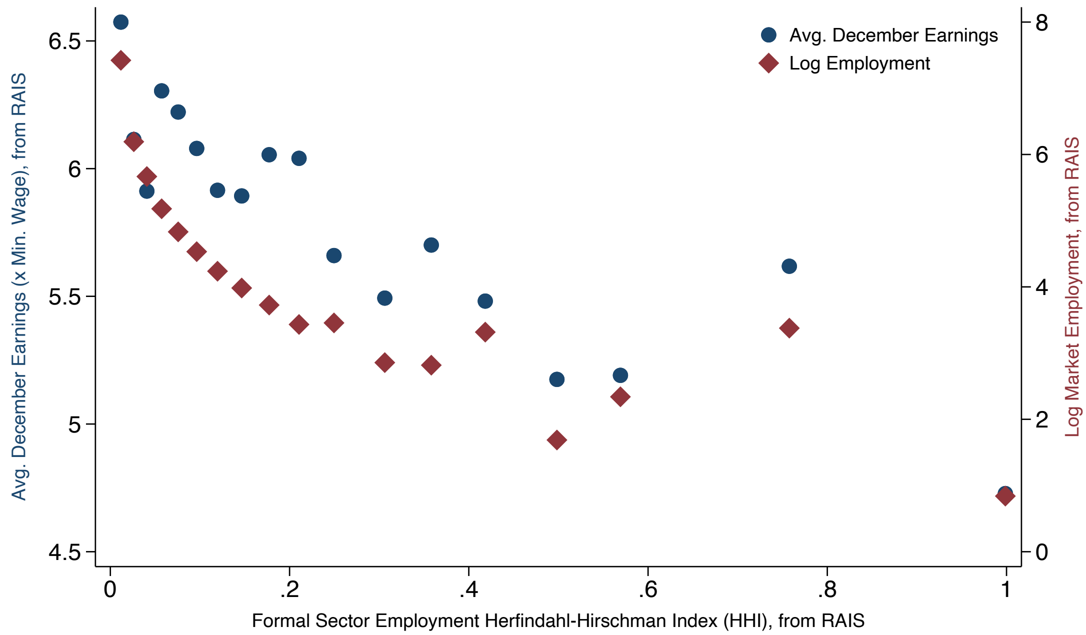
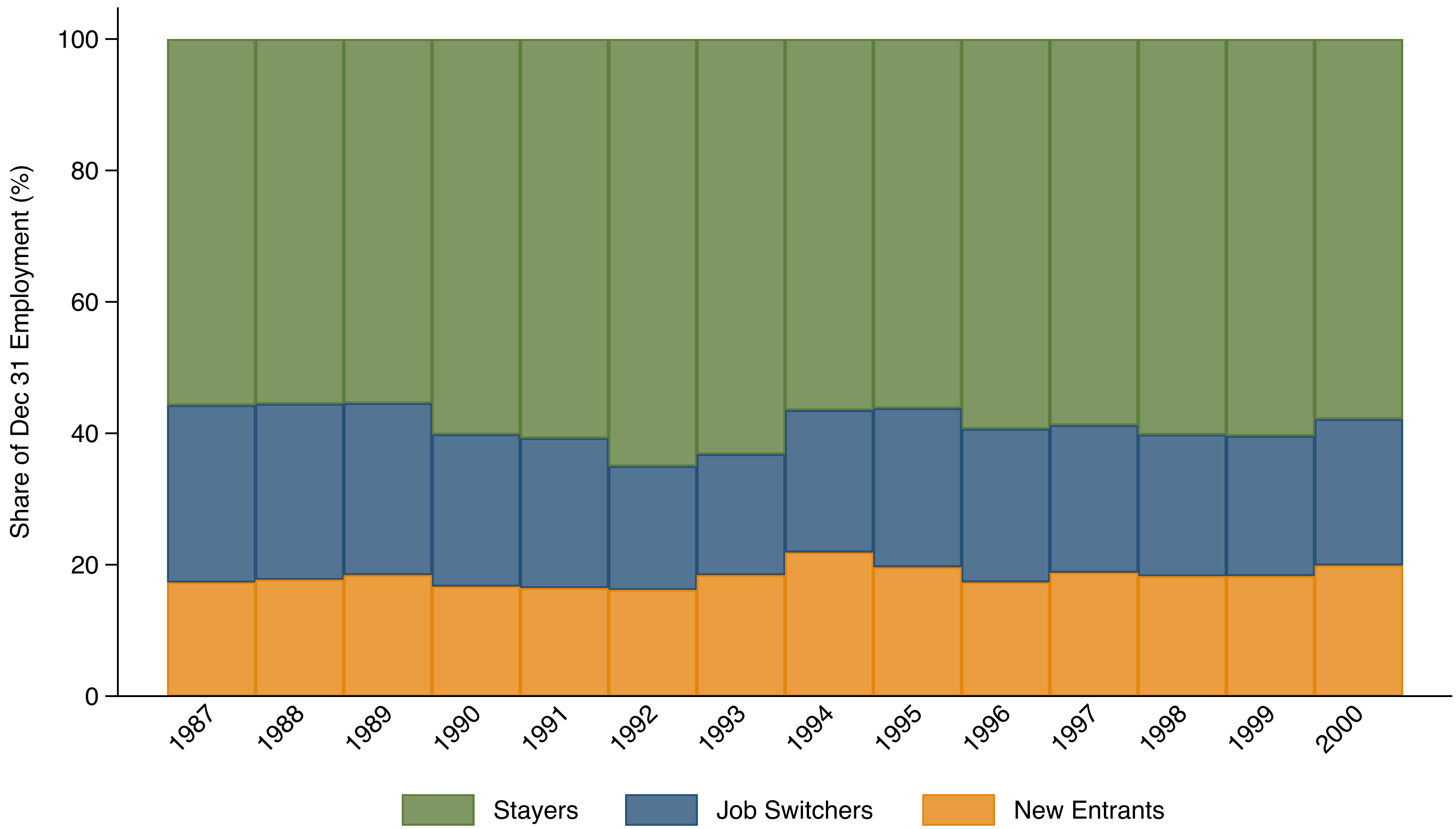
Online Appendix for
"Trade, Labor Market Concentration, and Wages,"
by Mayara Felix

