{a}, w_F, v_F)}{\partial w_F}}{l_F(\vec{a}, w_F, v_F)}$$

$$(C.37) \quad [\partial v_M]: \quad \frac{\partial c_M^v(v_M)}{v_M} = [p - w_M - c^a(\vec{a})] \frac{\partial l_M(\vec{a}, w_M, v_M)}{\partial v_M}$$

$$(C.38) \quad [\partial v_F]: \quad \frac{\partial c_F^v(v_F)}{v_F} = [p - w_F - c^a(\vec{a}) - z] \frac{\partial l_F(\vec{a}, w_F, v_F)}{\partial v_F}.$$

Note that the only FOC containing the term $\partial \tilde{c}^a(a_i) / \partial a_i$ is equation (C.34). All other FOCs depend on the level of the amenity cost $c^a(\vec{a})$ but not its derivative. Hence, we can scale the amenity cost function $c^a(\vec{a})$, the productivity level for both genders, or the gender wedge for women to satisfy the wage FOCs in equations (C.35)–(C.36) and vacancy posting in equations (C.37)–(C.38). By Assumption C.3, there are N equations (i.e., FOCs with respect to a_i for $i = 1, 2, \dots, N$) with N free parameters (i.e., amenity cost shifters $c_i^{a,0}$ for $i = 1, 2, \dots, N$), so there exists N amenity cost shifters $c_i^{a,0}$ that satisfy firm optimality and rationalize $\vec{a}$. Lemma C.2 establishes *observational equivalence* between our baseline model with gender-specific amenity values and the alternative formulation with an amenity vector and gender-specific utility weights. That is, any observed empirical pattern

of employer ranks and pay differences by gender can be rationalized by either of the two models.

Next, we characterize optimal amenity provision. An argument analogous to that in the main paper shows that under assumptions C.1–C.3, optimal amenities satisfy

$$(C.39) \quad [\partial a_i] : \quad c_i^{a,0} \times \frac{\partial \tilde{c}^a(a_i)}{\partial a_i} \sum_{g=M,F} l_g(x_g, v_g) = \beta_{M,i} l_M(x_M, v_M) + \beta_{F,i} l_F(x_F, v_F), \quad \forall i.$$

Under these assumptions, optimal amenity provision depends on the gender composition of a firm's workforce, which varies with firm fundamentals and counterfactual policies. However, the model under these assumptions is at odds with the empirical observation that amenity quantities differ significantly across men and women, as demonstrated in Section II.B. For example, in the data, women comprise the vast majority of beneficiaries of parental leaves. Thus, assuming that the cost of amenities that are enjoyed by a subset of workers is paid also for all other workers seems inconsistent with the empirical evidence. Motivated by this observation, we consider the following alternative assumption in lieu of Assumption C.3.

ASSUMPTION C.4: *Firms provide individual-specific amenities $\{\vec{a}_j\}_j$ for each worker j and the cost of amenity provision given by*

$$(C.40) \quad c^a(\{\vec{a}_j\}_j) = \int_{j=0}^{l_M+l_F} \sum_{i=1}^N c_i^{a,0} \tilde{c}^a(a_{i,j}) dj,$$

where j indexes workers, i indexes amenities, $c_i^{a,0}$ is an amenity-specific cost shifter that differs across firms, and $\tilde{c}^a(a_{i,j})$ is increasing convex such that $\tilde{c}^a(0) = 0$ and $\partial \tilde{c}^a / \partial a_{i,j}(0) = 0$.

An argument analogous to that in Lemma C.2 establishes observational equivalence between our baseline model and the alternative model under Assumption C.4. Next, we characterize the dependence of a firm's optimal amenity choice on amenity cost shifters $c_i^{a,0}$ for $i \in \{1, \dots, N\}$ and other model parameters.

LEMMA C.3 (Counterfactual equivalence): *Suppose Assumptions C.1, C.2, and C.4 hold. Then, a firm's optimal amenity policy $a_{i,j}^*(\cdot)$ for all (i, j) is strictly decreasing in its amenity cost shifter $c_i^{a,0}$, increasing in gender utility weights $\beta_{g,i}$ for $g \in \{M, F\}$, and invariant to all other parameters.*

PROOF:

Based on the insight that workers care only about the flow utility of a job, we can rewrite the problem of a firm as one of choosing in each market a flow utility

$x_g = w_g + \vec{a}'_g \vec{\beta}_g$ and vacancies v_g that solve the following problem:

(C.41)

$$\max_{\{x_g, v_g\}_{g=M,F}} \left\{ \sum_{g=M,F} [(1 - \tau_g)p] l_g(x_g, v_g) - c^x(x_M, x_F) - \sum_{g=M,F} c_g^v(v_g) \right\}, \quad \forall g,$$

where $c^x(x_M, x_F)$ is the solution to the following cost-minimization subproblem in each market (g, z) :

$$(C.42) \quad c^x(x_M, x_F) = \min_{w_M, w_F, \{\vec{a}_j\}_j} \left\{ w_M l_M(x_M, v_M) + w_F l_F(x_F, v_F) + \int_{j=0}^{l_M+l_F} \sum_{i=1}^N c^a(a_{i,j}) dj \right\}$$

$$\text{s.t.} \quad w_g + \vec{a}'_{g(j)} \vec{\beta}_g = x_g \quad \forall j$$

$$(C.43) \quad = \min_{\{\vec{a}_g\}_g} \left\{ (x_M - \vec{a}'_M \beta_M) l_M(x_M, v_M) + (x_F - \vec{a}'_F \beta_F) l_F(x_F, v_F) \right. \\ \left. + \sum_{i=1}^N [l_M(x_M, v_M) c^a(a_{M,i}) + l_F(x_F, v_F) c^a(a_{F,i})] \right\},$$

where we move from equation (C.42) to equation (C.43) by using Assumption C.4. Note that the cost of amenity production $c^a(\cdot)$ and the marginal amenity utility $\vec{\beta}_g$ are identical for individual workers j of the same gender g , but different across genders. Thus, a firm's optimal amenity choice is to offer the same vector of amenities $\vec{a}_g$ to workers of the same gender and different amenities to workers of different genders. This model prediction is consistent with the salient empirical fact that many job amenities (e.g., parental leave benefits) are differentially accessed by men and women at the same employer. The associated optimality conditions for amenities production are the following:

$$(C.44) \quad [\partial a_{g,i}] : \quad c_i^{a,0} \times \frac{\partial \tilde{c}^a(a_{g,i})}{\partial a_{g,i}} = \beta_{g,i}, \quad \forall (g, i).$$

Equation (C.44) pins down a firm's optimal amenity choice $a_{g,i}^*(c_i^{a,0}, \beta_{g,i})$ as a function of the heterogeneous amenity cost shifter $c_i^{a,0}$ as well as the set of gender-specific amenity-utility weights $\beta_{g,i}$. The optimal wage is then chosen to deliver the remainder of the flow utility x_g to workers of gender $g = M, F$.

Lemma C.3 is a powerful result because it establishes *counterfactual equivalence* with respect to the equilibrium decomposition of the gender pay gap, which lies at the heart of our analysis in Section VII. Specifically, it follows from Lemma C.3 that the gender-specific amenity vector $\vec{a}_g^*$ is independent of productivity p , the gender wedge τ , and other model parameters. Therefore, shutting down amenity cost differences across genders in counterfactual 1 of the main body of the paper has the same effects as equalizing gender preferences over the amenity vector in the alternative model.

MODEL CHOICE.

Both the baseline model with gender-specific amenity values and the alternative formulation with an amenity vector have attractive features. The alternative formulation seems realistic because it allows for a common set of amenities that are differentially accessible to men and women within a given employer.

A drawback of the alternative formulation is that the formulation with an amenity vector requires the strong assumption that we observe the full vector of amenities $\vec{a}$ or, alternatively, that the econometrician knows the gender-specific preference vectors $\vec{\beta}_g$ for $g = M, F$. In contrast, in the baseline model, we treat amenities as an unobserved, gender-employer-specific characteristic that we estimate without making further assumptions about the relevant set of amenities or gender-specific amenity preferences.

In the data, we find that a significant share of the estimated amenity values in our baseline model is accounted for by unobserved gender-firm-specific factors, which seems at odds with the assumption that we observe the full vector $\vec{a}$. Based on the observational equivalence (Lemma C.2) and counterfactual equivalence with respect to the equilibrium decomposition (Lemma C.3) of the two formulations under the stated assumptions, we adopt the baseline amenity-value formulation throughout the empirical analysis and for the equilibrium decomposition. When considering the equilibrium effects of policies, however, counterfactual equivalence between the two models does not generally hold. For this purpose, we proceed with our baseline model and consider robustness with regard to different parameterizations of the cost function.

6. Further Discussion of Model Assumptions

OUTPUT IS LINEAR WITHIN WORKER TYPES.

Additive separability of output in worker types allows the model to admit a log-linear wage equation, which we require to take the model to the data by pooling all workers of the same gender. Conceptually, there is no reason not to simultaneously allow for curvature in the ability-weighted number of workers of each type. However, if the marginal product of a given worker type were exceedingly high for small numbers of workers—as would be the case with standard constant-elasticity-of-substitution specifications—then we would see every firm employing a strictly positive mass of each worker type. This outcome would be clearly at odds with the presence of single-gender firms in the data.

Further supporting this assumption, [Fukui, Nakamura and Steinsson \(2023\)](#) find a small crowd-out between women and men in the U.S. Therefore, a natural starting point treats men and women as interchangeable inputs in production.

Here, we argue that our assumption of labor market segmentation by worker types is reasonable, given the empirical patterns we observe. That firms can direct wages and vacancies toward certain worker types may seem at odds with nondiscrimination laws. But of course a firm need not publicly post different wages or job openings by gender in order to discriminate. Such differences may arise in more subtle ways during résumé screenings, during interviews, and at the negotiation table (Goldin and Rouse, 2000). This assumption is consistent with evidence that men and women differentially apply to jobs based on the wording of vacancies (Abraham, Hallermeier and Stein, 2024) and accept jobs subject to deadlines (Cortés et al., 2023). Empirically, there are also good reasons to adopt market segmentation. First and foremost, our model must confront the significant differences in the gender-specific employer component of pay, amenity utilization, and employment of men and women within the same employer in the Brazilian data. In the previous section, we have already documented gender differences in pay and employment. Our model provides a natural way to rationalize these differences.

7. Alternative Modeling Assumptions on Vacancy Posting

MODEL ALTERNATIVE 1: DIRECTED VACANCY POSTING WITH JOINT COST FUNCTION.

As a first alternative to the benchmark model, suppose that instead of the vacancy cost being separable across genders, we assume that the vacancy cost is a function of the total number of vacancies posted. This model has the stark prediction that any firm will employ either only men or only women, except in knife-edge cases.

Each firm posts a number v_{Mz} of vacancies targeted at male workers and v_{Fz} vacancies targeted at women. The total cost of posting vacancies (v_{Mz}, v_{Fz}) for men and women is given by $c_z^v(v_{Mz} + v_{Fz})$, where the function c_z^v retains the properties laid out in the main text: $c_z^v(0) = 0$, $\partial c_z^v(\cdot)/\partial v > 0$, $\partial^2 c_z^v(\cdot)/\partial v^2 > 0$.

To see that this setup implies gender segregation except in knife-edge cases, note that the firm's problem can now be written as

$$(C.45) \quad \max_{x_{Mz}, x_{Fz}, v_{Mz}, v_{Fz}} \left\{ \sum_{g=M,F} (\tilde{p}_{gz} - x_{gz}) l_{gz}(x_{gz}, v_{gz}) - c_z^v(v_{Mz} + v_{Fz}) \right\}$$

The FOCs with respect to vacancy posting now read

$$(C.46) \quad [\partial v_{Mz}] : \quad c_z^{v'}(v_{Mz} + v_{Fz}) = T_{Mz}(\tilde{p}_{Mz} - x_{Mz}) \left(\frac{1}{\delta_{Mz} + \lambda_{Mz}^G + \lambda_{Mz}^E(1 - F_{Mz}(x_{Mz}))} \right)^2,$$

$$(C.47) \quad [\partial v_{Fz}] : \quad c_z^{v'}(v_{Mz} + v_{Fz}) = T_{Fz}(\tilde{p}_{Fz} - x_{Fz}) \left(\frac{1}{\delta_{Fz} + \lambda_{Fz}^G + \lambda_{Fz}^E(1 - F_{Fz}(x_{Fz}))} \right)^2.$$

Putting equations (C.46) and (C.47) into simple economic terms, the marginal cost of an additional vacancy (the left-hand side) is equated to the marginal benefit of an additional vacancy (the right-hand side). The latter consists of an increase in the employment of that worker type multiplied by the profits made per worker of that type, which is independent of the number of vacancies posted. This is because wages are set according to other FOCs, which do not depend on the number of vacancies posted by that firm.

Since the right-hand sides in equations (C.46) and (C.47) are generically not equal, except in knife-edge cases, it follows that both FOCs cannot hold. This means that the firm will be at a corner solution with regard to one of the two genders, and this must involve posting zero vacancies for that gender.

According to the above analysis, except in knife-edge cases, firms would hire only men or only women—whichever option yields the highest marginal benefit to the firm. This model implication is empirically counterfactual since the vast majority of firms in the real world employ a mix of men and women.

MODEL ALTERNATIVE 2: UNDIRECTED VACANCY POSTING.

As a second alternative to the benchmark model, suppose that instead of vacancies being directed separately to men and women, we assume that firms cannot discriminate between genders in their recruiting. While such a model can qualitatively account for dual-gender firms, it turns out that quantitatively, such a model clearly fails to replicate the empirical distribution of female employment shares across firms that we documented in Section II.A.

Each firm posts a number v_z of gender-neutral vacancies for workers of each ability level z at cost $c_z^v(v_z)$. In such a model, a firm's problem can be written as

$$(C.48) \quad \max_{x_{Mz}, x_{Fz}, v_a} \left\{ \sum_{g=M,F} (\tilde{p}_{gz} - x_{gz}) l_{gz}(x_{gz}, v_a) - c_z^v(v_z) \right\}.$$

Notice that we do not impose that firms hire both genders in each submarket: it is always possible for a firm to offer flow utility $x_{gz} < \phi_{gz}$ such that no worker of gender g will accept it. Consequently, while a total of V_z vacancies are posted in each submarket in the aggregate, only $V_{gz} \leq V_z = \int v_z(\tilde{p}_{gz}) d\Gamma_{gz}(\tilde{p}_{gz})$ vacancies are accepted in equilibrium by workers of type (g, z) . This implies that the number of matches produced in the labor market is given by

$$(C.49) \quad m_{gz} = \chi_{gz} [\mu_{gz}(u_{gz} + s_{gz}^E(1 - u_{gz}) + s_{gz}^G)]^\alpha V_z^{1-\alpha} \frac{V_{gz}}{V_z},$$

which already incorporates the probability that a worker of gender g will meet a vacancy that is associated with a wage below the reservation threshold, leading to a rejection. It is straightforward to show that this matching function exhibits all the properties of standard matching functions and that in particular, $f_{gz}/q_{gz} =$

$V_z/[u_{gz} + s_{gz}^E(1 - u_{gz}) + s_{gz}^G]$, where $f_{gz} = m_{gz}/[u_{gz} + s_{gz}^E(1 - u_{gz}) + s_{gz}^G]$ is the job-finding rate per effective job searcher and $q_{gz} = m_{gz}/V_z$ is the vacancy yield rate.

The following equation represents the law of motion of firm sizes:

(C.50)

$$\dot{l}_{gz}(x, v) = -\delta_{gz}l_{gz}(x, v) - s_{gz}\lambda_{gz}^E(1 - F_{gz}(x))l_{gz}(x, v) + vq_{gz} \left[\frac{u_{gz} + s_{gz}^G}{u_{gz} + s_{gz}^G + (1 - u_{gz})s_{gz}^E} + \frac{(1 - u_{gz})s_{gz}^E}{u_{gz} + s_{gz}^G + (1 - u_{gz})s_{gz}^E} G_{gz}(x) \right]. \quad (C.51)$$

Solving for the stationary solution,

$$l_{gz}(x_{gz}, v_z) = \left(\frac{1}{\delta_{gz} + \lambda_{gz}^G + \lambda_{gz}^E(1 - F_{gz}(x_{gz}))} \right)^2 \frac{v_z}{V_z} \mu_{gz}(u_{gz} + s_{gz}^G) \lambda_{gz}^U (\delta_{gz} + \lambda_{gz}^G + \lambda_{gz}^E). \quad (C.52)$$

To find the firm's policy functions, define $T_{gz} = \mu_{gz}[u_{gz}\lambda_{gz}^U(\delta_{gz} + s_{gz}\lambda_{gz}^U)]/V_z$ and composite productivity $\tilde{p}_{gz} = (1 - \tau_g)zp + a_{gz} - c_{gz}^a(a_{gz})$. We rewrite the firm's problem as a function of the steady state mass of employed workers as follows:

(C.53)

$$\max_{x_{Mz}, x_{Fz}, v_z} \left\{ T_{Mz}v_z(\tilde{p}_{Mz} - x_{Mz}) \left(\frac{1}{\delta_{Mz} + \lambda_{Mz}^G + \lambda_{Mz}^E(1 - F_{Mz}(x_{Mz}))} \right)^2 + T_{Fz}v_z(\tilde{p}_{Fz} - x_{Fz}) \left(\frac{1}{\delta_{Fz} + \lambda_{Fz}^G + \lambda_{Fz}^E(1 - F_{Fz}(x_{Fz}))} \right)^2 - c_z(v_z) \right\}. \quad (C.54)$$

The associated FOCs read

$$c'(v_z) = T_{Mz}(\tilde{p}_{Mz} - x_{Mz}) \left(\frac{1}{\delta_{Mz} + \lambda_{Mz}^G + \lambda_{Mz}^E(1 - F_{Mz}(x_{Mz}))} \right)^2 \quad (C.55)$$

$$+ T_{Fz}(\tilde{p}_{Fz} - x_{Fz}) \left(\frac{1}{\delta_{Fz} + \lambda_{Fz}^G + \lambda_{Fz}^E(1 - F_{Fz}(x_{Fz}))} \right)^2 \quad (C.56)$$

$$1 = (\tilde{p}_{Mz} - x_{Mz}) \frac{2\lambda_{Mz}^E f_{Mz}(x_{Mz})}{\delta_{Mz} + \lambda_{Mz}^G + \lambda_{Mz}^E(1 - F_{Mz}(x_{Mz}))} \quad (C.57)$$

$$1 = (\tilde{p}_{Fz} - x_{Fz}) \frac{2\lambda_{Fz}^E f_{Fz}(x_{Fz})}{\delta_{Fz} + \lambda_{Fz}^G + \lambda_{Fz}^E(1 - F_{Fz}(x_{Fz}))}. \quad (C.58)$$

Recall from Section II.A that firm-level female employment shares are dispersed, ranging from almost 0 to almost 1 in the data. It is this salient feature of the data that the undirected vacancy-posting model fails to replicate. To demonstrate this, we show that analytically derived expressions for the lowest and highest female

employment shares are inconsistent with the data for realistic calibrations of the labor market parameters guiding worker flows.

Using equation (C.52), we can write the female share of a firm as

$$\begin{aligned}
 (C.59) \quad s_f &= \frac{l_{Fz}(x_{Fz}, v_z)}{l_{Fz}(x_{Fz}, v_z) + l_{Mz}(x_{Mz}, v_z)} \\
 (C.60) \quad &= \frac{1}{1 + \frac{\left(\frac{1}{\delta_{Mz} + \lambda_{Mz}^G + \lambda_{Mz}^E (1 - F_{Mz}(x_{Mz}))} \right)^2 (u_{Mz} + s_{Mz}^G) \lambda_{Mz}^U (\delta_{Mz} + \lambda_{Mz}^G + \lambda_{Mz}^E)}{\left(\frac{1}{\delta_{Fz} + \lambda_{Fz}^G + \lambda_{Fz}^E (1 - F_{Fz}(x_{Fz}))} \right)^2 (u_{Fz} + s_{Fz}^G) \lambda_{Fz}^U (\delta_{Fz} + \lambda_{Fz}^G + \lambda_{Fz}^E)}}.
 \end{aligned}$$

On the right-hand side, we can substitute our empirical estimates of the U-E transition rates λ_{gz}^u , the E-U transition rates δ_{gz} , the compulsory offer arrival rate λ_{gz}^G , the voluntary offer arrival rate λ_{gz}^e , and the compulsory offer search units s_{gz}^G from Table 4. Thus, we obtain:

$$\begin{aligned}
 (C.61) \quad & (u_{Mz} + s_{Mz}^G) \lambda_{Mz}^U (\delta_{Mz} + \lambda_{Mz}^G + \lambda_{Mz}^E) \\
 &= (0.236 + 0.101) \times 0.104 \times (0.036 + 0.010 + 0.009) \approx 0.0019
 \end{aligned}$$

$$\begin{aligned}
 (C.62) \quad & (u_{Fz} + s_{Fz}^G) \lambda_{Fz}^U (\delta_{Fz} + \lambda_{Fz}^G + \lambda_{Fz}^E) \\
 &= (0.219 + 0.083) \times 0.091 \times (0.028 + 0.007 + 0.007) \approx 0.0012.
 \end{aligned}$$

Thus, this ratio is approximately equal to $0.0019/0.0012 = 1.583$, and the expression simplifies to

$$\begin{aligned}
 (C.63) \quad s_f &= \frac{1}{1 + 1.583 \times \left(\frac{\delta_{Fz} + \lambda_{Fz}^G + \lambda_{Fz}^E (1 - F_{Fz}(x_{Fz}))}{\delta_{Mz} + \lambda_{Mz}^G + \lambda_{Mz}^E (1 - F_{Mz}(x_{Mz}))} \right)^2} \\
 &= \frac{1}{1 + 1.583 \times \left(\frac{0.036 + 0.007 \times (1 - F_{Fz}(x_{Fz}))}{0.046 + 0.009 \times (1 - F_{Mz}(x_{Mz}))} \right)^2}.
 \end{aligned}$$

Since firm sizes are monotonically increasing in flow utility x offered by the firm, we can obtain expressions for the minimum female employment share $\underline{s}_f$ and the maximum female employment share $\bar{s}_f$ by focusing on employers that are at the very top of the job ladder for one gender and simultaneously at the very bottom of the job ladder for the other gender. Specifically, among all dual-gender firms, the firm with the highest female employment share has $F_{Fz} = 1$ and $F_{Mz} = 0$. Conversely, the firm with the lowest female employment share has $F_{Fz} = 0$ and $F_{Mz} = 1$.

Thus, we find that the minimum (maximum) female employment share in the

model is ≈ 0.419 (0.596), which is inconsistent with the minimum (maximum) female employment share being close to 0 (1) in the data.

8. Model with Heterogeneous Separation Rates to Nonemployment

In this subsection, we describe how our model would change if we allowed separation rates to nonemployment to be heterogeneous across firms. In particular, we study a case in which separation rates are a non-decreasing function of composite productivity $\tilde{p}$. We retain all other assumptions regarding demographics, market structure, matching function, utility of workers, production, firm heterogeneity, cost functions, and utility-posting behavior.

In this case, we are still able to rank firms by $\tilde{p}$ and, therefore, we can still write the offer distribution as $F(x) = F(x^*(\tilde{p}))$. Thus, with a slight abuse of notation, denote $\delta(x) = \delta(\tilde{p} : x^*(\tilde{p}) = x)$ as the separation rate of a firm with composite productivity $\tilde{p}$ optimally offering flow utility x .⁵⁶ There are a few key changes to the equilibrium structure of our model. First of all, The steady-state nonemployment rate for each worker type is now calculated as

$$(C.64) \quad u_{gz} = \frac{\bar{\delta}_{gz}}{\bar{\delta}_{gz} + \lambda_{gz}^U + \lambda_{gz}^G}.$$

where $\bar{\delta}_g = \int_x \delta_g(x) dG(x)$ is the average employment-weighted separation rate across firms and $G(x)$ is the cross-sectional distribution of flow utilities of employed workers.

Second, the cross-sectional distribution of flow utilities for each worker type is now written as

$$(C.65) \quad G_{gz}(x) = \frac{F_{gz}(x)}{1 + \kappa_{gz}^E(x) [1 - F_{gz}(x)]},$$

where $\kappa_{gz}^E(x) \equiv \lambda_{gz}^E / (\int_{x' < x} \delta_g(x') dG(x') / G(x) + \lambda_{gz}^G)$ governs the effective speed of climbing the firm ladder in utility space, which depends on the average separation rate faced by workers earning utility below x .

Third, solving for the stationary employment distribution in each market (g, z) , firm sizes are now given by

$$(C.66) \quad l_{gz}(x, v) = \left(\frac{1}{\delta_g(x) + \lambda_{gz}^G + \lambda_{gz}^E [1 - F_{gz}(x)]} \right) \frac{v}{V_{gz}} \mu_{gz} \lambda_{gz}^U [u_{gz} + (1 - u_{gz}) s_g^E G_{gz}(x) + s_g^G]$$

⁵⁶In a more general model where δ is distributed arbitrarily across firms, workers would rank firms in two dimensions: flow utility x and job security $(1 - \delta)$, leading to a substantially more complicated ladder model which at present we are not able to solve in general equilibrium.

where this time we could not simplify $G_{gz}(x)$ due to it including integral terms that are x -dependent. Nevertheless, we can still conclude that firm size will increase in x due to fewer workers quitting to join other firms, and a larger pool of effective searchers moving in from other firms. Also, firm size is decreasing in $\delta_g(x)$.

FIRST-ORDER CONDITIONS.

