The Gender Pay Gap: Micro Sources and Macro Consequences

Online Appendix—Not for Publication

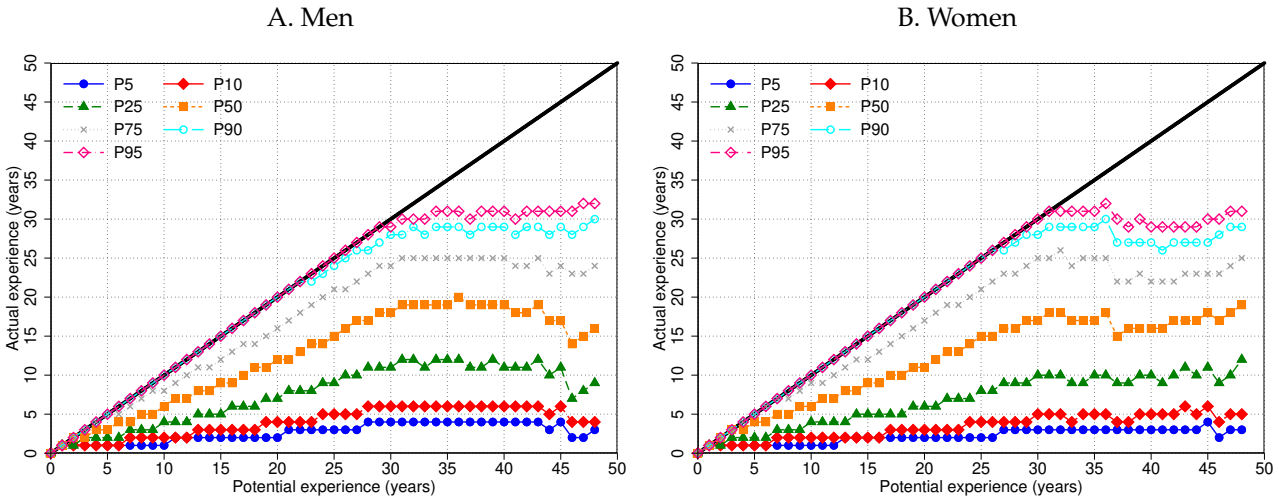
Iacopo Morchio
University of Bristol

Christian Moser
Columbia University, NBER, CEPR, and IZA

A Data Description Appendix

A.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.

A.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	29,455,332	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	3,420,973	2,623,620
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	2,064	4,332	2,775	1,757	4,402
(std. dev.)	(17,415)	(11,393)	(24,188)	(16,617)	(12,450)	(21,404)	(16,298)	(11,448)	(21,828)
Mean gender-employer size	1,744	698	3,485	1,852	1,154	2,920	1,743	926	3,048
(std. dev.)	(11,287)	(3,137)	(17,840)	(10,211)	(5,419)	(14,726)	(10,314)	(4,413)	(15,574)
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	27,609,184	17,248,420	10,360,763	35,134,432	21,238,264	13,896,165	271,400,512	166,964,240	104,436,280
Number of unique workers	27,609,184	17,248,420	10,360,763	35,134,432	21,238,264	13,896,165	56,868,704	33,975,716	22,892,988
Number of unique employers	425,173	295,134	130,039	492,155	334,905	157,250	640,639	420,451	220,188
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	0.350	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	4,393	2,815	1,774	4,497
(std. dev.)	(17,749)	(11,589)	(24,733)	(16,687)	(12,481)	(21,550)	(16,418)	(11,509)	(22,059)
Mean gender-employer size	1,805	719	3,635	1,868	1,160	2,961	1,768	936	3,114
(std. dev.)	(11,505)	(3,190)	(18,244)	(10,254)	(5,432)	(14,827)	(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.)	(5.6)	(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	26,545,820	16,656,529	9,889,290	34,829,856	21,129,484	13,700,373	267,318,328	165,149,632	102,168,696
Number of unique workers	26,545,820	16,656,529	9,889,290	34,829,856	21,129,484	13,700,373	56,297,308	33,761,656	22,535,652
Number of unique employers	396,269	278,455	117,814	477,942	328,628	149,314	607,029	403,585	203,444
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)	(21,828)	(16,264)	(11,559)	(21,410)	100.2	99.0	102.0
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)	(15,574)	(10,194)	(4,173)	(15,231)	101.2	105.8	102.3
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	271,400,512	166,964,240	104,436,280	449,390,272	269,897,824	179,492,464	60.4	61.9	58.2
Number of unique workers	56,868,704	33,975,716	22,892,988	77,397,648	44,436,752	32,960,898	73.5	76.5	69.5
Number of unique employers	640,639	420,451	220,188	6,044,593	3,420,973	2,623,620	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)

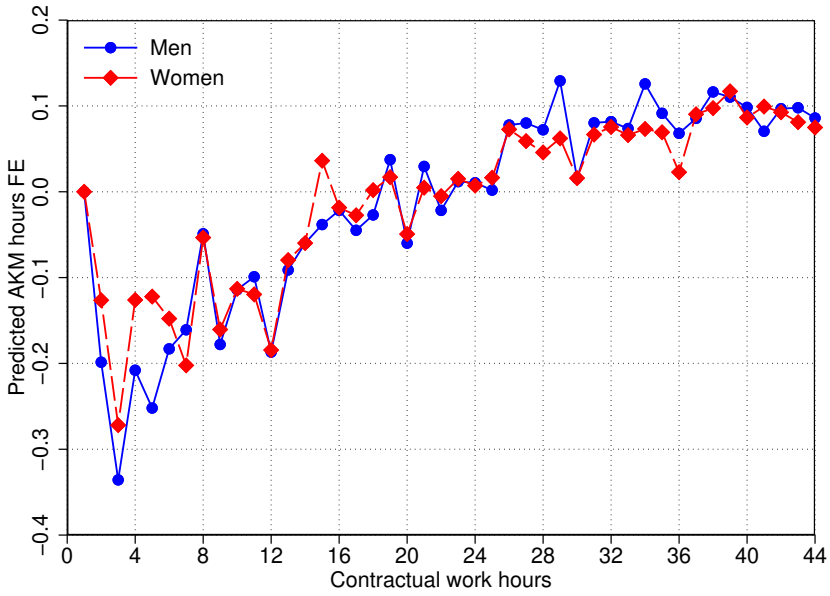
	Pooled 2007–2014, connected set			Pooled 2007–2014, selection set			Ratio: connected set to selection set		
	Overall	Men	Women	Overall	Men	Women	Overall	Men	Women
Share Nonwhite	0.378	0.409	0.327	0.378	0.409	0.329	100.0	100.0	99.4
Share primary school	0.088	0.116	0.042	0.088	0.117	0.043	100.0	99.1	97.7
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	4,497	2,775	1,757	4,402	101.4	101.0	102.2
(std. dev.)	(16,418)	(11,509)	(22,059)	(16,298)	(11,448)	(21,828)	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)	(15,574)	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	267,318,328	165,149,632	102,168,696	271,400,512	166,964,240	104,436,280	98.5	98.9	97.8
Number of unique workers	56,297,308	33,761,656	22,535,652	56,868,704	33,975,716	22,892,988	99.0	99.4	98.4
Number of unique employers	607,029	403,585	203,444	640,639	420,451	220,188	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

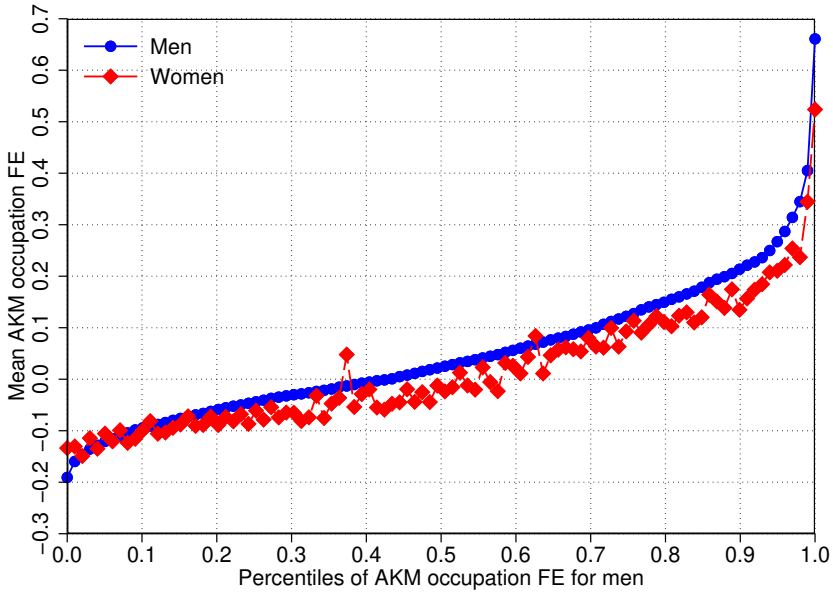
B.1 Detailed AKM Estimation Results

Figure B.1. Predicted AKM contractual work hours FEs, by gender



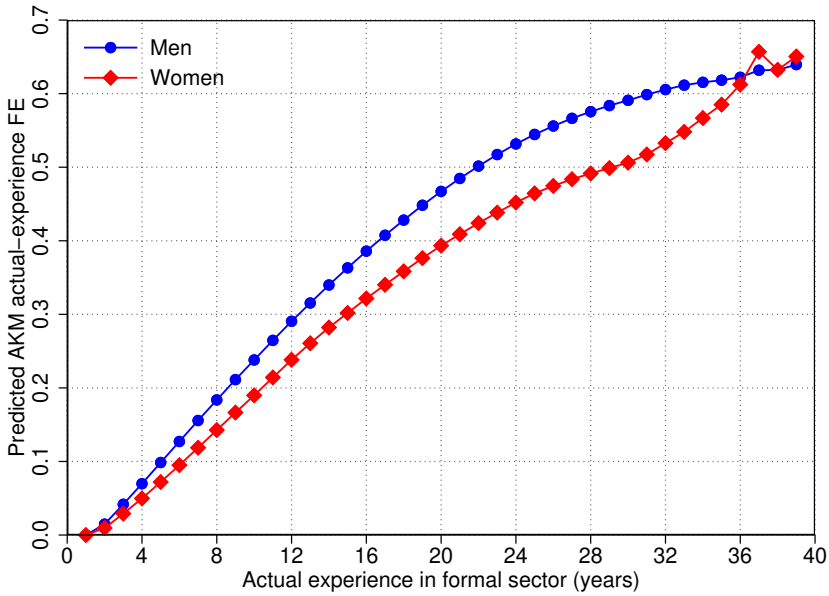
Note: Figure shows predicted AKM contractual work hours FEs separately for men and women based on estimating earnings equation (1). The omitted category is 1 hour, for which the FE value is normalized to 0. *Source:* RAIS, 2007–2014.

Figure B.2. Predicted AKM occupation FEs, by gender



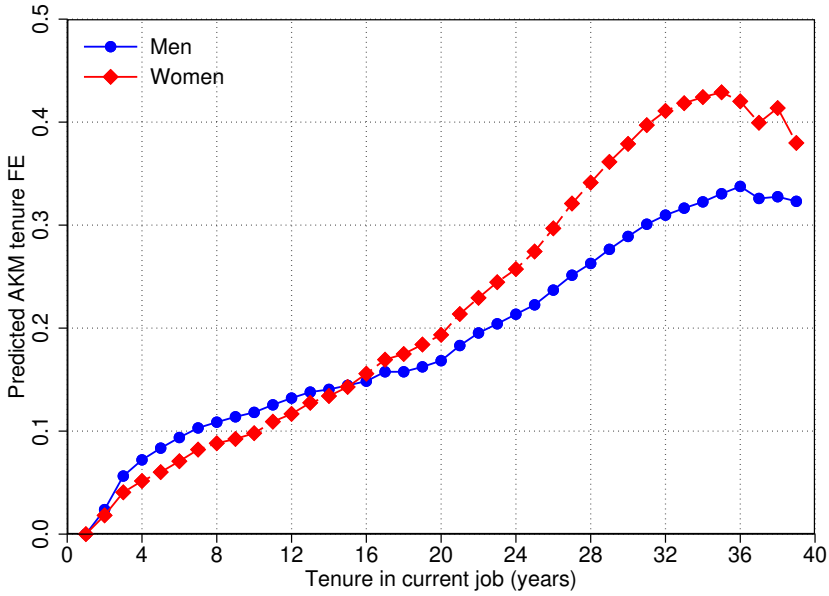
Note: Figure shows predicted AKM occupation FEs separately for men and women based on estimating earnings equation (1). Fixed effects of both genders are sorted by mean FEs of male FE quantiles. *Source:* RAIS, 2007–2014.

Figure B.3. Predicted AKM actual-experience FEs, by gender



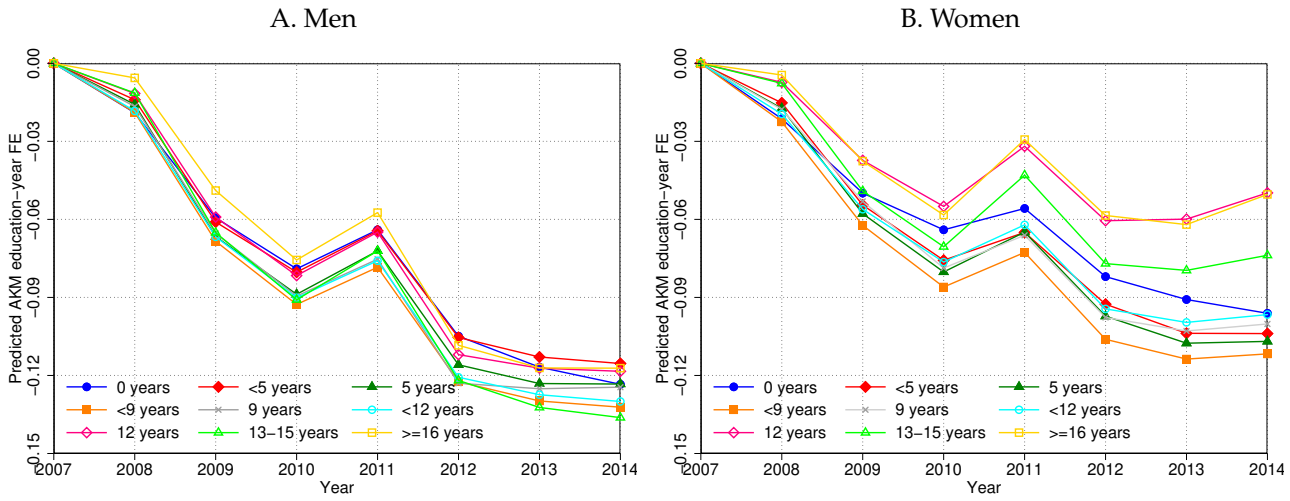
Note: Figure shows predicted AKM actual-experience FEs separately for men and women based on estimating earnings equation (1). *Source:* RAIS, 2007–2014.