Figure A.1: 1991 wages, employment, and formal sector concentration



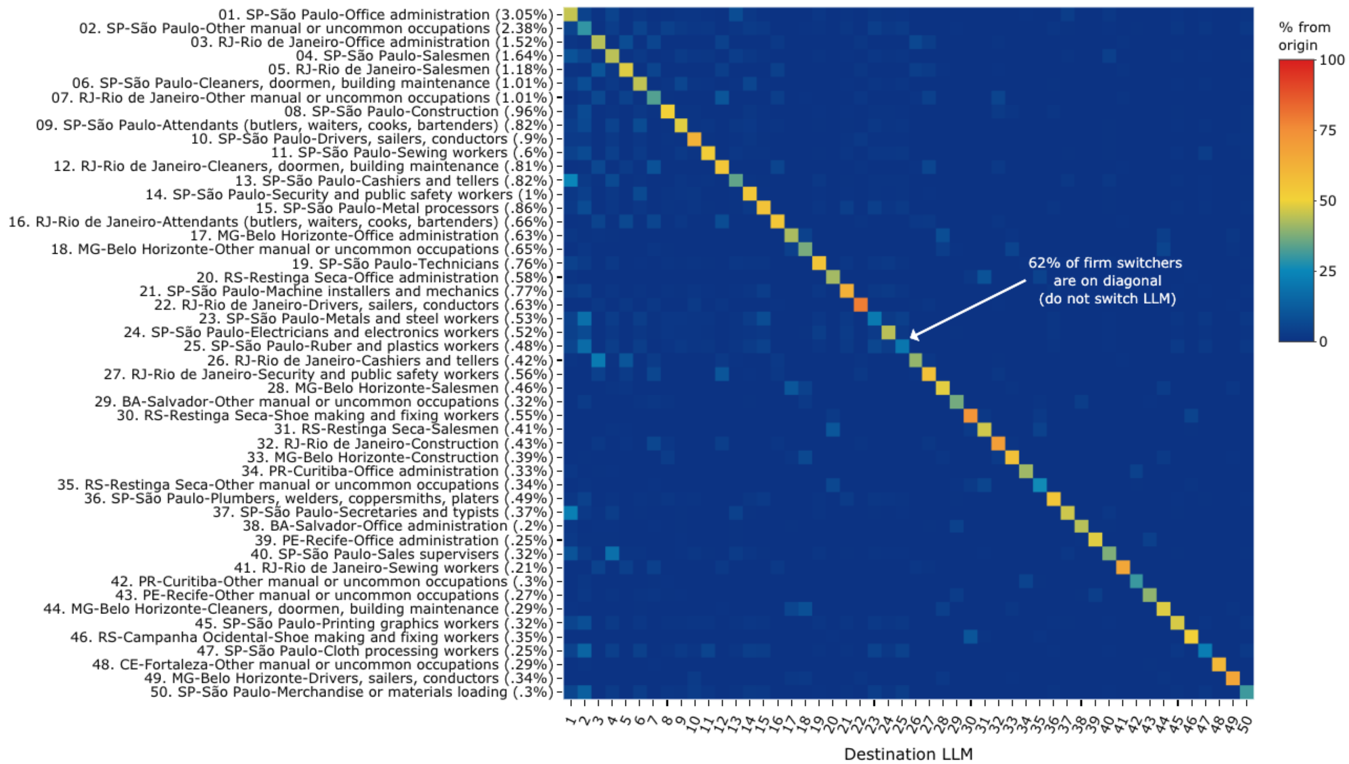
Notes: This figure plots wages and employment as a function of formal sector HHI in microregion x occupation pairs using RAIS data for 1991.