To solve the firm's problem, we still take FOCs with respect to utility x and vacancies posted v , bearing in mind that $\delta_g(\tilde{p}_{gz})$ is not a choice variable:

$$\begin{aligned}
\text{(C.67)} \quad \frac{d\Pi_{gz}(x, v)}{dv} : \quad c_g^{v,0} \frac{dc_g^v(v)}{dv} &= (\tilde{p}_{gz} - x) \frac{1}{V_{gz}} \left(\frac{u_{gz} + (1 - u_{gz})s_g^E G_{gz}(x) + s_g^G}{\delta_g(\tilde{p}_{gz}) + \lambda_{gz}^G + \lambda_{gz}^E [1 - F_{gz}(x)]} \right) \mu_{gz} \lambda_{gz}^U \\
\frac{d\Pi_{gz}(x, v)}{dx} : \quad l(x, v) &= (\tilde{p}_{gz} - x) \lambda_g^E f_{gz}(x) \frac{v}{V_{gz}} \left(\frac{u_{gz} + (1 - u_{gz})s_g^E G_{gz}(x) + s_g^G}{[\delta_g(\tilde{p}_{gz}) + \lambda_{gz}^G + \lambda_{gz}^E [1 - F_{gz}(x)]]^2} \right) \mu_{gz} \lambda_{gz}^U \\
\text{(C.68)} \quad &+ (\tilde{p}_{gz} - x) \frac{v}{V_{gz}} \left(\frac{(1 - u_{gz})s_g^E g_{gz}(x)}{\delta_g(\tilde{p}_{gz}) + \lambda_{gz}^G + \lambda_{gz}^E [1 - F_{gz}(x)]} \right) \mu_{gz} \lambda_{gz}^U
\end{aligned}$$

Equations (C.67) and (C.68) are more general versions of equations (C.23) and (C.23), where we allow $\delta_g(\tilde{p}_{gz})$ to be a decreasing function of $\tilde{p}_{gz}$. Equation (C.68) can be further simplified as:

$$\text{(C.69)} \quad 1 = (\tilde{p}_{gz} - x) \left(\frac{\lambda_g^E f_g(x)}{\delta_g(x) + \lambda_{gz}^G + \lambda_{gz}^E [1 - F_{gz}(x)]} + \frac{(1 - u_{gz})s_g^E g_{gz}(x)}{u_{gz} + (1 - u_{gz})s_g^E G_{gz}(x) + s_g^G} \right)$$

D. IDENTIFICATION APPENDIX

1. Proof of Proposition 1 (Equilibrium Wage Equation)

RESTATEMENT OF PROPOSITION 1 (EQUILIBRIUM WAGE EQUATION).

The equilibrium wage of a worker of gender g and ability z at a firm with composite productivity $\tilde{p}_g$, preference weight β_g and preference-adjusted amenity cost shifter $\tilde{c}_g^{a,0}$ is

$$(D.1) \quad \ln w_{gz}(\tilde{p}_g, \beta_g, \tilde{c}_g^{a,0}) = \underbrace{\alpha_z}_{\text{"worker wage FE"}} + \underbrace{\psi_g^w(\tilde{p}_g, \beta_g, \tilde{c}_g^{a,0})}_{\text{"gender-firm wage FE"}},$$

where

$$(D.2) \quad \alpha_z = \ln z,$$

$$(D.3)$$

$$\psi_g^w(\tilde{p}_g, \beta_g, \tilde{c}_g^{a,0}) = \ln \left(\tilde{p}_g - \beta_g a_g^*(\tilde{c}_g^{a,0}) - \int_{\tilde{p}' \geq \phi_g}^{\tilde{p}_g} \left[\frac{1 + \kappa_g^E [1 - F_g(x_g^*(\tilde{p}_g))]}{1 + \kappa_g^E [1 - F_g(x_g^*(\tilde{p}'))]} \right]^2 d\tilde{p}' \right).$$

PROOF:

We proceed in two steps. First, we prove the proposition under exogenous firm-level vacancies that are constant within but may differ across genders. Second, we prove that the same result applies under endogenous vacancy posting.

STEP 1.

Suppose firms differ in their exogenous number of vacancies for each gender, $\{v_g\}_g$. Define $T_{gz} = \mu_{gz}[(u_{gz} + s_{gz}^G)\lambda_{gz}^U(\delta_{gz} + \lambda_{gz}^G + \lambda_{gz}^E)]/V_{gz}$. First of all, we guess (and later verify) that $\lambda_{gz}^U = \lambda_g^U$ for all ability types z , and therefore also $\lambda_{gz}^G = \lambda_g^G$ and $\lambda_{gz}^E = \lambda_g^E$. Thus, our assumptions also imply that $T_{gz} = T_g$ for all z . Second, under exogenous vacancies, a firm's type is defined by its composite productivity $\tilde{p}_{gz}$ and its exogenous vacancies v_g , which are constant across ability markets. As a consequence, $V_{gz} = V_g$ in all z -markets. Using equation (16), the firm's problem can therefore be written as

$$(D.4) \quad x_{gz}^*(\tilde{p}_{gz}) = \arg \max_x (\tilde{p}_{gz} - x) \left(\frac{1}{\delta_g + \lambda_g^G + \lambda_g^E(1 - F_{gz}(x))} \right)^2 v_g T_g.$$

Thus, given fixed vacancies, equilibrium firm profits in equation (18) can be written as

$$(D.5) \quad \Pi_{gz}(\tilde{p}_{gz}, v_g) = (\tilde{p}_{gz} - x^*(\tilde{p}_{gz})) \left(\frac{1}{\delta_g + \lambda_g^G + \lambda_g^E(1 - F_{gz}(x^*(\tilde{p}_{gz})))} \right)^2 v_g T_g.$$

We can write the offer distribution as

$$(D.6) \quad F_{gz}(x^*(\tilde{p}_{gz})) = h_{gz}(\tilde{p}_{gz}) = \frac{1}{V_{gz}} \int_{p' > \phi_{gz}}^{\tilde{p}_{gz}} \int v_g \gamma(p', v') dp' dv',$$

where $\gamma_{gz}(p', v')$ is the joint density function of $\tilde{p}_{gz}$ and v_g , and $h_{gz}(\tilde{p}_{gz})$ is the CDF of the marginal distribution of $\tilde{p}_{gz}$, in which values are weighted by vacancies posted by firms with each particular $\tilde{p}_{gz}$. The expression for h_{gz} is equivalent to equation (C.27) in Lemma 3, except that here we are integrating over exogenous rather than endogenous vacancies. Applying the Envelope Theorem yields

$$(D.7) \quad \frac{\partial \Pi_{gz}(\tilde{p}_{gz}, v_g)}{\partial \tilde{p}_{gz}} = \left(\frac{1}{\delta_g + \lambda_g^G + \lambda_g^E(1 - h_{gz}(\tilde{p}))} \right)^2 v_g T_g.$$

When $\tilde{p}_{gz} = \phi_{gz}$, $\Pi(\phi_{gz}, v_g) = 0$ for all v_g , which gives us a boundary condition to solve the differential equation for profits. Rearranging (D.5) and integrating equation (D.7) yields

$$(D.8) \quad x_{gz}(\tilde{p}_{gz}) = \tilde{p}_{gz} - \int_{y \geq \phi_{gz}}^{\tilde{p}_{gz}} \left[\frac{1 + \kappa_g^E(1 - h_{gz}(\tilde{p}_{gz}))}{1 + \kappa_g^E(1 - h_{gz}(y))} \right]^2 dy,$$

where $\kappa_g^E = \lambda_g^E / (\delta_g + \lambda_g^G)$. This equation parallels equation (47) in Burdett and Mortensen (1998), where composite productivity $\tilde{p}_{gz}$ in the current model plays the role of job productivity differentials in their model.

Lemma 1 already proves that amenities are proportional to ability z , therefore we can express them as $a_{gz} = z a_g$, and the cost of producing amenities as $\tilde{c}_{gz}^a(a_{gz}) = z \tilde{c}_g^a(a_g)$.

Summing up, it follows that $a_{gz} = z a_g$ and $\tilde{c}_{gz}^a(a_{gz}) = z \tilde{c}_g^a(a_g)$. Therefore, composite productivity $\tilde{p}_{gz} = (1 - \tau_g)pz + \beta_g a_{gz} - \tilde{c}_{gz}^a(a_{gz})$ is proportional to z , and we can write $\tilde{p}_{gz} = z \tilde{p}_g$, where $\tilde{p}_g = (1 - \tau_g)p + \beta_g a_g - \tilde{c}_g^a(a_g)$ is distributed according to $h_g(\tilde{p}_g)$. By definition, $h_{gz}(\tilde{p}_{gz}) = h_{gz}(z \tilde{p}_g) = h_g(\tilde{p}_g)$. Thus, with a change of variables and using the fact that vacancies of each firm are constant across ability markets, we can rewrite equation (D.8) as

$$(D.9) \quad x_g(z, \tilde{p}_g) = \tilde{p}_g z - \int_{y \geq \phi_{gz}}^{\tilde{p}_g} z \left[\frac{1 + \kappa_g^E(1 - h_g(\tilde{p}_g))}{1 + \kappa_g^E(1 - h_g(y))} \right]^2 dy.$$

We still need to prove that ϕ_{gz} is also proportional to z under the assumption that

$b_{gz} = zb_g$. We use a guess-and-verify approach: we guess that the case in which ϕ_{gz} and equilibrium flow utility $x(\tilde{p}_g, v_g)$ are proportional to z is an equilibrium of the model, and we verify it below. From equation (5), we have

$$(D.10) \quad \phi_{gz} = zb_g + (\lambda_g^U - \lambda_g^E) \int_{x' \geq \phi_{gz}} \frac{1 - F_{gz}(x')}{\rho + \delta_g + \lambda_g^G + \lambda_g^E(1 - F_{gz}(x'))} dx'.$$

We proceed to show that if $\phi_{gz} = z\phi_g$, then $x(\tilde{p}_g)$ is also proportional to z . The proof follows trivially from equation (D.9) since $\phi_{gz} = z\phi_g$ implies that

$$(D.11) \quad x_g(z, \tilde{p}_g) = z\tilde{p}_g - z \int_{y \geq \phi_g} \left[\frac{1 + \kappa_g^E(1 - h_g(\tilde{p}_g))}{1 + \kappa_g^E(1 - h_g(y))} \right]^2 dy.$$

Next we show that if $x(z, \tilde{p}_g)$ is proportional to z , then ϕ_{gz} must be proportional to z . Consider the bijective mapping $\tilde{p}_g(x, z) = [x^*(z, \tilde{p}_g)]^{-1}$. We can rewrite the outside option as

$$(D.12) \quad \phi_{gz} = zb_g + (\lambda_g^U - \lambda_g^E) \int_{x' \geq \phi_{gz}} \frac{1 - h_{gz}([x'(z, \tilde{p}_g)]^{-1})}{\rho + \delta_g + \lambda_g^G + \lambda_g^E(1 - h_{gz}([x'(z, \tilde{p}_g)]^{-1}))} dx'$$

$$(D.13) \quad = zb_g + z(\lambda_g^U - \lambda_g^E) \int_{x' \geq \phi_{gz}} \frac{1 - h_g([x'(1, \tilde{p}_g)]^{-1})}{\rho + \delta_g + \lambda_g^G + \lambda_g^E(1 - h_g([x'(1, \tilde{p}_g)]^{-1}))} dx',$$

which implies that the only solution to this equation satisfies $\phi_{gz} = z\phi_g$.

Finally, recalling that $\tilde{p}_g = (1 - \tau)p + a_g - c_g^a(a_g)$ and that $w = x - \beta a$, we can write monetary wages as

$$(D.14) \quad w(z, \tilde{p}_g, \beta_g, \tilde{c}_g^{a,0}) = z \left[\tilde{p}_g - \beta_g a_g(\tilde{c}_g^{a,0}) - \int_{\tilde{p}' \geq \phi_g} \left[\frac{1 + \kappa_g^E(1 - h_g(\tilde{p}_g))}{1 + \kappa_g^E(1 - h_g(\tilde{p}'))} \right]^2 d\tilde{p}' \right],$$

which completes the proof that the desired equilibrium wage equation holds under exogenous vacancies that are constant across ability levels.

STEP 2.

All that remains to be shown for the desired result to follow is that in the model with endogenous vacancy posting, we have $v_{gz}^* = v_g^*$ for all z , so that the offer distribution h_{gz} is the same across all ability markets. We follow a guess-and-verify approach. Suppose that $x_{gz}^*(\tilde{p}_g)$ is proportional to ability z . Using that $F_{gz}(x_{gz}^*(\tilde{p}_{gz})) = h_{gz}(\tilde{p}_{gz})$, we can write the FOC for vacancy creation in equation (C.23) as

$$(D.15) \quad z c_{g,0}^{v,0} \frac{\partial \tilde{c}_g^v(v_{gz})}{\partial v_{gz}} = T_{gz}(\tilde{p}_{gz} - x_{gz}^*(\tilde{p}_{gz})) \left(\frac{1}{\delta_{gz} + \lambda_{gz}^G + \lambda_{gz}^E(1 - h_{gz}(\tilde{p}_{gz}))} \right)^2$$

$$(D.16) \quad c_{g,0}^{v,0} \frac{\partial \tilde{c}_g^v(v_{gz})}{\partial v_{gz}} = T_g(\tilde{p}_g - x_g^*(\tilde{p}_g)) \left(\frac{1}{\delta_g + \lambda_g^G + \lambda_g^E(1 - h_g(\tilde{p}_g))} \right)^2,$$

immediately proving that $v_{gz} = v_g$ for all z . Equation (12) thus implies that aggregate vacancies satisfy $V_{gz} = V_g$, which also implies that in equilibrium, $\lambda_{gz}^U = \lambda_g^U$ (which we previously guessed) and $u_{gz} = u_g$. As a consequence, all terms in the wage equation (D.14) scale linearly in ability. Therefore, log wages take the form of the desired equilibrium wage equation.

2. Equilibrium Amenity Equation

An analogous result shows that equilibrium amenities in this environment also have a log-additive structure, akin to the treatment of wages by the variant of the AKM framework introduced by [Card, Cardoso and Kline \(2016\)](#).

COROLLARY 1 (Equilibrium Amenity Equation): *The equilibrium amenity of a worker of gender g and ability z at a firm with preference weight β_g and preference-adjusted amenity cost shifter $\tilde{c}_g^{a,0}$ is*

$$(D.17) \quad \ln \beta_g a_{gz} (\tilde{c}_g^{a,0}) = \underbrace{\alpha_z}_{\text{"worker amenity FE"}} + \underbrace{\psi^a(\beta_g, \tilde{c}_g^{a,0})}_{\text{"gender-firm amenity FE"}},$$

where

$$(D.18) \quad \alpha_z = \ln z,$$

$$(D.19) \quad \tilde{c}_g^{a,0} = \frac{c_g^{a,0}}{\beta_g}$$

$$(D.20) \quad \psi^a(\beta_g, \tilde{c}_g^{a,0}) = \ln \beta_g + \frac{1}{1 - \eta^a} \ln \tilde{c}_g^{a,0}.$$

PROOF:

This result follows directly from equation (C.5) in the proof of Lemma 1 in Section C.2.

Corollary 1 shows that amenities, like wages, follow a specification that is log-additive between worker heterogeneity (α_z) and gender-specific firm heterogeneity (ψ^a). The worker amenity FE α_z is a strictly monotonic transformation of worker ability. The gender-firm amenity FE $\psi^a(\beta_g, \tilde{c}_g^{a,0})$ depends only on a firm's preference weight β_g and amenity cost shifter $c_g^{a,0}$ for each gender, scaled by a function of the economy-wide amenity cost elasticity η^a . However, an explicit treatment of amenities and compensating differentials was missing from the analysis in [Card, Cardoso and Kline \(2016\)](#). Our framework fills this gap by explicitly modeling firms' equilibrium wage and amenity choices.

3. Normalization of Gender-Specific Firm Pay

GENERAL MODEL WITH ENDOGENOUS AMENITIES.

Here, we discuss the normalization of gender-specific firm pay in our baseline model with endogenous amenities. Recall that we estimate gender-specific firm pay components by applying a two-way fixed effect model à la AKM separately for each gender:

$$(D.21) \quad \ln(w_{ijt}) = \alpha_i + \psi_{G(i)j} + X_{it}\beta_{G(i)} + \varepsilon_{ijt}.$$

From here on, with a slight abuse of notation and to be consistent with the notation in our structural model, we write $w_{gr} \equiv \exp(\psi_{gJ_g(r)})$ to denote the level pay of firm $j = J_g(r)$ at rank $r \in [0, 1]$ for gender g . As explained in [Abowd, Creedy and Kramarz \(2002\)](#) and [Card, Cardoso and Kline \(2016\)](#), equation (D.21) identifies the gender-specific firm fixed effects up to a constant within each connected set—i.e., one normalization must be imposed on a reference firm for each connected set. Because the connected sets for men and women are disconnected by construction, comparing firm pay across genders requires an additional normalization on the level of firm fixed effects for men compared to women. The baseline assumption in [Card, Cardoso and Kline \(2016\)](#) is that firms in a lower range of the value added per worker distribution have firm fixed effects equal to zero for both genders. [Card, Cardoso and Kline \(2016\)](#) show that similar results are obtained when imposing an alternative normalization that sets mean firm fixed effects equal to zero in the hotel and restaurant sector for both genders.

Through the lens of our equilibrium model, any normalization of gender-specific firm fixed effects must take into account the fact that amenities differ between firms for a given gender as well as within a firm across genders, in addition to the existence of firm-specific gender wedges. Thus, we derive a model-consistent normalization to the gender-specific estimates of fixed effects in equation (D.21).

Building on the equilibrium wage equation (19), we can derive a model-consistent normalization of gender-firm FEs that allows us to compare firm pay between men and women. Intuitively, our model suggests that such a normalization requires finding a set of firms j with the same $\tilde{p}_g$, β_g and $\tilde{c}_g^{a,0}$ across genders $g \in \{M, F\}$ so that $\psi_M(\tilde{p}_M, \beta_M, \tilde{c}_M^{a,0}) = \psi_F(\tilde{p}_F, \beta_F, \tilde{c}_F^{a,0})$. An advantage of our model is that it provides us guidance on how to find such a set of firms. To this end, let $\mathcal{B}_g \equiv [0, \hat{r}_g] \subseteq [0, 1]$ for some strictly positive but small $\hat{r}_g \approx 0$ be a set of firms with utility ranks near the bottom for gender g . Let $\mathcal{D}_g \subseteq [0, 1]$ be a set of firms with $\tau_{F_r} \approx 0$. Let $\mathcal{A}_g \subseteq [0, 1]$ be a set of dual-gender firms with $\beta_{gr}a_{gr} \approx \beta_{-gR_{-g}(J(r))}a_{-gR_{-g}(J(r))}$. Then the following result provides us a model-consistent normalization of gender-firm FEs:

PROPOSITION 6 (Firm Pay Normalization): *Firm pay ψ_{gj} can be equated across genders $g \in \{M, F\}$ for a set of firms j with rank $R_g(j) \in \mathcal{B}_g \cap \mathcal{D}_g \cap \mathcal{A}_g$ simulta-*

neously for both genders g :

$$(D.22) \quad \mathbb{E}_j[\psi_{Mj}] = \mathbb{E}_j[\psi_{Fj}] \quad \forall j : R_g(j) \in \mathcal{B}_g \cap \mathcal{D}_g \cap \mathcal{A}_g, \forall g \in \{M, F\}.$$

PROOF:

Recall the definition of composite productivity of a firm at rank r for gender g ,

$$(D.23) \quad \tilde{p}_{gr} \equiv (1 - \tau_{gr})p_r + \beta_{gr}a_{gr} - \tilde{c}_{gr}^a(a_{gr}).$$

Our model-consistent normalization applies to the intersection of three sets of firms.

First, we consider a set of firms with close-to-zero profits. Let $\mathcal{B}_g \equiv [0, \hat{r}_g] \subseteq [0, 1]$ for some strictly positive but small $\hat{r}_g \approx 0$ be a set of firms with utility ranks near the bottom for gender g . Bottom-ranked firms with $r \in \mathcal{B}_g$ provide worker utility $x_{gr} \equiv w_{gr} + \beta_{gr}a_{gr}$ approximately equal to their composite productivity $\tilde{p}_{gr}$ under the assumption that the lower bound of the support of (p, τ_g) extends low enough, which can always be guaranteed:

$$(D.24) \quad \tilde{p}_{gr} \approx x_{gr} \quad \forall r \in \mathcal{B}_g.$$

Second, we consider a set of firms that treat workers of both genders interchangeably in production. Let $\mathcal{D}_g \subseteq [0, 1]$ be a set of firms with $\tau_{Fr} \approx 0$. Also, recall that $\tau_{Mr} = 0$ by assumption. For those firms, we have

$$(D.25) \quad \tilde{p}_{gr} \approx p_r + \beta_{gr}a_{gr} - \tilde{c}_g^a(a_{gr}) \quad \forall r \in \mathcal{D}_g,$$

so the only gender-specific aspect of firms $r \in \mathcal{D}_g$ is $\beta_{gr}a_{gr} - \tilde{c}_g^a(a_{gr})$. Combining equations (D.24) and (D.25), we have $p_r + \beta_{gr}a_{gr} - \tilde{c}_g^a(a_{gr}) \approx w_{gr} + \beta_{gr}a_{gr}$ for $r \in \mathcal{B}_g \cap \mathcal{D}$ for both genders.

Third, we consider a set of firms that provide the same level of amenities to both genders. Let $\mathcal{A}_g \subseteq [0, 1]$ be a set of dual-gender firms with $\beta_{gr}a_{gr} \approx \beta_{-gR_{-g}(J(r))}a_{-gR_{-g}(J(r))}$. In words, the amenities enjoyed by gender g are enjoyed to the same level by gender $-g$ (e.g., both gender-specific amenities equal zero), where $-g \equiv \{M, F\} \setminus \{g\}$.

Due to our model result that a firm's optimal amenity choice a^* is strictly decreasing in the preference-adjusted amenity cost parameter $\tilde{c}^{a,0}$, a firm optimally provides the same level of amenities to both genders if and only if it faces the same cost function / preference combination for both genders. As a result,

$$(D.26) \quad \beta_{gr}a_{gr} - \tilde{c}_g^a(a_{gr}) = \beta_{-gR_{-g}(J(r))}a_{-gR_{-g}(J(r))} - \tilde{c}_{-g}^a(a_{-gR_{-g}(J(r))}) \quad \forall r \in \mathcal{A}_g,$$

so the only gender-specific aspect of firms $r \in \mathcal{A}_g$ is τ_{gr} .

Combining the above insights, the definition of $\tilde{p}_{gr}$ in equation (D.23) yields that $w_{MR_M(j)} = w_{FR_F(j)}$ for any firm j with rank $R_g(j) \in \mathcal{B}_g \cap \mathcal{D}_g \cap \mathcal{A}_g$ for both genders g at the same time. In words, for a subset of firms that simultaneously (i) are located near the bottom of both genders' utility ranks, (ii) treat men and

women as perfect substitutes in production, and (iii) provide the same amenity level to men and women, we can equalize the gender-specific AKM firm fixed effects in equation (D.21) across men and women according to equation (D.22).

In words, Proposition 6 states that we can equalize the gender-specific AKM firm fixed effects in equation (19) across men and women for a subset of firms that simultaneously (i) are located near the bottom of both genders' utility ranks, (ii) treat men and women as perfect substitutes in production, and (iii) provide a fixed amenity level to men and women.

The normalization in equation (D.22) embeds two untestable assumptions, namely that (i) men and women are treated as perfect substitutes (i.e., $\tau_{Fr} \approx 0$) at firms with rank $r \in \mathcal{D}_F$ for women, and (ii) the same amenity value is provided to both men and women (i.e., $\beta_{MR_M(j)} a_{MR_M(j)} \approx \beta_{FR_F(j)} a_{FR_F(j)}$) at firms j such that $R_g(j) \in \mathcal{A}_g$ for both genders g . Note that the fact that we are looking for firms near the bottom of both genders' utility ranks (i.e., $R_g(j) \in \mathcal{B}_g$ for both genders g) is a verifiable condition given our revealed-preference measures of gender-specific firm ranks.

To summarize, while the original normalization of gender-firm FEs based on value added per worker proposed by Card, Cardoso and Kline (2016) is not valid in our environment, equation (D.22) provides a normalization that extends the argument in Card, Cardoso and Kline (2016) to our environment with gender-specific amenities and compensating differentials.

SPECIAL CASE OF THE MODEL WITH EXOGENOUS AMENITIES.

Now suppose that gender-specific firm amenities a_{gr} are exogenous. Then $c_g^a(a_{gr}) = 0$ without loss of generality. In this case, equation (D.24) reduces to

$$(D.27) \quad (1 - \tau_{gr})p_r - w_{gr} \approx 0 \quad \forall r \in \mathcal{B}.$$

Therefore, if we can find a subset of firms in $\mathcal{B} \cap \mathcal{D}$ that simultaneously (i) are located near the bottom of the utility ranks for both genders (i.e., $r \in \mathcal{B}$) and (ii) treat men and women as perfect substitutes in production net of employers' taste for gender (i.e., $r \in \mathcal{D}$), then equalizing firm pay across genders for this set of firms constitutes a normalization of the gender-specific AKM firm fixed effects in equation (D.21) that is consistent with our model featuring exogenous amenities:

$$(D.28) \quad \mathbb{E}_r[\psi_{Fr}] = \mathbb{E}_r[\psi_{Mr}] \quad \forall r \in \mathcal{B} \cap \mathcal{D}.$$

The normalization in equation (D.28) embeds only one untestable assumption, namely that men and women are treated as perfect substitutes (i.e., $\tau_r \approx 0$) at firms with rank $r \in \mathcal{D}$. Note that the same normalization as imposed in the case with endogenous amenities, discussed in Section D.3 is still valid in this case.