Figure B.4. Predicted AKM tenure FEs, by gender



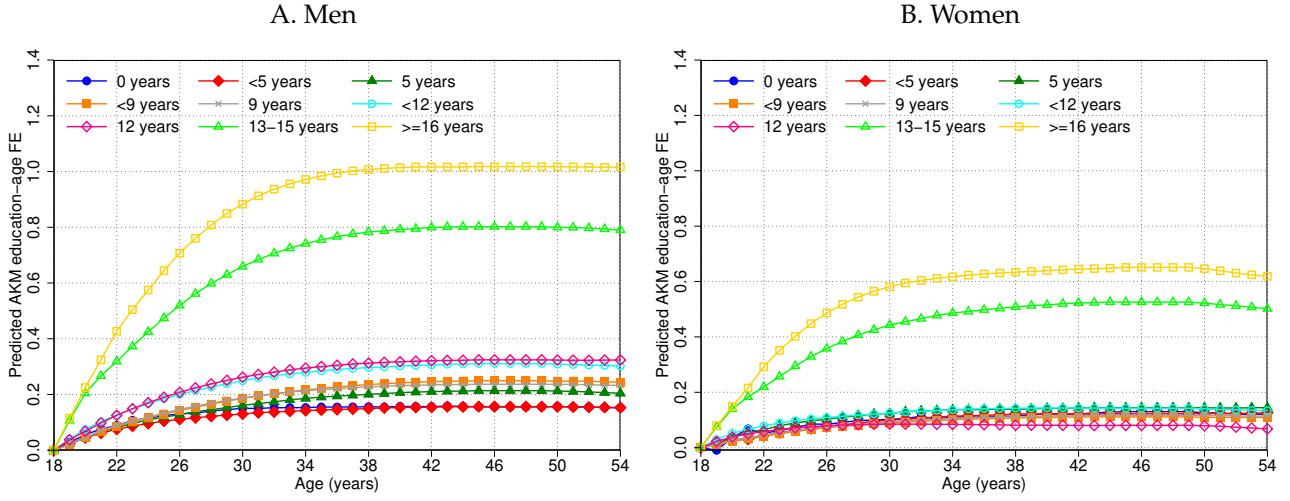
Note: Figure shows predicted AKM tenure FEs separately for men and women based on estimating earnings equation (1). *Source:* RAIS, 2007–2014.

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 (2022) 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.

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

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

$$\mathbb{E}_{i,t}\left[\psi_{MJ(i,t)}\middle|\, G(i)=M\right]-\mathbb{E}_{i,t}\left[\psi_{FJ(i,t)}\middle|\, G(i)=F\right] \tag{B.1}$$

$$=\underbrace{\left(\mathbb{E}_{i,t}\left[\psi_{MJ(i,t)}\middle|\, G(i)=M\right]-\mathbb{E}_{i,t}\left[\psi_{MJ(i,t)}\middle|\, G(i)=F\right]\right)}_{\text{between-employer gap}}+\underbrace{\mathbb{E}_{i,t}\left[\psi_{MJ(i,t)}-\psi_{FJ(i,t)}\middle|\, G(i)=F\right]}_{\text{within-employer gap}} \tag{B.2}$$

$$=\underbrace{\left(\mathbb{E}_{i,t}\left[\psi_{FJ(i,t)}\middle|\, G(i)=M\right]-\mathbb{E}_{i,t}\left[\psi_{FJ(i,t)}\middle|\, G(i)=F\right]\right)}_{\text{between-employer gap}}+\underbrace{\mathbb{E}_{i,t}\left[\psi_{MJ(i,t)}-\psi_{FJ(i,t)}\middle|\, G(i)=M\right]}_{\text{within-employer gap}}. \tag{B.3}$$

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).

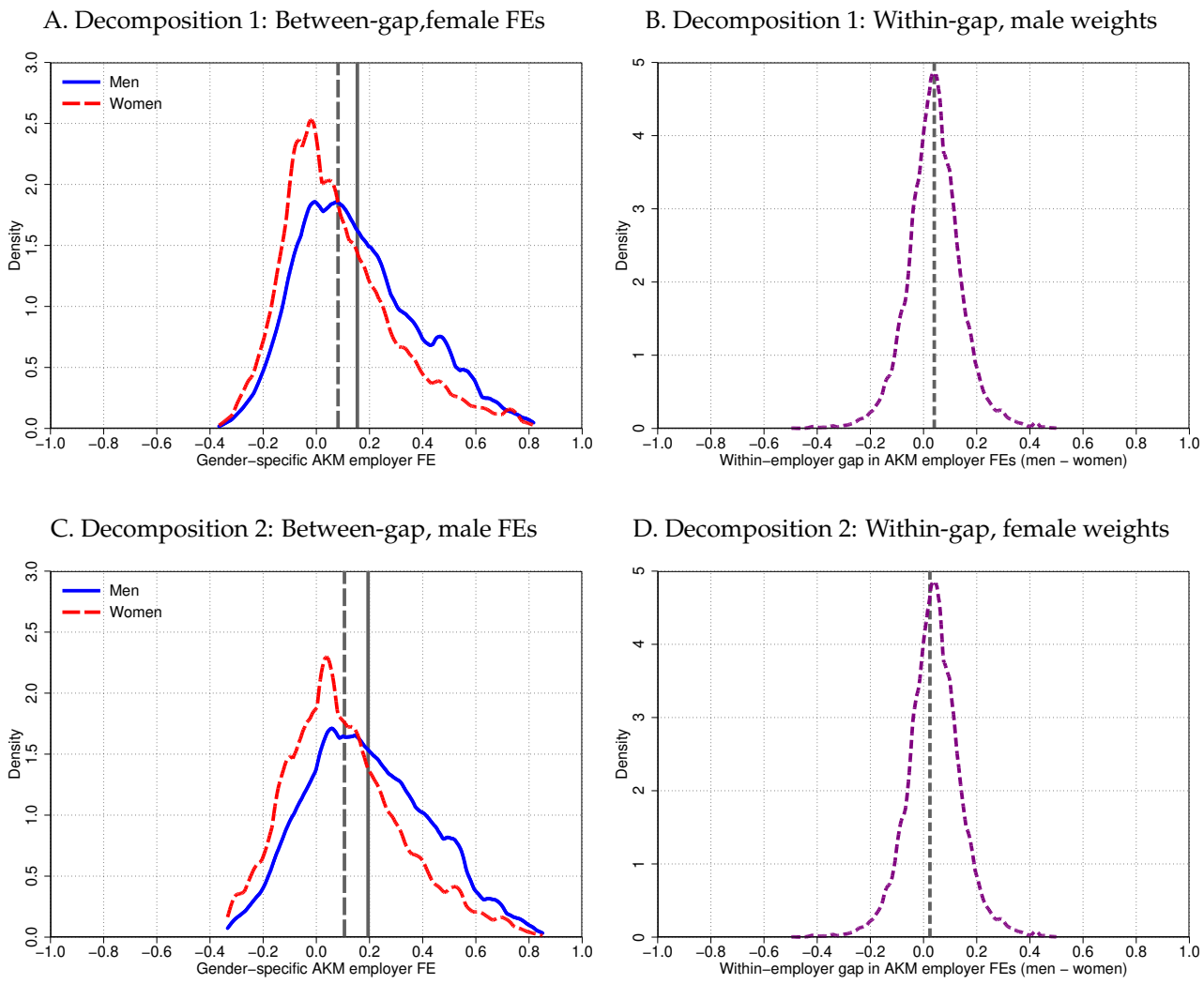
Table B.1. Alternative Kitagawa-Oaxaca-Blinder decompositions of the gender log pay gap

		Between-employer gap		Within-employer gap	
	Gender log pay gap	Level	Share (%)	Level	Share (%)
Decomposition 1	0.113	0.089	78.7	0.024	21.3
Decomposition 2	0.113	0.073	64.3	0.040	35.7

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 Oaxaca-Blinder decompositions



Note: Figure shows pay distributions underlying the Oaxaca-Blinder decompositions—specifically, the between-gap using female FEs (Panel A) and using male FEs (Panel C). as well as the within-gap using male weights (Panel B) and using female weights (Panel D) Decomposition 1 (Panels A and B) and decomposition 2 (Panels C and D) correspond to equations (B.2) and (B.3) of the main text, respectively. Dashed vertical line shows mean of the distribution. *Source:* RAIS, 2007–2014.

B.3 Life-Cycle Profiles by Gender and Parent Status

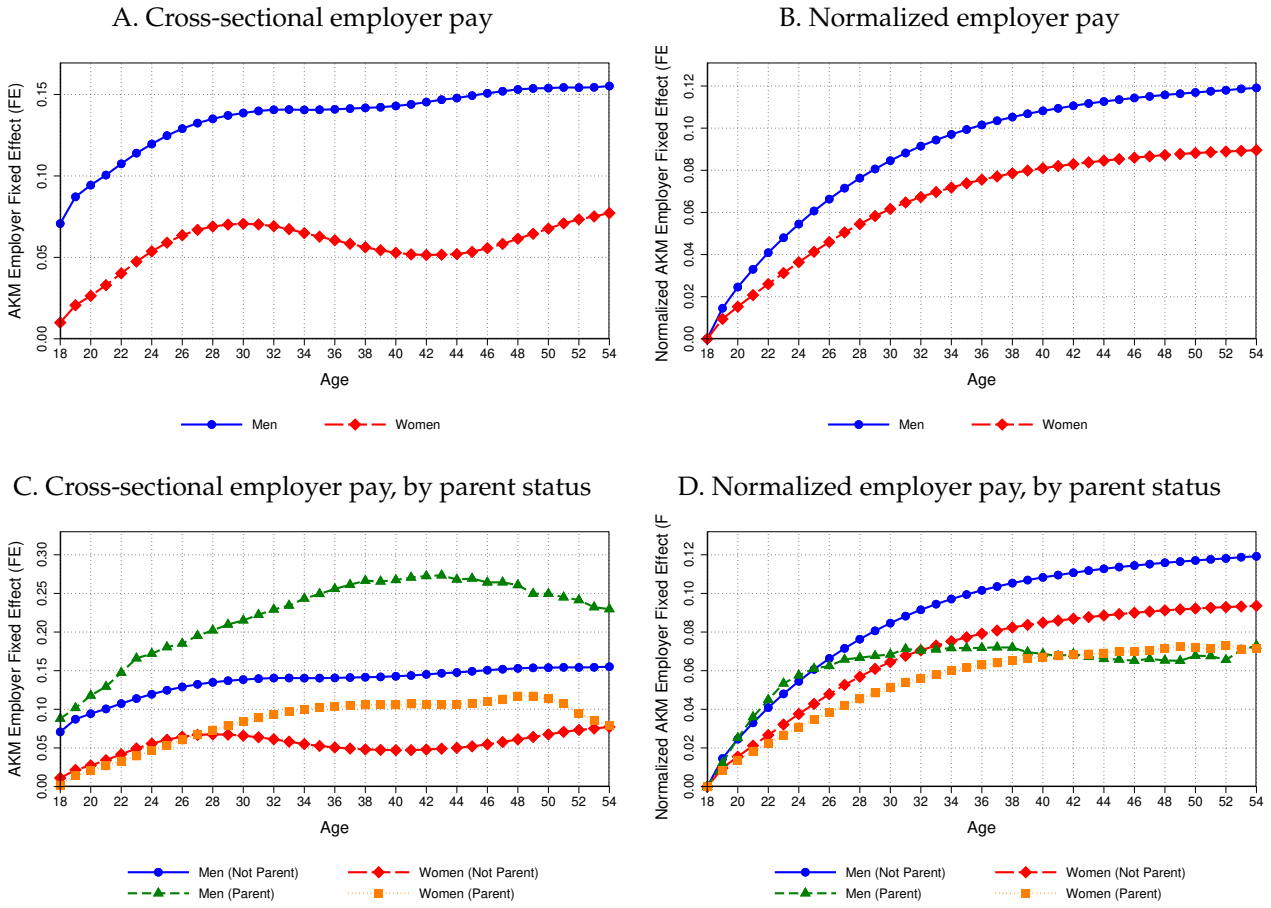
In this section, we are interested in life-cycle patterns in employer heterogeneity and how they differ by gender and parental status. Here, we classify individuals as "parent" if they ever went on registered parental leave from their employer during the sample period 2007–2014 and as "not parent" if they did not.[48] We compute two types of life-cycle statistics. The first set of statistics comprises raw, cross-sectional binned means. The second set of statistics comprises binned means of differenced variables, which we normalize to 0 at age 18.

Figure B.8 shows estimated gender-specific employer FEs by gender and parent status. A few things are worth noting. First, employer pay for women is less than that for men over the entire life-cycle. Second, cross-sectional life cycles (Panels A and C) can be quite different from the normalized life cycles (Panels B and D), plausibly owing to cohort effects and other dimensions of permanent individual heterogeneity that is differenced out in the normalized statistics. Third, both men and women see marked growth in employer FEs over their life cycle, although men significantly more so than women (Panel B). Fourth, parent men look more similar to women in general, and to women with children in particular, compared with nonparent men, although nonparent women still differ from nonparent men (Panel D).

Altogether, these life-cycle patterns suggest that childbirth could play some role in explaining parts of the gender pay gap, consistent with findings from similar studies in other contexts, such as Coudin et al. (2018) for France.

[48] For comparison, the birth rate in Brazil from official birth statistics is 12.9 per 1,000 people (World Bank, 2021a). In our setting, approximately 11.1 workers per 1,000 in the RAIS data go on a registered parental leave during 2014. At face value, this means that our data cover approximately 86.0% of births from vital statistics records, suggesting that there is a high take-up rate of parental leaves among new parents. The difference between the official statistic and that computed on the RAIS data might be due to multiple births per parent (e.g., twins or two separate births within the same calendar year), less than perfect take-up of parental leaves (e.g., continuing office duties in spite of the federally mandated parental leave) or different fertility across workers in formal-sector jobs covered by RAIS (e.g., lower fertility among relatively high-income workers in the formal sector).

Figure B.8. Life-cycle mean gender-specific employer FEs, by gender and parent status



Note: Figure shows cross-sectional (Panels A and C) and normalized (Panels B and D) employer pay across age separately by gender (Panels A and B) and separately by gender and ever-parent status (Panels C and D). Cross-sectional estimates represent binned means. Normalized estimates are binned means of differenced variable, normalized to 0 at age 18. *Source:* RAIS, 2007–2014.