Figure A.2: Formal sector job stayers versus job switchers as of Dec 31 of each year



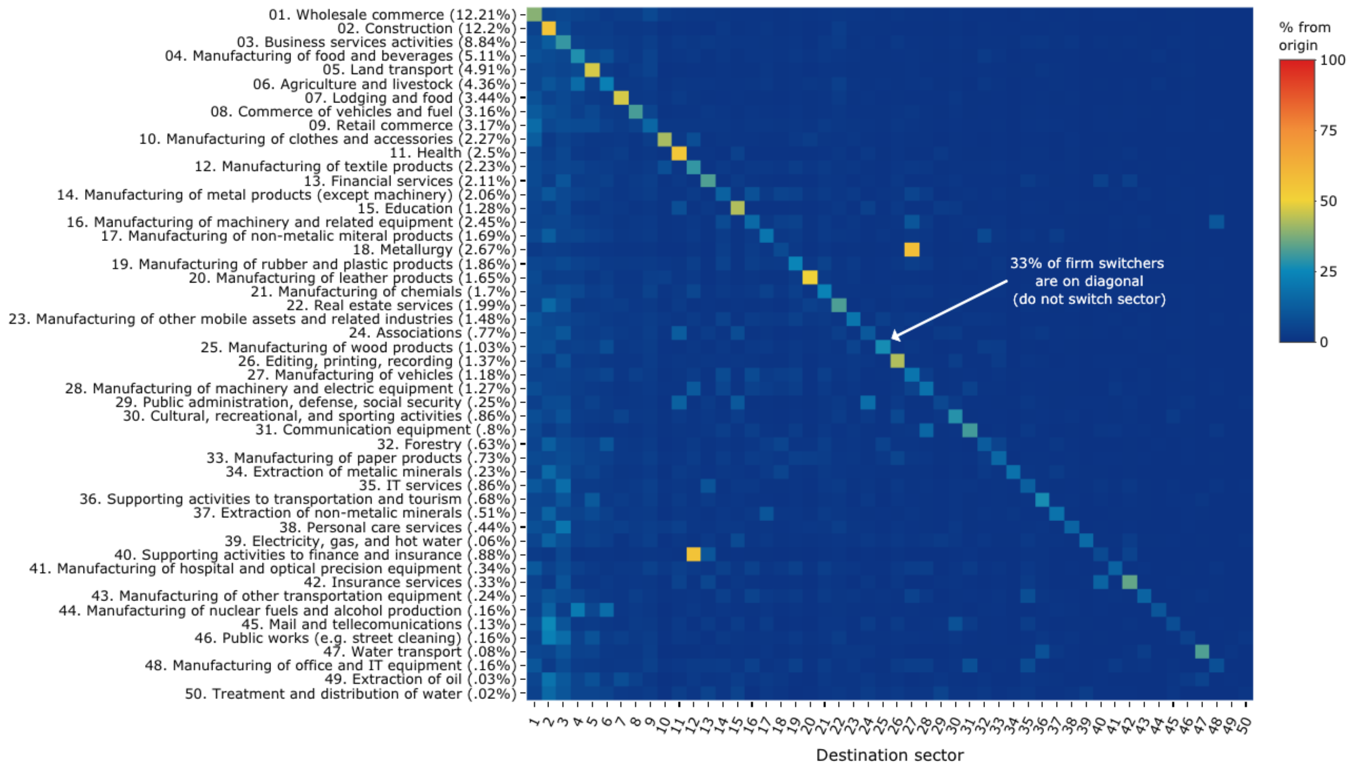
Notes: This figure plots formal private sector employment composition in Brazil from 1986 to 2000 using matched employer-employee data from RAIS. Each bar reports December 31 employment (in millions) decomposed into job stayers (workers employed at the same firm as of December 31 of the previous year) and job switchers (workers employed at a different firm or entering the formal sector). "Hired from Outside" counts workers appearing in the formal sector on December 31 who were not employed in any formal firm on December 31 of the previous year. "Hired from Other" counts workers who switched from one formal firm to another within the year. Flows exclude transitions to/from public administration, retirements, and deaths.

Figure A.3: 1990-1991 local labor market transitions conditional on switching firms (Top 50)



Note: This figure plots worker local labor market to local labor market transitions, among workers who switched employers between 1990 and 1991, for the top 50 local labor markets by number of workers at origin. A local labor market is a microregion × occupational group pair. Each row lists the origin microregion (with percent of total workers indicated in parentheses), while each column lists the destination microregion.