4. Proof of Proposition 2 (Employer Ranks)

RESTATEMENT OF PROPOSITION 2 (EMPLOYER RANKS)

All workers of the same gender share a common employer ranking $r \in [0, 1]$, which is identified given employer sizes $l(r)$. PROOF:

First, note that Propositions 1 and Corollary 1 together imply that the flow consumption that a worker of gender g and ability z receives at firm j can be written as

$$(D.29) \quad x_{gz}(j) = w_{gz}(j) + \beta_g(j)a_{gz}(j)$$

$$(D.30) \quad = \exp(\alpha_z + \psi_g^w(j)) + \exp(\alpha_z + \psi^a(j))$$

$$(D.31) \quad = \exp(\alpha_z) \exp(\psi_g^w(j)) + \exp(\alpha_z) \exp(\psi^a(j))$$

$$(D.32) \quad = \exp(\alpha_z) [\exp(\psi_g^w(j)) + \exp(\psi^a(j))]$$

$$(D.33) \quad = z [\exp(\psi_g^w(j)) + \exp(\psi^a(j))]$$

Fixing gender g , equation (D.33) is linear in worker ability z . In particular, the firm-specific term, $\exp(\psi_g^w(j)) + \exp(\psi^a(j))$, is independent of worker ability z . Therefore, in the eyes of a worker of type (g, z) , a firm j is higher-ranked than another firm j' if and only if it provides higher flow utility $x_{gz}(j) > x_{gz}(j')$ if and only if $\exp(\psi_g^w(j)) + \exp(\psi^a(j)) > \exp(\psi_g^w(j')) + \exp(\psi^a(j'))$. This proves that all workers of the same gender share a common ranking of all firms in the economy. Therefore, for the remainder of the proof we drop the ability subscript z , and we focus on one gender at a time, thus dropping the gender subscript g too for readability.

As we have proved in Lemma 3 that utility $x(\tilde{p})$ is increasing in composite productivity $\tilde{p}$, and that higher utility x means that an employer is higher-ranked, it follows that employers with higher $\tilde{p}$ are higher-ranked. Therefore, higher-ranked employers:

1. Post higher utility: $x(r') > x(r)$ whenever $r' > r$, from Lemma 3;
2. Retain more workers: by definition $F(x(r')) > F(x(r))$ whenever $x(r') > x(r)$;
3. Post more vacancies: from Lemma 2, $v(r') > v(r)$ whenever $r' > r$;
4. Have larger size: as a result of posting more vacancies and offering higher utility, $l(r') > l(r)$ whenever $r' > r$.

These results demonstrate that firm sizes $l(r)$ are monotonically increasing in a firm's rank, and therefore we can recover firm ranks r by ordering firms by size, thereby proving the proposition.

5. Proof of Proposition 3 (Labor Market Objects)

RESTATEMENT OF PROPOSITION 3 (LABOR MARKET OBJECTS)

Gender-firm-specific recruiting intensities $f(r)$ and vacancies $v(r)$ as well as gender-specific separation hazards δ , job offer hazards from nonemployment λ^U , involuntary job offer hazards λ^G , voluntary on-the-job offer hazards λ^E , and aggregate vacancies V are identified given employer ranks and data on worker flows between employment states. PROOF:

We proceed in steps, with each step linking one empirical object to one model object (i.e., either a model parameter or an equilibrium outcome of interest).

Empirical Firm Nonemployment Hiring Shares $\leftrightarrow$ Model Recruiting Intensities. We obtain the empirical gender-specific distribution of recruiting intensities $f(r)$ by inverting the gender-specific distribution of employment G_g across ranks by solving equation (7) for the offer distribution F_g . We then obtain the hiring distribution as the change in F_g across ranks.

This empirical object directly corresponds to firms' recruiting intensities v_r/V in the model, where v_r denotes the vacancies posted by firm r and $V \equiv \int_r v_r dr$ denotes the total number of vacancies in the economy.⁵⁷

Empirical Rate of Moving into Nonemployment $\leftrightarrow$ Model Rate of Exogenous Separations. We identify δ off rates of workers i moving into nonemployment:

$$(D.34) \quad \hat{\delta} = \mathbb{E}_i \mathbf{1} [\text{nonemployed}_{i,t+1} | \text{employed}_{i,t}].$$

Empirical Job Finding Rate $\leftrightarrow$ Model Rate of Job Offers from Nonemployment. A simple log-hazard model of worker E - N - E transitions, where E denotes employment at some firm and N denotes nonemployment, can be used to recover $\lambda^U + \lambda^G$ separately by gender.⁵⁸

Empirical Share of Moves Down the Firm Ranks $\leftrightarrow$ Model Rate of Involuntary Job Offer Shocks. Two insights allow us to use information on worker transitions between employers to identify λ^G . First, we focus on transitions in rank space, not pay space. Second, the share of rank-increasing transitions due to involuntary on-the-job offers declines in F_r . Formally, the total number of job-to-job transitions from employer rank r is

$$(D.35) \quad J2J_r = l_r [\lambda^E (1 - F^r) + \lambda^G],$$

where l_r is the gender-specific number of workers at firm r . Rearranging and

⁵⁷Since $f_r = v_r/V$ refers to shares (i.e., not levels), the mass of aggregate vacancies V does not matter for its computation. At the end of this section, we spell out how to deduce the mass of aggregate vacancies V , which will be useful later on.

⁵⁸Since in our model, the job-finding rate from nonemployment is $\lambda^U + \lambda^G$, we use our estimate $\hat{\lambda}^G$ below to obtain $\hat{\lambda}^U$.

averaging across all firms, we have

$$(D.36) \quad \hat{\lambda}^G = \mathbb{E}_r \left[\frac{J2J_{r\downarrow}}{l_r F_r} \right],$$

where $J2J_{r\downarrow} = J2J_r - l_r(\lambda^E + \lambda^G)(1 - F_r)$ is the number of job-to-job transitions to lower ranks. Based on this, we derive the parameter estimate $\hat{s}^G \equiv \hat{\lambda}^G / \hat{\lambda}^U$.

Empirical Share of Moves up the Firm Ranks $\leftrightarrow$ Model Rate of Voluntary On-the-Job Offers. On-the-job offers not associated with involuntary transitions must have been voluntary. Hence, once we know $\hat{\lambda}^G$, we can use equation (D.35) to estimate λ^E as

$$(D.37) \quad \hat{\lambda}^E = \frac{J2J_r/n_r - \hat{\lambda}^G}{1 - F_r}.$$

Notice that all of these parameters are over-identified, as in principle, we could use just a fraction of the firms and job-to-job moves to identify them. We choose to use the overall sample average of these two moments. Based on this, we derive the parameter estimate $\hat{s}^E \equiv \hat{\lambda}^E / \hat{\lambda}^U$.

Aggregate Vacancies. Given estimates of δ , λ^U , λ^E , and λ^G , in addition to the relative mass of workers of a given gender, μ , we are equipped to deduce aggregate vacancies V . To this end, recall that a worker's job-finding probability due to the aggregate matching function is

$$(D.38) \quad \frac{m}{u} = \lambda^U + \lambda^G = \chi [\mu (u + s^E(1 - u) + s^G)]^\alpha V^{1-\alpha},$$

where m is the number of matches and $u = \delta / (\delta + \lambda^U + \lambda^G)$ is the nonemployment rate. Given the normalization $\chi = 1$, we can solve equation (D.38) for aggregate vacancies V :⁵⁹

$$(D.39) \quad V = \left(\frac{\lambda^U + \lambda^G}{[\mu (u + s^E(1 - u) + s^G)]^\alpha} \right)^{\frac{1}{1-\alpha}}.$$

6. Proof of Proposition 4 (Firm-Level Parameters)

RESTATEMENT OF PROPOSITION 4 (FIRM-LEVEL PARAMETERS)

The following gender-firm-specific parameters as functions of r are point identified given employer ranks and labor market objects: productivity $p(r)$, the gender wedge $\tau(r)$, the amenity valuation $\beta(r)a^*(r)$, and amenity costs $c^a(a^*(r))$. PROOF:

Here, we present an identification result based on continuous firm types, as in the model developed in Section III and further discussed in Section IV. In

⁵⁹As discussed in the main text, this normalization is inconsequential for our purposes.

Supplemental Appendix D.7, we extend this argument to the case of discrete firm types, as in the data.

First, by the results of Proposition 1 and Corollary 1, we abstract from heterogeneity in ability and replace $z = 1$ in all functional forms without loss of generality. Therefore, the vacancy cost function can be written as $c^v(v) = c^{v,0}v^{\eta^v}/\eta^v$ and the amenity cost function can be written as $c^a(a) = c^{a,0}a^{\eta^a}/\eta^a$.

Recall that the composite productivity of firm r is given by $\tilde{p}(r) = (1 - \tau(r))p(r) + \beta(r)a(r) - c^a(a(r); r)$. Let

$$(D.40) \quad T \equiv \frac{\mu[(u + s^G)\lambda^u(\delta + \lambda^G + \lambda^E)]}{V}$$

be a constant which only depends on previously estimated labor market rates and aggregate vacancies. We start from the firm's FOCs for optimality in equations (C.23) and (C.24). We substitute the functional form of $c^v(v)$ in equation (C.23) and express the FOCs as a coupled pair of differential equations:

$$(D.41) \quad h'(\tilde{p}(r)) = \frac{1}{V} \left[\frac{T(\tilde{p}(r) - x(\tilde{p}(r)))}{c^{v,0}[\delta + \lambda^G + \lambda^E(1 - h(\tilde{p}(r)))]^2} \right]^{\frac{1}{\eta^v - 1}} \gamma(\tilde{p}(r)),$$

$$(D.42) \quad x'(\tilde{p}(r)) = \frac{1}{V} \frac{2\lambda^E(\tilde{p}(r) - x(\tilde{p}(r)))}{\delta + \lambda^G + \lambda^E(1 - h(\tilde{p}(r)))} \left[\frac{T(\tilde{p}(r) - x(\tilde{p}(r)))}{c^{v,0}[\delta + \lambda^G + \lambda^E(1 - h(\tilde{p}(r)))]^2} \right]^{\frac{1}{\eta^v - 1}} \gamma(\tilde{p}(r)).$$

We first derived the differential equation (D.42) as equation (C.28) in Supplemental Appendix C.4, and here we substituted therein the explicit solution for vacancies in equation (D.41). In equations (D.41)–(D.42) above, $h(\tilde{p}(r)) \equiv F(x(\tilde{p}(r))) = F(r)$ is the CDF of *composite productivity offers* (i.e., weighted by firms' vacancies) and $h'(\tilde{p}(r)) = \partial h(\tilde{p}(r))/\partial \tilde{p}$ is the derivative thereof, while $\Gamma(\tilde{p}(r))$ and $\gamma(\tilde{p}(r)) \equiv \Gamma'(\tilde{p}(r))$ are the CDF of *composite productivity* and its derivative. Note that we can obtain $h(\tilde{p}(r))$ by integrating the hiring density $f(r)$ over observed ranks r , since $F(r) = \int_{r'=0}^r f(r') dr'$.

Next, we perform a change of variables using the fact that, by definition,

$$(D.43) \quad \Gamma(\tilde{p}(r)) = \int_{\tilde{p}'=\tilde{p}}^{\tilde{p}(r)} \gamma(\tilde{p}') d\tilde{p}' = \int_{r'=0}^r 1 dr' = r,$$

That is, the share of firms with composite productivity up to $\tilde{p}(r)$ equals the share of firms with rank up to r . Therefore, $\gamma(\tilde{p}(r)) d\tilde{p}/dr = 1 \forall r$. Equivalently, $h'(\tilde{p}(r)) d\tilde{p}/dr = f(r)$. Using the definition that $F(r) = h(\tilde{p}(r))$, we can rewrite equation (D.41) to obtain an expression for the firm's *flow payoff per matched worker*,

$$(D.44) \quad P(r) \equiv \tilde{p}(r) - x(\tilde{p}(r)) = [f(r)V]^{\eta^v - 1} \frac{c^{v,0}}{T} [\delta + \lambda^G + \lambda^E(1 - F(r))]^2,$$

In equation (D.44), $f(r)$, $c^{v,0}$, and $F(r)$ are all known quantities identified by

applying Proposition 3 to the data. Thus, equation (D.44) can be rearranged to yield an explicit expression for $P(r)$ as a function of known objects. Finally, we perform a similar change of variables to write the derivative of utility $x'(r)$ as a function of ranks:

$$(D.45) \quad x'(r) = \frac{1}{V} \frac{2\lambda^E P(r)}{\delta + \lambda^G + \lambda^E(1 - F(r))} \left[\frac{TP(r)}{c^{v,0} [\delta + \lambda^G + \lambda^E(1 - F(r))]^2} \right]^{\frac{1}{\eta^v - 1}}.$$

Plugging $P(r)$ from equation (D.44) into this expression allows us to identify flow utility across ranks, $x(r) = K + \int_{r'=0}^r x'(r') dr'$, up to a constant of integration K . Intuitively, the model helps us pin down differences in utilities between rungs of the job ladder. We choose a value for the constant of integration K such that the equilibrium distribution of amenities—derived in equation (D.46) below—attains a lower bound strictly above but arbitrarily close to zero.⁶⁰

This allows us to identify gender-firm-specific amenity valuations, $\beta(r)a(r)$, by simply taking the difference between utility and wages at each firm, and to identify productivity net of the gender wedge, $(1 - \tau(r))p(r)$, by adding wages and the cost of amenities to the firm's flow payoff per matched worker:

$$(D.46) \quad \beta(r)a(r) = x(r) - w(r),$$

$$(D.47) \quad (1 - \tau(r))p(r) = P(r) + w(r) + c^a(a(r); r).$$

On the right-hand side of equation (D.46), $x(r)$ is known from integrating equation (D.45) and $w(r)$ is known from the data, so we can infer $\beta(r)a(r)$. On the right-hand side of equation (D.47), $P(r)$ is known from equation (D.44), $w(r)$ is known from the data, and $c^a(a(r); r) = \beta(r)a(r)/\eta^a$ is known from equation (C.9) of Lemma 1, with the known value for $\beta(r)a(r)$ substituted from equation (D.46).

If, in addition, we fix the preference weights $\beta(r)$ at a known value, say $\beta(r) = 1$, then we can obtain gender- and firm-specific preference-adjusted amenity cost shifters by inverting the firm's FOC for optimal amenity creation—see Lemma 1 and the functional form in equation (8)—which yields

$$(D.48) \quad \tilde{c}^{a,0}(r) = [a(r)]^{1-\eta^a}.$$

Regardless of whether or not the preference weights $\beta(r)$ are known, after having identified productivity $p(r)$ for men, under the normalization that $\tau(r) = 0$ for them, and $(1 - \tau(r))p(r)$ for women from equation (D.47), we recover the gender

⁶⁰Our choice of amenities starting strictly above but arbitrarily close to zero seems natural given the interpretation of amenities being endogenously produced by firms. It can be rationalized through an assumption that some employers put a vanishingly small weight on the (unmeasured by the tax authorities) amenity-related welfare of their workforce. At the same time, this choice minimizes the share of amenity valuations in total compensation across firms and maximizes the remaining gender gap in log total compensation. In contrast, if we modeled amenities as exogenous firm characteristics, then the choice of the constant of integration K would be inconsequential for our analysis as the absolute level of amenities would not be pinned down in that case.

wedge $\tau(r)$ based on the ratio of estimated productivities net of the gender wedge at dual-gender firms. It is worth noting that, for this last step, the exact functional form by which $\tau(r)$ enters firms' payoff function (e.g., multiplicatively or additively) is immaterial for identification and estimation.

7. Identification of Firm-Level Parameters With Discrete Firm Types

Here, we adapt the identification proof in Section D.6 to discrete data where we observe a finite number of firms N . We observe their ranks, defined as $r \in \{1/N, 2/N, \dots, 1\}$, their empirical hiring intensities f_r and their wage w_r .⁶¹ The FOCs read:

$$(D.49) \quad h'(\tilde{p}_r) = \frac{1}{V} \left[\frac{T(\tilde{p}_r - x(\tilde{p}_r))}{c^{v,0}} \left(\frac{1}{\delta + \lambda^G + \lambda^E(1 - h(\tilde{p}_r))} \right)^2 \right]^{\frac{1}{\eta^{v-1}}} \gamma(\tilde{p}_r),$$

$$(D.50) \quad x'(\tilde{p}_r) = \frac{1}{V} \frac{2\lambda^E(\tilde{p}_r - x(\tilde{p}_r))}{\delta + \lambda^G + \lambda^E(1 - h(\tilde{p}_r))} \left[\frac{T(\tilde{p}_r - x(\tilde{p}_r))}{c^{v,0}} \left(\frac{1}{\delta + \lambda^G + \lambda^E(1 - h(\tilde{p}_r))} \right)^2 \right]^{\frac{1}{\eta^{v-1}}} \gamma(\tilde{p}_r).$$

In what follows, we want to move from model objects as functions of continuous firm ranks r (e.g., $\tilde{p}_r$) to empirical objects as functions of discrete firm ranks, r . For example, the change in the CDF of recruiting intensities between two adjacent firm ranks $r - 1$ and r is

$$(D.51) \quad \Delta h(\tilde{p}_r) \equiv \int_{\tilde{p}_{r-1}}^{\tilde{p}_r} \frac{v(\tilde{p})}{V} \gamma(\tilde{p}) d\tilde{p}$$

$$(D.52) \quad \approx \frac{v(\tilde{p}_{r-1})}{V} \gamma(\tilde{p}_{r-1}) \times [\tilde{p}_r - \tilde{p}_{r-1}],$$

where Δ is an operator that takes differences between the current and the previous (discrete) rank. We write a discretized version of equation (D.43) as follows:

$$(D.53) \quad \Gamma(\tilde{p}_r) = \int_{\tilde{p}_1}^{\tilde{p}_r} \gamma(\tilde{p}) d\tilde{p}$$

and interpreting the empirical (discrete) distribution of firms as representative of the theoretical (continuous) distribution of firms, the CDF of composite productivity $\tilde{p}$ of the first n firms in the ranking is simply⁶²

$$(D.54) \quad \Gamma(\tilde{p}_r) = r \in \left\{ \frac{1}{N}, \frac{2}{N}, \dots, 1 \right\}.$$

⁶¹Instead of functional notation, here we denote firm-specific objects by subscript r to highlight that there is a discrete number of firms in the data.

⁶²Recall that $r \in \{0, 1/(N-1), \dots, 1\}$, where N is the total number of firms, so $n \equiv rN \in \{1, 2, \dots, N\}$ is the rescaled rank of a firm, with workers of a given gender preferring firms with higher values of n .

Therefore, a good approximation for the change in the CDF of composite productivity is simply

$$(D.55) \quad \gamma(\tilde{p}_{r-1})(\tilde{p}_r - \tilde{p}_{r-1}) \approx 1/N$$

at all rungs of the ladder.⁶³ Now we rewrite the change in the CDF of composite productivity offers using a finite-difference approximation and equation (D.49):

$$(D.56) \quad \Delta F(x(\tilde{p}_r)) = h'(\tilde{p}_{r-1})(\tilde{p}_r - \tilde{p}_{r-1})$$

$$(D.57) \quad = \frac{1}{V} \left[\frac{T(\tilde{p}_{r-1} - x(\tilde{p}_{r-1}))}{c^{v,0}} \left(\frac{1}{\delta + \lambda^G + \lambda^E(1 - h(\tilde{p}_{r-1}))} \right)^2 \right]^{\frac{1}{\eta^{v-1}}} \gamma(\tilde{p}_{r-1})(\tilde{p}_r - \tilde{p}_{r-1})$$

Now replace $\Delta F(x(\tilde{p}_r))$ by $\hat{f}_r$, replace $\gamma(\tilde{p}_{r-1})(\tilde{p}_r - \tilde{p}_{r-1})$ by $1/N$, and replace $h(\tilde{p}_{r-1})$ by F_{r-1} . It then follows that

$$(D.58) \quad \hat{f}_r = \left[\frac{T(\tilde{p}_{r-1} - x(\tilde{p}_{r-1}))}{c^{v,0}} \left(\frac{1}{\delta + \lambda^G + \lambda^E(1 - F_{r-1})} \right)^2 \right]^{\frac{1}{\eta^{v-1}}} \frac{1}{VN}$$

All of the elements of equation (D.58) are known from the data, except for the firm's flow payoff per matched worker, $P_{r-1} \equiv (\tilde{p}_{r-1} - x(\tilde{p}_{r-1}))$. Thus, we have one equation (D.58) in one unknown (P_r) for each firm rank r , which can be rewritten as

$$(D.59) \quad P_{r-1} \equiv (\tilde{p}_{r-1} - x_{r-1}(\tilde{p}_{r-1})) = \left(\hat{f}_r VN \right)^{\eta^{v-1}} \frac{V c^{v,0}}{T} \left(\delta + \lambda^G + \lambda^E(1 - F_{r-1}) \right)^2.$$

Intuitively, a firm posts more vacancies if it receives a greater flow payoff per matched worker.

Going back to equation (D.50), we apply a similar finite-difference approximation:

$$(D.60) \quad \Delta x_r = x'(\tilde{p}_{r-1})(\tilde{p}_r - \tilde{p}_{r-1})$$

$$(D.61) \quad = \frac{2\lambda^E(P_{r-1})}{\delta + \lambda^G + \lambda^E(1 - F_{r-1})} \left[\frac{TP_{r-1}}{c^{v,0}} \left(\frac{1}{\delta + \lambda^G + \lambda^E(1 - F_{r-1})} \right)^2 \right]^{\frac{1}{\eta^{v-1}}} \frac{1}{VN}.$$

Recall that equation (D.59) above already identifies P_r . Therefore, we can iteratively apply equation (D.61) through ranks $r > 0$ to deduce x_r as follows:

$$(D.62) \quad x_r = x_0 + \sum_{i=1}^r \Delta x_i.$$

⁶³Note that our data covers the universe (i.e., a very large number) of firms in Brazil. If we consider a small sample of firms randomly drawn from the population, then for the first few firms in the left tail of the distribution of ranks, our approximation may be relatively poor. However, as we sum over increasingly higher ranks, our approximation becomes increasingly precise since errors vanish rather than accumulate.

For a given initial condition x_0 , based on equation (D.62), we deduce

$$(D.63) \quad \beta_r a_r = x_r - w_r,$$

where $\beta_r a_r$ is the amenity valuation to be pinned down, x_r is the total compensation recursively defined in equation (D.62), and w_r is the known firm component of wages at firm rank r . Similarly to the continuous case, we set x_0 such that $\min\{\beta_r a_r\}_r \approx 0$. This allows us to pin down the distribution of firm-level amenity valuations, $\beta_r a_r$, based on equation (D.63). From this, we deduce the amenity cost, $c^a(a_r)$, by plugging back $\beta_r a_r$ into the expression for the optimal amenity cost in equation (C.9). Finally, as we have already identified $P_r = \tilde{p}_r - x_r = p_r - w_r - c^a(a_r)$, we can deduce p_r firm by firm as

$$(D.64) \quad p_r = P_r + w_r + c^a(a_r),$$

proceeding exactly as in the continuous case in Supplemental Appendix D.6.

Having identified p_{Mj} for men and p_{Fj} for women at the same firm j , we can recover the gender wedge τ_j as

$$(D.65) \quad \tau_j = 1 - \frac{p_{Fj}}{p_{Mj}}.$$

In summary, equations (D.63), (D.64), and (D.65) jointly identify the unobserved multidimensional type $(p_{g,r}, \beta_{g,r} a_{g,r}, \tau_{g,r})$ of firms at any rank r and for both genders g . Intuitively, what our identification proof exploits is the fact that firm surplus and utility offers—consisting of productivity, the gender wedge, and the amenity value net of amenity creation costs—are identified by firms' labor demand as revealed through hires from nonemployment, while information on wage offers together with the already-revealed utility offers identifies firm amenity values. Finally, comparing outcomes across men and women at the same firm allows us to deduce the gender wedge, which stands in for differences in firm surplus of employing otherwise identical male and female workers at the same wage rate and amenity value net of amenity creation costs.

8. Proof of Proposition 5 (Economy-Wide Parameters)

RESTATEMENT OF PROPOSITION 5 (ECONOMY-WIDE PARAMETERS).

(i) The vacancy cost shifter $c^{v,0}$ is identified based on the aggregate labor share;
(ii) the elasticity of the vacancy cost function η^v is identified based on the firm pay-profit gradient; (iii) the elasticity of the amenity cost function η^a is identified based on the aggregate amenity cost share in the data. PROOF:

We proceed in three parts.

Part (i): vacancy cost shifter. First, we find an expression for firm profits $\rho\Pi(r)$.

Recall the definition of profits per matched worker,

$$(D.66) \quad p(r) - w(r) - c^a(a_r) = [f(r)V]^{\eta^v-1} \frac{c^{v,0}}{T} [\delta + \lambda^G + \lambda^E(1 - F(r))]^2.$$

From this expression, we multiply by size $l(r) = f(r)VT/[\delta + \lambda^G + \lambda^E(1 - F(r))]^2$ and subtract vacancy posting costs to obtain flow profits $\rho\Pi(r)$,

$$(D.67) \quad \rho\Pi(r) = (p(r) - w(r) - c^a(a_r))l(r) - c^v(v(r))$$

$$(D.68) \quad = [f(r)V]^{\eta^v} c^{v,0} - c^v(v(r)).$$

We now substitute the functional form of the cost of posting vacancies $c^v(v(r)) = c^{v,0}v(r)^{\eta^v}/\eta^v$ and $f(r)V = v(r)$ to obtain

$$(D.69) \quad \rho\Pi(r) = [v(r)]^{\eta^v} c^{v,0} - c^{v,0} \frac{[v(r)]^{\eta^v}}{\eta^v}$$

$$(D.70) \quad = \left(1 - \frac{1}{\eta^v}\right) [v(r)]^{\eta^v} c^{v,0}.$$

From this expression, it is immediately clear that flow profits $\rho\Pi(r)$ are proportional to $c^{v,0}$. Specifically, profits are always positive but they can be arbitrarily large as $c^{v,0} \rightarrow \infty$ and arbitrarily small as $c^{v,0} \rightarrow 0^+$, and therefore the profit share of the economy can range from 0 to 1 depending on the value of $c^{v,0}$.