B.4 Event Study Analysis around Parental Leaves by Gender

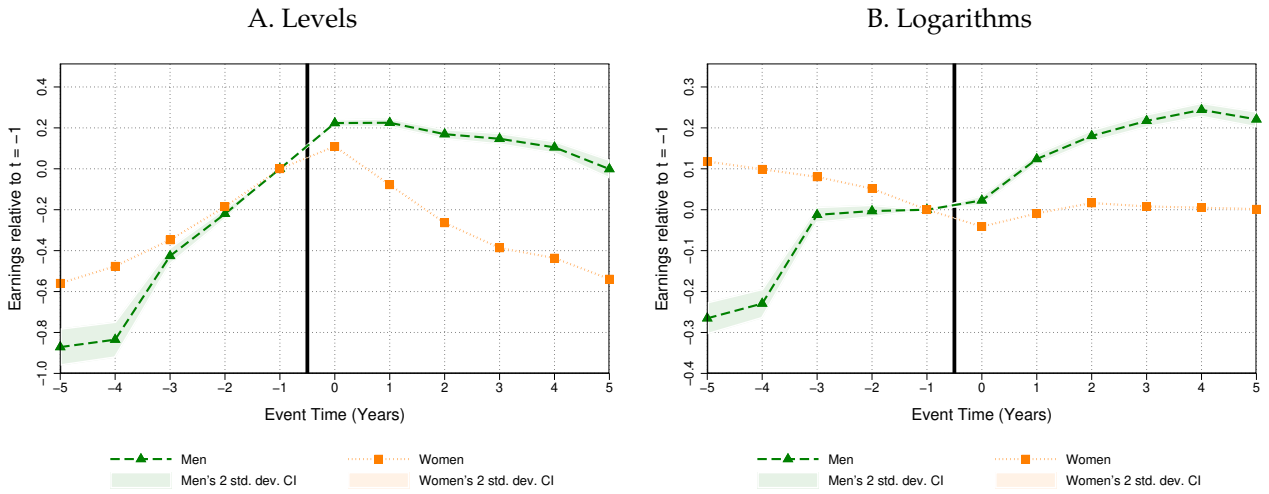
Following Kleven et al. (2019), we estimate the following event-study regression for individual i of gender g in year s and at event time t:

$$y_{ist} = \sum_{t' \neq 1} \alpha_{t'}^{g} \mathbf{1}[t' = t] + \sum_{a} \beta_{a}^{g} \mathbf{1}[a = \text{age}_{is}] + \sum_{s'} \gamma_{s'}^{g} \mathbf{1}[s' = s] + \nu_{ist}^{g}, \tag{B.4}$$

where y_{ist} is the outcome variable of interest, $\alpha_{t'}^{g}$ denotes a set of gender-specific event time controls, β_{a}^{g} denotes a set of gender-specific age controls, $\gamma_{s'}^{g}$ denotes a set of gender-specific time controls, and ν_{ist}^{g} is an error term. As dependent variables, we will use the level of earnings (filling in zero earnings for missing observations) or, alternatively, log earnings (dropping missing observations). Our focus will be on estimates of the coefficients α_{t}^{g}, based on equation (B.4), for men and women in an 11-year window around the birth of individuals' first child.

Figure B.9 plots the resulting event study graph, including gender-specific point estimates and confidence intervals. Panel A of the figure shows the event study for earnings in levels. Men and women are on comparable earnings paths leading up to their first child's birth, marked by the vertical black solid line. After their first child, women's earnings markedly decline, both in absolute value and compared with men's earnings, which remain relatively more stable. Panel B shows the event study for earnings in logarithms. Women's earnings show a declining pretrend in the five years leading up to their first child, both in absolute terms and compared to men, whose earnings increase over the same preperiod. After the birth of their child, women's earnings take a one-year dip and then remain relatively constant over the next five years. In contrast, men's earnings grow over the five years following their first child.

Figure B.9. Event-study plot of earnings relative to the year before the first child



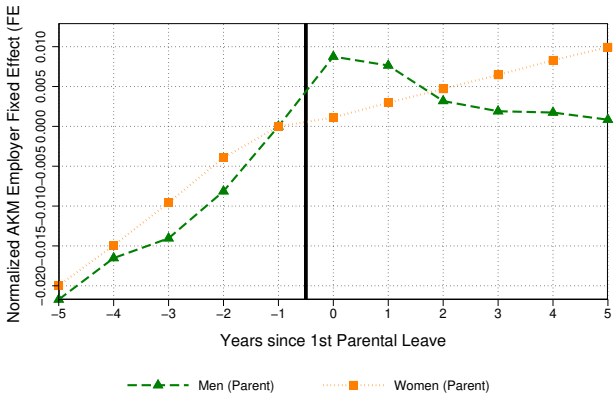
Note: Figure shows event study tracking earnings levels (Panel A) and log earnings (Panel B) in a time window of 10 years around first childbirth. The vertical solid black line separates years before and after first childbirth. *Source:* RAIS, 2007–2014.

Taken at face value, these results suggest that having children has an effect on the earnings and participation of women relative to men. However, it seems that women's earnings losses around having children are not systematically related to changes in the employer component of earnings. Figure B.10 illustrates this point by plotting an analogous event study with estimated gender-specific employer FEs from equation (1) as the dependent variable. Men and women follow a similar trend before the birth of a child. In the first two years after the birth of a child, women's gender-specific

employer FE falls behind that for men, but the pattern reverses during years 3 through 5. At any time in the event study, the gender gap in gender-specific employer FEs is less than 1 log point.

Altogether, this suggests that, while childbirth has a significant effect on overall earnings and participation of women compared to men, firm pay heterogeneity is *not* not a very important factor behind women's pay penalty for having children.

Figure B.10. Event-study plot of gender-specific employer FEs rel. to the year before the birth of the first child



Note: Figure shows event study tracking gender-specific employer FEs based on estimating earnings equation (1) on the 2007–2014 sample in a time window of 10 years around the birth of a first child. The vertical solid black line separates years before and after the birth. *Source:* RAIS, 2007–2014.

C Equilibrium Model of Employer Pay, Amenities, and Size Appendix

C.1 Definition of a Stationary Equilibrium

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

Definition. *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 $\{\lambda^U_{gz}, \lambda^E_{gz}, \lambda^G_{gz}\}_{gz}$; and firm sizes $\{l_{gz}\}_{gz}$ such that, for all (g, z),*

- *given $F_{gz}(x)$ and $\{\lambda^U_{gz}, \lambda^E_{gz}, \lambda^G_{gz}\}$, 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 $\{\lambda^U_{gz}, \lambda^E_{gz}, \lambda^G_{gz}\}$ satisfy equation (14);*
- *given $F_{gz}(x)$, $\{\lambda^U_{gz}, \lambda^E_{gz}\}_{gz}, \lambda^G_{gz}$, and V_{gz}, firm sizes satisfy equation (16);*
- *the offer distribution satisfies $F_{gz}(x) = \int_j v_{gz}(j)\mathbf{1}[x_{gz}(j) \leq x]\, d\Gamma(j)/V_{gz}$.*

C.2 Proof of Lemma 1 (Optimal Amenities)

Restatement of Lemma 1 (Optimal Amenities). *A firm's optimal amenity policy function $a^*_{gz}(\cdot)$ is linear in worker ability z, strictly decreasing in its amenity cost shifter $c^{a,0}_g$, and invariant to all other firm parameters (i.e., p and τ_g). In equilibrium, $a^*_{gz}(c^{a,0}_g, z) = (c^{a,0}_g)^{1/(1-\eta^a)} z$.*

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 $x = w + a$ and vacancies v that solve the following problem:

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

where $c^x_{gz}(x)$ is the solution to the following cost-minimization subproblem in each market:

$$c^x_{gz}(x) = \min_{w,a} \left\{ w + c^a_{gz}(a) \right\} \quad \text{s.t.} \quad w + a = x. \tag{C.2}$$

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 optimality conditions:

$$\frac{\partial c^a_{gz}(a^*)}{\partial a} = 1 \tag{C.3}$$

$$w^* = x - a^*, \quad \forall (g, z). \tag{C.4}$$

Equation (C.3) uniquely pins down a firm's optimal amenity choice $a^*_{gz}(c_g^{a,0})$ for every market (g,z) as a function of only the heterogeneous amenity cost shifter $c_g^{a,0}$ and z. Obviously, $\partial a^*/\partial c_g^{a,0} < 0$ by the chain rule. The optimal wage is then chosen to deliver the remainder of flow utility x.[49] Since $c^a_{gz}(0) = 0$ and $\partial c^a_{gz}/\partial a(0) = 0$, a firm will always produce some amount of amenities $P_{gz} > 0$. Finally, from the functional form of the cost of amenities in equation (8) we can further solve the FOC in equation (C.3) to show that:

$$c_g^{a,0}\left[\frac{a^*}{z}\right]^{\eta^a-1} = 1 \tag{C.6}$$

$$a^* = z\left[c_g^{a,0}\right]^{\frac{1}{1-\eta^a}}. \tag{C.7}$$

Equation (C.7) shows that optimal amenities are proportional to ability z and, as $\eta^a > 1$, they are decreasing in the amenity cost shifter $c_g^{a,0}$. Also, optimal amenities are invariant to all other parameters.

We can make further progress and obtain an explicit expression for the cost of amenities in equilibrium by rearranging equation (C.7) as:

$$c_g^{a,0} = \left[\frac{a^*}{z}\right]^{1-\eta^a}. \tag{C.8}$$

By plugging equation (C.8) into the amenity cost function from equation (8), we can rewrite the cost of amenities at the optimum as:

$$c^a(a^*) = \frac{c_g^{a,0}}{z^{\eta^a-1}}\frac{(a^*)^{\eta^a}}{\eta^a} \tag{C.9}$$

$$= z^{1-\eta^a}\left[\frac{a^*}{z}\right]^{1-\eta^a}\frac{(a^*)^{\eta^a}}{\eta^a} \tag{C.10}$$

$$= \frac{a^*}{\eta^a}, \tag{C.11}$$

showing that, in equilibrium, the cost of producing amenities is inversely proportional to $\eta^a > 1$. This completes the proof.

□

C.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 amenity cost shifter $c_g^{a,0}$.*

[49] Taking into account possible corner solutions, the optimal wage-amenity combination takes the following form:

$$P^{**}_{gz}\left(x, c_g^{a,0}\right) = \begin{cases} x & \text{if } x < \overline{x}\left(c_g^{a,0}\right) \\ P^*_{gz}\left(c_g^{a,0}\right) & \text{if } x \geq \overline{x}\left(c_g^{a,0}\right) \end{cases}, \quad w^{**}_{gz}\left(x, c_g^{a,0}\right) = \begin{cases} 0 & \text{if } x < \overline{x}\left(c_g^{a,0}\right) \\ x - P^{**}_{gz}\left(c_g^{a,0}, x\right) & \text{if } x \geq \overline{x}\left(c_g^{a,0}\right), \end{cases} \tag{C.5}$$

where $\overline{x}\left(c_g^{a,0}\right)$ solves $\partial c^a_{gz}(\overline{x}\left(c_g^{a,0}\right))/\partial a = 1$. Note, however, that in such corner solutions the optimal wage is $w^{**} = 0$, which is empirically not relevant. Going forward, we focus on the more natural case of an interior solution.

Proof. We first reformulate the firm's problem. Expected profits per worker contacted by a firm is

$$P_{gz}(\tilde{p}_{gz}, x) = h_{gz}(x) J_{gz}(\tilde{p}_{gz}, x), \tag{C.12}$$

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

$$J_{gz}(\tilde{p}_{gz}, x) = \frac{\tilde{p}_{gz} - x}{\delta_{gz} + \lambda^E_{gz}(1 - F_{gz}(x)) + \lambda^G_{gz}} \tag{C.13}$$

$$= \frac{\left(\tilde{p}_{gz} - x\right) / \left(\delta_{gz} + \lambda^G_{gz}\right)}{1 + \kappa^E_{gz}\left(1 - F_{gz}\left(x\right)\right)}, \tag{C.14}$$

where $\kappa^e_{gz} = \lambda^e_{gz}/(\delta_{gz} + \lambda^G_{gz})$. The acceptance probability for a firm offering x is

$$h_{gz}(x) = \frac{u_{gz} + s^E_{gz}\left(1 - u_{gz}\right) G_{gz}\left(x\right) + s^G_{gz}}{u_{gz} + s^E_{gz}\left(1 - u_{gz}\right) + s^G_{gz}} \tag{C.15}$$

$$= \frac{\delta_{gz} + s^E_{gz}\left(\lambda^U_{gz} + \lambda^G_{gz}\right) G_{gz}\left(x\right) + s^G_{gz}\left(\delta_{gz} + \lambda^U_{gz} + \lambda^G_{gz}\right)}{\delta_{gz} + s^E_{gz}\left(\lambda^U_{gz} + \lambda^G_{gz}\right) + s^G_{gz}\left(\delta_{gz} + \lambda^U_{gz} + \lambda^G_{gz}\right)} \tag{C.16}$$

$$= \frac{1 + s^E_{gz}\kappa^U_{gz} G_{gz}\left(x\right) + s^G_{gz}\left(1 + \kappa^U_{gz}\right)}{1 + s^E_{gz}\kappa^U_{gz} + s^G_{gz}\left(1 + \kappa^U_{gz}\right)} \tag{C.17}$$

$$= \frac{1 + s^E_{gz}\kappa^U_{gz}\left[\frac{F_{gz}(x)}{1+\kappa^E_{gz}\left[1-F_{gz}(x)\right]}\right] + s^G_{gz}\left(1 + \kappa^U_{gz}\right)}{1 + s^E_{gz}\kappa^U_{gz} + s^G_{gz}\left(1 + \kappa^U_{gz}\right)} \tag{C.18}$$

$$= \frac{1 + \kappa^E_{gz}\left[1 - F_{gz}\left(x\right)\right] + s^E_{gz}\kappa^U_{gz} F_{gz}\left(x\right) + s^G_{gz}\left(1 + \kappa^U_{gz}\right)\left[1 + \kappa^E_{gz}\left[1 - F_{gz}\left(x\right)\right]\right]}{\left[1 + s^E_{gz}\kappa^U_{gz} + s^G_{gz}\left(1 + \kappa^U_{gz}\right)\right]\left[1 + \kappa^E_{gz}\left[1 - F_{gz}\left(x\right)\right]\right]}, \tag{C.19}$$

where $\kappa^U_{gz} = (\lambda^U_{gz} + \lambda^G_{gz})/\delta_{gz}$ and u_{gz} is substituted with its expression in equation (6). Combining expressions, expected profits per contacted worker are

$$z\left(\tilde{p}_{gz}, x\right) = h\left(x\right) J\left(\tilde{p}, x\right) \tag{C.20}$$

$$= \frac{\left\{1 + \kappa^E_{gz}\left[1 - F_{gz}\left(x\right)\right] + s^E_{gz}\kappa^U_{gz} F_{gz}\left(x\right) + s^G_{gz}\left(1 + \kappa^U_{gz}\right)\left[1 + \kappa^E_{gz}\left[1 - F_{gz}\left(x\right)\right]\right]\right\}\left(\tilde{p}_{gz} - x\right)}{\left[1 + s^E_{gz}\kappa^U_{gz} + s^G_{gz}\left(1 + \kappa^U_{gz}\right)\right]\left[1 + \kappa^E_{gz}\left(1 - F_{gz}\left(x\right)\right)\right]^2\left(\delta_{gz} + \lambda^G_{gz}\right)}. \tag{C.21}$$

Then, the firm's problem becomes

$$\max_{x,v}\left\{P_{gz}\left(\tilde{p}_{gz}, x\right) v q_{gz} - c^v_{gz}\left(v\right)\right\}, \tag{C.22}$$

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

tions satisfy

$$x^*_{gz}\left(\tilde{p}_{gz},\cdot\right) = \arg\max_x P_{gz}\left(\tilde{p}_{gz},x\right)$$
$$\frac{\partial c^v_{gz}\left(v^*\left(\tilde{p}_{gz},\cdot\right)\right)}{\partial v} = \max_x P_{gz}\left(\tilde{p}_{gz},x\right). \tag{C.23}$$

Since the vacancy cost function $c^v(\cdot)$ is convex, and $z\left(\tilde{p}_{gz},x\right)$ in equation (C.21) is strictly increasing in $\tilde{p}_{gz}$, then it follows from an application of the envelope theorem to equation (C.23) that $v^*\left(\tilde{p}_{gz},\cdot\right)$ 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 amenity cost shifter $c^{a,0}_g$ due to the result in Lemma 1. □