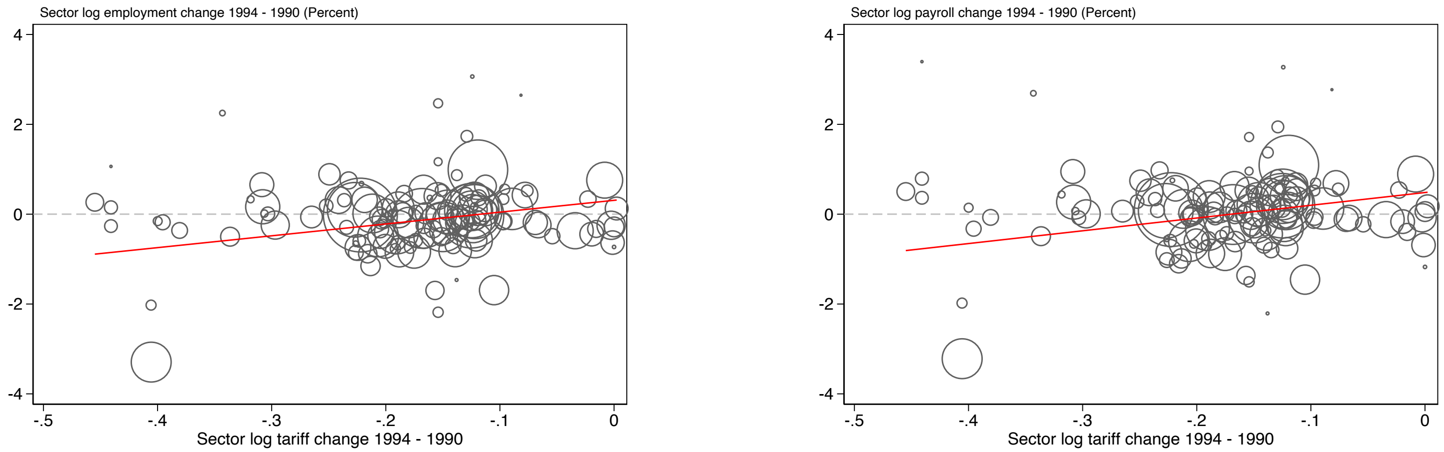
Figure A.4: 1990-1991 sector transitions conditional on switching firms (Top 50)



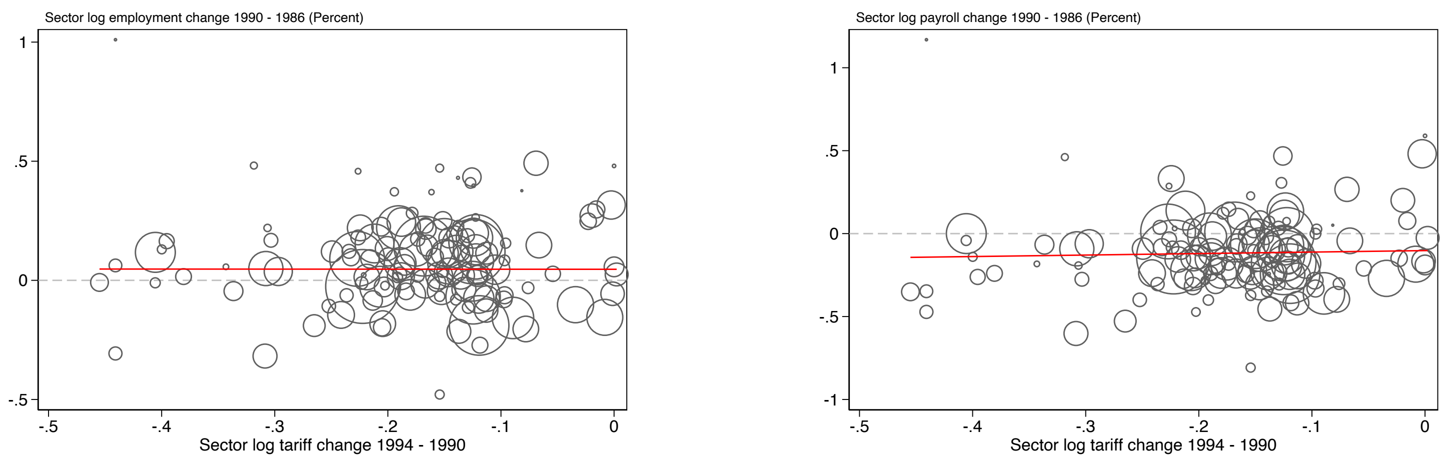
Note: This figure plots worker sector group to sector group transitions, among workers who switched employers between 1990 and 1991, for the top 50 sector (2-digit CNAE95) groups by number of workers at origin. Each row lists the origin sector group (with percent of total workers indicated in parentheses), while each column lists the destination sector group.