Substituting back $v(r) = f(r)V$ to make clear the dependency on observed hiring intensities $f(r)$ we obtain

$$(D.71) \quad \rho\Pi(r) = \left(1 - \frac{1}{\eta^v}\right) [f(r)V]^{\eta^v} c^{v,0}$$

In equation (D.71), η^v is treated as unknown. However, its value does not matter for any of our arguments concerning the identification of the vacancy cost shifter $c^{v,0}$. We can write the labor share as

$$(D.72) \quad \mathcal{L} \equiv \frac{\int w(r)l(r) dr}{\int [p(r)l(r) - c^a(a_r))l(r) - c^v(v(r))] dr}$$

$$(D.73) \quad = \frac{\int w(r)l(r) dr}{\int [w(r)l(r) + \rho\Pi(r)] dr}$$

$$(D.74) \quad = 1 - \frac{\int \rho\Pi(r) dr}{\int [w(r)l(r) + \rho\Pi(r)] dr}.$$

Since flow profits $\rho\Pi(r)$ in equation (D.71) are strictly increasing in the vacancy cost shifter $c^{v,0}$ and the labor share in equation (D.74) is strictly decreasing in profits $\rho\Pi(r)$, it follows that the labor share is strictly decreasing in the vacancy

cost shifter $c^{v,0}$. Furthermore, the labor share in equation (D.74) can take on any value in $(0, 1)$ by choosing an appropriate value of the vacancy cost shifter $c^{v,0}$. As a result, the vacancy cost shifter $c^{v,0}$ is identified based on the aggregate labor share in the data. This proves part (i).

Part (ii): elasticity of the vacancy cost function. Next, we show that the variance of log profits is monotonically increasing in η^v . Applying natural logarithms to both sides of (D.71) yields

$$(D.75) \quad \ln(\rho\Pi(r)) = \ln\left(1 - \frac{1}{\eta^v}\right) + \eta^v \ln[f(r)V] + \ln(c^{v,0}).$$

Taking variances on both sides of equation (D.75), we get

$$(D.76) \quad Var[\ln(\rho\Pi(r))] = Var[(\eta^v \times \ln f(r)) + \eta^v \times \ln V + \ln(c^{v,0})]$$

$$(D.77) \quad = (\eta^v)^2 \times Var[\ln f(r)]$$

Except for $f(r)$, all terms in equation (D.76) are constant across firms, so they drop out of the calculation of the variance. As a result, equation (D.77) shows that the variance of log profits is proportional to $(\eta^v)^2$ and thus monotonically increasing in η^v . Now consider a regression of log firm pay on log profits,

$$(D.78) \quad \ln w(r) = \alpha + \beta \ln \Pi(r) + \varepsilon(r),$$

where the regression coefficient β in equation (D.78) captures the elasticity of firm pay with respect to firm profits. The regression coefficient

$$(D.79) \quad \beta = \frac{Cov[\ln w(r), \ln(\rho\Pi(r))]}{Var[\ln(\rho\Pi(r))]},$$

is inversely proportional to the variance of log profits, which scales in $(\eta^v)^2$, and proportional to the covariance between log profits and observed log pay, which scales in η^v , we know that the regression coefficient is strictly decreasing in η^v . Thus, if an empirical regression coefficient β is attained for some parameter value η^v , then this is the unique value of η^v that rationalizes this empirical β . As a result, η^v is identified based on the elasticity of firm pay to firm profits. This proves part (ii).

Part (iii): amenity cost elasticity. Finally, in equation (C.9) in Lemma 1 we demonstrated that amenity costs are inversely proportional to η^a . Therefore, we can write the economy-wide cost share of amenities as

$$(D.80) \quad \mathcal{A} \equiv \frac{\int c^a(a^*(r))l(r) dr}{\int [p(r)l(r) - c^a(a^*(r))l(r) - c^v(v(r))] dr}.$$

Obviously, the aggregate amenity cost share in equation (C.9) is monotonically

decreasing in η^a , as the numerator is monotonically decreasing in η^a and the denominator is monotonically increasing in η^a . Furthermore, as $\eta^a \rightarrow \infty$, the amenities cost share monotonically tends to zero. Thus, if an empirical value of the aggregate amenity cost share $\mathcal{A}$ is attained for some parameter value η^a , then this is the unique value of η^a that rationalizes this empirical $\mathcal{A}$. This proves that the elasticity of the amenity cost function η^a is identified based on the aggregate amenity cost share in the data.

9. Recovering Parameter Values in Monte Carlo Simulations

In this subsection, we perform Monte Carlo simulations of our model and use our estimation algorithm to recover the underlying distribution of firm-level parameters, based only on the same information we observe in the data, as detailed in Section IV. As our proof shows that all parameters are point-identified, this exercise is not strictly necessary, but we view it as a proof of concept and as further validation of our strategy.

For the purpose of this exercise, we only need to focus on one gender, with the understanding that the simulations recover p if the algorithm is run on men's data and $(1 - \tau)p$ if the algorithm is run on women's data. We start by drawing 100,000 firms, characterized by productivity p and amenities a , jointly normally distributed with correlation $\rho(p, a)$. We then transform productivity to have a Pareto marginal distribution.

We then feed the nonparametric joint distribution of $\{p, a\}$ to our discrete numerical solution algorithm, detailed in Supplemental Appendix G.1, to solve our model. The outputs of the simulation are firm-level wages, amenities, ranks, and hiring intensities. We use this data to construct our estimates of rank r , hiring intensities f_r , and offer CDF F_r as explained in Proposition 3. Finally, in order to test whether our algorithm is successful at uncovering the true firm-specific parameters, we use only data on firm-level ranks r , wages w_r , hiring intensities f_r , and CDF F_r to estimate amenities and productivity at the firm level.

Our results are summarized in Table D.1, which shows moments of the distribution of recovered estimates under different parameterizations of the underlying amenities distribution. Under different parameterizations of the data-generating process, shown in columns (1)–(5) of the table, our algorithm recovers estimates of the amenity values and productivities that are *identical* to the true values up to machine precision. In all experiments, we keep the marginal distribution of productivity fixed, but we alter the economy-wide parameters η^v and η^a , as well as the underlying variance of amenities $Var(a)$ and the correlation of amenities and productivity $\rho(p, a)$ in the initial joint normal distribution. In Column (1), $Var(a) = 0.1$, $\rho(p, a) = 0$, $\eta^v = 2$ and $\eta^a = 8$. In Column (2), $Var(a) = 0.15$, $\rho(p, a) = 0$, $\eta^v = 3$ and $\eta^a = 4$. In Column (3), $Var(a) = 0.125$, $\rho(p, a) = 0$, $\eta^v = 3.5$ and $\eta^a = 7$. In Column (4), $Var(a) = 0.2$, $\rho(p, a) = -0.5$, $\eta^v = 2$ and $\eta^a = 9$. In Column (5), $Var(a) = 0.15$, $\rho(p, a) = 0.5$, $\eta^v = 2.5$ and $\eta^a = 10$.

To evaluate the goodness of fit of our procedure, Panel E shows the correlations

and mean squared error (MSE) between our amenity estimates and true amenity values, and those between our productivity estimates and true productivity values. It also shows the MSE for the estimates of economy-wide parameters. The correlation for all estimates is equal to 1 approximated at the seventh decimal digit. The MSE for amenities, productivities, and economy-wide parameters is close to machine zero.

Figure [D.1](#) visualizes the fit of our estimation routine with respect to the main objects of interest—namely, the amenity values. Across all five simulations, estimates lie on the 45-degree line, showing how our procedure is robust to different joint distributions of amenities and productivity, and different values of economy-wide parameters.

TABLE D.1—MONTE CARLO SIMULATION AND ESTIMATION ON 100,000 FIRMS

	(1)	(2)	(3)	(4)	(5)
<i>Panel A. Properties of true wages w</i>					
Correlation between w and r	0.232	0.154	0.199	0.311	-0.080
<i>Panel B. Properties of true amenity values a</i>					
Standard deviation of a	0.316	0.388	0.354	0.449	0.387
Correlation between a and w	-0.661	-0.676	-0.667	-0.735	-0.613
Correlation between a and r	0.550	0.613	0.587	0.408	0.816
<i>Panel C. Properties of true productivity p</i>					
Standard deviation of p	0.733	0.739	0.735	0.710	0.750
Correlation between p and w	0.659	0.564	0.615	0.831	0.332
Correlation between p and r	0.710	0.723	0.700	0.583	0.738
Correlation between p and a	0.048	0.126	0.064	-0.358	0.455
<i>Panel D. Properties of estimated amenity values $\hat{a}$ and productivity $\hat{p}$</i>					
Standard deviation of $\hat{a}$	0.316	0.388	0.354	0.449	0.387
Correlation between $\hat{a}$ and w	-0.661	-0.676	-0.667	-0.735	-0.613
Correlation between $\hat{a}$ and r	0.549	0.613	0.587	0.408	0.816
Standard deviation of $\hat{p}$	0.733	0.739	0.735	0.710	0.750
Correlation between $\hat{p}$ and w	0.659	0.564	0.615	0.831	0.332
Correlation between $\hat{p}$ and r	0.710	0.723	0.700	0.583	0.738
Correlation between $\hat{p}$ and $\hat{a}$	0.048	0.126	0.064	-0.358	0.455
<i>Panel E. Goodness of fit</i>					
Correlation between $\hat{a}$ and a	1.000	1.000	1.000	1.000	1.000
Mean squared error of $\hat{a}$	0.000	0.000	0.000	0.000	0.000
Correlation between $\hat{p}$ and p	1.000	1.000	1.000	1.000	1.000
Mean squared error of $\hat{p}$	0.000	0.000	0.000	0.000	0.000
Mean squared error of η^v	0.000	0.000	0.000	0.000	0.000
Mean squared error of η^a	0.000	0.000	0.000	0.000	0.000

Note: Table reports estimation results using simulated data from 100,000 firms under different parameterizations of the underlying distribution of firm heterogeneity, including amenity values a , productivity p , wages w , and employer ranks r . MSE denotes the mean squared error. Columns (1)–(5) show separate simulations under different parameterizations of the data-generating process described in the text of Supplemental Appendix D.9.

Source: Model simulations.

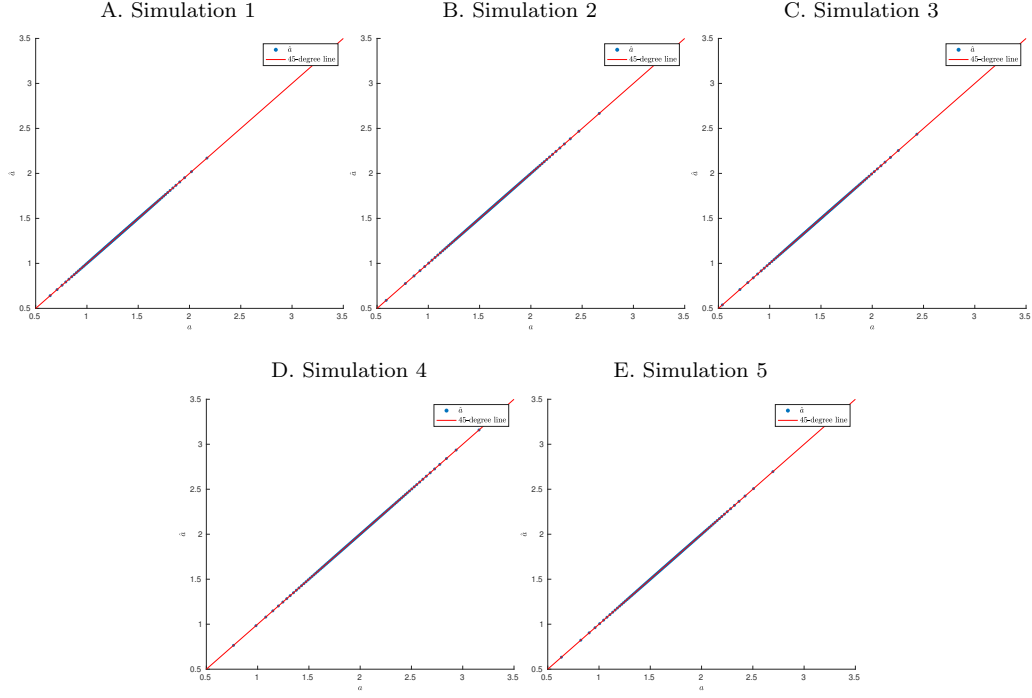


FIGURE D.1. AMENITY ESTIMATES AGAINST TRUE AMENITIES IN MONTE CARLO SIMULATIONS

Note: Figure plots average amenity estimates against percentile bins of true amenity values. The simulations in Panels A–E correspond to columns (1)–(5) in Table D.1.

Source: Model simulations.

10. Robustness to Model Misspecification and Subgroup Heterogeneity

To assess the robustness of our estimation results to model misspecification, we use a series of Monte Carlo simulations using a more flexible DGP that allows for heterogeneity in idiosyncratic worker preferences and firm attributes across population subgroups, even within gender. In what follows, we describe the details of how we estimate our (misspecified) model based on simulated data from such a DGP, and the results from this analysis. The bottom line is that we find our model to be remarkably robust to a wide range of misspecifications in the form of idiosyncratic worker and firm attributes.

To address the concern that preferences may be heterogeneous in the worker population and that, therefore, firm rankings may be heterogeneous, we run the following experiment. We assume that there are two groups of workers labeled A and B . Each firm is characterized by six objects: a shared productivity fundamental, p , a shared amenity cost shifter fundamental, $c^{a,0}$, two group-

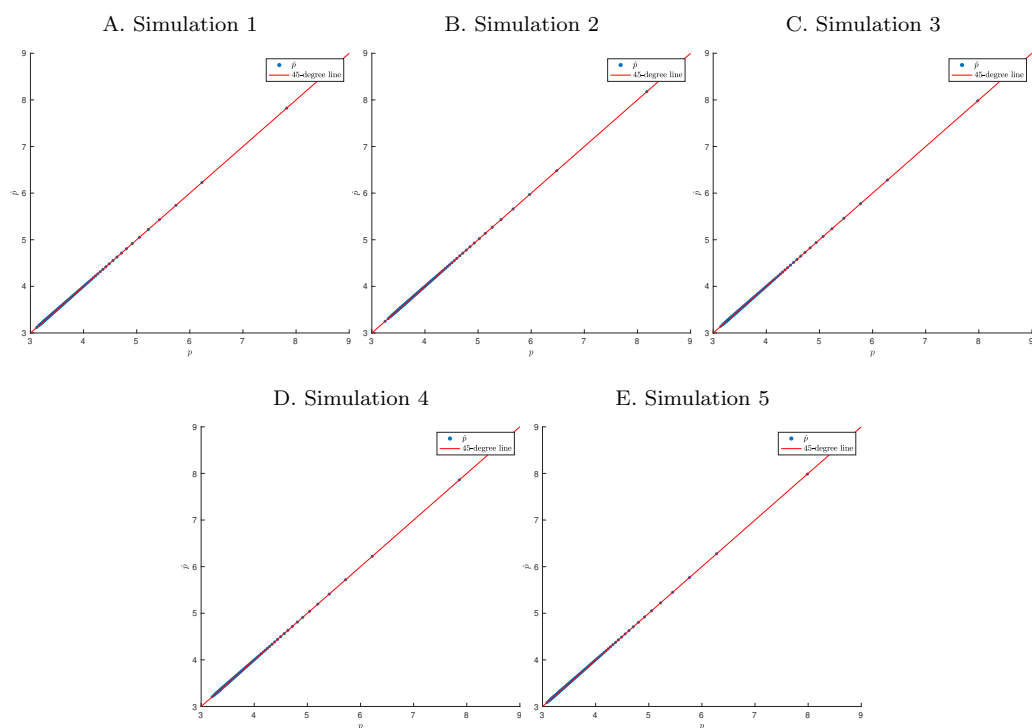


FIGURE D.2. PRODUCTIVITY ESTIMATES AGAINST TRUE PRODUCTIVITIES IN MONTE CARLO SIMULATIONS

Note: Figure plots average productivity estimates against percentile bins of true productivity values. The simulations in Panels A–E correspond to columns (1)–(5) in Table D.1.

Source: Model simulations.

specific productivity components, $\{\varepsilon_i\}_{i \in \{A, B\}}$, and two group-specific amenity cost components $\{\zeta_i\}_{i \in \{A, B\}}$. We define the group-specific productivity and amenity cost shifter, respectively, for population group $i \in \{A, B\}$ at firm j as:

$$\begin{aligned} p_{ij} &= p_j + \varepsilon_i, \\ c_{ij}^{a,0} &= c_j^{a,0} + \zeta_i. \end{aligned}$$

We refer to p_{ij} and $c_{ij}^{a,0}$ —and thus the implied equilibrium amenities a_{ij} —as the *group-specific fundamentals*, in contrast to the fundamentals p_j and $c_j^{a,0}$ that are shared between worker groups at the same firm j . Starting from these two sets of group-specific fundamentals, we fix all labor market parameters ($\lambda^U, \lambda^E, \lambda^G, \delta$) across groups and across simulations, and we fix the economy-wide parameters η^v and η^a across simulations. We then solve the general equilibrium across firms j in our model, separately for each group $i \in \{A, B\}$, as if these were two separate economies populated by workers who climb separate job ladders. This is the sense in which this simulation exercise features idiosyncratic worker preferences and firm rankings. Importantly, we can link workers within the same firm j by use of the simulated firm identifier.

Next, we treat the group identity $i \in \{A, B\}$ as unobserved by the econometrician. Instead, the econometrician observes pooled population averages as we measure them in our paper. To this end, we construct firm-level total employment g_j as the sum of employment of workers of type i , g_{ij} , and firm-level mean wages w_j as the weighted average of wages of workers of type i , w_{ij} , for each firm j by aggregating group-level quantities across model simulations for worker groups $i \in \{A, B\}$:

$$\begin{aligned} g_j &= g_{Aj} + g_{Bj}, \\ w_j &= \frac{g_{Aj}}{g_j} w_{Aj} + \frac{g_{Bj}}{g_j} w_{Bj}. \end{aligned}$$

We then use data on firm-level employment g_j and firm-level wages w_j as inputs to our estimation procedure, again assuming that the econometrician cannot separately observe outcomes of individual worker groups $i \in \{A, B\}$. Specifically, we use firm-level average employment g_j to infer firm ranks, we back out firm-level hires following the steps described in the paper, and we assign each firm their average wages as computed above. Importantly, our model with shared worker preferences across firms is generally misspecified vis-à-vis the DGP featuring two separate firm rankings of worker groups $i \in \{A, B\}$.

We construct three sets of experiments. In the first experiment, the group-specific components ε_i and ζ_i are uncorrelated across groups. In the second experiment, the group-specific components are perfectly negatively correlated across groups. In the last experiment, the group-specific components are perfectly positively correlated. For each set of experiments, we simulate six different scenarios in which we let the variance of the group-specific components become an increasingly

large multiple of the variance of the shared fundamental. In our most extreme case study, group-specific components of p_{ij} and $c_{ij}^{a,0}$ feature 200% of the dispersion as that of the shared fundamentals, including firm productivity and amenities. We then evaluate the performance of our algorithm by comparing the estimated values of productivity $\hat{p}$ and estimated amenities $\hat{a}$ with the true employment-weighted average productivity at the same firm $\bar{p}$ and the true employment-weighted average amenity valuation at the same firm $\bar{a}$, which we define as the employment-weighted average of group-specific fundamentals.

The simulation results from our first experiment, in which group-specific fundamentals are uncorrelated across groups, are summarized in Table D.2. We find that our estimation algorithm recovers well the *average* firm-level fundamentals in all exercises. As we increase the importance of group-specific fundamentals, the performance of our estimation procedure worsens, but it remains strong throughout, where the correlation between estimates and true averages never drops below 0.97. For comparison, we also display the correlation of the firm-specific average fundamentals that we estimate and the group-specific fundamentals. Obviously, as we increase the importance of group-specific drivers, the firm-level average becomes a worse predictor of the group-specific fundamentals. However, it remains highly reliable at estimating the average fundamental.

For our second experiment, we assume that group-specific components are perfectly negatively correlated. Intuitively, it is more difficult for our estimation algorithm to accurately approximate the objects of interest when group-specific components are perfectly negatively correlated.⁶⁴ The simulation results from this second experiment are summarized in Table D.3. While the performance of our estimation algorithm slightly worsens, it is still surprisingly effective at recovering the firm-specific average productivities and amenity valuations, even when the relative variance of group-specific components is large. For productivity, the correlation between the estimated value and the true average remains at or above 0.976 through all exercises. For amenities, it decreases to 0.908, highlighting the challenge of identifying group-specific determinants of utility and employment when they are negatively correlated. Obviously, the relationship between our estimated average fundamentals and the group-specific fundamentals worsens.

As a third experiment and sanity check, we let the group-specific components be perfectly positively correlated across groups. In this case, there is, of course, a common ranking for both worker groups $i \in \{A, B\}$. As a result of this shared ranking among all workers in the population, our estimation should deliver the exact values of productivity and amenities for any variance of the different components. The simulation results from this third experiment are summarized

⁶⁴To see why a negative correlation causes difficulties, consider the case in which the shared fundamental is irrelevant so that group-specific components are the *only* determinant of a firm's rank, wages, and employment—i.e., $p_{ij} = \varepsilon_i$ and $c_{ij}^{a,0} = \zeta_i$. If the group-specific components are perfectly negatively correlated across groups, then average employment at each firm is constant in expectation once aggregated across worker groups. This is despite the existence of clear firm rankings and meaningful firm size distributions within each worker group.

TABLE D.2—ESTIMATION UNDER UNCORRELATED GROUP-SPECIFIC COMPONENTS

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A. Properties of group-specific components</i>						
Relative variance of productivity component, ε_i	0.000	0.100	0.200	0.500	1.000	2.000
Relative variance of amenity component, ζ_i	0.000	0.100	0.200	0.500	1.000	2.000
<i>Panel B. Goodness of fit of pooled estimates vs. population average</i>						
Correlation between $\hat{p}$ and $\bar{p}$	1.000	0.998	0.995	0.986	0.979	0.976
Correlation between $\hat{a}$ and $\bar{a}$	1.000	1.000	0.998	0.995	0.988	0.980
Mean squared error of $\hat{p} - \bar{p}$	0.000	0.007	0.021	0.064	0.107	0.144
Mean squared error of $\hat{a} - \bar{a}$	0.000	0.001	0.002	0.002	0.004	0.030
<i>Panel C. Goodness of fit of pooled estimates vs. group-specific fundamentals</i>						
Correlation between $\hat{p}$ and p_A	1.000	0.969	0.944	0.890	0.839	0.785
Correlation between $\hat{p}$ and p_B	1.000	0.967	0.938	0.875	0.811	0.758
Correlation between $\hat{a}$ and a_A	1.000	0.977	0.955	0.905	0.847	0.784
Correlation between $\hat{a}$ and a_B	1.000	0.975	0.954	0.903	0.851	0.795
Mean squared error of $\hat{p} - p_A$	0.000	0.114	0.196	0.353	0.468	0.557
Mean squared error of $\hat{p} - p_B$	0.000	0.111	0.197	0.359	0.495	0.593
Mean squared error of $\hat{a} - a_A$	0.000	0.009	0.015	0.027	0.043	0.071
Mean squared error of $\hat{a} - a_B$	0.000	0.007	0.014	0.028	0.042	0.078

Note: This table reports results from estimation in the case of a misspecified model in which two groups of workers have different, uncorrelated group-specific components. Column 1 is the baseline estimation with no group-specific components. Between columns 2 and 6, the relative variances of group-specific components progressively increase from 10% to 200% of those of the shared components.

TABLE D.3—ESTIMATION UNDER PERFECTLY NEGATIVELY CORRELATED GROUP-SPECIFIC COMPONENTS

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A. Properties of group-specific components</i>						
Relative variance of productivity component, ε_i	0.000	0.100	0.200	0.500	1.000	2.000
Relative variance of amenity component, ζ_i	0.000	0.100	0.200	0.500	1.000	2.000
<i>Panel B. Goodness of fit of pooled estimates vs. population average</i>						
Correlation between $\hat{p}$ and $\bar{p}$	1.000	0.994	0.989	0.977	0.976	0.980
Correlation between $\hat{a}$ and $\bar{a}$	1.000	0.998	0.995	0.980	0.954	0.908
Mean squared error of $\hat{p} - \bar{p}$	0.000	0.024	0.048	0.124	0.187	0.234
Mean squared error of $\hat{a} - \bar{a}$	0.000	0.005	0.011	0.053	0.153	0.303
<i>Panel C. Goodness of fit of pooled estimates vs. group-specific fundamentals</i>						
Correlation between $\hat{p}$ and p_A	1.000	0.943	0.888	0.774	0.647	0.538
Correlation between $\hat{p}$ and p_B	1.000	0.934	0.892	0.785	0.716	0.679
Correlation between $\hat{a}$ and a_A	1.000	0.949	0.904	0.791	0.679	0.555
Correlation between $\hat{a}$ and a_B	1.000	0.951	0.906	0.792	0.637	0.443
Mean squared error of $\hat{p} - p_A$	0.000	0.194	0.318	0.537	0.737	0.921
Mean squared error of $\hat{p} - p_B$	0.000	0.205	0.321	0.533	0.641	0.709
Mean squared error of $\hat{a} - a_A$	0.000	0.017	0.033	0.093	0.242	0.468
Mean squared error of $\hat{a} - a_B$	0.000	0.019	0.035	0.094	0.169	0.264

Note: This table reports results from estimation in the case of a misspecified model in which two groups of workers have different, perfectly negatively correlated group-specific components. Column 1 is the baseline estimation with no group-specific components. Between columns 2 and 6, the relative variances of group-specific components progressively increase from 10% to 200% of those of the shared components.

in Table D.4. Reassuringly, in this case, our estimation procedure successfully recovers the exact fundamentals—see also the additional Monte Carlo simulation results in Supplemental Appendix D.9 of the revised paper.