C.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 amenity cost shifter $c^{a,0}_g$. 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}:

$$[\partial v_{gz}]: \quad c^{v,0}_{gz}\frac{\partial \tilde{c}^v(v_{gz})}{\partial v_{gz}} = T_{gz}(\tilde{p}_{gz}-x_{gz})\left(\frac{1}{\delta_{gz}+\lambda^G_{gz}+\lambda^E_{gz}(1-F_{gz}(x_{gz}))}\right)^2, \tag{C.24}$$

$$[\partial x_{gz}]: \quad 1 = (\tilde{p}_{gz}-x_{gz})\frac{2\lambda^E_{gz}f_{gz}(x_{gz})}{\delta_{gz}+\lambda^G_{gz}+\lambda^E_{gz}(1-F_{gz}(x_{gz}))}, \tag{C.25}$$

where $T_{gz} = \mu_{gz}[(u_{gz}+s^G_{gz})\lambda^U_{gz}(\delta_{gz}+\lambda^G_{gz}+\lambda^E_{gz})]/V_{gz}$. Now consider equation (C.24); because the term on the right-hand side is always positive for $\tilde{p}_{gz} > \phi_{gz}$, it follows that optimal vacancies $v^*_{gz}(\tilde{p}_{gz}, c^{v,0}_{gz})$ 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

$$h_{gz}(\tilde{p}_{gz}) = \frac{\int_{\tilde{p}'\geq\phi_{gz}}^{\tilde{p}_{gz}} v^*_{gz}(\tilde{p}_{gz})\gamma_{gz}(\tilde{p}_{gz})}{V_{gz}}\, d\,\tilde{p}' \tag{C.26}$$

$$h'_{gz}(\tilde{p}_{gz}) = f_{gz}(x^*_{gz}(\tilde{p}_{gz}))\, x^{*\prime}_{gz}(\tilde{p}_{gz}) \tag{C.27}$$

$$f_{gz}(x^*_{gz}(\tilde{p}_{gz})) = h'_{gz}(\tilde{p}_{gz})/x^{*\prime}_{gz}(\tilde{p}_{gz}), \tag{C.28}$$

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 $\partial x^*_{gz}(\tilde{p}_{gz})/\partial\tilde{p}_{gz} = x^{*\prime}_{gz}(\tilde{p}_{gz})$ 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.28) 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.25) as

$$\frac{\partial x^*_{gz}(\tilde{p}_{gz})}{\partial \tilde{p}_{gz}} = (\tilde{p}_{gz} - x^*_{gz})\frac{2\lambda^E_{gz}}{\delta_{gz} + \lambda^G_{gz} + \lambda^E_{gz}(1 - h_{gz}(\tilde{p}_{gz}))}\frac{v^*_{gz}(\tilde{p}_{gz})}{V_{gz}}\gamma_{gz}(\tilde{p}_{gz}). \tag{C.29}$$

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 amenity cost shifter $c^{a,0}_g$ 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 $c^{a,0}_g$.

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

C.5 Alternative Modeling Assumption on Amenity Production

We here present an alternative formulation of firm's amenity production technology. While in the baseline formulation of the model firms produce gender-specific amenity values for each worker, in the alternative formulation 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 $\vec{a} = (a_1, a_2, \ldots, a_N) \in \mathbb{R}^{N\times 1}$ subject to cost $c^a(\vec{a})$. Workers derive gender-specific utility from amenities given a preference vector $\vec{\beta}_g = (\beta_{g,1}, \beta_{g,2}, \ldots, \beta_{g,N}) \in \mathbb{R}^{N\times 1}$. 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 $sign(\beta_M) = sign(\beta_F)$. Otherwise, if $sign(\beta_M) \neq sign(\beta_F)$, 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:*

$$N \geq 2. \tag{C.30}$$

Assumption C.2. *The gender-specific preference vectors $\vec{\beta}^a_M$ for men and $\vec{\beta}^a_F$ for women are linearly independent:*

$$\nexists c \in \mathbb{R} \quad s.t. \quad \vec{\beta}_M = c\vec{\beta}_F. \tag{C.31}$$

Under these assumptions, the following result obtains, 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^a_M, U^a_F) at a given employer, there exists an amenity vector $\vec{a} = (a_1, a_2, \ldots, a_N)$ such*

that

$$\underset{2\times 1}{\begin{bmatrix} U_M^a \\ U_F^a \end{bmatrix}} = \underset{2\times N}{\begin{bmatrix} \vec{\beta}_M' \\ \vec{\beta}_F' \end{bmatrix}} \underset{N\times 1}{\vec{a}}. \tag{C.32}$$

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.32) 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 to that in the baseline model:

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

$$c^a\left(\vec{a}, l_M, l_F\right) = \int_{j=0}^{l_M+l_F} \sum_{i=1}^{N} c_i^{a,0} \tilde{c}^a\left(a_i\right) dj, \tag{C.33}$$

where j indexes workers, i indexes amenities, $c_i^{a,0}$ is an amenity-specific cost shifter that differs across firms, and $\tilde{c}^a\left(a_i\right)$ 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 $i \in \{1, \ldots, N\}$ 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 $c^a\left(\vec{a}\right)$ such that $\vec{a}$ solves*

$$\left(\vec{a}, w_M, w_F, v_M, v_F\right) = \underset{\tilde{\vec{a}}, \tilde{w}_M, \tilde{w}_F, \tilde{v}_M', \tilde{v}_F}{\arg\max} \left\{ \sum_{g=M,F} \left[(1-\tau_g)p - \tilde{w}_g - c^a\left(\tilde{\vec{a}}\right)\right] l_g\left(\tilde{\vec{a}}, \tilde{w}_g, \tilde{v}_g\right) - \sum_{g=M,F} c^v\left(\tilde{v}_g\right) \right\} \tag{C.34}$$

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

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

$$[\partial a_i]: \quad c_i^{a,0} \frac{\partial \tilde{c}^a\left(a_i\right)}{\partial a_i} = \frac{\left[p - w_M - c^a\left(\vec{a}\right)\right] \frac{\partial l_M(\vec{a}, w_M, v_M)}{\partial a_i} + \left[(1-\tau)p - w_F - c^a\left(\vec{a}\right)\right] \frac{\partial l_F(\vec{a}, w_F, v_F)}{\partial a_i}}{l_M\left(\vec{a}, w_g, v_g\right) + l_F\left(\vec{a}, w_F, v_F\right)} \tag{C.35}$$

$$[\partial w_M]: \quad 1 = \left[p - w_M - c^a\left(\vec{a}\right)\right] \frac{\frac{\partial l_M(\vec{a}, w_M, v_M)}{\partial w_M}}{l_M\left(\vec{a}, w_M, v_M\right)} \tag{C.36}$$

$$[\partial w_F]: \quad 1 = \left[(1-\tau)p - w_F - c^a\left(\vec{a}\right)\right] \frac{\frac{\partial l_F(\vec{a}, w_F, v_F)}{\partial w_F}}{l_F\left(\vec{a}, w_F, v_F\right)} \tag{C.37}$$

$$[\partial v_M]: \quad \frac{\partial c_M^v(v_M)}{v_M} = \left[p - w_M - c^a\left(\vec{a}\right)\right] \frac{\partial l_M\left(\vec{a}, w_M, v_M\right)}{\partial v_M} \tag{C.38}$$

$$[\partial v_F]: \quad \frac{\partial c_F^v(v_F)}{v_F} = \left[p - w_F - c^a\left(\vec{a}\right) - z\right] \frac{\partial l_F\left(\vec{a}, w_F, v_F\right)}{\partial v_F}. \tag{C.39}$$

Note that the only FOC containing the term $\partial \tilde{c}^a(a_i)/\partial a_i$ is equation (C.35). 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.36)–(C.37) and vacancy posting in equations (C.38)–(C.39). By Assumption C.3, there are N equations (i.e., FOCs with respect to a_i for $i = 1, 2, \ldots, N$) with N free parameters (i.e., amenity cost shifters $c_i^{a,0}$ for $i = 1, 2, \ldots, N$), 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

$$[\partial a_i]: \quad c_i^{a,0} \times \frac{\partial \tilde{c}^a(a_i)}{\partial a_i} \sum_{g=M,F} l_g(x_g, v_g) = \beta_{M,i} l_M(x_M, v_M) + \beta_{F,i} l_F(x_F, v_F), \quad \forall i. \tag{C.40}$$

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 3.2. For example, in the data, women make up 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*

$$c^a\left(\{\vec{a}_j\}_j\right) = \int_{j=0}^{l_M+l_F} \sum_{i=1}^{N} c_i^{a,0} \tilde{c}^a(a_{i,j})\, dj, \tag{C.41}$$

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 $i \in \{1, \ldots, N\}$ 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:

$$\max_{\{x_g, v_g\}_{g=M,F}} \left\{ \sum_{g=M,F} \left[(1-\tau_g)p\right] l_g(x_g, v_g) - c^x(x_M, x_F) - \sum_{g=M,F} c_g^v(v_g) \right\}, \quad \forall (g), \tag{C.42}$$

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

$$c^x(x_M, x_F) = \min_{w_M, w_F, \{\vec{a}_j\}_j} \left\{ w_M l_M(x_M, v_M) + w_F l_F(x_F, v_F) + \int_{j=0}^{l_M+l_F} \sum_{i=1}^{N} c^a\left(a_{i,j}\right) dj \right\} \tag{C.43}$$

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

$$= \min_{\{\vec{a}_g\}_g} \left\{ \left(x_M - \vec{a}'_M \beta_M\right) l_M\left(x_M, v_M\right) + \left(x_F - \vec{a}'_F \beta_F\right) l_F\left(x_F, v_F\right) \right. \tag{C.44}$$

$$\left. + \sum_{i=1}^{N} \left[l_M\left(x_M, v_M\right) c^a\left(a_{M,i}\right) + l_F\left(x_F, v_F\right) c^a\left(a_{F,i}\right) \right] \right\},$$

where we move from equation (C.43) to equation (C.44) 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:

$$\left[\partial a_{g,i}\right]: \quad c_i^{a,0} \times \frac{\partial \tilde{c}^a\left(a_{g,i}\right)}{\partial a_{g,i}} = \beta_{g,i}, \quad \forall\,(g,i)\,. \tag{C.45}$$

Equation (C.45) 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 $\beta_{g,i}$. Obviously, $\partial a_i^* / \partial c_i^{a,0} < 0$ and $\partial a_i^* / \partial \beta_{g,i} > 0$ by the chain rule. The optimal wage is then chosen to deliver the remainder of 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 8. 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.

Choice of Model. 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 further assumptions on 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 are 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 formu-

lations 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 parametrizations of the cost function.

C.6 Further Discussion of Model Assumptions

Output Separability in 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. Supporting our assumption, Fukui et al. (2023) find a small crowd-out between women and men in the U.S. Therefore, a natural starting point treats men and women as perfect substitutes in production.

Labor Market Segmentation. Here we argue that our assumption of labor market segmentation by worker types is reasonable give 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 and Stein, 2020) 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. In Section 3.2, we show that there are also large differences in amenity utilization across genders. Our model provides a natural way of rationalizing these differences.

C.7 Alternative Modeling Assumptions on Vacancy Posting

C.7.1 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 strong prediction that any firm will employ either only men or only women, except in knife-edge cases.

Setup. 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: $c_z^v(0) = 0$, $\partial c_z^v(\cdot)/\partial v > 0$, $\partial^2 c_z^v(\cdot)/\partial v^2 > 0$.

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

$$\max_{x_{Mz},x_{Fz},v_{Mz},v_{Fz}} \left\{ \sum_{g=M,F} (\tilde{p}_{gz} - x_{gz}) l_{gz}(x_{gz}, v_{gz}) - c_z^v(v_{Mz} + v_{Fz}). \right\} \tag{C.46}$$

The FOCs with respect to vacancy posting now read

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

$$[\partial v_{Fz}]: \quad c_z^{v\prime}(v_{Mz} + v_{Fz}) = T_{Fz}(\tilde{p}_{Fz} - x_{Fz}) \left(\frac{1}{\delta_{Fz} + \lambda_{Fz}^G + \lambda_{Fz}^E (1 - F_{Fz}(x_{Fz}))} \right)^2. \tag{C.48}$$

Putting equations (C.47) and (C.48) 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 amount of vacancies posted. This is because wages are set according to other first-order conditions, which do not depend on the amount of vacancies posted by that firm.

Since the right-hand sides in equations (C.47) and (C.48) 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.

Empirical Shortcomings. According to the above analysis, except for knife-edge cases, firms would hire only men or only women—whichever gives 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.

C.7.2 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 3.1.

Setup. 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

$$\max_{x_{Mz},x_{Fz},v_a} \left\{ \sum_{g=M,F} (\tilde{p}_{gz} - x_{gz}) l_{gz}(x_{gz}, v_a) - c_z^v(v_z) \right\}. \tag{C.49}$$

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

$$m_{gz} = \chi_{gz}[\mu_{gz}(u_{gz} + s^E_{gz}(1-u_{gz}) + s^G_{gz})]^\alpha V_z^{1-\alpha} \frac{V_{gz}}{V_z}, \tag{C.50}$$

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

Equilibrium Characterization. The following equation represents the law of motion of firm sizes:

$$\dot{l}_{gz}(x,v) = -\delta_{gz} l_{gz}(x,v) - s_{gz}\lambda^E_{gz}(1-F_{gz}(x))l_{gz}(x,v) + \tag{C.51}$$

$$vq_{gz}\left[\frac{u_{gz} + s^G_{gz}}{u_{gz} + s^G_{gz} + (1-u_{gz})s^e_{gz}} + \frac{(1-u_{gz})s^e_{gz}}{u_{gz} + s^G_{gz} + (1-u_{gz})s_{gz})} G_{gz}(x)\right]. \tag{C.52}$$

Solving for the stationary solution,

$$l_{gz}(x_{gz}, v_z) = \left(\frac{1}{\delta_{gz} + \lambda^G_{gz} + \lambda^E_{gz}(1-F_{gz}(x_{gz}))}\right)^2 \frac{v_z}{V_z}\mu_{gz}(u_{gz} + s^G_{gz})\lambda^U_{gz}(\delta_{gz} + \lambda^G_{gz} + \lambda^E_{gz}). \tag{C.53}$$

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

$$\max_{x_{Mz}, x_{Fz}, v_z} \left\{ T_{Mz} v_z(\tilde{p}_{Mz} - x_{Mz})\left(\frac{1}{\delta_{Mz} + \lambda^G_{Mz} + \lambda^E_{Mz}(1-F_{Mz}(x_{Mz}))}\right)^2 \right. \tag{C.54}$$

$$\left. + T_{Fz} v_z(\tilde{p}_{Fz} - x_{Fz})\left(\frac{1}{\delta_{Fz} + \lambda^G_{Fz} + \lambda^E_{Fz}(1-F_{Fz}(x_{Fz}))}\right)^2 - c_z(v_z)\right\}. \tag{C.55}$$

The associated FOCs read

$$c'(v_z) = T_{Mz}(\tilde{p}_{Mz} - x_{Mz})\left(\frac{1}{\delta_{Mz} + \lambda^G_{gz} + \lambda^E_{Mz}(1-F_{Mz}(x_{Mz}))}\right)^2 \tag{C.56}$$

$$+ T_{Fz}(\tilde{p}_{Fz} - x_{Fz})\left(\frac{1}{\delta_{Fz} + \lambda^G_{gz} + \lambda^E_{Fz}(1-F_{Fz}(x_{Fz}))}\right)^2 \tag{C.57}$$

$$1 = (\tilde{p}_{Mz} - x_{Mz})\frac{2\lambda^E_{Mz} f_{Mz}(x_{Mz})}{\delta_{Mz} + \lambda^G_{Mz} + \lambda^E_{Mz}(1-F_{Mz}(x_{Mz}))} \tag{C.58}$$

$$1 = (\tilde{p}_{Fz} - x_{Fz})\frac{2\lambda^E_{Fz} f_{Fz}(x_{Fz})}{\delta_{Fz} + \lambda^G_{Fz} + \lambda^E_{Fz}(1-F_{Fz}(x_{Fz}))}. \tag{C.59}$$