Figure A.5: Changes in sector-level outcomes Before vs. After liberalization

(a) After liberalization (1990–1994)

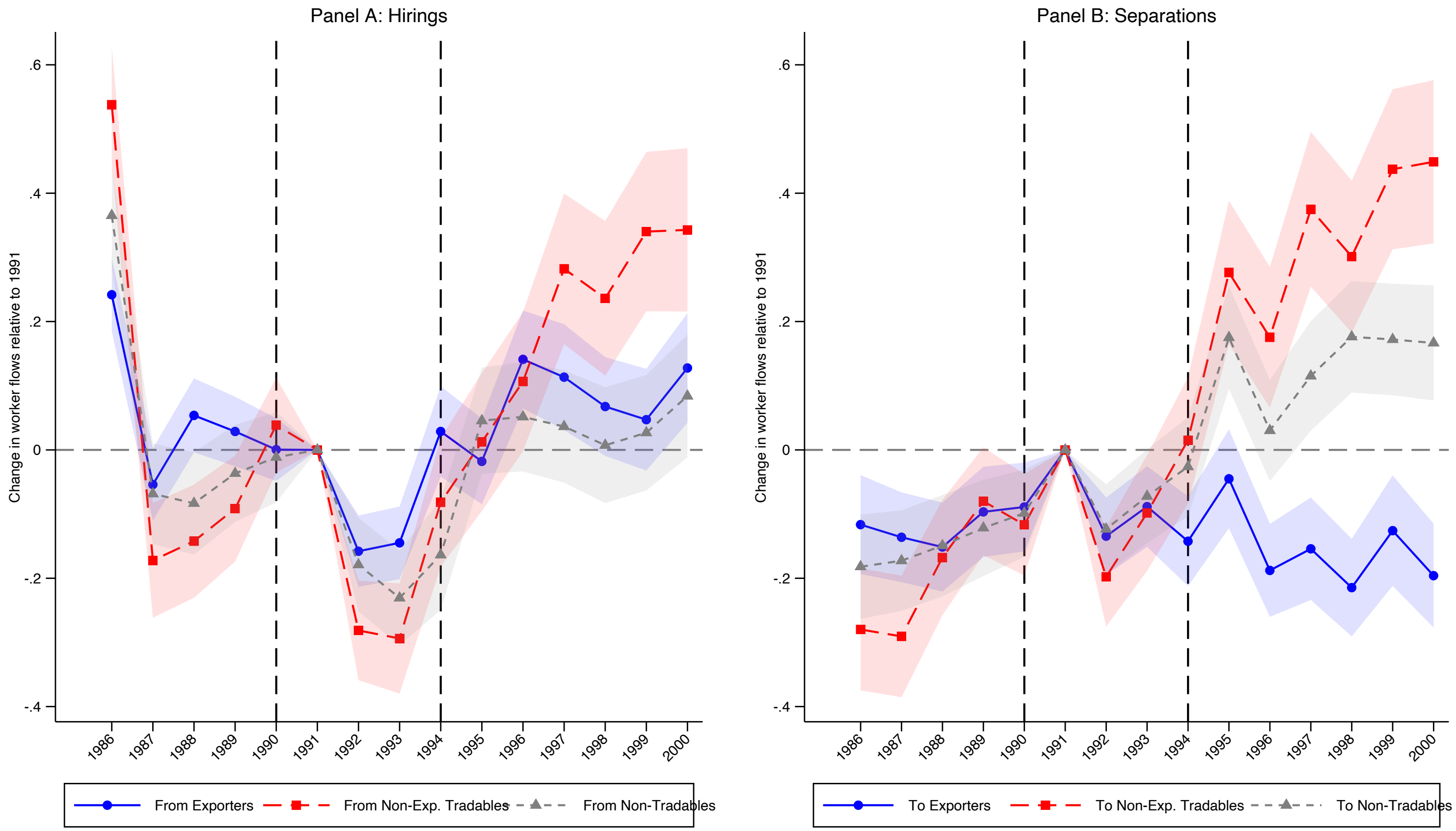


(b) Before liberalization (1986–1990)



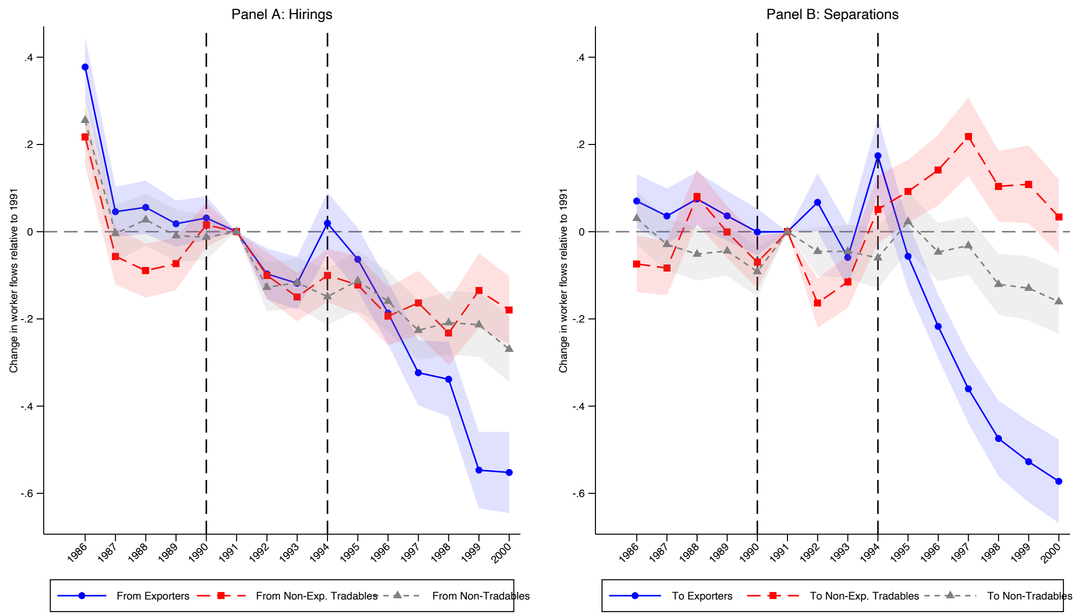
Note: This figure plots changes in sector-level local total employment and sector-level log total payroll for the 4-year periods before versus after liberalization, against the 1990–1994 import tariff reduction on each sector's output. Sector totals are based on data for the entire country and are aggregated into 148 tradable sector codes based on RAIS's 4-digit "ibgesubatividade" sector variable, consistently reported throughout the period.