TABLE D.4—ESTIMATION UNDER PERFECTLY POSITIVELY CORRELATED GROUP-SPECIFIC COMPONENTS

	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A. Properties of group-specific components</i>						
Relative variance of productivity component, ε_i	0.000	0.100	0.200	0.500	1.000	2.000
Relative variance of amenity component, ζ_i	0.000	0.100	0.200	0.500	1.000	2.000
<i>Panel B. Goodness of fit of pooled estimates vs. population average</i>						
Correlation between $\hat{p}$ and $\bar{p}$	1.000	1.000	1.000	1.000	1.000	1.000
Correlation between $\hat{a}$ and $\bar{a}$	1.000	1.000	1.000	1.000	1.000	1.000
Mean squared error of $\hat{p} - \bar{p}$	0.000	0.000	0.000	0.000	0.000	0.000
Mean squared error of $\hat{a} - \bar{a}$	0.000	0.000	0.000	0.000	0.000	0.000
<i>Panel C. Goodness of fit of pooled estimates vs. group-specific fundamentals</i>						
Correlation between $\hat{p}$ and p_A	1.000	1.000	1.000	1.000	1.000	1.000
Correlation between $\hat{p}$ and p_B	1.000	1.000	1.000	1.000	1.000	1.000
Correlation between $\hat{a}$ and a_A	1.000	1.000	1.000	1.000	1.000	1.000
Correlation between $\hat{a}$ and a_B	1.000	1.000	1.000	1.000	1.000	1.000
Mean squared error of $\hat{p} - p_A$	0.000	0.000	0.000	0.000	0.000	0.000
Mean squared error of $\hat{p} - p_B$	0.000	0.000	0.000	0.000	0.000	0.000
Mean squared error of $\hat{a} - a_A$	0.000	0.000	0.000	0.000	0.000	0.000
Mean squared error of $\hat{a} - a_B$	0.000	0.000	0.000	0.000	0.000	0.000

Note: This table reports results from estimation in the case of a misspecified model in which two groups of workers have different, perfectly positively correlated group-specific components. Column 1 is the baseline estimation with no group-specific components. Between columns 2 and 6, the relative variances of group-specific components progressively increase from 10% to 200% of those of the shared components.

Our simulation-based findings that individual heterogeneity and other sources of noise approximately wash out in our model estimation can be rationalized based on an interesting property of our model. To see this, we start by revisiting equation (D.44) from the identification proof of Proposition 4 in Supplemental Appendix D.6:

$$P_{r-1} \equiv (\tilde{p}_{r-1} - x_{r-1}(\tilde{p}_{r-1})) = \left(\hat{f}_r V N \right)^{\eta^v - 1} \frac{V c^{v,0}}{T} \left(\delta + \lambda^G + \lambda^E (1 - F_{r-1}) \right)^2,$$

where P_r is the flow payoff per matched worker, $\tilde{p}_r$ is the composite productivity, $x_r(\tilde{p}_r)$ is the flow utility delivered to workers, $\hat{f}_r$ is the recruiting intensity, and F_r is the ladder rank, all for a firm at rank r , and V is a constant reflecting aggregate vacancies, N is the number of firms, η^v is the elasticity of the vacancy cost function, $c^{v,0}$ is the economy-wide vacancy cost intercept, T is a constant defined in equation (D.40), and $(\delta, \lambda^G, \lambda^E)$ are labor market transition rates.

What is key is that there are two idiosyncratic terms in the expression for P_r above: the recruiting intensity f_r and the firm rank F_r . Hence, both f_r and F_r are potential sources for estimation error based on idiosyncratic variation due to preference heterogeneity or other sources of noise. Let us first inspect the term reflecting a firm's estimated rank, F_r . Importantly, we see that F_r is premultiplied by $\lambda^E \approx 1\%$ and the whole term containing F_r is of magnitude $\approx 5\%$, which means that the squared term becomes of magnitude $\approx (5\%)^2 \ll 1\%$. Therefore, idiosyncratic variation in the estimate of a firm's rank, F_r , has limited influence on our estimates of P_r and thus firm productivity p and amenities a .

This leaves us with only one source of idiosyncratic variation in the estimation, namely, firms' recruiting intensity f_r . In general, the impact of noise in f_r can be substantial in our model through the term $(\hat{f}_r VN)^{\eta^v-1}$ on the right-hand side of the equation above. However, a result of our estimation procedure is that our estimated elasticity of the vacancy cost function is $\eta^v = 2.062 \approx 2$, which implies that $(\hat{f}_r VN)^{\eta^v-1} \approx \hat{f}_r VN$. Therefore, P_r is approximately linear in $\hat{f}_r$, which implies that any idiosyncratic variation in $\hat{f}_r$ washes out in expectation when estimating P_r .⁶⁵ As a result, our estimation procedure recovers estimates of P_r that are unbiased. As a result of having an unbiased estimate of the firm's flow payoff from a matched worker, P_r , we obtain the change in utility offered across rungs of the firm ladder, Δx , defined in equation (D.61). Furthermore, when computing the levels of flow utility $x_r = x_0 + \sum_{i=1}^r \Delta x_i$ offered at each firm by use of equation (D.62), idiosyncratic noise in individual firms' P_r washes out because we are summing over many such terms across rungs of the ladder.⁶⁶ Therefore, our estimation procedure yields unbiased and stable estimates of flow utility x_r , which is robust to idiosyncratic variation or noise in f_r . As a result of plugging x_r back into equation (D.63), we also see that our amenity estimates are unbiased and stable for a given observation of the empirical firm pay w_r . Plugging all of these objects back into the definition of P_r above, we see that the model object that should pick up most of the idiosyncratic variation in f_r should be firm productivity p_r .

Finally, it is interesting to assess how much the data clearly violate the assumption of homogeneous employer preferences across population subgroups within each gender. To this end, we have split our sample by three worker characteristics, one at a time. These worker characteristics are: low education (i.e., at most high school) vs. high education (i.e., at least high school), ever parents (i.e., registered at least one parental leave between 2007 and 2018) vs. never parents (i.e., did not register any parental leaves between 2007 and 2018), and young cohorts (i.e., year of birth starting in 1968) vs. old cohorts (i.e., year of birth up to 1967). We then compute firm ranks as employment at a given firm, separately for each population

⁶⁵The same logic also predicts that idiosyncratic variation in f_r should have more substantial effects on the accuracy of our model estimates, which we confirm in Monte Carlo simulations as described above.

⁶⁶Again, the same logic also predicts that we should have more accurate estimates in the center of the distribution compared to the tails, which we confirm in additional Monte Carlo simulations.

subgroup. With this in hand, we are interested in how related the two population subgroups' firm ranks are within each gender. For example, this allows us to study the extent to which the firm ranking of women with kids differs from that of women without kids.

Figure D.3 provides a first look at the results from this exercise by showing binscatter plots of empirical employer ranks across population subgroups, separately by gender. The key takeaway is that there is a strong and close-to-linear relationship between employer ranks of different subpopulations within each gender. In particular, the ranks are close to monotonic in each other. Of course, the linearity and monotonicity of the plotted relationships are not perfect. For example, bottom-ranked firms for population subgroup 1 (on the horizontal axis of each panel) tend to be higher-ranked for members of population subgroup 2 (on the vertical axis of each panel). These imperfections could either reflect preference heterogeneity or measurement error, both of which might be particularly pronounced in the tails of the distribution of firm ranks. Note also that the parent vs. nonparent comparison for men shows a relatively worse fit than for the other population subgroups and genders. This may be partly due to misclassification in what we group into the “parent” category, which consists of all workers who took a registered parental leave from work between 2007 and 2018. Since women are orders of magnitude more likely to go on paid parental leave than men, our sample of men who are on parental leave might be a special population subgroup and also subject to measurement error. Overall, though, we take these results to be reassuring that firm ranks strongly tend to be shared across population subgroups within workers of a given gender.

11. Identification Under Heterogeneous Separation Rates

To discuss identification of our model when separation rates are heterogeneous, as sketched in Supplemental Appendix C.8, we start as in Supplemental Appendix D.6, expressing the first-order conditions (C.67) and (C.68) as a coupled pair of differential equations. We use the same changes of variables as in Supplemental Appendix C.4, but we apply them to this more complex case in which $\delta_g(\tilde{p}_{gz})$ is a function of composite productivity $\tilde{p}_{gz}$. From now on, we drop the productivity

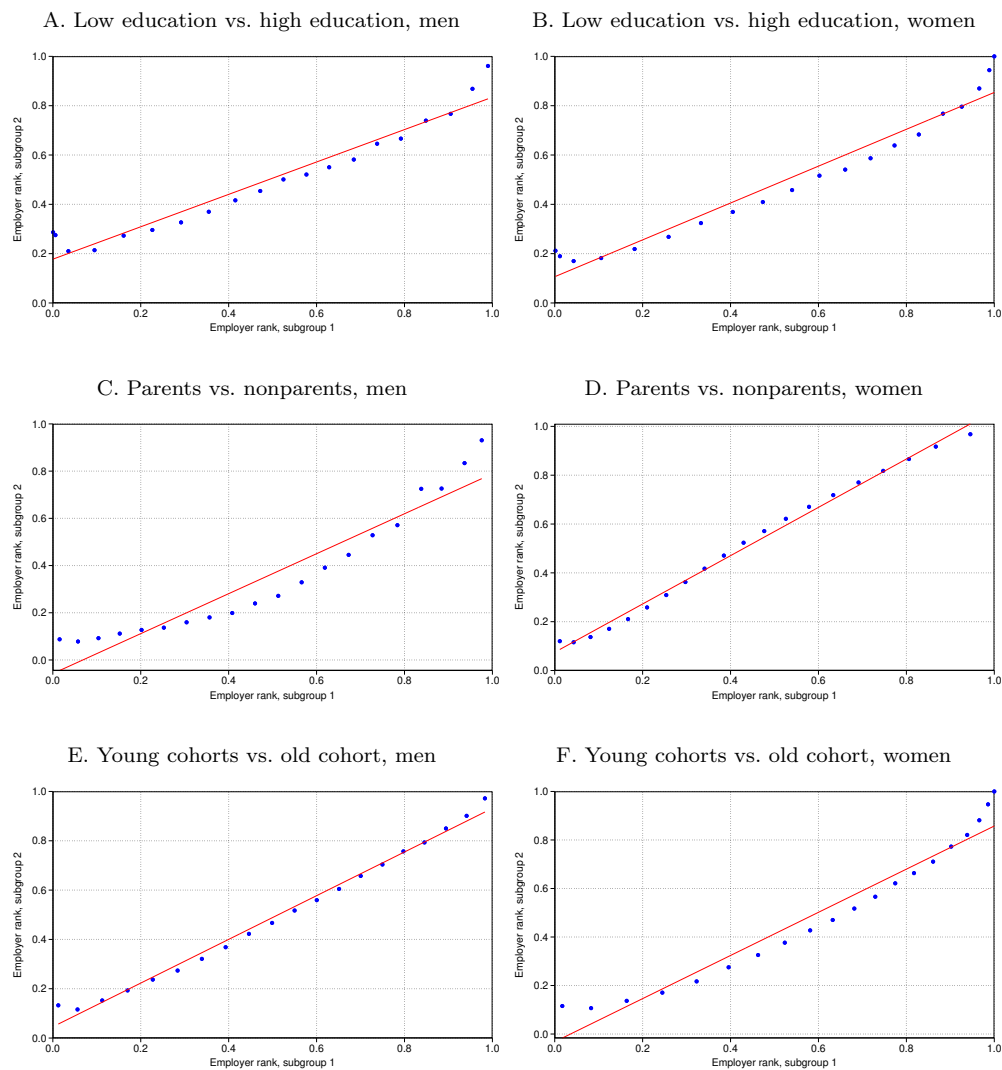


FIGURE D.3. BINSKATTER PLOTS OF EMPLOYER RANKS ACROSS POPULATION SUBGROUPS, BY GENDER

Note: This figure shows binscatter plots of empirical employer ranks across population subgroups, separately by gender. In each panel, the vertical axis shows the mean employer rank for population subgroup 2 corresponding to a set of firms in a given quantile of the distribution of employer ranks for population subgroup 1. Here, population subgroups 1 and 2 refer to the two population subgroups referenced in the relevant panel's caption.

Source: RAIS, 2007–2014.

and gender subscripts for readability. Define:

$$(D.81) \quad h(\tilde{p}) = F(x^*(\tilde{p})) = \frac{\int_{\tilde{p}' \geq \phi} v^*(\tilde{p}') \gamma(\tilde{p}') d\tilde{p}'}{V}$$

$$(D.82) \quad j(\tilde{p}) = G(x^*(\tilde{p})) = \frac{\int_{\tilde{p}' \geq \phi} l^*(\tilde{p}') \gamma(\tilde{p}') d\tilde{p}'}{\mu(1-u)}$$

$$(D.83) \quad h'(\tilde{p}) = f(x^*(\tilde{p}))x^{*'}(\tilde{p})$$

$$(D.84) \quad j'(\tilde{p}) = g(x^*(\tilde{p}))x^{*'}(\tilde{p})$$

$$(D.85) \quad f(x^*(\tilde{p})) = h'(\tilde{p})/x^{*'}(\tilde{p})$$

$$(D.86) \quad g(x^*(\tilde{p})) = j'(\tilde{p})/x^{*'}(\tilde{p})$$

Thus, by differentiating equations (D.81) and (D.82) using Leibniz's integral rule, we can write $f(x^*(\tilde{p})) = (v^*(\tilde{p}')/V)\gamma(\tilde{p}) d\tilde{p}/dx^*(\tilde{p})$, and $g(x^*(\tilde{p})) = (l^*(\tilde{p}')/(\mu(1-u))\gamma(\tilde{p}) d\tilde{p}/dx^*(\tilde{p})$. Therefore, we can rewrite the first-order conditions as:

$$(D.87) \quad h'(\tilde{p}) = \frac{1}{V} \left[(\tilde{p} - x) \frac{1}{c^{v,0}V} \left(\frac{u + (1-u)s^E j(\tilde{p}) + s^G}{\delta(\tilde{p}) + \lambda^G + \lambda^E [1 - h(\tilde{p})]} \right) \mu \lambda^U \right]^{\frac{1}{\eta^{v-1}}} \gamma(\tilde{p})$$

$$(D.88) \quad x'(\tilde{p}) = (\tilde{p} - x) \left(\frac{\lambda^E h'(\tilde{p})}{\delta(x) + \lambda^G + \lambda^E [1 - h(x)]} + \frac{(1-u)s^E j'(\tilde{p})}{u + (1-u)s^E j(\tilde{p}) + s_g^G} \right).$$

Once again, equations (D.87) and (D.88) are more general versions of equations (D.49) and (D.50), where we allow separation rates to be a decreasing function of $\tilde{p}$. Using similar changes of variables as we did in Supplemental Appendix D.6, we can write an expression for profits per worker depending on ranks by rearranging equation (D.87):

$$(D.89) \quad P(r) = \tilde{p}(r) - x(\tilde{p}(r)) = [f(r)V]^{\eta^{v-1}} \frac{c^{v,0}V (\delta(r) + \lambda^G + \lambda^E [1 - h(r)])}{\mu \lambda^U (u + (1-u)s^E j(r) + s_g^G)},$$

which allows us to extract profits per worker from the data using a modified version of our original equation (D.44).

We now turn to the identification of x . Further substituting $j'(\tilde{p}) = l^*(\tilde{p})\gamma(\tilde{p})/(\mu(1-u))$, then $l^*(\tilde{p})$ with its solution from equation (C.66) and $h'(\tilde{p})$ from equation (D.87) yields:

$$(D.90) \quad x'(\tilde{p}) = (\tilde{p} - x) \left(\frac{\lambda^E}{\delta(\tilde{p}) + \lambda^G + \lambda^E [1 - h(\tilde{p})]} \frac{1}{V} \left[(\tilde{p} - x) \frac{1}{c^{v,0}V} \left(\frac{u + (1-u)s^E j(\tilde{p}) + s^G}{\delta(\tilde{p}) + \lambda^G + \lambda^E [1 - h(\tilde{p})]} \right) \mu \lambda^U \right]^{\frac{1}{\eta^{v-1}}} \right. \\ \left. + \frac{s^E}{u + (1-u)s^E j(\tilde{p}) + s_g^G} \left(\frac{1}{\delta(\tilde{p}) + \lambda^G + \lambda^E [1 - h(\tilde{p})]} \right) \frac{v^*(\tilde{p})}{V} \frac{\mu}{\mu} \lambda^U [u + (1-u)s_g^E j(\tilde{p}) + s_g^G] \right) \gamma(\tilde{p})$$

which can be further simplified as:

$$(D.91) \quad x'(\tilde{p}) = (\tilde{p} - x) \left(\frac{\lambda^E}{\delta(\tilde{p}) + \lambda^G + \lambda^E [1 - h(\tilde{p})]} \frac{1}{V} \left[(\tilde{p} - x) \frac{1}{c^{v,0}V} \left(\frac{u + (1-u)s^E j(\tilde{p}) + s^G}{\delta(\tilde{p}) + \lambda^G + \lambda^E [1 - h(\tilde{p})]} \right) \mu \lambda^U \right]^{\frac{1}{\eta^{v-1}}} \right. \\ \left. + \left(\frac{s^E \lambda^U}{\delta(\tilde{p}) + \lambda^G + \lambda^E [1 - h(\tilde{p})]} \right) \frac{v^*(\tilde{p})}{V} \right) \gamma(\tilde{p}).$$

This further yields

$$(D.92) \quad x'(\tilde{p}) = \frac{2\lambda^E(\tilde{p} - x)}{\delta(\tilde{p}) + \lambda^G + \lambda^E [1 - h(\tilde{p})]} \frac{1}{V} \left[(\tilde{p} - x) \frac{1}{c^{v,0}V} \left(\frac{u + (1-u)s^E j(\tilde{p}) + s^G}{\delta(\tilde{p}) + \lambda^G + \lambda^E [1 - h(\tilde{p})]} \right) \mu \lambda^U \right]^{\frac{1}{\eta^{v-1}}} \gamma(\tilde{p})$$

which is identical to our expression in (D.45), except for a change in the way firm sizes (and therefore vacancies) are computed. Interestingly, the optimality condition for utility turns out to be unaffected by heterogeneity in $\delta(\tilde{p})$: the reason is that, while the distribution of workers across firms in the economy is affected by how the separation rate varies with composite productivity, at the margin the employment distribution only varies with $f(x^*(\tilde{p}))$ and $\delta(x(\tilde{p}))$, the local competition of a firm offering utility x . Thus, with a final change of variables and calling $\delta(r) = \delta(\tilde{p}(r))$, we can write:

$$(D.93) \quad \Delta x(r) = \frac{2\lambda^E P(r)}{\delta(r) + \lambda^G + \lambda^E [1 - F(r)]} \frac{1}{V} \left[P(r) \frac{1}{c^{v,0}V} \left(\frac{u + (1-u)s^E G(r) + s^G}{\delta(r) + \lambda^G + \lambda^E [1 - h(x)]} \right) \mu \lambda^U \right]^{\frac{1}{\eta^{v-1}}}$$

Therefore, we proceed as in Supplemental Appendix D.7, adapting our equations to discrete data. We first identify profits per worker $P(r)$ by applying equation D.89, plugging in the data on the employment distribution $G(r)$, the implied hiring intensities $f(r)$, and the associated empirical offer distribution $F(r)$. We then identify utility offers $x(r)$ by applying equation (D.93) iteratively, plugging in previously obtained profits per worker $P(r)$, the empirical offer distribution $F(r)$, and the empirical employment distribution $G(r)$. Finally, we proceed exactly as in Supplemental Appendix D.7 to identify amenities, productivity, and gender wedges.

E. ESTIMATION RESULTS APPENDIX

1. Constructing Aggregate Statistics

We use three aggregate statistics for identification in our model: the aggregate labor share, the elasticity of firm pay with respect to firm value added per worker, and the aggregate amenity cost share.

AGGREGATE LABOR SHARE.

Part (i) of our identification result in Proposition 5 of Section IV links the vacancy cost shifter $c^{v,0}$ to the aggregate labor share in the data. To bridge the model with the data, we define the labor share as the share of value added accruing to workers in pay,

$$(E.1) \quad \mathcal{L} \equiv \frac{\sum_g \sum_j w_{gj} l_{gj}}{\sum_g \sum_j [p_j l_{gj} - c_g^a(a_{gj}) l_{gj} - c_g^v(v_{gj})]} = 48\%,$$

where the numerator contains the wage bill, $\sum_g \sum_j w_{gj} l_{gj}$, while the denominator contains value added, $\sum_g \sum_j [p_j l_{gj} - c_g^a(a_{gj}) l_{gj} - c_g^v(v_{gj})]$. The labor share value of 0.48 is the 2007 value of the share of labor compensation of employees (i.e., excluding the self-employed, who are outside of our model) for Brazil, based on [Feenstra, Inklaar and Timmer \(2015\)](#) as retrieved via FRED.

ELASTICITY OF FIRM PAY WITH RESPECT TO FIRM VALUE ADDED PER WORKER.

Part (ii) of our identification result in Proposition 5 of Section IV links the elasticity of the vacancy cost function η^v to the elasticity of firm pay with respect to firm profits in the data. While firm profits are readily measured in firm financial data, we instead rely on existing estimates of the elasticity of firm pay with respect to value added per worker,

$$(E.2) \quad \tilde{\beta} = \frac{Cov(\ln w(r), \ln \left(\frac{\Pi(r) + w(r)l(r)}{l(r)} \right))}{Var(\ln \left(\frac{\Pi(r) + w(r)l(r)}{l(r)} \right))},$$

To put a number on the elasticity $\tilde{\beta}$ in equation (E.2), we take the coefficient from a regression of firm fixed effects in wages on log value added per worker at the firm level in Brazil from [Alvarez et al. \(2018\)](#). The estimated coefficient is 0.179 in a balanced panel of Brazilian manufacturing firms from 2004–2008 (see Table E1 of [Alvarez et al., 2018](#)), which matches the beginning of our sample period.

AGGREGATE AMENITY COST SHARE.

Part (iii) of our identification result in Proposition 5 of Section IV links the elasticity of the amenity cost function η^a to the aggregate amenity cost share in

the data. A challenge we face is that for the share of amenity costs in value added, no precise estimate exists in the literature, much less so for the Brazilian context. Therefore, any assumed value necessarily comes with significant uncertainty. With this caveat in mind, related estimates on the value of local amenities (Bieri, Kuminoff and Pope, 2023) suggest that the approximate cost share of amenities in value added is

$$(E.3) \quad \mathcal{A} \equiv \frac{\sum_g \sum_j c_g^a(a_{gj})l_{gj}}{\sum_g \sum_j [p_j l_{gj} - c_g^a(a_{gj})l_{gj} - c_g^v(v_{gj})]} = 8\%.$$

Due to the significant uncertainty surrounding this estimate, we use it as a baseline and conduct robustness checks around it.

2. Details on Covariates Related to Productivity Estimates

To obtain productivity estimates for a subset of the firms in the RAIS linked employer-employee data, we merge in firm financial data from Bureau van Dijk’s Orbis Historical database. The Orbis Historical data is the largest cross-country firm-level database containing information on public and private firms’ balance sheets and income statements—see Kalemli-Özcan et al. (2024) for a detailed description. We merge the Orbis Historical data into RAIS using Brazilian firms’ tax identifiers (*Cadastro Nacional de Pessoas Jurídicas*, or *CNPJ*). We then use as our empirical productivity proxy revenue per worker in the Orbis Historical data, which we relate to our model notion of firm productivity p .

3. Details on Covariates Related to Gender Wedge Estimates

We include as covariates in Z_j in equation (25) the following six variables that we construct using the RAIS data, in addition to two sets of FEs.

FEMALE MANAGER.

We create an indicator for whether the employer has a woman in the highest-paid position.

NONROUTINE MANUAL TASK INTENSITY.

We compute the mean z -score for nonroutine manual task intensity measured by linking 5-digit occupation codes from the Brazilian *Classificação Brasileira de Ocupações (CBO)* 1994 occupation classification to United States 1990 Census occupation codes based on the occupational crosswalk of de Souza (2024) extending previous work of Autor and Dorn (2009) and Acemoglu and Autor (2011).

NONROUTINE INTERPERSONAL TASK INTENSITY.

We compute mean z -score for nonroutine interpersonal task intensity measured by linking 5-digit occupation codes from the Brazilian *Classificação Brasileira de*

Ocupações (CBO) 1994 occupation classification to United States 1990 Census occupation codes based on the occupational crosswalk of [de Souza \(2024\)](#) extending previous work of [Autor and Dorn \(2009\)](#) and [Acemoglu and Autor \(2011\)](#).

MEAN WORKING HOURS.

We compute the mean log number of contractual work hours.

NO MAJOR FINANCIAL STAKEHOLDERS.

We use an indicator for whether an employer has no major financial stakeholder, as proxied by their participation in the small-business tax regime *Simples Nacional*.⁶⁷

EMPLOYER SIZE.

We compute total employer size is measured as the log number of full-time equivalent employees during a year.

MUNICIPALITY FES.

We construct dummies for 4,733 municipalities represented in our sample.

SECTOR FES.

We construct dummies for 661 five-digit sectors represented in our sample.

4. *Details on Covariates Related to Amenity Estimates*

We include in our analysis in Section II.B and also as covariates in Z_{gj} in equation (26) of Section V.C the following eight variables that we construct using the RAIS data, in addition to two sets of FEs.

PART-TIME WORK INCIDENCE.

We compute the share of workers with contractual work hours below 40.

WORKING HOURS FLEXIBILITY.

We compute the share of workers who change contractual work hours between any two consecutive years.

⁶⁷Eligibility for the Simples Nacional tax regime requires that the enterprise is a micro- or small business with annual revenues below BRL 1,200,00 (around USD 200,000), that it has no other companies as stakeholders, that it is not internationally owned, that it has no shareholder or partner with significant financial stakes in other companies, and that the enterprise itself has no stake in other companies.