Empirical Shortcomings. Recall from Section 3.1 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.53), we can write the female share of a firm as

$$s_f = \frac{l_{Fz}(x_{Fz}, v_z)}{l_{Fz}(x_{Fz}, v_z) + l_{Mz}(x_{Mz}, v_z)} \tag{C.60}$$

$$= \frac{1}{1 + \frac{\left(\frac{1}{\delta_{Mz}+\lambda^G_{Mz}+\lambda^E_{Mz}(1-F_{Mz}(x_{Mz}))}\right)^2 (u_{Mz}+s^G_{Mz})\lambda^U_{Mz}(\delta_{Mz}+\lambda^G_{Mz}+\lambda^E_{Mz})}{\left(\frac{1}{\delta_{Fz}+\lambda^G_{Fz}+\lambda^U_{Fz}(1-F_{Fz}(x_{Fz}))}\right)^2 (u_{Fz}+s^G_{Fz})\lambda^U_{Fz}(\delta_{Fz}+\lambda^G_{Fz}+\lambda^E_{Fz})}}. \tag{C.61}$$

On the right-hand side, we can substitute our empirical estimates of the U-E transition rates λ^u_{gz}, the E-U transition rates δ_{gz}, the compulsory offer arrival rate λ^G_{gz}, the voluntary offer arrival rate λ^e_{gz}, and the compulsory offer search units s^G_{gz} from Table 5. Thus, we obtain:

$$(u_{Mz} + s^G_{Mz})\lambda^U_{Mz}(\delta_{Mz} + \lambda^G_{Mz} + \lambda^E_{Mz}) = (0.236 + 0.101) \times 0.104 \times (0.036 + 0.010 + 0.009) \approx 0.0019 \tag{C.62}$$

$$(u_{Fz} + s^G_{Fz})\lambda^U_{Fz}(\delta_{Fz} + \lambda^G_{Fz} + \lambda^E_{Fz}) = (0.219 + 0.083) \times 0.091 \times (0.028 + 0.007 + 0.007) \approx 0.0012. \tag{C.63}$$

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

$$s_f = \frac{1}{1 + 1.583 \times \left(\frac{\delta_{Fz}+\lambda^G_{Fz}+\lambda^E_{Fz}(1-F_{Fz}(x_{Fz}))}{\delta_{Mz}+\lambda^G_{Mz}+\lambda^E_{Mz}(1-F_{Mz}(x_{Mz}))}\right)^2} = \frac{1}{1 + 1.583 \times \left(\frac{0.036+0.007\times(1-F_{Fz}(x_{Fz}))}{0.046+0.009\times(1-F_{Mz}(x_{Mz}))}\right)^2}. \tag{C.64}$$

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 $\overline{s}_f$ 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's being close to 0 (1) in the data.

C.8 Comparison with Sorkin (2018)

Table C.1. Comparison with Sorkin (2018)

	Sorkin (2018)	Morchio and Moser (2023)
Equilibrium concept	Partial equilibrium: does not model underlying sources of firm heterogeneity	General equilibrium: firm wages, amenities, and vacancies are determined in equilibrium
Worker heterogeneity	None	Ability, gender
Firm heterogeneity	Amenity cost parameter, separation shock rate, relocation shock rate	Productivity, gender wedge, gender-specific amenity cost parameters
Idiosyncratic heterogeneity	Type-I extreme value distribution with scale parameter 1	None
Utility function	$V = \omega_1[\ln(w) + \ln(a)]$	$V = \omega_0 + \omega_1[w + a]$
Determination of utility	Exogenous firm values	Endogenous firm values
Firm optimization	Choose pay and amenities to maximize profit s.t. exogenous firm value	Choose pay, amenities, and vacancies s.t. profit maximization
Steady state?	No	Yes
Endogenous job destruction?	Yes	No
Time	discrete	Continuous
Unemployed reject offers?	Yes	No
Reasons to move to lower pay	Amenities, idiosyncratic utility shock, relocation shock	Amenities, relocation shock
Reasons to move to lower rank	Idiosyncratic utility shock, relocation shock	Relocation shock
Compensating differentials	Only "Rosen" motive, cannot identify "Mortensen" motive	Whole joint distribution of (w, a)
Variance of amenities	Identify "Rosen" component but not "Mortensen" component of $Var(\ln(a_j))$	Identify both "Rosen" and "Mortensen" components of $Var(a_j)$
Variance of pay	Decompose Var$(\ln w_j)$ into rents and compensating differentials	Decompose $Var(\ln w_j)$ into $Var(\ln x_j)$, $Var(\tilde{a}_j)$, and $Cov(x_j, \tilde{a}_j)$
Country	U.S.	Brazil
Data source	Linked employer-employee data (LEHD)	Linked employer-employee data (RAIS)
Data coverage	Employees in 27 U.S. states	All formal-sector employees
Start date	2000Q4	January 2007
End date	2008Q1	December 2014
Period length	29 quarters	84 months
Data frequency	Quarterly	Monthly

Note: This table compares the theoretical framework in Sorkin (2018) with that in Section 4 of the current paper.

D Identification Appendix

D.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 $\tilde{p}_g$ and amenity cost shifter $c_g^{a,0}$ is*

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

where

$$\alpha_z = \ln z, \tag{D.2}$$

$$\psi_g^w\left(\tilde{p}_g, c_g^{a,0}\right) = \ln\left(\tilde{p}_g - a_g^*\left(c_g^{a,0}\right) - \int_{\tilde{p}' \geq \phi_g}^{\tilde{p}_g} \left[\frac{1+\kappa_g^E\left[1-F_g\left(x_g^*\left(\tilde{p}_g\right)\right)\right]}{1+\kappa_g^E\left[1-F_g\left(x_g^*\left(\tilde{p}'\right)\right)\right]}\right]^2 d\tilde{p}'\right). \tag{D.3}$$

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

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

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

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

We can write the offer distribution as

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

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.28) in Lemma 3, except that here we are integrating over exogenous rather than endogenous vacancies. Applying the Envelope Theorem yields

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

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

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

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 $a_{gz} = a_g z$, and the cost of producing amenities as $c_{gz}^a(a_{gz}) = z c_g^a(a_g)$.

Summing up, it follows that $a_{gz} = z a_g$ and $c_{gz}^a(a_{gz}) = z c_g^a(a_g)$. Therefore, composite productivity $\tilde{p}_{gz} = (1 - \tau_g)pz + a_{gz} - c_{gz}^a(a_{gz})$ is proportional to z, and we can write $\tilde{p}_{gz} = z\tilde{p}_g$, where $\tilde{p}_g = (1 - \tau_g)p + a_g - c_g^a(a_g)$ 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 that vacancies of each firm are constant across ability markets, we can rewrite equation (D.8) as

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

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

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

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

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

Next we show that if $x\left(z, \tilde{p}_g\right)$ 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

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

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

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 $w = x - a$, we can write monetary wages as

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

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 first-order condition for vacancy creation in equation (C.24) as

$$zc_g^{v,0}\frac{\partial \tilde{c}_g^v(v_{gz})}{\partial v_{gz}} = T_{gz}(\tilde{p}_{gz} - x^*_{gz}(\tilde{p}_{gz}))\left(\frac{1}{\delta_{gz} + \lambda^G_{gz} + \lambda^E_{gz}(1 - h_{gz}(\tilde{p}_{gz}))}\right)^2 \tag{D.15}$$

$$c_g^{v,0}\frac{\partial \tilde{c}_g^v(v_{gz})}{\partial v_{gz}} = T_g(\tilde{p}_g - x^*_g(\tilde{p}_g)\left(\frac{1}{\delta_g + \lambda^G_g + \lambda^E_g(1 - h_g(\tilde{p}_g))}\right)^2, \tag{D.16}$$

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, $\lambda^U_{gz} = \lambda^U_g$ (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. □

D.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 Card et al.'s (2016) variant of the AKM framework.

Corollary 1 (Equilibrium Amenity Equation)**.** *The equilibrium amenity of a worker of gender g and ability z at a firm with amenity cost shifter $c_g^{a,0}$ is*

$$\ln a_{gz}\left(c_g^{a,0}\right) = \underbrace{\alpha_z}_{\text{"worker amenity FE"}} + \underbrace{\psi^a\left(c_g^{a,0}\right)}_{\text{"gender-firm amenity FE"}}, \tag{D.17}$$

where

$$\alpha_z = \ln z, \tag{D.18}$$

$$\psi^a\left(c_g^{a,0}\right) = \frac{1}{1-\eta^a}\ln c_g^{a,0}. \tag{D.19}$$

Proof. This result follows directly from equation (C.7) 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 $\psi^a(c_g^{a,0})$ depends only on a firm's 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 et al. (2016). Our framework fills this gap by explicitly modeling firms' equilibrium wage and amenity choices.

D.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:

$$\ln(w_{ijt}) - \alpha_i + \psi_{G(i)j} + X_{it}\beta_{G(i)} + \varepsilon_{ijt}. \tag{D.20}$$

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 et al. (2002) and Card et al. (2016), equation (D.20) 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 et al. (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 et al. (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.20).

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 $\tilde{p}_g$ and $c_g^{a,0}$ across genders $g \in \{M, F\}$ so that $\psi_M(\tilde{p}_M, c_M^{a,0}) = \psi_F(\tilde{p}_F, c_F^{a,0})$. 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 $\tau_{Fr} \approx 0$. Let $\mathcal{A}_g \subseteq [0,1]$ be a set of dual-gender firms with $a_{gr} \approx a_{-gR_{-g}(J(r))}$. 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 $R_g(j) \in \mathcal{B}_g \cap \mathcal{D}_g \cap \mathcal{A}_g$ simultaneously for both genders g:*

$$\mathbb{E}_j[\psi_{Mj}] = \mathbb{E}_j[\psi_{Fj}] \quad \forall j : R_g(j) \in \mathcal{B}_g \cap \mathcal{D}_g \cap \mathcal{A}_g, \forall g \in \{M, F\}. \tag{D.21}$$

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

$$\tilde{p}_{gr} \equiv (1 - \tau_{gr})p_r + a_{gr} - c_{gr}^a(a_{gr}). \tag{D.22}$$

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

First, we consider a set of close-to-zero profit firms. 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 $x_{gr} \equiv w_{gr} + a_{gr}$ 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:

$$\tilde{p}_{gr} \approx x_{gr} \qquad \forall r \in \mathcal{B}_g. \tag{D.23}$$

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

$$\tilde{p}_{gr} \approx p_r + a_{gr} - c_g^a(a_{gr}) \qquad \forall r \in \mathcal{D}_g, \tag{D.24}$$

so the only gender-specific aspect of firms $r \in \mathcal{D}_g$ is $a_{gr} - c_g^a(a_{gr})$. Combining equations (D.23) and (D.24), we have $p_r + a_{gr} - c_g^a(a_{gr}) \approx w_{gr} + a_{gr}$ 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 $a_{gr} \approx a_{-gR_{-g}(J(r))}$. 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 amenity cost parameter $c^{a,0}$, a firm optimally provides the same level of amenities to both genders if and only if it faces the same cost function for both genders. As a result,

$$a_{gr} - c_g^a(a_{gr}) = a_{-gR_{-g}(J(r))} - c_{-g}^a(a_{-gR_{-g}(J(r))}) \qquad \forall r \in \mathcal{A}_g, \tag{D.25}$$

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.22) 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.20) across men and women according to equation (D.21). □

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.21) 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., $a_{MR_M(j)} \approx a_{FR_F(j)}$) 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 et al. (2016) is not valid in our environment, equation (D.21) provides a normalization that extends the argument in Card et al. (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.23) reduces to

$$(1 - \tau_{gr})p_r - w_{gr} \approx 0 \qquad \forall r \in \mathcal{B}. \tag{D.26}$$

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.20) that is consistent with our model featuring exogenous amenities:

$$\mathbb{E}_r[\psi_{Fr}] = \mathbb{E}_r[\psi_{Mr}] \qquad \forall r \in \mathcal{B} \cap \mathcal{D}. \tag{D.27}$$

The normalization in equation (D.27) 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.

D.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 can be identified from 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

$$\begin{aligned} x_{gz}(j) &= w_{gz}(j) + a_{gz}(j) && \text{(D.28)}\\ &= \exp\left(\alpha_z + \psi_g^w(j)\right) + \exp\left(\alpha_z + \psi^a(j)\right) && \text{(D.29)}\\ &= \exp\left(\alpha_z\right)\exp\left(\psi_g^w(j)\right) + \exp\left(\alpha_z\right)\exp\left(\psi^a(j)\right) && \text{(D.30)}\\ &= \exp\left(\alpha_z\right)\left[\exp\left(\psi_g^w(j)\right) + \exp\left(\psi^a(j)\right)\right] && \text{(D.31)}\\ &= z\left[\exp\left(\psi_g^w(j)\right) + \exp\left(\psi^a(j)\right)\right] && \text{(D.32)} \end{aligned}$$

Fixing gender g, equation (D.32) 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 prove trivially 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, proving the proposition.

□

D.5 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., 1998; 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

$$\text{Poaching rank}_j = \frac{\text{number of E-to-E hires}_j}{\text{number of all hires}_j} \tag{D.33}$$

$$= \frac{\text{number of E-to-E hires}_j}{\text{number of E-to-E hires}_j + \text{number of U-to-E hires}_j}, \tag{D.34}$$

which in our model can be rewritten as

$$\text{Poaching rank}_j = \frac{\lambda^E G_j + \lambda^G (1-u)}{\lambda^E G_j + \lambda^G + \lambda^U u} \tag{D.35}$$

The poaching rank in equation (D.35) 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.[50] 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., 1998; 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

$$h_j := \left(1 - \delta - \lambda^G - \lambda^E (1 - F_j)\right) l_j. \tag{D.36}$$

Here, we have used the fact that firms are atomistic. This fact implies, for example, that workers hit with an involuntary job offer at 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

$$o_{i,j} = \lambda^G l_i f_j, \tag{D.37}$$

or else $r_i < r_j$, in which case

$$o_{i,j} = (\lambda^E + \lambda^G) l_i f_j. \tag{D.38}$$

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

[50] Note 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.