Figure A.6: Effect of import competition exposure on within-market flows into (Panel A) and out of (Panel B) non-exporting tradable sector firms



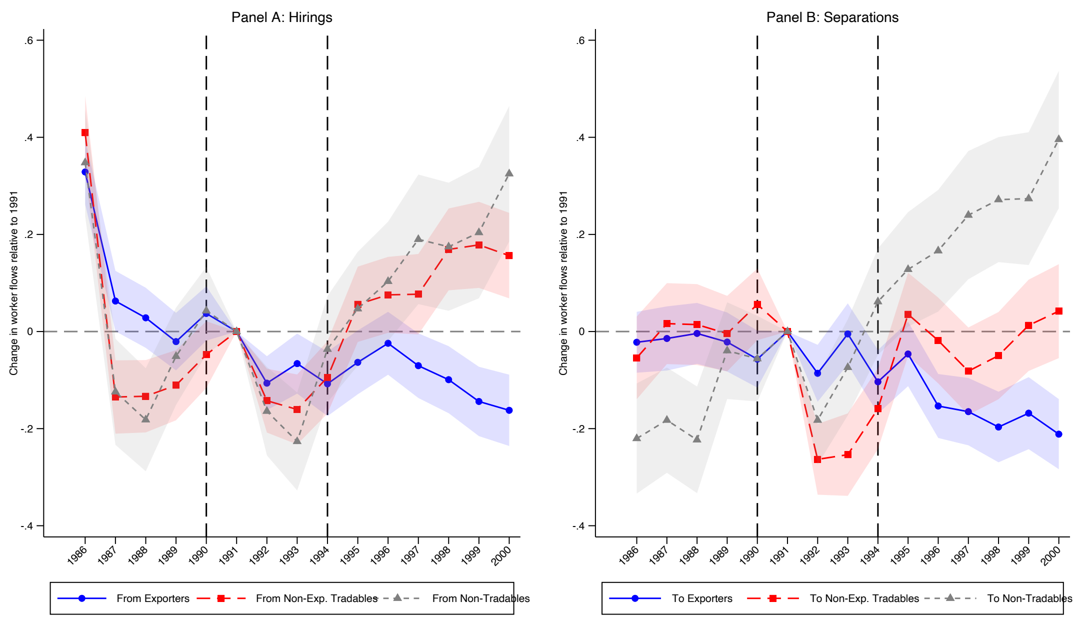
Note: This figure plots year-by-year coefficients from stacked differences-in-differences regressions of within-market cross-firm worker flows on import competition exposure (ICE), with local labor market (mmc × cbo942d) fixed effects, clustering at the local labor market level. Panel A shows hirings into non-exporting tradable firms from exporters, other non-exporting tradable firms, and non-tradable firms within the same local labor market. Panel B shows separations from non-exporting tradable firms to each of these firm types. Outcomes are IHS-transformed long differences relative to 1991. Shaded areas report 95% confidence intervals. Appendix Figures A.7–A.11 present corresponding results for flows into and out of other firm types, including flows into and out of the formal sector.

Figure A.7: Effect of import competition on within-market flows into and out of Exporters