PARENTAL LEAVE GENEROSITY.

We create an indicator for whether workers at an employer have parental leave duration above the national median.

INCOME FLUCTUATIONS.

We compute the share of workers who change mean earnings between any two consecutive years.

WORKPLACE HAZARDS.

We compute the share of workers who report absence from work due to work-related illness, multiplied by 100.

INCIDENCE OF UNJUST FIRINGS.

We compute the share of workers who report ending their job due to an employer-induced firing for no officially recognized cause.

INCIDENCE OF WORKPLACE DEATHS.

We compute the share of jobs who report having ended due to the worker's death at the workplace.

EMPLOYER SIZE.

We compute the total employer size measured as the log number of full-time equivalent employees during a year.

MUNICIPALITY FES.

We construct dummies for 4,733 municipalities represented in our sample.

SECTOR FES.

We construct dummies for 661 five-digit sectors represented in our sample.

5. Detailed Results from the Estimation of Labor Market Objects

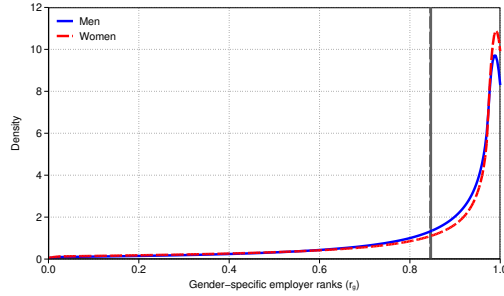


FIGURE E.1. EMPLOYMENT-WEIGHTED DENSITY OF EMPLOYER RANKS, BY GENDER

Note: This figure shows the distributions over employer ranks r_g using gender-specific employment weights separately for men (blue solid line) and women (red dashed line). Employer ranks r_g are defined to be uniformly distributed across firms. Grey vertical patterned lines represent mean values for workers of a given gender.

Source: Model estimates based on RAIS, 2007–2014.

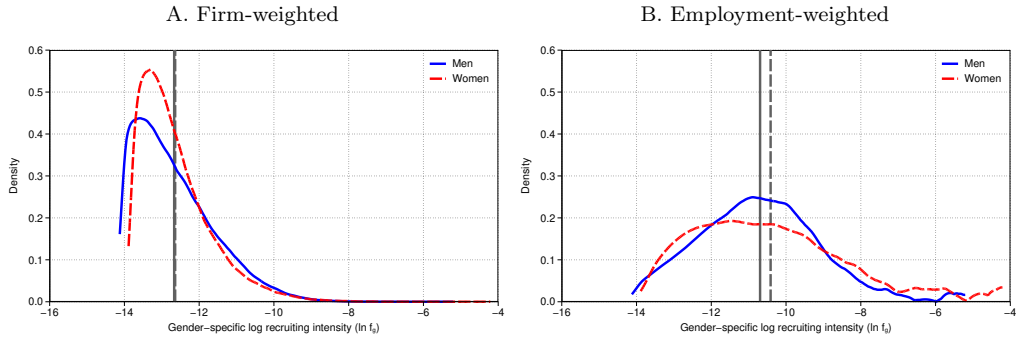


FIGURE E.2. DENSITY OF RECRUITING INTENSITIES, BY GENDER

Note: This figure shows density estimates of logarithmic firm recruiting intensities $\ln f_g = \ln(v_g/V_g)$ separately for men (blue solid line) and women (red dashed line). Panel A is firm-weighted, while Panel B uses gender-specific employment distributions. Grey vertical patterned lines represent mean values for workers of a given gender.

Source: Model estimates based on RAIS, 2007–2014.

6. Detailed Results from the Estimation of on Gender-Specific Firm Types

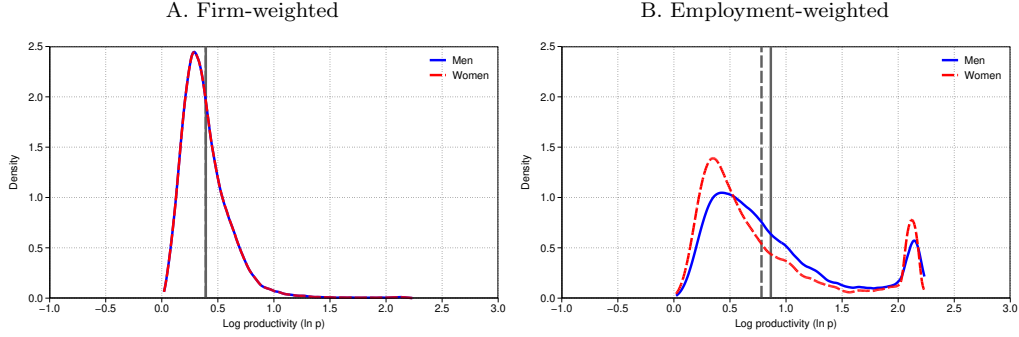


FIGURE E.3. DENSITIES OF LOG ESTIMATED PRODUCTIVITY, BY GENDER

Note: This figure shows densities of log estimated productivity (p) separately for men (blue solid line) and women (red dashed line). Panel A is firm-weighted, while Panel B uses gender-specific employment distributions. Grey vertical patterned lines represent mean values for workers of a given gender.
Source: Model estimates based on RAIS, 2007–2014.

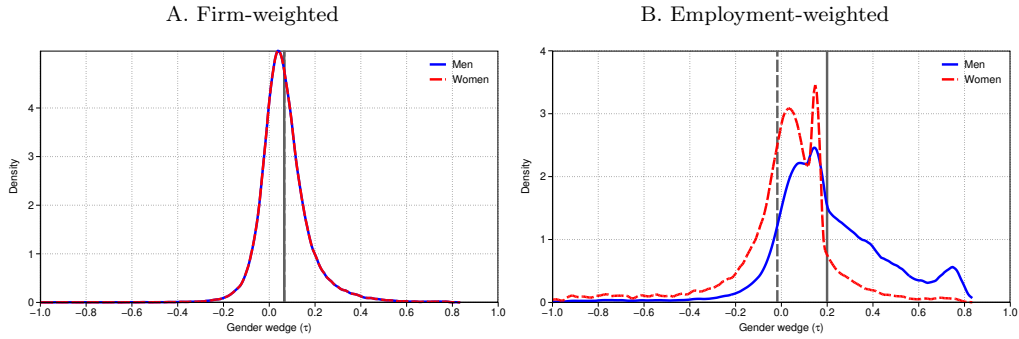


FIGURE E.4. DENSITIES OF ESTIMATED GENDER WEDGES, BY GENDER

Note: This figure shows densities of women's gender wedges (τ) separately for men (blue solid line) and women (red dashed line). Panel A is firm-weighted, while Panel B uses gender-specific employment distributions. Grey vertical patterned lines represent mean values for workers of a given gender.
Source: Model estimates based on RAIS, 2007–2014.

7. Additional Model-Fit Statistics

Figure E.6 shows the model versus data predictions for firm size and pay across 2-digit CNAE industry codes. These results suggest that the model captures well the distribution of pay and firm size across industries for both genders.

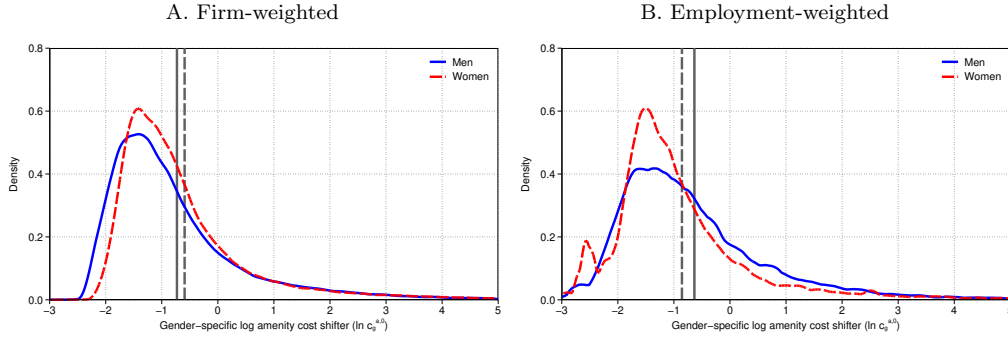


FIGURE E.5. DENSITIES OF LOG AMENITY COST SHIFTERS, BY GENDER

Note: This figure shows densities of log amenity cost shifters ($c_g^{a,0}$) separately for men (blue solid line) and women (red dashed line). Panel A is firm-weighted, while Panel B uses gender-specific employment distributions. Grey vertical patterned lines represent mean values for workers of a given gender.
Source: Model estimates based on RAIS, 2007–2014.

TABLE E.1—CORRELATION TABLE FOR ESTIMATED EMPLOYER PARAMETERS

A. Men							B. Women						
	w_M	a_M	x_M	p	l_M	r_M		w_F	a_F	x_F	$(1-\tau)p$	l_F	r_F
w_M	1.000						w_F	1.000					
a_M	-0.915	1.000					a_F	-0.896	1.000				
x_M	0.246	0.167	1.000				x_F	0.020	0.425	1.000			
p	0.342	0.062	0.985	1.000			$(1-\tau)p$	0.135	0.310	0.974	1.000		
l_M	0.096	0.132	0.552	0.504	1.000		l_F	-0.085	0.332	0.577	0.480	1.000	
r_M	0.225	-0.026	0.486	0.457	0.160	1.000	r_F	0.010	0.173	0.410	0.426	0.161	1.000

C. Cross-gender correlations

	w_g	a_g	x_g	$(1-\tau_g)p_g$	l_g	r_g
Cross-gender correlation	0.909	0.893	0.806	0.770	0.891	0.576

Note: This table reports employment-weighted pairwise correlations across employers between gender-specific pay (w_g), gender-specific amenities (a_g), gender-specific flow utilities (x_g), productivity net of the gender wedge ($(1-\tau_g)p$), gender-specific employment (l_g), and gender-specific employer ranks (r_g) separately for workers of gender $g \in \{M, F\}$. Panel A shows these correlations within the set of employers for men, while Panel B shows the same correlations for women. Panel C shows cross-gender correlations within the same employers.
Source: Model estimates based on RAIS, 2007–2014.

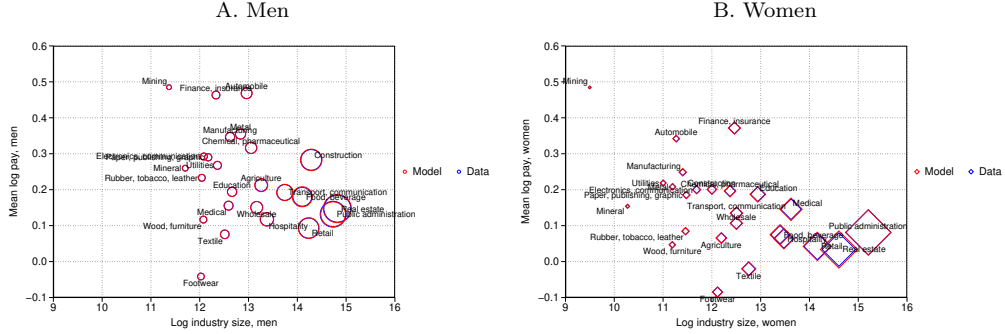


FIGURE E.6. MODEL FIT IN TERMS OF MEAN LOG PAY AGAINST LOG INDUSTRY SIZE, BY GENDER

Note: This figure plots the model fit in terms of mean log pay against log industry size across 25 sectors, separately for men in Panel A and for women in Panel B. Marker sizes represent employment weights.
Source: Model estimates based on RAIS, 2007–2014.

8. Relationship to Other Employer Rank Measures

In this Appendix, we show that our model is consistent with other methods to measure employer ranks, such as poaching ranks (Bagger and Lentz, 2019) or PageRanks (Page et al., 1999; Sorkin, 2018). To this end, we provide the following Proposition.

PROPOSITION 7 (Employer Ranks using Worker Flows): *The ranking $r \in (0, 1)$ can also be identified from worker flows between employers.*

PROOF:

We go through each alternative rank measure.

POACHING RANKS.

The *poaching rank* (Bagger and Lentz, 2019) of every firm j is defined as

$$(E.4) \quad \text{Poaching rank}_j = \frac{\text{number of E-to-E hires}_j}{\text{number of all hires}_j}$$

$$(E.5) \quad = \frac{\text{number of E-to-E hires}_j}{\text{number of E-to-E hires}_j + \text{number of U-to-E hires}_j},$$

which in our model can be rewritten as

$$(E.6) \quad \text{Poaching rank}_j = \frac{\lambda^E G_j + \lambda^G (1 - u)}{\lambda^E G_j + \lambda^G + \lambda^U u}$$

The poaching rank in equation (E.6) is a monotonic transformation of the utility rank of a firm because it is strictly increasing in the cumulative employment distribution G_j , which is precisely the employment-weighted rank of firm j in the pool of all firms.⁶⁸ This proves that, in our model, the Poaching rank is monotonically increasing in firm rank.

PAGERANKS.

Another alternative is to exploit the full pattern of worker flows between employers in order to construct each employer's *PageRank* (Page et al., 1999; Sorkin, 2018), which represents the utility rank of a firm among the pool of all firms.

Let r_j be the rank of firm j . Let $f_j := v_j/V$ be the recruiting intensity of firm j .

Retentions: Let h_j be the mass of workers that firm j retains from one period to the next, defined as

$$(E.7) \quad h_j := (1 - \delta - \lambda^G - \lambda^E(1 - F_j)) l_j.$$

Here, we have used the fact that firms are atomistic. This fact implies, for example, that workers hit with an involuntary job offer at a rate λ^G leave their current employer with probability one (as opposed to the complement of a firm's share of all vacancies in the economy, $1 - f_j$).

Separations: Let $o_{i,j}$ be the *outflow* of workers from firm i to firm j . There are two cases: Either $r_i > r_j$, in which case

$$(E.8) \quad o_{i,j} = \lambda^G l_i f_j,$$

or else $r_i < r_j$, in which case

$$(E.9) \quad o_{i,j} = (\lambda^E + \lambda^G) l_i f_j.$$

Markov Transition Matrix: The Markov transition matrix of the model can be written as follows. Rows are indexed by i and represent workers' current employer. Columns are indexed by j and represent workers' future employers. Fixing the row corresponding to a given firm i , the entries of this row are made up of the diagonal retention probability

$$(E.10) \quad \rho_i = \frac{h_i}{l_i}, \quad \forall j \in \{1, \dots, N\}$$

and the off-diagonal separation probabilities

$$(E.11) \quad p_{i,j} = \frac{o_{i,j}}{l_i} \quad \forall j \neq i, j > 0.$$

⁶⁸Note also that G_j itself is a monotonic transformation of the cumulative offer distribution F_j , which is the vacancy-weighted rank of firm j in the pool of all firms.

Finally, we add the nonemployment state to the Markov transition matrix as firm 0, where the probability of transiting from any firm i to nonemployment is $p_{i,0} = p_0 = \delta$; the probability of finding a job in any firm j when nonemployed is $p_{0,j} = f_j(\lambda^G + \lambda^U)$; and the probability of remaining in nonemployment (i.e., nonemployment's retention probability) is $\rho_0 = p_{0,0} = (1 - \lambda^G - \lambda^U)$. To summarize, the Markov transition matrix looks as follows:

$$(E.12) \quad P = \begin{bmatrix} \rho_0 & p_{0,1} & p_{0,2} & \cdots & p_{1,N} \\ p_0 & \rho_1 & p_{1,2} & \cdots & p_{1,N} \\ p_0 & p_{2,1} & \rho_2 & \cdots & p_{2,N} \\ \vdots & \vdots & \ddots & \vdots & \\ p_0 & p_{N,1} & p_{N,2} & \cdots & \rho_N \end{bmatrix}$$

If we write the system $P' \times l = l$, where l is the $N + 1 \times 1$ vector of firm sizes that includes nonemployment as an additional “firm”, it is easy to see that the above system is the discrete-version equivalent of the continuous Kolmogorov forward equation for the evolution of firm size in the model:

$$(E.13) \quad \dot{l}_j = [-\delta - \lambda^E [1 - F_j] - \lambda^G] l_j + \left[\frac{u + (1 - u)s^E G_r + s^G}{u + (1 - u)s^E + s^G} \right] v_j q.$$

where we can use the discrete time, discrete firms equivalents:

$$(E.14) \quad (1 - u) = \sum_{j \in \{1/N, \dots, 1\}} l_j,$$

$$(E.15) \quad G_j = \frac{1}{1 - u} \sum_{i \in \{1/N, \dots, r\}} l_i,$$

$$(E.16) \quad q = m(\theta)/V = \lambda^U \left[\frac{u + (1 - u)s^E + s^G}{V} \right]$$

$$(E.17) \quad F_j = \sum_{k \in \{1, \dots, N\}} (f_k \mathbf{1}[r(j) > r(k)])$$

$$(E.18) \quad f_j = \frac{v_j}{V}.$$

Therefore, in a steady state,

$$(E.19) \quad 0 = [-\delta - \lambda^E (1 - F_r) - \lambda^G] l_j + [u + (1 - u)s^E G_r + s^G] \lambda^U \frac{v}{V}.$$

$$(E.20) \quad [\delta + \lambda^G + \lambda^E (1 - F_r)] l_j = \frac{v}{V} [\lambda^U u + \lambda^E (1 - u) G_j + \lambda^G]$$

Using the definition of F_j and G_j for discrete firms, we can write:

$$(E.21) \quad \left[\delta + \lambda^G + \lambda^E \left(1 - \sum_{k \in \{1, \dots, N\}} (f_k \mathbf{1}[r(j) > r(k)]) \right) \right] l_j = f_j \left[\lambda^U u + \lambda^G + \lambda^E \sum_{k \in \{1, \dots, N\}} (l_k \mathbf{1}[r(j) > r(k)]) \right].$$

Focusing on a single row, the system $P' \times l = l$ can be written as

$$(E.22) \quad f_j \left[\lambda^U l_0 + \lambda^G + \lambda^E \sum_{k \in \{1, \dots, N\}} (l_k \mathbf{1}[r(j) > r(k)]) \right]$$

$$(E.23) \quad + (1 - \delta - \lambda^G - \lambda^E (1 - \sum_{k \in \{1, \dots, N\}} [f_k \mathbf{1}[r(j) > r(k)]]) l_j = l_j.$$

By substituting $l_0 = u$ by definition, it is immediately apparent that equations (E.21) and (E.23) are identical, proving that this matrix implementation of PageRank solves simultaneously for model sizes l_j that are implied by worker flows, and that are consistent with the model ranking. Solving for l_j yields

$$(E.24) \quad l_j = f_j \left[\frac{\lambda^U u + \lambda^G + \lambda^E \sum_{k \in \{1, \dots, N\}} (l_k \mathbf{1}[r(j) > r(k)])}{\delta + \lambda^G + \lambda^E \left(1 - \sum_{k \in \{1, \dots, N\}} (f_k \mathbf{1}[r(j) > r(k)]) \right)} \right]$$

which makes explicit that the steady-state size l_j of every firm j depends on the relative rank in the ladder, where firms higher up the job ladder hire from a larger set of firms. Further rearranging and substituting for G_j leads us to the discrete equivalent of the steady state firm size in equation (16):

$$(E.25) \quad l_j = f_j \left[\frac{1}{\delta + \lambda^G + \lambda^E (1 - F_j)} \right]^2 (u + s^G) \lambda^U (\delta + \lambda^G + \lambda^E)$$

In short, if we construct the matrix of firm-to-firm flows in the data and solve the following equation,

$$(E.26) \quad P' l = l,$$

where the vector l contains the steady-state firm sizes of equation (E.25), which include hiring intensity f_j . Therefore, dividing firm size l by hiring intensity gives us the rank-implied firm size $\hat{r}_j$. This approach is similar in spirit to Sorkin (2018), who also uses worker flows between firms to identify firm ranks. However, there is an important distinction between our construction of PageRanks and that of Sorkin (2018). Specifically, our definition of PageRank is closer to the original one due to Page et al. (1999), in the sense that our modified adjacency matrix is a Markov transition matrix (i.e., the elements in each row sum to 1), whereas this

is not the case in the discrete-choice setting of [Sorkin \(2018\)](#). Therefore, we have proved that we can use the pattern of worker flows to identify ranks in our model.

9. Comparison across Different Employer Rank Measures

For robustness, we computed a variant of PageRanks ([Page et al., 1999](#); [Sorkin, 2018](#)) and poaching ranks ([Bagger and Lentz, 2019](#)) as described in Supplemental Appendix E.8.⁶⁹ The two measures—PageRanks and poaching ranks—are related insofar as poaching ranks can be viewed as the result of a single application of the PageRanks’ transition matrix to an initial guess of ones for all employers and zero for nonemployment. Viewed this way, PageRanks are obtained as the infinite (or converged) iteration of the same transition matrix that generated poaching ranks in the first iteration. See also the discussion in [Lentz \(2024\)](#). To be transparent, we will compare both PageRanks and poaching ranks to our baseline rank measure, which is employer size. Tests of comparing our baseline rank measure of size to both PageRanks and poaching ranks are presented in Table E.2 below. We find that our model-consistent PageRank estimates and firm sizes align quite closely in the data, with correlations always above 0.87. Poaching Ranks are still positively correlated to both size and PageRank, albeit more weakly.

TABLE E.2—RANK CORRELATIONS BETWEEN ALTERNATIVE EMPLOYER RANK MEASURES, BY GENDER

	Men			Women		
	Size	PageRank	Poaching rank	Size	PageRank	Poaching rank
Size	1.000			1.000		
PageRank	0.872	1.000		0.886	1.000	
Poaching rank	0.258	0.263	1.000	0.199	0.218	1.000

Note: This table reports Spearman rank correlations between three alternative employer rank measures, separately by gender. The three rank measures are employer size as used in the paper’s baseline analysis, PageRanks ([Page et al., 1999](#); [Sorkin, 2018](#)), and poaching ranks ([Bagger and Lentz, 2019](#)). See Supplemental Appendix E.8 for details of how these are constructed.

Source: RAIS, 2007–2014.

To see what factors load differently onto each employer rank measure, we

⁶⁹Importantly, the variant of PageRanks that we use is consistent with our labor market model and closely parallels the original PageRanks proposed by [Page et al. \(1999\)](#). It is worth noting that our variant of PageRanks differs from that introduced by [Sorkin \(2018\)](#) in that our measure reflects the fixed-point weights assigned to each employer based on the empirical transition matrix—i.e., the adjacency matrix normalized such that all rows sum to unity. This is in contrast to the approach in [Sorkin \(2018\)](#) that constructs an alternative set of PageRanks defined based on the adjacency matrix normalized such that all columns (i.e., not rows) sum to unity. While the latter is the appropriate normalization in the discrete-choice setting of [Sorkin \(2018\)](#), we show in Supplemental Appendix E.8 of our paper that the original definition of PageRanks due to [Page et al. \(1999\)](#) is the one consistent with our labor market model. In addition, we compute poaching ranks following the exact definition in ([Bagger and Lentz, 2019](#)).

compare our alternative employer rank measures across sectors. Specifically, we compute the mean PageRank (poaching index) ranks onto mean employer size by two-digit sector in the RAIS data. We then assess to what extent the two different rank measures align overall, and how the degree of alignment varies across sectors. Figure E.7 shows the results of this exercise. Panel A projects Pagerank ranks onto employer size ranks by sector for men, while Panel B does the same for women. Overall, there is a strong correlation between the two measures, as evidenced by the steep slope and relatively little dispersion around the linear best-fit line. PageRanks and employer size also agree on the relative ranking of employers in several key sectors—e.g., the public sector for women is ranked near the very top according to both PageRanks and employer size. Panels C and D repeat the same analysis for the relationship between poaching ranks and employer size ranks across sectors for men and women, respectively. The slope is notably flatter and dispersion is notably greater, compared to the comparison between PageRanks and employer size in the panels A–B, consistent with the rank correlations reported in Table E.2 above. Nevertheless, there is a strong correlation between the two measures. Still, there are some important exceptions—e.g., the public sector for women is ranked near the top according to employer size but only above average according to poaching ranks. This may reflect differences in the quantity of worker outflows as well as the quality of worker in- and outflows of various sectors. It is reassuring that PageRanks, which take into account the entire empirical transition matrix of worker flows between employers, line up remarkably closely with our baseline analysis based on employer size.

We re-estimate our entire model based on a cardinal interpretation of PageRanks instead of employer size. To this end, we leverage the interpretation of PageRanks as the steady-state employer size implied by worker flows between firms (Page et al., 1999). This may be a preferable measure of employer rank if the cross-sectional distribution of employer sizes is contaminated by empirical factors such as history dependence—e.g., some employers used to be desirable to work and thus large, but they are no longer desirable to work at now after the realization of some shocks, though they remain large during a transition to their new steady-state size. Table E.3 compares the estimation results based on PageRanks to our baseline estimation results based on employer size as a measure of employer ranks. We find that the main properties of our estimates are remarkably stable. In particular, amenities estimated using PageRanks are strongly correlated to our baseline amenity estimates. Productivity is also strongly correlated, though it is estimated to be more dispersed when using PageRanks than under our baseline estimation. Similarly, gender wedges are estimated to be more dispersed when using PageRanks to rank firms.