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

$$\rho_i = \frac{h_i}{l_i}, \quad \forall j \in \{1, \ldots, N\} \tag{D.39}$$

and the off-diagonal separation probabilities

$$p_{i,j} = \frac{o_{i,j}}{l_i} \quad \forall j \neq i, j > 0. \tag{D.40}$$

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:

$$P = \begin{bmatrix} \rho_0 & p_{0,1} & p_{0,2} & \cdots & p_{1,N} \\ p_0 & \rho_1 & p_{1,2} & \cdots & p_{1,N} \\ p_0 & p_{2,1} & \rho_2 & \cdots & p_{2,N} \\ \vdots & \vdots & \ddots & \vdots & \\ p_0 & p_{N,1} & p_{N,2} & \cdots & \rho_N \end{bmatrix} \tag{D.41}$$

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:

$$\dot{l}_j = \left[-\delta - \lambda^E \left[1 - F_j\right] - \lambda^G\right] l_j + \left[\frac{u + (1-u)s^E G_r + s^G}{u + (1-u)s^E + s^G}\right] v_j q. \tag{D.42}$$

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

$$(1-u) = \sum_{j \in \{1/N, \ldots, 1\}} l_j, \tag{D.43}$$

$$G_j = \frac{1}{1-u} \sum_{i \in \{1/N, \ldots, r\}} l_i, \tag{D.44}$$

$$q = m(\theta)/V = \lambda^U \left[\frac{u + (1-u)s^E + s^G}{V}\right] \tag{D.45}$$

$$F_j = \sum_{k \in \{1, \ldots, N\}} (f_k \mathbf{1}[r(j) > r(k)]) \tag{D.46}$$

$$f_j = \frac{v_j}{V} \tag{D.47}$$

so we can rewrite that, in steady state,

$$0 = \left[-\delta - \lambda^E (1 - F_r) - \lambda^G\right] l_j + \left[u + (1-u)s^E G_r + s^G\right] \lambda^U \frac{v}{V}. \tag{D.48}$$

$$\left[\delta + \lambda^G + \lambda^E (1 - F_r)\right] l_j = \frac{v}{V} \left[\lambda^U u + \lambda^E (1-u) G_j + \lambda^G\right] \tag{D.49}$$

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

$$\left[\delta + \lambda^G + \lambda^E \left(1 - \sum_{k\in\{1,\dots,N\}} (f_k \mathbf{1}[r(j) > r(k)])\right)\right] l_j = f_j \left[\lambda^U u + \lambda^G + \lambda^E \sum_{k\in\{1,\dots,N\}} (l\mathbf{1}[r(j) > r(k)])\right]. \tag{D.50}$$

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

$$f_j \left[\lambda^U l_0 + \lambda^G + \lambda^E \sum_{k\in\{1,\dots,N\}} (l_k \mathbf{1}[r(j) > r(k)])\right] \tag{D.51}$$

$$+(1 - \delta - \lambda^G - \lambda^E (1 - \sum_{k\in\{1,\dots,N\}} [f_k \mathbf{1}[r(j) > r(k)]]))l_j = l_j. \tag{D.52}$$

By substituting $l_0 = u$ by definition, it is immediately apparent that equations (D.50) and (D.52) 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

$$l_j = f_j \left[\frac{\lambda^U u + \lambda^G + \lambda^E \sum_{k\in\{1,\dots,N\}} (l_k \mathbf{1}[r(j) > r(k)])}{\delta + \lambda^G + \lambda^E \left(1 - \sum_{k\in\{1,\dots,N\}} (f_k \mathbf{1}[r(j) > r(k)])\right)}\right] \tag{D.53}$$

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):

$$l_j = f_j \left[\frac{1}{\delta + \lambda^G + \lambda^E (1 - F_j)}\right]^2 (u + s^G)\lambda^U (\delta + \lambda^G + \lambda^E) \tag{D.54}$$

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

$$P' l = l, \tag{D.55}$$

where the vector l contains the steady-state firm sizes of equation (D.54), which include hiring intensity f_j. Therefore, dividing firm size l by hiring intensity will give 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. (1998), 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.

□

D.6 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 ↔ 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.[51]

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

$$\hat{\delta} = \mathbb{E}_i \mathbf{1}\left[\text{nonemployed}_{i,t+1} \middle| \text{employed}_{i,t}\right]. \tag{D.56}$$

Empirical Job Finding Rate ↔ 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 $\widehat{\lambda^U + \lambda^G}$ separately by gender.[52]

Empirical Share of Moves Down the Firm Ranks ↔ 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

$$J2J_r = l_r[\lambda^E(1 - F^r) + \lambda^G], \tag{D.57}$$

where l_r is the gender-specific number of workers at firm r. Rearranging and averaging across all firms, we have

$$\hat{\lambda}^G = \mathbb{E}_r\left[\frac{J2J_{r\downarrow}}{l_r F_r}\right], \tag{D.58}$$

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 ↔ 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.57) to estimate λ^E as

$$\hat{\lambda}^E = \frac{J2J_r/n_r - \hat{\lambda}^G}{1 - F_r}. \tag{D.59}$$

Notice that all of these parameters are over-identified, as in principle we could use just a fraction of the firms and of the 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

[51] Since $f_r = v_r/V$ 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.

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

that a worker's job-finding probability due to the aggregate matching function is

$$\frac{m}{u} = \lambda^U + \lambda^G = \chi \left[\mu \left(u + s^E(1-u) + s^G\right)\right]^{\alpha} V^{1-\alpha}, \tag{D.60}$$

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.60) for aggregate vacancies V:[53]

$$V = \left(\frac{\lambda^U + \lambda^G}{\left[\mu\left(u + s^E(1-u) + s^G\right)\right]^{\alpha}}\right)^{\frac{1}{1-\alpha}}. \tag{D.61}$$

□

D.7 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: productivity $p(r)$, the gender wedge $\tau(r)$, and the amenity cost shifter $c^{a,0}(r)$.*

Proof. Here, we present an identification result based on continuous firm types, as in the model developed in Section 4 and further discussed in Section 5. In Appendix D.8, 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 $\tilde{p}(r) = (1-\tau(r))p(r) + a(r) - c^a(a(r); r)$. Let

$$T \equiv \frac{\mu[(u + s^G)\lambda^u(\delta + \lambda^G + \lambda^E)]}{V} \tag{D.62}$$

be a constant which only depends on previously estimated labor market rates and aggregate vacancies. We start from the firms' first-order conditions (FOCs) for optimality in equations (C.24) and (C.25). We substitute the functional form of $c^v(v)$ in equation (C.24) and express the FOCs as a coupled pair of differential equations:

$$h'(\tilde{p}(r)) = \frac{1}{V}\left[\frac{T(\tilde{p}(r) - x(\tilde{p}(r)))}{c^{v,0}\left[\delta + \lambda^G + \lambda^E(1 - h(\tilde{p}(r)))\right]^2}\right]^{\frac{1}{\eta^v - 1}} \gamma(\tilde{p}(r)), \tag{D.63}$$

$$x'(\tilde{p}(r)) = \frac{1}{V}\frac{2\lambda^E(\tilde{p}(r) - x(\tilde{p}(r)))}{\delta + \lambda^G + \lambda^E(1 - h(\tilde{p}(r)))}\left[\frac{T(\tilde{p}(r) - x(\tilde{p}(r)))}{c^{v,0}\left[\delta + \lambda^G + \lambda^E(1 - h(\tilde{p}(r)))\right]^2}\right]^{\frac{1}{\eta^v - 1}} \gamma(\tilde{p}(r)). \tag{D.64}$$

We first derived the differential equation (D.64) as equation (C.29) in Appendix C.4, and here we substituted therein the explicit solution for vacancies in equation (D.63). In equations (D.63)–(D.64) 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 $F(r) = \int_{r'=0}^{r} f(r')\, dr'$.

[53] As discussed in the main text, this normalization is inconsequential for our purposes.

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

$$\Gamma(\tilde{p}(r)) = \int_{\tilde{p}'=\underline{\tilde{p}}}^{\tilde{p}(r)} \gamma(\tilde{p}')\, d\tilde{p}' = \int_{r'=0}^{r} 1\, dr' = r, \tag{D.65}$$

That is, 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.63) to obtain an expression for *flow profits per matched worker*,

$$P(r) \equiv \tilde{p}(r) - x(\tilde{p}(r)) = [f(r)V]^{\eta^v - 1} \frac{c^{v,0}}{T} \left[\delta + \lambda^G + \lambda^E(1 - F(r))\right]^2, \tag{D.66}$$

In equation (D.66), $f(r)$, $c^{v,0}$, and $F(r)$ are all known quantities identified by applying Proposition 3 to the data. Thus, equation (D.66) 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:

$$x'(r) = \frac{1}{V} \frac{2\lambda^E P(r)}{\delta + \lambda^G + \lambda^E(1 - F(r))} \left[\frac{TP(r)}{c^{v,0}\left[\delta + \lambda^G + \lambda^E(1 - F(r))\right]^2}\right]^{\frac{1}{\eta^v - 1}}. \tag{D.67}$$

Plugging $P(r)$ from equation (D.66) 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.68) below—attains a lower bound strictly above but arbitrarily close to zero.[54]

This allows us to identify gender-firm-specific amenities by simply taking the difference between utility and wages at each firm, and productivity by adding wages and the cost of amenities to flow profits per matched worker:

$$a(r) = x(r) - w(r), \tag{D.68}$$

$$(1 - \tau(r))p(r) = P(r) + w(r) + c^a(a(r); r). \tag{D.69}$$

Consequently, we obtain gender-firm-specific 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

$$c^{a,0}(r) = [a(r)]^{1-\eta^a}. \tag{D.70}$$

Finally, 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.69), we recover the gender wedge $\tau(r)$ based on the ratio of estimated productivities net of the gender wedge at dual-gender firms. □

[54] Our choice of amenities starting strictly above but arbitrarily close to zero seems natural given the interpretation of amenities being endogenously produced by firms. At the same time, this choice tends to minimize the importance of amenities in total utility across firms. 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.

D.8 Identification of Firm-Level Parameters with Discrete Firm Types

Here, we adapt the identification proof in Section D.7 to discrete data where we observe a finite number of firms N. We observe their ranks, defined as $r \in \{1/N, 2/N, \ldots, 1\}$, their empirical hiring intensities f_r and their wage w_r.[55] The FOCs read:

$$h'(\tilde{p}_r) = \frac{1}{V}\left[\frac{T(\tilde{p}_r - x(\tilde{p}_r))}{c^{v,0}}\left(\frac{1}{\delta + \lambda^G + \lambda^E(1-h(\tilde{p}_r))}\right)^2\right]^{\frac{1}{\eta^v - 1}}\gamma(\tilde{p}_r), \tag{D.71}$$

$$x'(\tilde{p}_r) = \frac{1}{V}\frac{2\lambda^E(\tilde{p}_r - x(\tilde{p}_r))}{\delta + \lambda^G + \lambda^E(1-h(\tilde{p}_r))}\left[\frac{T(\tilde{p}_r - x(\tilde{p}_r))}{c^{v,0}}\left(\frac{1}{\delta + \lambda^G + \lambda^E(1-h(\tilde{p}_r))}\right)^2\right]^{\frac{1}{\eta^v - 1}}\gamma(\tilde{p}_r). \tag{D.72}$$

How do we move from the continuous representation to the discrete case? In what follows, we want to move from functions of model objects (e.g., $\tilde{p}_r$) to functions of ranks, r. For instance, the change in the CDF of recruiting intensities between two (discrete) ranks $r-1$ and r is

$$\Delta h(\tilde{p}_r) \equiv \int_{\tilde{p}_{r-1}}^{\tilde{p}_r} \frac{v(\tilde{p})}{V}\gamma(\tilde{p})\, d\tilde{p} \tag{D.73}$$

$$\approx \frac{v(\tilde{p}_{r-1})}{V}\gamma(\tilde{p}_{r-1}) \times [\tilde{p}_r - \tilde{p}_{r-1}], \tag{D.74}$$

where Δ is an operator that takes differences between the current and the previous (discrete) rank. We write a discretized version of equation (D.65) as follows:

$$\Gamma(\tilde{p}_r) = \int_{\tilde{p}_1}^{\tilde{p}_r} \gamma(\tilde{p})\, d\tilde{p} \tag{D.75}$$

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 simply[56]

$$\Gamma(\tilde{p}_r) = r \in \left\{\frac{1}{N}, \frac{2}{N}, \ldots, 1\right\}. \tag{D.76}$$

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

$$\gamma(\tilde{p}_{r-1})(\tilde{p}_r - \tilde{p}_{r-1}) \approx 1/N \tag{D.77}$$

[55] Instead 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.

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

at all rungs of the ladder.[57] Now we rewrite the change in the CDF of composite productivity offers using a finite-difference approximation and equation (D.71):

$$\Delta F(x(\tilde{p}_r)) = h'(\tilde{p}_{r-1})(\tilde{p}_r - \tilde{p}_{r-1}) \tag{D.78}$$

$$= \frac{1}{V}\left[\frac{T(\tilde{p}_{r-1} - x(\tilde{p}_{r-1}))}{c^{v,0}}\left(\frac{1}{\delta + \lambda^G + \lambda^E(1 - h(\tilde{p}_{r-1}))}\right)^2\right]^{\frac{1}{\eta^v - 1}} \gamma(\tilde{p}_{r-1})(\tilde{p}_r - \tilde{p}_{r-1}) \tag{D.79}$$

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

$$\hat{f}_r = \left[\frac{T(\tilde{p}_{r-1} - x(\tilde{p}_{r-1}))}{c^{v,0}}\left(\frac{1}{\delta + \lambda^G + \lambda^E(1 - F_{r-1})}\right)^2\right]^{\frac{1}{\eta^v - 1}} \frac{1}{VN} \tag{D.80}$$

All of the elements of equation (D.80) are known from the data, except for flow profits per matched worker, $P_{r-1} \equiv (\tilde{p}_{r-1} - x(\tilde{p}_{r-1}))$. Thus, we have one equation (D.80) in one unknown (P_r) for each firm rank r, which can be rewritten as

$$P_{r-1} \equiv (\tilde{p}_{r-1} - x_{r-1}(\tilde{p}_{r-1})) = \left(\hat{f}_r VN\right)^{\eta^v - 1} \frac{V c^{v,0}}{T}\left(\delta + \lambda^G + \lambda^E(1 - F_{r-1}))\right)^2. \tag{D.81}$$

Intuitively, a firm posts more vacancies if it makes greater flow profits per matched worker.

Going back to equation (D.72), we apply a similar finite-difference approximation:

$$\Delta x_r = x'(\tilde{p}_{r-1})(\tilde{p}_r - \tilde{p}_{r-1}) = \frac{2\lambda^E(P_{r-1})}{\delta + \lambda^G + \lambda^E(1 - F_{r-1})}\left[\frac{TP_{r-1}}{c^{v,0}}\left(\frac{1}{\delta + \lambda^G + \lambda^E(1 - F_{r-1})}\right)^2\right]^{\frac{1}{\eta^v - 1}} \frac{1}{VN}. \tag{D.82}$$

Recall that equation (D.81) above already identifies P_r. Therefore, without loss of generality, we impose the initial condition $x_0 = 0$, and then we can iteratively apply equation (D.82) through ranks $r > 0$ to deduce x_r as follows:

$$x_r = x_0 + \sum_{i=1}^{r} \Delta x_i. \tag{D.83}$$

Once x_r is known from equation (D.83), then by definition we deduce

$$a_r = x_r - w_r, \tag{D.84}$$

where w_r is the firm component of wages at firm rank r. Given a_r from equation (D.84), we can deduce $c^a(a_r)$ simply by plugging back a_r into the cost function $c^a(a)$. Finally, as we have already identified $P_r = \tilde{p}_r - x_r = p_r - c^a(a_r) - w_r$, we can deduce p_r firm by firm as

$$p_r = P_r + c^a(a_r) + w_r. \tag{D.85}$$

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

[57] Note that our data covers the universe (i.e., a very large number) of firms in Brazil. If we considered 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.

τ_j as

$$\tau_j = 1 - \frac{p_{Fj}}{p_{Mj}}. \tag{D.86}$$

In summary, equations (D.85), (D.84), and (D.86) jointly identify the unobserved multidimensional type (p_r^g, a_r^g, τ_r^g) 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.