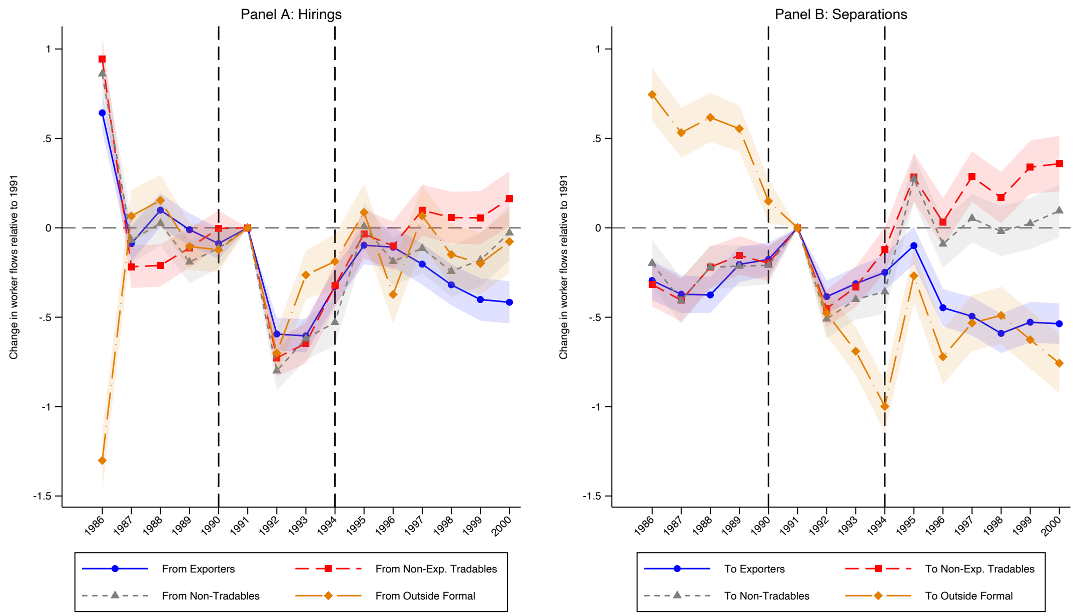
Note: See notes to Figure A.6.

Figure A.8: Effect of import competition on within-market flows into and out of Non-tradables



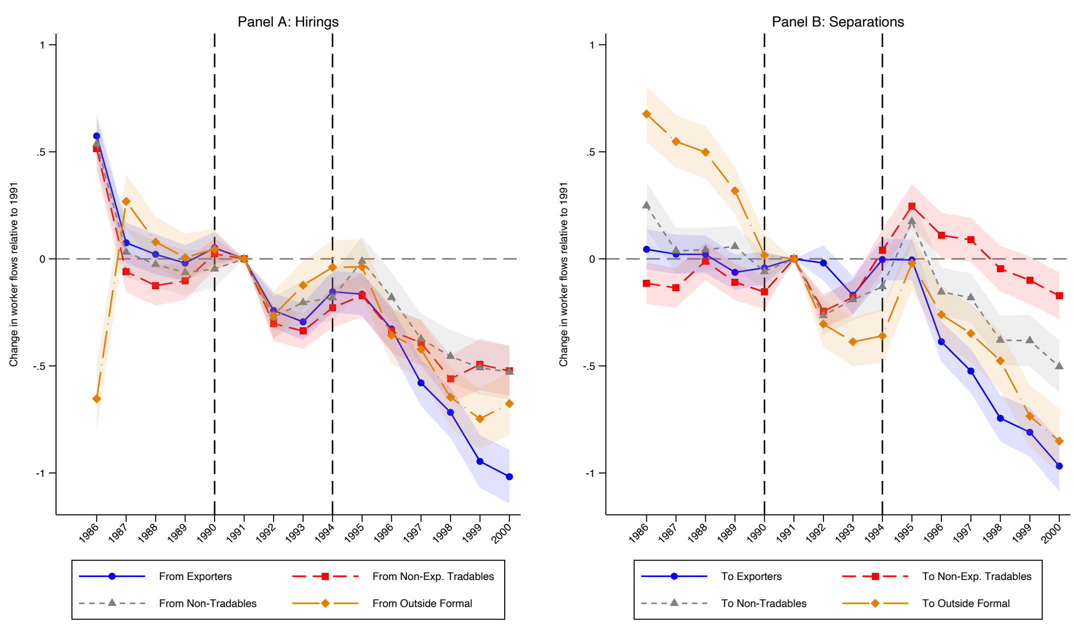
Note: See notes to Figure A.6.

Figure A.9: Effect of import competition on all flows into and out of Non-Exporting Tradable sector firms



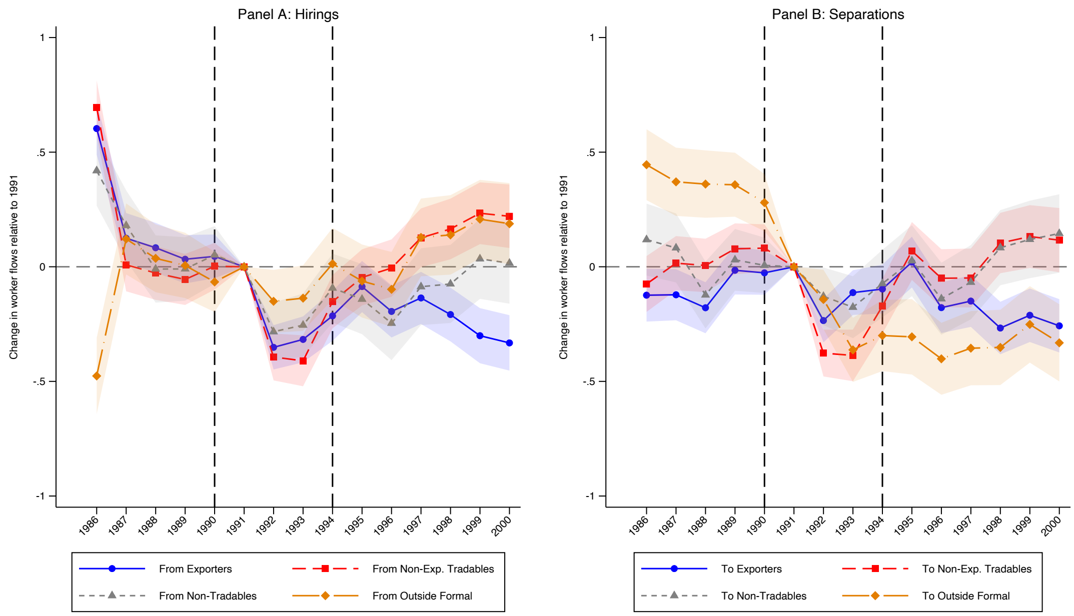
Note: See notes to Figure A.6.