Building on these new model estimates, Table E.4 recomputes the Kitaga-Oaxaca-Blinder decomposition of gender gaps in pay, amenities, and total compensation based on PageRanks as an alternative employer rank.⁷⁰

⁷⁰It is worth noting that poaching ranks (Bagger and Lentz, 2019) have no cardinal interpretation in

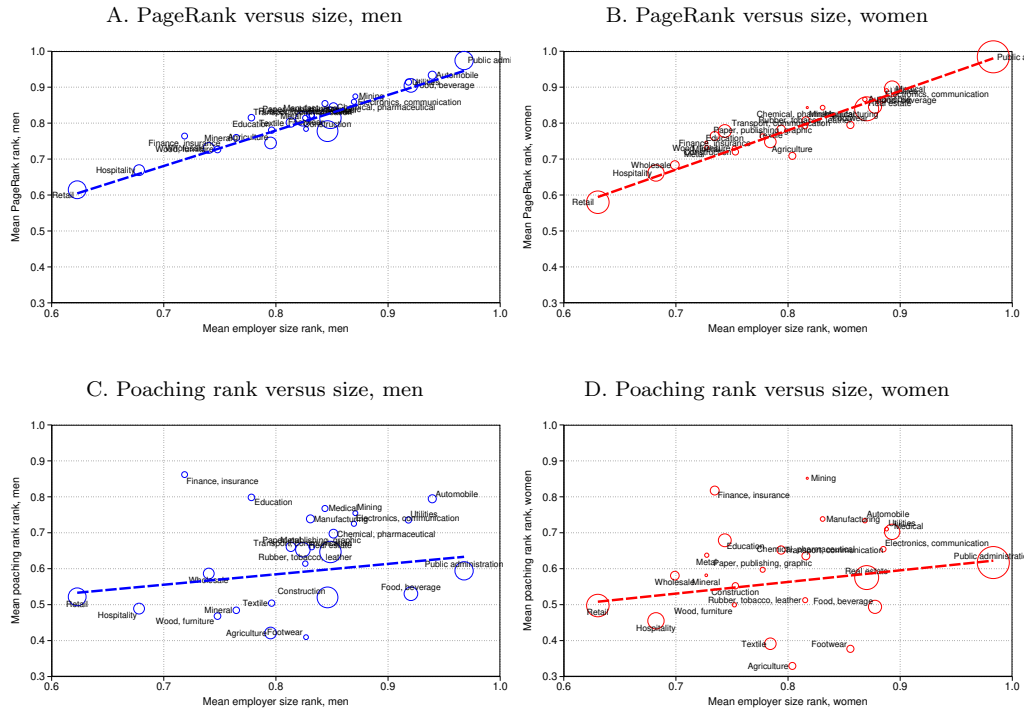


FIGURE E.7. COMPARING ALTERNATIVE EMPLOYER RANK MEASURES ACROSS SECTORS, BY GENDER

Note: This figure shows the means of alternative employer ranks—either PageRanks (Page et al., 1999; Sorkin, 2018) or poaching ranks (Bagger and Lentz, 2019)—against employer size ranks, which is the baseline rank measure in the paper. See Supplemental Appendix E.8 for details of how these are constructed.
Source: RAIS, 2007–2014.

TABLE E.3—ESTIMATION RESULTS UNDER PAGERANKS AS ALTERNATIVE MEASURE OF EMPLOYER RANKS

	Baseline	PageRank-based
η_a	5.721	6.960
η_v	2.062	2.500
Mean of p	2.883	3.158
Variance of p	4.648	8.481
Mean of a_m	1.193	1.514
Variance of a_m	0.089	0.128
Mean of a_f	1.245	1.445
Variance of a_f	0.081	0.141
Mean of τ	−0.018	−0.125
Variance of τ	0.097	0.276
Correlation between p_{baseline} and alternative		0.859
Correlation between $a_{m,\text{baseline}}$ and alternative		0.998
Correlation between $a_{f,\text{baseline}}$ and alternative		0.999
Correlation between τ_{baseline} and alternative		0.801

Note: This table compares estimation results from our baseline model to one based on PageRanks as an alternative measure of employer ranks.

Source: Model estimates based on RAIS, 2007–2014.

TABLE E.4—KITAGAWA-OAXACA-BLINDER DECOMPOSITIONS—ALTERNATIVE EMPLOYER RANKS BASED ON PAGERANKS

	Gender gap	Between-employer gap		Within-employer gap	
		Level	Share (%)	Level	Share (%)
Pay	0.122	0.109	90.0	0.012	10.0
Amenity-valuation	−0.039	−0.099	255.1	0.060	−155.1
Total compensation	0.083	0.010	12.6	0.072	87.4

Note: This table shows results from the Kitagawa-Oaxaca-Blinder decomposition of the gender gaps in log pay ($\ln w$), log amenity valuations ($\ln \tilde{a} = \ln(x/w)$), and log total compensation ($\ln x$) when re-estimating our model based on PageRanks as an alternative employer rank measure.

Source: Model estimates based on RAIS, 2007–2014.

The results based on PageRanks as a measure of employer ranks are in line with our baseline results shown in Table 11 of the main text. As a starting point, we have a slightly higher gender pay gap, due to the fact that our selection criteria depend on the rank measure. Comparing the amenity and total-compensation estimates, we find that all of our results are qualitatively unchanged across the two estimation methods. In particular, we continue to find an amenity-valuation gap in favor of women. Quantitatively, we find that men’s within-employer amenity advantage is more pronounced than in our baseline, thereby leading to a lower magnitude of the negative gender gap in amenity valuations. Still, this results in a total-compensation gap that is significantly lower than the gender pay gap and mostly within (rather than between) employers. This suggests that our baseline analysis based on employer size as a measure of employer rank is robust to alternative, worker-flows-based measures of employer ranks.

10. Estimation Results Under Heterogeneous Separation Rates

To estimate our model with separation rates that are heterogeneous across firms, we use the modified procedure detailed in Supplemental Appendix D.11. Our results are summarized in Table E.5. We find that the estimated productivity distribution is more dispersed when separation rates are allowed to be heterogeneous across firms. Estimated amenities are also more dispersed; however, their estimated mean is lower, suggesting that a share of utility that is explained by amenities in the baseline model can be explained by heterogeneity in separation rates. We also find that heterogeneous separation rates differ more for men than for women between firms, which affects men’s results more than women’s. Thus, the gap in amenities between men and women widens compared to our baseline estimation. Finally, all estimates are strongly correlated to our baseline estimates, suggesting that though heterogeneity in separation rates qualitatively affects our results, quantitatively, the difference is relatively small.

To assess the extent to which differences in estimation results affect the substantive insights from our analysis, Table E.6 below shows the results of Kitagawa-Oaxaca-Blinder decompositions based on the model with heterogeneous firm-level separation rates ($\delta(r)$).

Comparing the two sets of results—i.e., those in the model with heterogeneous firm-level separation rates in Table E.6 compared to those in our baseline model in Table 11—we find two main results. First, the amenity-valuation gap is slightly more pronounced, implying that amenities play a slightly more important role

our model. The reason is intuitive: firm size and PageRanks both reflect different notions of the *number* of workers that a firm manages to attract and retain in steady state, while poaching ranks reflect the *share* of workers that a firm poaches from competitors *relative to all hires at that firm*. As such, poaching ranks normalize by total hiring, which removes cardinal information from this employer rank, making it impossible to recover firms’ labor demand that can be mapped into composite productivity in our equilibrium framework. Thus, the same model inversion that we use for firm size in the baseline analysis (Table 11) and for PageRanks in a robustness exercise (Table E.4) is technically not possible for poaching ranks.

TABLE E.5—ESTIMATION RESULTS UNDER HETEROGENEOUS SEPARATION RATES

	Baseline	Heterogeneous separations
η_a	5.721	5.306
η_v	2.062	1.927
Mean of p	2.883	2.850
Variance of p	4.648	4.408
Mean of a_m	1.193	1.105
Variance of $\log a_m$	0.089	0.091
Mean of a_f	1.245	1.168
Variance of a_f	0.081	0.067
Mean of τ	−0.018	0.001
Variance of τ	0.097	0.084
Correlation between p_{baseline} and alternative		1.000
Correlation between $a_{m,\text{baseline}}$ and alternative		0.995
Correlation between $a_{f,\text{baseline}}$ and alternative		0.998
Correlation between τ_{baseline} and alternative		0.999

Note: This table compares estimation results from our baseline analysis with homogeneous separation rate (δ) across firms to the case in which separation rates ($\delta(r)$) are allowed to be nonincreasing in firm ranks.
Source: Model estimates based on RAIS, 2007–2014.

TABLE E.6—KITAGAWA-OAXACA-BLINDER DECOMPOSITIONS—
HETEROGENEOUS FIRM-LEVEL SEPARATION RATES ($\delta(r)$)

	Gender gap	Between-employer gap		Within-employer gap	
		Level	Share (%)	Level	Share (%)
Pay	0.112	0.089	79.5	0.023	20.5
Amenity-valuation	−0.080	−0.089	111.7	0.009	−11.7
Total compensation	0.032	0.000	0.1	0.032	99.9

Note: This table shows results from the Kitagawa-Oaxaca-Blinder decomposition of the gender gaps in $\log$ pay ($\ln w$), $\log$ amenity valuations ($\ln \tilde{a} = \ln(x/w)$), and $\log$ total compensation ($\ln x$) when re-estimating our model with heterogeneous firm-level separation rates ($\delta(r)$) that are nonincreasing across firm ranks r .
Source: Model estimates based on RAIS, 2007–2014.

in explaining gender differences in total compensation. Second, the relative contribution of between- versus within-components of all three compensation concepts is similar across both models. Altogether, this analysis suggests that a model with heterogeneous firm-level separation rates yields results that are largely in line and, if anything, slightly more pronounced than what we find using our baseline model.

11. *Sensitivity of Estimation Results to Alternative Amenity Cost Shares*

Here, we further investigate the robustness of our results with respect to our estimation target of the cost share of amenities in value added, taken from [Bieri, Kuminoff and Pope \(2023\)](#). We acknowledge that this is a notoriously difficult moment to find. While an amenity cost share of value added of 8 percent is maybe a reasonable target—absent a better target from the literature—it is useful to check the sensitivity of our results to alternative targets by re-estimating our model to different amenity cost shares and evaluating the resulting model predictions.

We repeated our estimation for three different targets for the cost share of amenities: 4 percent, 12 percent and 16 percent. Our results are summarized in Table E.7, where we compare all these results to the baseline. Obviously, the assumed value of the cost share of amenities makes a difference for the estimated value of η_a , which is decreasing in the assumed cost share. However, we find that different values of the cost share of amenities—and therefore different values of the elasticity of amenity costs η_a —make relatively little difference for our estimates of productivity. We also find that the different values of the cost share of amenities make no difference at all for our estimates of amenities, which we discuss below. A higher cost share of amenities leads to a higher estimated average productivity, with the biggest difference in mean productivity from our baseline being less than 14 percent when the cost share of amenity doubles, a case we do not believe to be empirically relevant. The variance of gender wedges is slightly more affected in relative terms, but the impact remains small. All sets of estimates remain strongly correlated across estimation sets. We take this as evidence that our results are not strongly influenced by the specific choice of target for the economy-wide cost share of amenities.

It is important to note that the distribution of amenities remains exactly unchanged across estimation sets for both genders, as the elasticity of the amenity creation cost is not used in the estimation of firm-level profits per worker, which determine estimated utilities and, in turn, estimated amenities. Intuitively, the elasticity of the amenity cost function does not affect our estimation of composite productivity $\tilde{p}$ and thus does not affect our estimation of total compensation x and amenity valuations βa .⁷¹ To confirm this intuition, after re-estimating our model under alternative targets for the amenity share, we assess the robustness of our

⁷¹The elasticity of the amenity cost function matters only for the split of composite productivity $\tilde{p}$ into productivity p and the net surplus due to amenity valuations net of the amenity cost $\beta a - c^a(a)$.

TABLE E.7—ROBUSTNESS: ESTIMATION RESULTS UNDER DIFFERENT ASSUMED COST SHARES OF AMENITIES

Assumed cost share value	0.08 (Baseline)	0.04	0.12	0.16
η_a	5.721	11.443	3.814	2.861
η_v	2.062	2.062	2.062	2.062
Mean of p	1.527	1.421	1.632	1.738
Variance of p	0.245	0.256	0.234	0.225
Mean of a_m	1.209	1.209	1.209	1.209
Variance of a_m	0.071	0.071	0.071	0.071
Mean of a_f	1.173	1.173	1.173	1.173
Variance of a_f	0.062	0.062	0.062	0.062
Mean of τ	0.065	0.066	0.064	0.063
Variance of τ	0.015	0.017	0.013	0.011
Correlation between p_{baseline} and alternative		0.999	0.999	0.996
Correlation between $a_{m,\text{baseline}}$ and alternative		1.000	1.000	1.000
Correlation between $a_{f,\text{baseline}}$ and alternative		1.000	1.000	1.000
Correlation between τ_{baseline} and alternative		0.998	0.998	0.992

Note: This table shows summary statistics for the distributions of estimated productivity p , amenities a_m and a_f , gender wedge τ and economy-wide parameters η_a and η_v , depending on the assumed value of the cost share of amenities.

Source: Model estimates based on RAIS, 2007–2014.

Kitagawa-Oaxaca-Blinder decomposition of pay, amenities, and total compensation. Table E.8 shows the results from this decomposition based on half of the baseline’s amenity share target (i.e., 4 percent). The results are identical to those in our baseline.

TABLE E.8—KITAGAWA-OAXACA-BLINDER DECOMPOSITIONS—HALF BASELINE’S AMENITY SHARE TARGET

	Gender gap	Between-employer gap		Within-employer gap	
		Level	Share (%)	Level	Share (%)
Pay	0.112	0.089	79.5	0.023	20.5
Amenity-valuation	−0.077	−0.087	113.5	0.010	−13.5
Total compensation	0.035	0.002	6.1	0.033	93.9

Note: This table shows results from the Kitagawa-Oaxaca-Blinder decomposition of the gender gaps in log pay ($\ln w$), log amenity valuations ($\ln \tilde{a} = \ln(x/w)$), and log total compensation ($\ln x$) when re-estimating our model with half the baseline’s amenity share target (i.e., 0.04).

Source: Model estimates based on RAIS, 2007–2014.

Analogously, Table E.9 shows the results from this decomposition based on double the baseline’s amenity share target (i.e., 16 percent). Again, the results are identical to those in our baseline.

TABLE E.9—KITAGAWA-OAXACA-BLINDER DECOMPOSITIONS—DOUBLE BASE-
LINE’S AMENITY SHARE TARGET

	Gender gap	Between-employer gap		Within-employer gap	
		Level	Share (%)	Level	Share (%)
Pay	0.112	0.089	79.5	0.023	20.5
Amenity-valuation	−0.077	−0.087	113.5	0.010	−13.5
Total compensation	0.035	0.002	6.1	0.033	93.9

Note: This table shows results from the Kitagawa-Oaxaca-Blinder decomposition of the gender gaps in log pay ($\ln w$), log amenity valuations ($\ln \tilde{a} = \ln(x/w)$), and log total compensation ($\ln x$) when re-estimating our model with double the baseline’s amenity share target (i.e., 0.16).

Source: Model estimates based on RAIS, 2007–2014.

Altogether, this gives us confidence that—although there is significant uncertainty around the estimate of the amenity cost elasticity η_a based on the assumed amenity cost share—our main conclusions are robust to its exact choice.

F. GENDER-SPECIFIC COMPENSATION STRUCTURES ACROSS EMPLOYERS APPENDIX

1. Alternative Kitagawa-Oaxaca-Blinder Decompositions

Recall that in Table B.1 of Supplemental Appendix B.2, we presented alternative Kitagawa-Oaxaca-Blinder decompositions of the gender log pay gap. Here, we present analogous decompositions for the gender gaps in amenities (Table F.1) and utility (Table F.2).

TABLE F.1—ALTERNATIVE KITAGAWA-OAXACA-BLINDER DECOMPOSITIONS OF THE GENDER GAP IN AMENITIES

	Gender log amenities gap	Between-employer gap		Within-employer gap	
		Level	Share (%)	Level	Share (%)
Decomposition 1	−0.077	−0.087	113.5	0.010	−13.5
Decomposition 2	−0.077	−0.090	117.3	0.013	−17.3

Note: This table shows results from the Kitagawa-Oaxaca-Blinder decomposition of the overall gender log amenities gap into a between-employer log amenities gap and a within-employer log amenities gap. Decomposition 1 corresponds to equation (B.2) and uses women’s estimates of log amenities for computing the between-employer component. Decomposition 2 corresponds to equation (B.3) and uses men’s estimates of log amenities for computing the between-employer component.

Source: Model estimates based on RAIS, 2007–2014.

TABLE F.2—ALTERNATIVE KITAGAWA-OAXACA-BLINDER DECOMPOSITIONS OF THE GENDER GAP IN UTILITY

	Gender log total compensation gap	Between-employer gap		Within-employer gap	
		Level	Share (%)	Level	Share (%)
Decomposition 1	0.035	0.002	6.1	0.033	93.9
Decomposition 2	0.035	−0.017	−48.4	0.053	148.4

Note: This table shows results from the Kitagawa-Oaxaca-Blinder decomposition of the overall gender log utility gap into a between-employer log utility gap and a within-employer log utility gap. Decomposition 1 corresponds to equation (B.2) and uses women’s estimates of log utility for computing the between-employer component. Decomposition 2 corresponds to equation (B.3) and uses men’s estimates of log utility for computing the between-employer component.

Source: Model estimates based on RAIS, 2007–2014.

2. Different Margins of Gender Discrimination

In classical theories of taste-based discrimination (e.g., Becker, 1971), employers receive disutility from employing certain population groups, which affects the employer’s recruiting decisions vis-à-vis these groups. In our framework, the gender wedge τ plays precisely this role. In addition, our framework features frictional pay

dispersion across employers and allows employers to choose amenities separately for each gender.⁷² This highlights amenities as a novel margin of employer “discrimination” (i.e., unequal treatment) across genders. Under the hypothesis that employers use amenities to differentiate between workers, one would expect gender wedges and women’s amenity values a_F to be negatively related. Indeed, we confirm such a negative relationship, shown in Figure F.1.

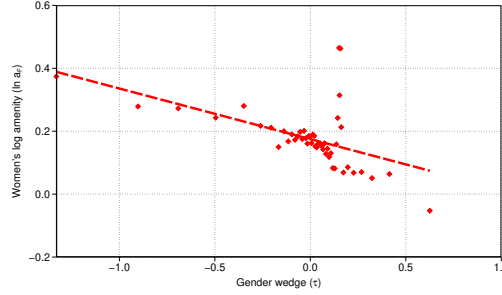


FIGURE F.1. NEGATIVE RELATION BETWEEN WOMEN’S AMENITIES AND GENDER WEDGES

Note: This figure shows a binned scatter plot of women’s log amenities ($\ln a_F$) against gender wedges (τ), with the linear best fit line in dashed red being weighted by female employment.
Source: Model estimates based on RAIS, 2007–2014.

3. The Role of the Public Sector

To investigate the role of the public sector in our analysis, we have classified each employer in our sample into public vs. private. We then complement Figure 6 from the main text by computing the shares of employment in the public sector across firm ranks separately by gender. The results are shown in Figure F.2 below.

From this, see that the public sector plays an important role in accounting for high-ranked employers, and more so for women. Note that the top 4% of employment for women and the top 2% of employment for men are all concentrated in the public sector. Below those thresholds, public- and private-sector firms are mixed. In other words, 96% of the firm ladder for women and 98% of the firm ladder for men feature a combination of public and private-sector employers. This result echoes the findings from Section VI.A and, in particular, Figure 3 of the main text.

These findings suggest that the public sector clearly plays an important role, complementing existing academic studies on the role of public-sector employment

⁷²We have in mind the costly provision of amenities such as parental leaves, which have large gender differences in take-up rates.

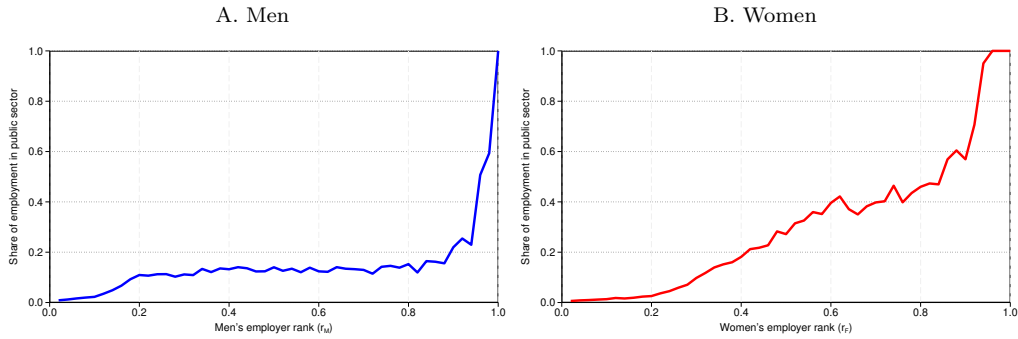


FIGURE F.2. SHARE OF EMPLOYMENT IN PUBLIC SECTOR ACROSS FIRM RANKS, BY GENDER

Note: This figure plots the share of employment in the public sector across firm ranks, separately for men in Panel A and for women in Panel B.
Source: Model estimates based on RAIS, 2007–2014.

in relation to gender in the labor market. As such, we investigate to what extent our results are shaped by the public vs. private sector distinction by recomputing the same Kitagawa-Oaxaca-Blinder decompositions separately for the whole economy, the private sector only, and the public sector only. Table 11 in the main text showed that amenities played an important role in understanding gender gaps in firm pay in the overall economy. Table F.3 below estimates the same Kitagawa-Oaxaca-Blinder decompositions on the subset of workers and firms in Brazil’s private sector. We find a baseline gender pay gap that is approximately the same as for the whole economy, an amenity-valuation gap of the same sign and only slightly smaller magnitude than for the whole economy, and a total-compensation gap similar to that for the whole economy. The split into between versus within components is qualitatively unchanged.

TABLE F.3—KITAGAWA-OAXACA-BLINDER DECOMPOSITIONS—PRIVATE SECTOR ONLY

	Gender gap	Between-employer gap		Within-employer gap	
		Level	Share (%)	Level	Share (%)
Pay	0.112	0.074	66.3	0.038	33.7
Amenity-valuation	−0.065	−0.062	95.6	−0.003	4.4
Total compensation	0.047	0.012	25.4	0.035	74.6

Note: This table shows results from the Kitagawa-Oaxaca-Blinder decomposition of the gender gaps in log pay ($\ln w$), log amenity valuations ($\ln \tilde{a} = \ln(x/w)$), and log total compensation ($\ln x$) for the private sector only.

Source: Model estimates based on RAIS, 2007–2014.

Table F.4 repeats the same decomposition for the subset of workers and firms in Brazil’s public sector. We find a baseline gender pay gap that is a bit smaller than for the whole economy, an amenity-valuation gap that is of the same sign but a bit smaller in magnitude than in the whole economy, and a total-compensation gap that is a bit larger than in the whole economy. The split into between versus within components is qualitatively unchanged, with the exception of the amenity-valuation gap, which is more concentrated between rather than within employers in Brazil’s public sector. These results are intuitive since we naturally expect pay gaps within the same public-sector organizations to be more muted by design. The public sector is also commonly viewed as a high-amenity sector, especially for women, rationalizing the larger within-employer gap in amenity valuations.

TABLE F.4—KITAGAWA-OAXACA-BLINDER DECOMPOSITIONS—PUBLIC SECTOR ONLY

	Gender gap	Between-employer gap		Within-employer gap	
		Level	Share (%)	Level	Share (%)
Pay	0.090	0.099	110.4	−0.009	−10.4
Amenity-valuation	−0.055	−0.094	171.8	0.039	−71.8
Total compensation	0.035	0.005	13.7	0.030	86.3

Note: This table shows results from the Kitagawa-Oaxaca-Blinder decomposition of the gender gaps in log pay ($\ln w$), log amenity valuations ($\ln \tilde{a} = \ln(x/w)$), and log total compensation ($\ln x$) for the public sector only.

Source: Model estimates based on RAIS, 2007–2014.

4. Implications for Productivity

Panel A of Figure F.3 shows the mean productivity p at different rungs of men’s and women’s firm ladders. Productivity is more steeply increasing in employer ranks for men than for women. The differences are meaningful. For example, men’s median-ranked employer is 35 log points more productive than bottom-ranked employers, while the same statistic is only 21 log points for women. Thus, for men more so than for women, improvements in labor market efficiency yield productivity gains. Panel B plots women’s productivity net of the gender wedge $((1 - \tau)p)$ against men’s productivity (p). Among the least productive firms, gender wedges are close to zero but they steadily increase toward higher productivity levels. Again, the magnitudes are noteworthy. Near the top of the productivity distribution, the average gender wedge accounts for up to 50 log points of productivity.

While not targeted, our model generates a lower elasticity of pay with respect to productivity for women (0.094) than for men (0.174). In our model, rent sharing is determined endogenously due to the combination of three gender-specific factors

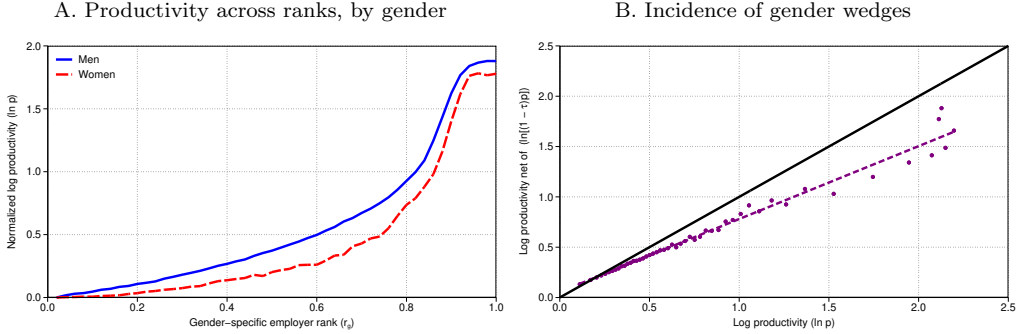


FIGURE F.3. PRODUCTIVITY, RANKS, AND GENDER WEDGES

Note: Panel A of this figure shows gender-specific employment-weighted log productivity ($\ln p$) across employer ranks (r_g) separately by gender $g \in \{M, F\}$. Panel B shows a binned scatter plot of women's productivity net of gender wedges ($\ln[(1-\tau)p]$) against men's log productivity ($\ln p$), with the linear best fit line in dashed purple being weighted by total (i.e., male plus female) employment.