D.9 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,

$$p(r) - w(r) - c^a(a_r) = [f(r)V]^{\eta^v - 1} \frac{c^{v,0}}{T} \left[\delta + \lambda^G + \lambda^E(1 - F(r))\right]^2. \tag{D.87}$$

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 $\rho\Pi(r)$,

$$\rho\Pi(r) = (p(r) - w(r) - c^a(a_r))\, l(r) - c^v(v(r)) \tag{D.88}$$

$$= [f(r)V]^{\eta^v} c^{v,0} - c^v(v(r)). \tag{D.89}$$

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

$$\rho\Pi(r) = [v(r)]^{\eta^v} c^{v,0} - c^{v,0} \frac{[v(r)]^{\eta^v}}{\eta^v} \tag{D.90}$$

$$= \left(1 - \frac{1}{\eta^v}\right) [v(r)]^{\eta^v} c^{v,0}. \tag{D.91}$$

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 $c^{v,0} \to \infty$ and arbitrarily small as $c^{v,0} \to 0^+$, 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

$$\rho\Pi(r) = \left(1 - \frac{1}{\eta^v}\right) [f(r)V]^{\eta^v} c^{v,0} \tag{D.92}$$

In equation (D.92), η^v is treated as unknown, however its value does not matter for any of our argument concerning identification of the vacancy cost shifter $c^{v,0}$. We can write the labor share as

$$\mathcal{L} \equiv \frac{\int w(r)l(r)\,dr}{\int [p(r)l(r) - c^a(a(r))l(r) - c^v(v(r))]\,dr} \tag{D.93}$$

$$= \frac{\int w(r)l(r)\,dr}{\int [w(r)l(r) + \rho\Pi(r)]\,dr} \tag{D.94}$$

$$= 1 - \frac{\int \rho\Pi(r)\,dr}{\int [w(r)l(r) + \rho\Pi(r)]\,dr}. \tag{D.95}$$

Since flow profits $\rho\Pi(r)$ in equation (D.92) are strictly increasing in the vacancy cost shifter $c^{v,0}$ and the labor share in equation (D.95) 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.95) 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.92) yields

$$\ln(\rho\Pi(r)) = \ln\left(1 - \frac{1}{\eta^v}\right) + \eta^v \ln[f(r)V] + \ln(c^{v,0}). \tag{D.96}$$

Taking variances on both sides of equation (D.96), we get

$$Var\left[\ln(\rho\Pi(r))\right] = Var\left[(\eta^v \times \ln f(r)) + \eta^v \times \ln V + \ln\left(c^{v,0}\right)\right] \tag{D.97}$$

$$= (\eta^v)^2 \times Var\left[\ln f(r)\right] \tag{D.98}$$

Except for $f(r)$, all terms in equation (D.97) are constant across firms, so they drop out of the calculation of the variance. As a result, equation (D.98) 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,

$$\ln w(r) = \alpha + \beta \ln \Pi(r) + \varepsilon(r), \tag{D.99}$$

where the regression coefficient β in equation (D.99) captures the elasticity of firm pay with respect to firm profits. The regression coefficient

$$\beta = \frac{Cov[\ln w(r), \ln(\rho\Pi(r))]}{Var[\ln(\rho\Pi(r))]}, \tag{D.100}$$

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.11) 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

$$\mathcal{A} \equiv \frac{\int c^a(a^*(r))l(r)\,dr}{\int \left[p(r)l(r) - c^a(a^*(r))l(r) - c^v(v(r))\right]dr} \tag{D.101}$$

$$= \frac{\int \frac{a^*(r)}{\eta^a}l(r)\,dr}{\int \left[p(r)l(r) - \frac{a^*(r)}{\eta^a}l(r) - c^v(v(r))\right]dr}. \tag{D.102}$$

Obviously, the aggregate amenity cost share in equation (C.11) 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 $\eta^a \to \infty$, 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. □

D.10 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 5. 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 Pareto marginal distribution.

We then feed the nonparametric joint distribution of $\{p,a\}$ to our discrete numerical solution algorithm, detailed in 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 parametrizations of the underlying amenities distribution. Under different parametrizations 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. η^v and η^a can be easily read in Panel D of Table D.1. In Column (1), $Var(a) = 0.1$ and $\rho(p,a) = 0$. In Column (2), $Var(a) = 0.15$ and $\rho(p,a) = 0$. In Column (3), $Var(a) = 0.125$ and $\rho(p,a) = 0$. In Column (4), $Var(a) = 0.2$ and $\rho(p,a) = -0.5$. In Column (5), $Var(a) = 0.15$ and $\rho(p,a) = 0.5$

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. The correlation for all estimates is equal to 1 approxi-

mated at the seventh decimal digit. The MSE for amenities is close to machine zero.

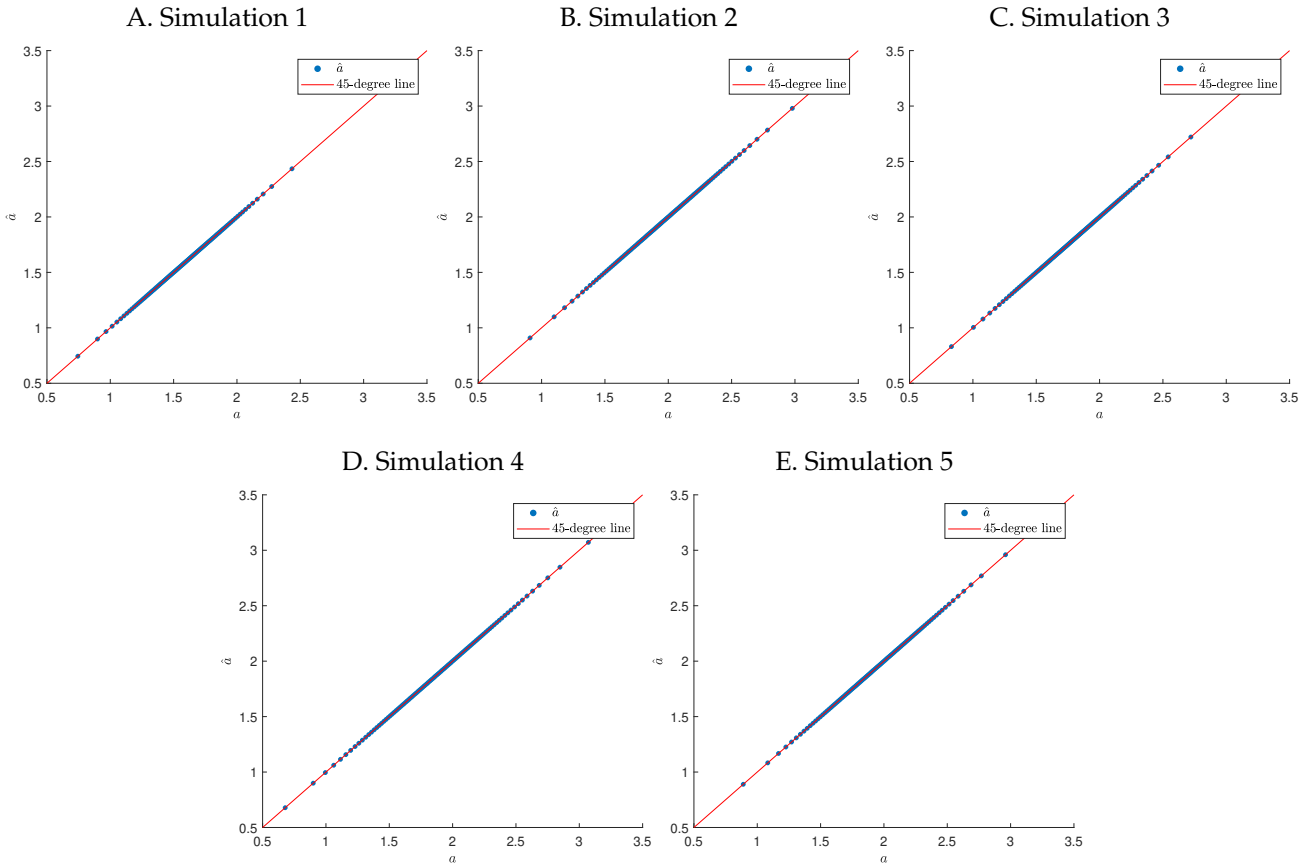
Figure D.1 visualizes the fit of our estimation routine with respect to the main objects of interest—namely, the amenity values. Across all five simulations, estimates lie on the 45 degrees line, showing how our procedure is robust to different joint distributions of amenities and productivity, and different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

Table D.1. Monte Carlo simulation and estimation on 100,000 Firms

	(1)	(2)	(3)	(4)	(5)
Panel A. Properties of amenity values a					
Standard deviation, true values	0.312	0.377	0.347	0.443	0.365
Standard deviation, estimates	0.312	0.377	0.347	0.443	0.365
Correlation with wages w, true values	-0.645	-0.629	-0.628	-0.812	-0.557
Correlation with wages w, estimates	-0.645	-0.629	-0.628	-0.812	-0.557
Correlation with ranks r, true values	0.561	0.616	0.595	0.475	0.805
Correlation with ranks r, estimates	0.561	0.616	0.594	0.475	0.805
Panel B. Properties of productivity p					
Standard deviation, true values	0.514	0.519	0.515	0.496	0.529
Standard deviation, estimates	0.514	0.519	0.515	0.496	0.529
Correlation with wages w, true values	0.720	0.663	0.713	0.824	0.457
Correlation with wages w, estimates	0.720	0.663	0.713	0.824	0.457
Correlation with ranks r, true values	0.785	0.807	0.776	0.602	0.831
Correlation with ranks r, estimates	0.785	0.807	0.776	0.602	0.831
Correlation with amenities a, true values	0.040	0.131	0.056	-0.368	0.457
Correlation with amenities $\hat{a}$, estimates	0.040	0.131	0.056	-0.368	0.457
Panel C. Goodness of fit					
Correlation($\hat{a}, a$)	1.000	1.000	1.000	1.000	1.000
Correlation($\hat{p}, p$)	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
Mean squared error of $\hat{p}$	0.000	0.000	0.000	0.000	0.000
Panel D. economy-wide parameters					
Elasticity of amenity cost function η_a, true value	8.000	4.000	7.000	9.000	10.000
Elasticity of amenity cost function $\hat{\eta_a}$, estimate	8.000	4.000	7.000	9.000	10.000
Elasticity of vacancy cost function η_v, true value	2.000	3.000	3.500	2.000	2.500
Elasticity of vacancy cost function $\hat{\eta_v}$, estimate	2.000	3.000	3.500	2.000	2.500

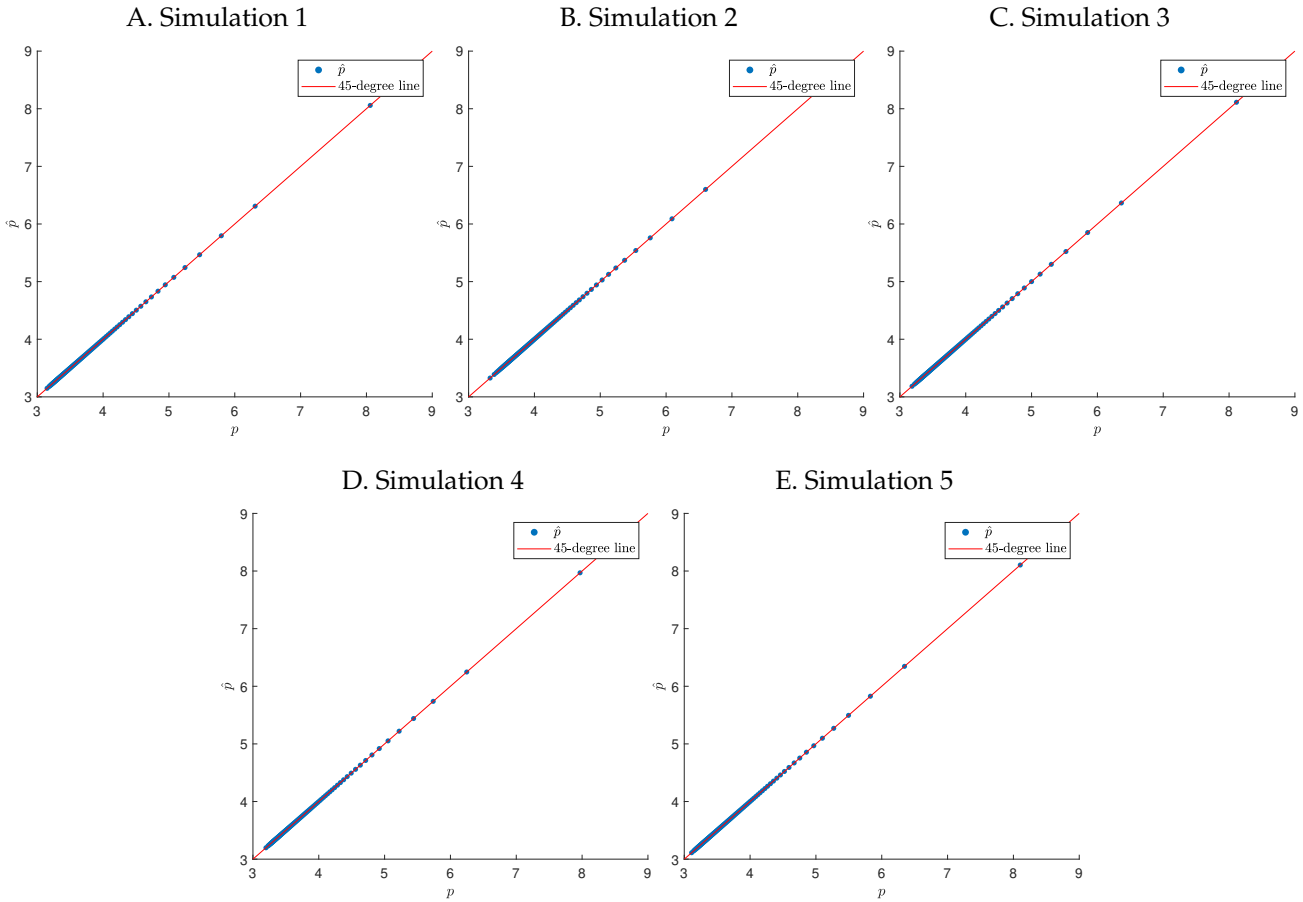
Note: Table reports estimation results using simulated data from 100,000 firms under different parametrizations of the underlying distribution of firm heterogeneity, including amenity values a, productivity p, wages w, and employer ranks r. MSE denotes the mean squared error. Columns (1)–(5) show separate simulations under different parametrizations of the data-generating process described in the text of Appendix D.10. *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.

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.

E Estimation Results Appendix

E.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 5 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,

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

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 et al. (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 5 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,

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

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 5 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 et al., 2023) suggest that the approximate cost share of amenities in value added is

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

Because there is significant uncertainty around this estimate, we use this number as a baseline and conduct robustness checks around it.

E.2 Details on Covariates Related to Gender Wedge Estimates

We include as covariates in Z_j in equation (24) the following six variables that we construct using the RAIS data, in addition to two sets of FEs.

Female Manager: An indicator for whether the employer has a woman in the highest-paid position.

Nonroutine Manual Task Intensity: 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 (2022) extending previous work of Autor and Dorn (2009) and Acemoglu and Autor (2011).

Nonroutine Interpersonal Task Intensity: The 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 (2022) extending previous work of Autor and Dorn (2009) and Acemoglu and Autor (2011).