Source: Model estimates based on RAIS, 2007–2014.

that we separately identify: compensating differentials (Rosen, 1986), taste-based discrimination (Becker, 1971), and labor market frictions (Manning, 2003).

5. Switching Employment and Compensation Policies Across Genders

Next, we relocate all workers of one gender into the other gender's employers, while keeping compensation policies constant, and vice versa. In an accounting sense, we then ask whether either men or women would prefer the other gender's employment distribution or compensation policies.

We first shift men's employment distribution to that of women: $l_M \mapsto l_F$, the results of which are presented in Panels A–C of Figure F.4. In the baseline, men's CDF of employment is represented by the 45-degree line in each panel. After moving men into women's firms, the solid blue line indicates the simulated CDF of employment. Such a move reduces men's welfare by 0.2 log points overall, consisting of a 8.9 log points loss in pay and a 9.2 log points gain in amenities. Interestingly, workers in the top 15% of the utility distribution actually prefer women's employment distribution. If, instead, we keep men's employment constant but change firm pay and amenities to those for women, men's total compensation decreases by 5.3 log points, consisting of a decrease in amenities by 7.7 log points and a decrease in pay by 3.9 log points. This experiment demonstrates that there are large differences in the treatment of men and women within employers.

Next, we shift women's employment distribution to that of men: $l_F \mapsto l_M$, the results of which are presented in Panels D–F of Figure F.4. As a result of this shift, women's welfare decreases by 1.7 log points, consisting of a 7.3 log points increase in pay but a 12.4 log points decrease in amenities. If, instead, we keep women's employment constant but change firm pay and amenities to those of

men, women's amenities increase by 4.5 log points and their pay increases by 2.3 log points. As a result, women's total compensation increases by 3.3 log points, accounting for almost all of the gender utility gap, which again highlights the importance of unequal treatment within employers.

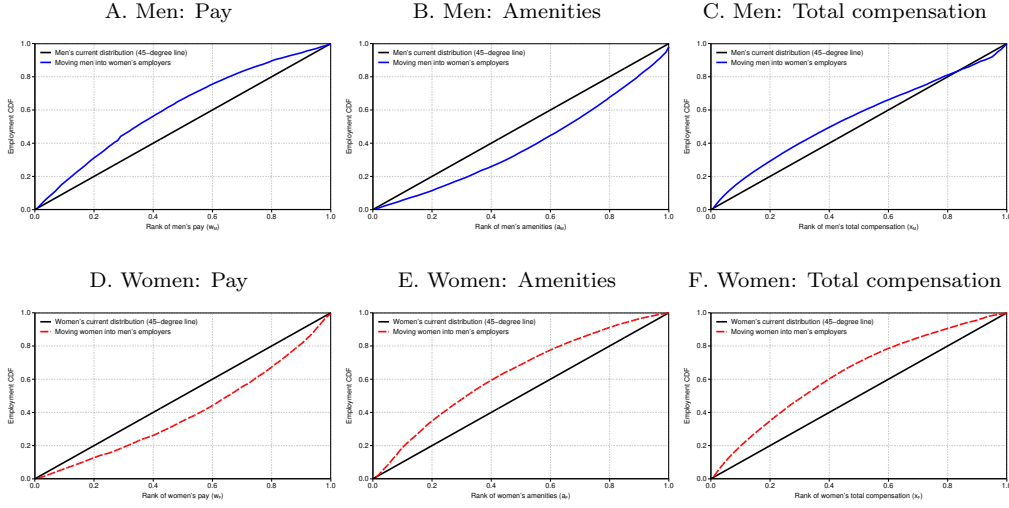


FIGURE F.4. OUTCOMES ASSOCIATED WITH MOVING MEN INTO WOMEN'S EMPLOYERS AND VICE VERSA

Note: This figure shows men's CDF over pay (w_M) in Panel A, amenities (a_M) in Panel B, and total compensation (x_M) in Panel C in the baseline as the diagonal solid black line and after moving men into women's employers as the solid blue line. Analogously, it shows women's CDF over pay (w_F) in Panel D, amenities (a_F) in Panel E, and total compensation (x_F) in Panel F in the baseline as the diagonal solid black line and after moving women into men's employers as the dashed red line.

Source: Model estimates based on RAIS, 2007–2014.

G. EQUILIBRIUM COUNTERFACTUALS APPENDIX

1. Numerical Solution Algorithm to Solve Baseline Equilibrium

Firstly, we feed to the model the estimated labor market parameters $\{\lambda_M^U, \lambda_F^U, s_M^E, s_F^E, s_M^G, s_F^G, \delta_M, \delta_F\}$ and the firm-level estimates of $\{p, a_M, a_F, \tau, c_M^{v,0}, c_F^{v,0}\}$. Then, we rank firms according to composite productivity $\tilde{p}_g$ for each gender. This is useful because, as stated in Lemma 3, firms with higher composite productivity $\tilde{p}_g$ offer higher utility x_g .

We must first find the equilibrium level of aggregate vacancies V_g . We invert the equation for the offer arrival rate from unemployment in (14) to obtain:

$$(G.1) \quad V_g = U_g \left(\frac{\lambda_g^U}{\chi_g} \right)^{1/\alpha}.$$

We recursively apply the discrete version of the FOCs of the firm, as in equations (D.58) and (D.61):

(G.2)

$$\Delta F_g(x(\tilde{p}_{gr})) = \left[\frac{1}{c_g^{v,0}} \frac{T_g(\tilde{p}_{gr-1} - x(\tilde{p}_{gr-1}))}{(\delta_g + \lambda_g^G + \lambda_g^E(1 - F_g(x(\tilde{p}_{gr-1}))))^2} \right]^{\frac{1}{\eta^v - 1}} \frac{1}{V_g N}$$

(G.3)

$$\Delta x_g(\tilde{p}_{gr}) = \frac{2\lambda_g^E(\tilde{p}_{gr-1} - x_g(\tilde{p}_{gr-1}))}{\delta_g + \lambda_g^G + \lambda_g^E(1 - F_g(x_g(\tilde{p}_{gr-1})))} \left[\frac{1}{c_g^{v,0}} \frac{T_g(\tilde{p}_{gr-1} - x_g(\tilde{p}_{gr-1}))}{(\delta_g + \lambda_g^G + \lambda_g^E(1 - F_g(x_g(\tilde{p}_{gr-1}))))^2} \right]^{\frac{1}{\eta^v - 1}} \frac{1}{V_g N},$$

where N is the total number of firms in our data. Then, we calculate total vacancies obtained in equilibrium as $V_g^* = \sum_r v_g(\tilde{p}_{gr})/N$. We solve the algorithm setting the initial conditions $x_g(\tilde{p}_{g0}) = \phi_g$ and $F_g(\tilde{p}_{g0}) = 0$, and we loop over the vacancy cost shifter $c_g^{v,0}$ until we obtain that $V_g^* = V_g$.⁷³ In the baseline simulation in which we plug in the parameter estimates from the data, our solution algorithm produces gender-specific firm-level flow utility $x_g(\tilde{p}_{gr})$, easily converted to wages $w_g(\tilde{p}_{gr}) = x_g(\tilde{p}_{gr}) - a_g(\tilde{p}_{gr})$, gender-specific firm-level recruiting intensities $v_g(\tilde{p}_{gr})$ and gender-specific firm-level ranks in the offer distribution $F_g(x_g(\tilde{p}_{gr}))$ that are identical to those observed in the data, up to machine precision.

In all counterfactual simulations, we keep constant all firm-level parameters and the economy-wide parameters that are not explicitly mentioned as modified in the exercise, including the gender-specific cost shifters $c_g^{v,0}$. Also, we recalculate the workers' outside option in the counterfactuals by finding a new solution to equation (4) by gender. This allows the equilibrium job-finding probability λ_g^U , and therefore equilibrium employment, to respond to changes in the economy in counterfactual simulations.

⁷³Alternatively, we could plug in the cost shifter we have estimated in Step 5 of our identification strategy, but this approach is equivalent in the baseline solution and allows us to generalize to cases in which we set λ_g^U to a value that is different from the one we estimated in the data.

2. Numerical Solution Algorithm to Solve Policy Counterfactuals

When we simulate the equal pay policy and the equal-hiring policy, we can no longer rely on the model prediction that firms with higher composite productivity $(1 - \tau_g)p + a_g - c_g^a(a_g)$ will post higher flow utility and vacancies separately by submarket. The reason is that under the policies we consider, the effective productivity levels of both men and women matter for wages, amenities, and vacancies to be posted for either gender as firms now maximize total profits across markets. Instead, we solve the following firm profit-maximization problem for the equal pay policy:

(G.4)

$$\begin{aligned} \max_{w, a_M, a_F, v_M, v_F} & \left\{ T_M v_M (p - w - c_M^a(a_M)) \left(\frac{1}{\delta_M + \lambda_M^G + \lambda_M^E (1 - F_M(w + a_M))} \right)^2 \right. \\ & + T_F v_F ((1 - \tau)p - w - c_F^a(a_F)) \left(\frac{1}{\delta_F + \lambda_F^G + \lambda_g^E (1 - F_F(w + a_F))} \right)^2 \\ & \left. - c_M^v(v_M) - c_F^v(v_F) \right\}, \end{aligned} \quad \begin{aligned} \text{(G.5)} \\ \text{(G.6)} \end{aligned}$$

and the following for the equal-hiring policy:

$$\begin{aligned} \max_{x_M, x_F, a_M, a_F, v} & \left\{ T_M v (\tilde{p}_M - x_M) \left(\frac{1}{\delta_M + \lambda_M^G + \lambda_M^E (1 - F_M(x_M))} \right)^2 \right. \\ & + T_F v (\tilde{p}_F - x_F) \left(\frac{1}{\delta_F + \lambda_F^G + \lambda_g^E (1 - F_F(x_F))} \right)^2 \\ & \left. - c_M^v(v) - c_F^v(v) \right\}, \end{aligned} \quad \begin{aligned} \text{(G.7)} \\ \text{(G.8)} \\ \text{(G.9)} \end{aligned}$$

where the definitions of T_g , F_g and V_g are as in the standard solution algorithm. The only unknowns in this problem are F_M and F_F , two endogenous objects to be determined in the counterfactual policy equilibrium. In the equal-pay policy, firms can still hire both genders, only one gender, or neither. In the equal-hiring policy, dual-gender firms are forced to post the same number of vacancies across genders.

Denote by $\mathcal{F}_g$ the mapping from $\{F_M, F_F\}$ to the offer distribution for gender $g \in \{M, F\}$ implied by firms' optimizing behavior. We solve the following system of functional equations:

$$\mathcal{F}_M(F_M^*, F_F^*) = F_M^* \quad \text{(G.10)}$$

$$\mathcal{F}_F(F_M^*, F_F^*) = F_F^*, \quad \text{(G.11)}$$

where $\mathcal{F}_g(F_M^*, F_F^*)$ represents the offer distributions implied by the optimal choices of firms that are a function of the offer distributions in the economy. It's worth

noticing that the offer distributions of both genders implicitly depend on the offer distributions of both men and women. In the equal pay policy case, this is because when firms decide which wage to set, they have to take into account the effects this will have for attracting both genders with respect to the competition they face along the job ladder. In the equal-hiring policy case, this is because when firms decide how many vacancies to post, they balance the profits per contacted worker of both genders.

Therefore, we solve for the equilibrium offer distributions F_M and F_F as follows:

1. Start with a guess for F_M and F_F ; compute the firm's policy functions for optimal wages, amenities, and vacancies, taking F_M and F_F as given.
2. Using equation (12), aggregate optimal vacancies of firms to calculate V_g .
3. Compute the offer distributions F_M and F_F that are implied by the firms' policy functions.
4. Find F_M and F_F such that the offer distributions taken as given by firms and the offer distributions implied by the firms' behavior are identical.

3. Other Counterfactual Equilibrium Simulations

We expand the analysis in Section VII by performing three more structural counterfactuals that illustrate the relative contributions of taste-based discrimination (Becker, 1971), labor market frictions (Manning, 2003), and their interplay. First, we remove employer heterogeneity in gender wedges. Second, we remove gender differences in labor market efficiency. Third, we remove all gender differences. Table G.1 summarizes our results, alongside the counterfactual we performed in Section VII where we removed amenity differences across employers.

TABLE G.1—STRUCTURAL DECOMPOSITION OF THE GENDER PAY GAP

	Baseline	Counterfactuals			
	(0)	(1)	(2)	(3)	(4)
Differences in Amenities	✓		✓	✓	
Gender wedges	✓	✓		✓	
Labor market parameters	✓	✓	✓		
Gender log pay gap	0.107	0.054	−0.044	0.106	−0.000
between employers	0.082	0.020	−0.030	0.082	−0.000
within employers	0.025	0.034	−0.014	0.024	0.000
Gender log amenities gap	−0.068	−0.011	0.049	−0.067	0.000
between employers	−0.075	−0.009	0.032	−0.075	0.000
within employers	0.007	−0.002	0.016	0.008	0.000
Gender log utility gap	0.039	0.042	0.005	0.038	−0.000
between employers	0.007	0.010	0.003	0.007	−0.000
within employers	0.032	0.032	0.002	0.032	0.000
Output	1.000	1.016	1.081	0.992	1.048
Worker welfare	1.000	1.014	1.013	1.000	1.014
for men	1.000	1.014	1.000	1.000	1.000
for women	1.000	1.013	1.032	1.000	1.035
Total employment	0.771	0.783	0.786	0.764	0.764
for men	0.764	0.772	0.764	0.764	0.764
for women	0.781	0.799	0.818	0.764	0.764

Note: This table reports results from equilibrium counterfactuals. Log amenities are defined as the log amenity valuations ($\ln \tilde{a} = \ln(x/w)$). The baseline economy (column 0) is compared to counterfactuals without differences in amenities across employers (column 1), without differences in gender wedges across employers (column 2), without gender differences in labor market efficiency (column 3), and without any gender differences in the economy (column 4).

Source: Model estimates based on RAIS, 2007–2014.

THE ROLE OF EMPLOYER PREFERENCES OVER GENDER.

Next, we simulate gender inequality if there were no heterogeneity in gender wedges by setting gender wedges at all firms to the average gender wedge, $\tau \mapsto \bar{\tau}$. As a result, the gender pay gap disappears completely and even turns negative, largely because women relocate toward higher-productivity employers who initially have higher gender wedges and, once removed, compete more for women by increasing their recruiting intensity and pay. As a result, the between-employer pay gap closes and output increases by 8.1 percent. However, more productive employers also offer lower amenities to women, leading to an increase in the gender amenities gap by 11.6 log points. Overall, the gender utility gap is reduced by 3.4 out of the original 3.9 log points. While worker welfare of men is unchanged, worker welfare of women increases by 3.2 percent. Taken together, this counterfactual makes clear that amenities interact with gender wedges: low-paying low-productivity firms require both higher amenities and lower gender wedges in order to attract a substantial female workforce.

THE ROLE OF GENDER-SPECIFIC LABOR MARKET FRICTIONS POWER.

Next, we study the role of gender-specific labor market frictions in shaping gender inequality by setting the parameters guiding women’s labor market efficiency so that $\lambda_F^U \mapsto \lambda_M^U$, $s_F^E \mapsto s_M^E$, $s_F^G \mapsto s_M^G$, and $\delta_F \mapsto \delta_M$. As a result, the gender pay gap declines by only 0.1 log points and the gender utility gap by 0.1 log points. The reason behind the small effect is that women receive more on-the-job offers but also separate more frequently from their employers. Thus, their overall speed of climbing the firm ladder barely changes. All in all, while there is large monopsony power in Brazil, differences in labor market fluidity are not a key driver of gender gaps.⁷⁴

THE EFFECTS OF MOVING TO A GENDER-NEUTRAL LABOR MARKET.

Finally, we simulate a labor market with no gender differences in amenities (i.e., $a_F \mapsto a_M$), no gender wedges (i.e., $\tau = 0$), and no gender differences in labor market efficiency (i.e., $\lambda_F^U \mapsto \lambda_M^U$, $s_F^E \mapsto s_M^E$, $s_F^G \mapsto s_M^G$, and $\delta_F \mapsto \delta_M$). Obviously, this closes the gender gaps in pay and amenities. Strikingly, this also increases aggregate output by 4.8 percent and overall worker welfare by 1.4 percent, as women experience greater flow utility by 3.9 log points associated with them moving to more productive firms with higher total compensation. Relative to the counterfactuals that shut down one ingredient at a time, we find strong nonlinearities that reflect the interaction between different dimensions of firm heterogeneity affecting pay, amenities, and employment in equilibrium. For

⁷⁴This contrasts with [Bowlus \(1997\)](#) and [Flinn, Todd and Zhang \(2025\)](#) who attribute a significant fraction of the U.S. gender wage gap to different labor market behaviors across the sexes using an equilibrium search model *without* compensating differentials.

example, changing the wage-amenity composition of total compensation at a firm initially employing few women has a greater impact when also increasing the same firm's composite productivity in a way that leads to more women being hired. In this sense, our equilibrium model features strong complementarities between the different dimensions of firm heterogeneity. In summary, moving to a gender-neutral labor market yields significant output and welfare gains.

4. Counterfactual Simulations for the Equal-Amenities Policy

We simulate a mandate for all firms to provide equal amenity valuations across workers of both genders: $\beta_M(j)a_M(j) = \beta_F(j)a_F(j)$ for all j . The results from this counterfactual policy simulation are summarized in Table G.2 below.

TABLE G.2—EFFECTS OF SIMULATED EQUAL AMENITIES POLICY.

	Baseline (0)	Equal-amenities policy (1)
Gender log pay gap	0.107	0.088
between employers	0.082	0.031
within employers	0.025	0.057
Gender total amenities gap	−0.037	−0.041
between employers	−0.084	−0.041
within employers	0.047	0.000
Gender log utility gap	0.039	0.026
between employers	0.007	−0.001
within employers	0.032	0.027
Output	1.000	0.979
Worker welfare	1.000	0.989
for men	1.000	0.990
for women	1.000	0.986
Total employment	0.771	0.763
for men	0.764	0.760
for women	0.781	0.769

Note: Table reports results from a counterfactual policy experiments. Total amenities are defined as amenity valuations ($\beta_g a_g$). Baseline results (column 0) are compared against the economy with an equal-amenities policy (column 1).

Source: Model estimates based on RAIS, 2007–2014.

We find that, when firms are mandated to offer the same amenities to men and women, both the gender pay gap and the gender utility gap decrease. As firms need now to compromise on the amenities they offer to men and women, those firms that were previously attractive to women due to high amenities become less so. Thus, women move to higher-paying firms, resulting in a lower between-firms gap but a larger within-firms gap, as higher-paying firms tend to be more discriminatory. We find this policy to be even more detrimental to welfare, output and employment than the equal-hiring policy. Output declines by 2.1 percent, while welfare declines by 1.1 percent, and more so for women.

5. *Counterfactual Simulations for a Restricted Equal-Hiring Policy*

We simulate a scenario in which an equal-hiring policy is imposed only on a subset $\mathbb{J}$ of firms: $v_M(j) = v_F(j)$ for $j \in \mathbb{J}$. The idea is to capture a policy that might be imposed, for instance, only on the public sector of the economy. To this end, we construct a dummy variable that takes the value of 1 if a firm is classified as operating in the public sector and zero otherwise, and we simulate the equilibrium of our model when only this subset of firms is constrained by the equal-hiring policy. Our results are summarized in Table G.3.

We find that, when limited to the public sector, the equal-hiring policy leads to a decline in the gender pay gap by 0.9 percent, exclusively driven by a decline in the between-firms gender pay gap. However, this is accompanied by a reduction of the amenities gap, which is in favor of women, by 0.9 percent. As a result, the overall utility gap remains unchanged. The reason is that the public sector already disproportionately hires women: imposing an equal-hiring policy on the public sector increases the representation of men in these firms, while women relocate to other firms that offer higher wages but lower amenities.

TABLE G.3—EFFECTS OF SIMULATED EQUAL-HIRING POLICY, RESTRICTED TO THE PUBLIC SECTOR.

	Baseline	Equal-hiring	Equal-hiring, public only
	(0)	(1)	(2)
Gender log pay gap	0.107	0.032	0.098
between employers	0.082	0.006	0.073
within employers	0.025	0.026	0.025
Gender log amenities gap	−0.068	0.010	−0.059
between employers	−0.075	−0.006	−0.068
within employers	0.007	0.016	0.009
Gender log utility gap	0.039	0.042	0.039
between employers	0.007	0.000	0.005
within employers	0.032	0.041	0.034
Output	1.000	0.997	0.999
Worker welfare	1.000	0.992	0.999
for men	1.000	0.991	0.999
for women	1.000	0.993	0.999
Total employment	0.771	0.764	0.771
for men	0.764	0.722	0.763
for women	0.781	0.825	0.782

Note: Table reports results from counterfactual policy experiments. Log amenities are defined as the log amenity valuations ($\ln \bar{a} = \ln(x/w)$). Baseline results (column 0) are compared against the economy with an equal-hiring policy (column 1) and an economy where the equal-hiring policy is enforced on the public sector only (column 1).

Source: Model estimates based on RAIS, 2007–2014.

SUPPLEMENTAL APPENDIX REFERENCES

- Abowd, John M., Robert H. Creecy, and Francis Kramarz.** 2002. "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data." *U.S. Census Bureau Technical Report 2002-06*.
- Abraham, Lisa, Johannes Hallermeier, and Alison Stein.** 2024. "Words Matter: Experimental Evidence from Job Applications." *Journal of Economic Behavior and Organization*, 225: 348–391.
- Acemoglu, Daron, and David Autor.** 2011. "Skills, Tasks and Technologies: Implications for Employment and Earnings." In *Handbook of Labor Economics*. Vol. 4, , ed. O. Ashenfelter and D. Card, Chapter 12, 1043–1171. Amsterdam: Elsevier-North.
- Alvarez, Jorge, Felipe Benguria, Niklas Engbom, and Christian Moser.** 2018. "Firms and the Decline in Earnings Inequality in Brazil." *American Economic Journal: Macroeconomics*, 10(1): 149–189.
- Autor, David, and David Dorn.** 2009. "This Job Is "Getting Old": Measuring Changes in Job Opportunities Using Occupational Age Structure." *American Economic Review*, 99(2): 45–51.
- Bagger, Jesper, and Rasmus Lentz.** 2019. "An Empirical Model of Wage Dispersion with Sorting." *Review of Economic Studies*, 86(1): 153–190.
- Becker, Gary S.** 1971. *The Economics of Discrimination*. . Second ed., Chicago:University of Chicago Press.
- Bieri, David S., Nicolai V. Kuminoff, and Jaren C. Pope.** 2023. "National Expenditures on Local Amenities." *Journal of Environmental Economics and Management*, 117: 102717.
- Bowlus, Audra J.** 1997. "A Search Interpretation of Male-Female Wage Differentials." *Journal of Labor Economics*, 15(4): 625–657.
- Burdett, Kenneth, and Dale T. Mortensen.** 1998. "Wage Differentials, Employer Size, and Unemployment." *International Economic Review*, 39(2): 257–273.
- Card, David, Ana Rute Cardoso, and Patrick Kline.** 2016. "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women." *Quarterly Journal of Economics*, 131(2): 633–686.
- Cortés, Patricia, Jessica Pan, Laura Pilossoph, Ernesto Reuben, and Basit Zafar.** 2023. "Gender Differences in Job Search and the Earnings Gap: Evidence from the Field and Lab." *Quarterly Journal of Economics*, 138(4): 2069–2126.

- de Souza, Gustavo.** 2024. “The Labor Market Consequences of Appropriate Technology.” *Federal Reserve Bank of Chicago Working Paper No. 2022-53*.
- Engbom, Niklas, and Christian Moser.** 2022a. “Earnings Inequality and the Minimum Wage: Evidence from Brazil.” *American Economic Review*, 112(12): 3803–3847.
- Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer.** 2015. “The Next Generation of the Penn World Table.” *American Economic Review*, 105(10): 3150–3182.
- Flinn, Christopher J., Petra E. Todd, and Weilong Zhang.** 2025. “Labor Market Returns to Personality: A Job Search Approach to Understanding Gender Gaps.” *Journal of Political Economy*, 133(4): 1169–1234.
- Fukui, Masao, Emi Nakamura, and Jón Steinsson.** 2023. “Women, Wealth Effects, and Slow Recoveries.” *American Economic Journal: Macroeconomics*, 15(1): 269–313.
- Goldin, Claudia, and Cecilia Rouse.** 2000. “Orchestrating Impartiality: The Impact of “Blind” Auditions on Female Musicians.” *American Economic Review*, 90(4): 715–741.
- Kalemlı-Özcan, Şebnem, Bent E. Sørensen, Carolina Villegas-Sanchez, Vadym Volosovych, and Sevcen Yeşiltas.** 2024. “How to Construct Nationally Representative Firm-Level Data from the Orbis Global Database: New Facts on SMEs and Aggregate Implications for Industry Concentration.” *American Economic Journal: Macroeconomics*, 16(2): 353–374.
- Lentz, Rasmus.** 2024. “Models of Wages and Mobility in Frictional Labor Markets with Random Search.” *Revue Economique*, 51(1): 113–145.
- Manning, Alan.** 2003. *Monopsony in Motion: Imperfect Competition in Labor Markets*. Princeton, NJ: Princeton University Press.
- Page, Lawrence, Sergey Brin, Rajeev Motwani, and Terry Winograd.** 1999. “The PageRank Citation Ranking: Bringing Order to the Web.” *Stanford InfoLab Working Paper*.
- Rosen, Sherwin.** 1986. “The Theory of Equalizing Differences.” In *Handbook of Labor Economics*. Vol. 1, , ed. O. Ashenfelter and R. Layard, Chapter 12, 641–692. Amsterdam: Elsevier-North.
- Sorkin, Isaac.** 2018. “Ranking Firms Using Revealed Preference.” *Quarterly Journal of Economics*, 133(3): 1331–1393.