Mean Working Hours: The mean log number of contractual work hours.

No Major Financial Stakeholders: An indicator for whether an employer has no major financial stakeholder, as proxied by their participation in the small-business tax regime *Simples Nacional*.[58]

Employer Size: Total employer size measured as the log number of full-time equivalent employees during a year.

Municipality FEs: Dummies for 4,733 municipalities represented in our sample.

Sector FEs: Dummies for 661 five-digit sectors represented in our sample.

E.3 Details on Covariates Related to Amenity Estimates

We include in our analysis in Section 3.2 and also as covariates in Z_{gj} in equation (25) of Section 6.3 the following eight variables that we construct using the RAIS data, in addition to two sets of FEs.

Part-Time Work Incidence: The share of workers with contractual work hours below 40.

Working Hours Flexibility: The share of workers who change contractual work hours between any two consecutive years.

Parental Leave Generosity: An indicator for whether workers at an employer have parental leave duration above the national median.

[58] Eligibility 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.

Income Fluctuations: The share of workers who change mean earnings between any two consecutive years.

Workplace Hazards: The share of workers who report absence from work due to work-related illness, multiplied by 100.

Incidence of Unjust Firings: The share of workers who report ending their job due to an employer-induced firing for no officially recognized cause.

Incidence of Workplace Deaths: The share of workers who report ending their job due to death at the workplace.

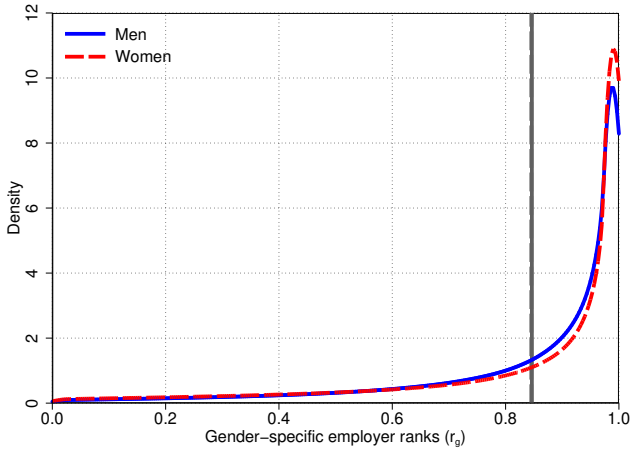
Employer Size: Total employer size measured as the log number of full-time equivalent employees during a year.

Municipality FEs: Dummies for 4,733 municipalities represented in our sample.

Sector FEs: Dummies for 661 five-digit sectors represented in our sample.

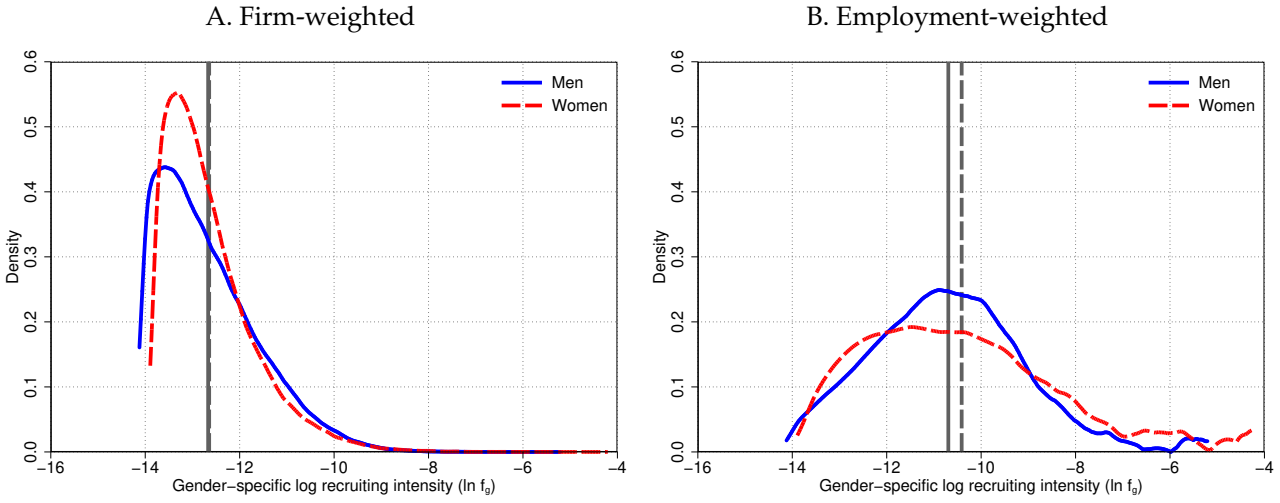
E.4 Detailed Results from the Estimation of Labor Market Objects

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.

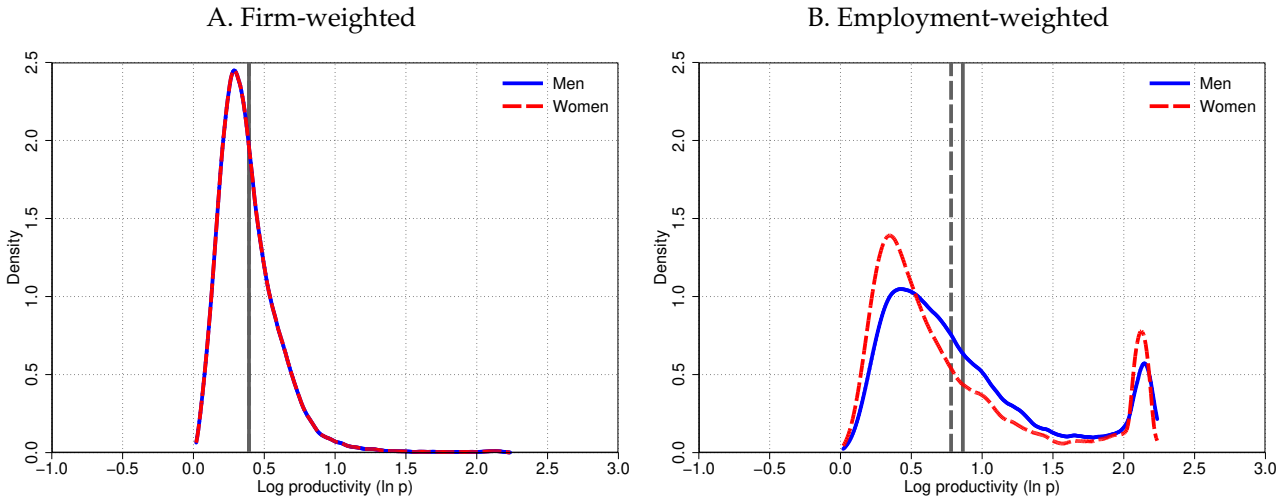
Figure E.2. Density estimates of recruiting intensities, by gender



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.

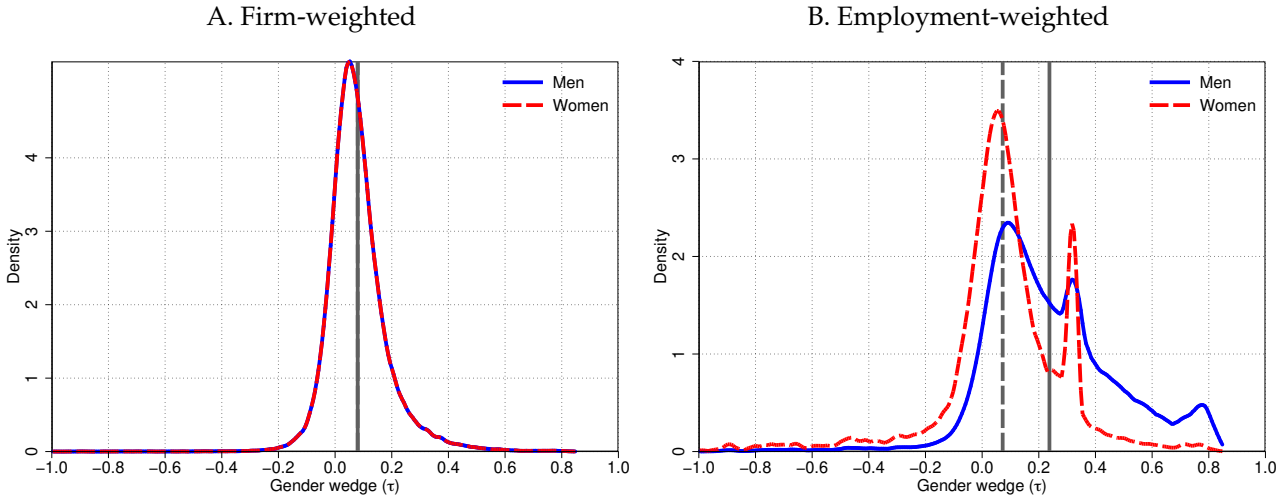
E.5 Detailed Results from the Estimation of on Gender-Specific Firm Types

Figure E.3. Employment-weighted densities of log estimated productivity, by gender



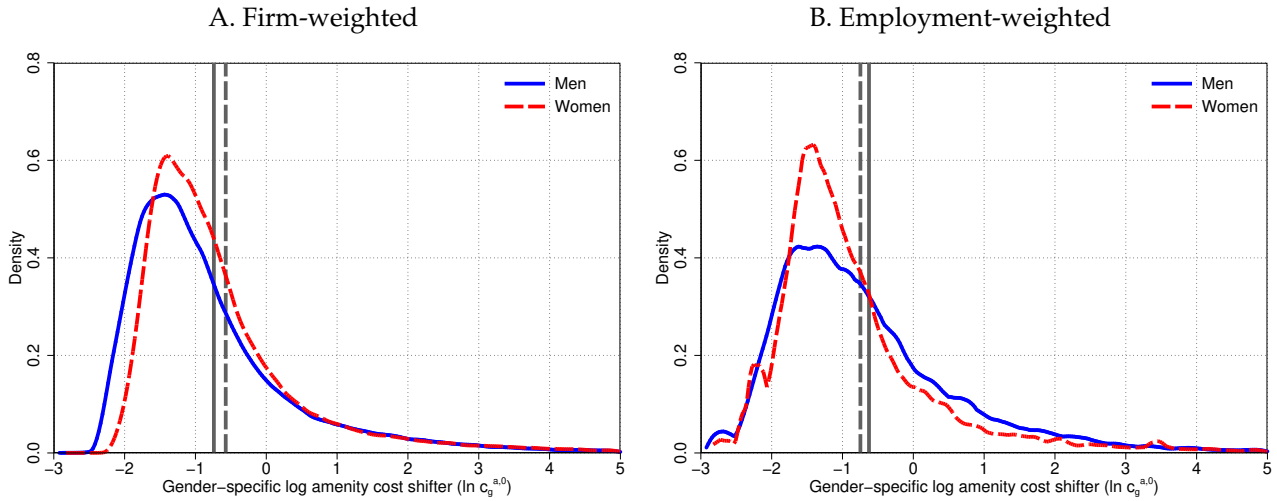
Note: This figure shows employment-weighted densities of log estimated productivity (p) separately for men (blue solid line) and women (red dashed line). Panel A is firm-weighted, while Panel B uses gender-specific employment distributions. Grey vertical patterned lines represent mean values for workers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

Figure E.4. Employment-weighted densities of estimated gender wedges, by gender



Note: This figure shows employment-weighted 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.

Figure E.5. Employment-weighted densities of log amenity cost shifters, by gender



Note: This figure shows employment-weighted 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

	w_M	a_M	x_M	p	l_M	r_M
w_M	1.000					
a_M	−0.914	1.000				
x_M	0.246	0.168	1.000			
p	0.342	0.064	0.985	1.000		
l_M	0.097	0.133	0.552	0.504	1.000	
r_M	0.225	−0.025	0.486	0.456	0.160	1.000

B. Women

	w_F	a_F	x_F	$(1-\tau)p$	l_F	r_F
w_F	1.000					
a_F	−0.937	1.000				
x_F	0.020	0.331	1.000			
$(1-\tau)p$	0.162	0.187	0.970	1.000		
l_F	−0.085	0.282	0.578	0.476	1.000	
r_F	0.009	0.134	0.408	0.424	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	0.884	0.806	0.776	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.

F Gender-Specific Compensation Structures Across Employers Appendix

F.1 Alternative Kitagawa-Oaxaca-Blinder Decompositions

Recall that in Table B.1 of Appendix B.2, we presented alternative Kitagawa-Oaxaca-Blinder decompositions of the gender log pay gap. Here, we present analogous decomposisions 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

		Between-employer gap		Within-employer gap	
	Gender log amenities gap	Level	Share (%)	Level	Share (%)
Decomposition 1	−0.026	−0.092	348.6	0.065	−248.6
Decomposition 2	−0.026	−0.118	447.2	0.091	−347.2

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

		Between-employer gap		Within-employer gap	
	Gender log utility gap	Level	Share (%)	Level	Share (%)
Decomposition 1	0.046	0.002	4.6	0.044	95.4
Decomposition 2	0.046	−0.013	−28.5	0.059	128.5

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.

G Equilibrium Counterfactuals Appendix

G.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:

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

We now apply recursively the discrete version of the first order conditions of the firm, as in equations (D.80) and (D.82):

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

$$\Delta x_g(\tilde{p}_{gr}) = \frac{2\lambda_g^E(\tilde{p}_{gr-1} - x_g(\tilde{p}_{gr-1}))}{\delta_g + \lambda_g^G + \lambda_g^E(1 - F_g(x_g(\tilde{p}_{gr-1})))} \left[\frac{1}{c_g^{v,0}} \frac{T_g(\tilde{p}_{gr-1} - x_g(\tilde{p}_{gr-1}))}{\left(\delta_g + \lambda_g^G + \lambda_g^E(1 - F_g(x_g(\tilde{p}_{gr-1})))\right)^2} \right]^{\frac{1}{\eta^v - 1}} \frac{1}{V_g N}. \tag{G.3}$$

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 $V_g^* = V_g$.[59] 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.

G.2 Alternative 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

[59] Alternatively, we could plug in the cost shifter we have estimated in Step 5 of our identification strategy, but this alternative approach is equivalent in the baseline solution, and allows 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.

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:

$$\max_{w,a_M,a_F,v_M,v_F} \quad \Bigg\{ T_M v_M (p - w - c_M^a(a_M)) \left(\frac{1}{\delta_M + \lambda_M^G + \lambda_M^E (1 - F_M(w + a_M))} \right)^2 \tag{G.4}$$

$$+ T_F v_F ((1-\tau)p - w - c_F^a(a_F)) \left(\frac{1}{\delta_F + \lambda_F^G + \lambda_g^E (1 - F_F(w + a_F))} \right)^2 \tag{G.5}$$

$$- c_M^v(v_M) - c_F^v(v_F) \Bigg\}, \tag{G.6}$$

and the following for the equal hiring policy:

$$\max_{x_M,x_F,a_M,a_F,v} \quad \Bigg\{ T_M v (\tilde{p}_M - x_M) \left(\frac{1}{\delta_M + \lambda_M^G + \lambda_M^E (1 - F_M(x_M))} \right)^2 \tag{G.7}$$

$$+ T_F v (\tilde{p}_F - x_F) \left(\frac{1}{\delta_F + \lambda_F^G + \lambda_g^E (1 - F_F(x_F))} \right)^2 \tag{G.8}$$

$$- c_M^v(v) - c_F^v(v) \Bigg\}, \tag{G.9}$$

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:

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

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

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 in the 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.