Figure A.10: Effect of import competition on all flows into and out of Exporters



Note: See notes to Figure A.6.

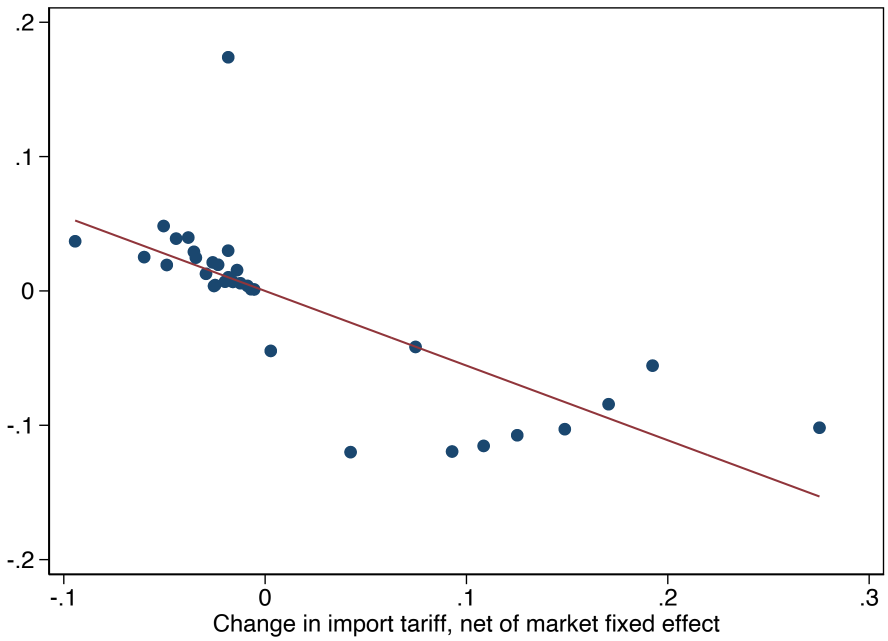
Figure A.11: Effect of import competition on all flows into and out of Non-tradables



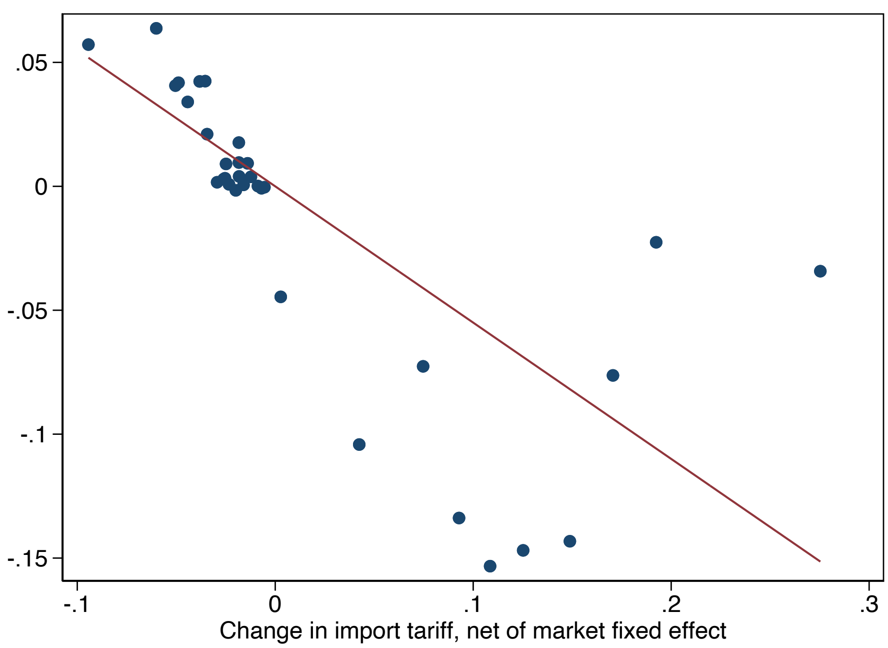
Note: See notes to Figure A.6.

Figure A.12: Identifying variation for within-market cross-firm inverse elasticity of substitution

(a) Panel A: First stage



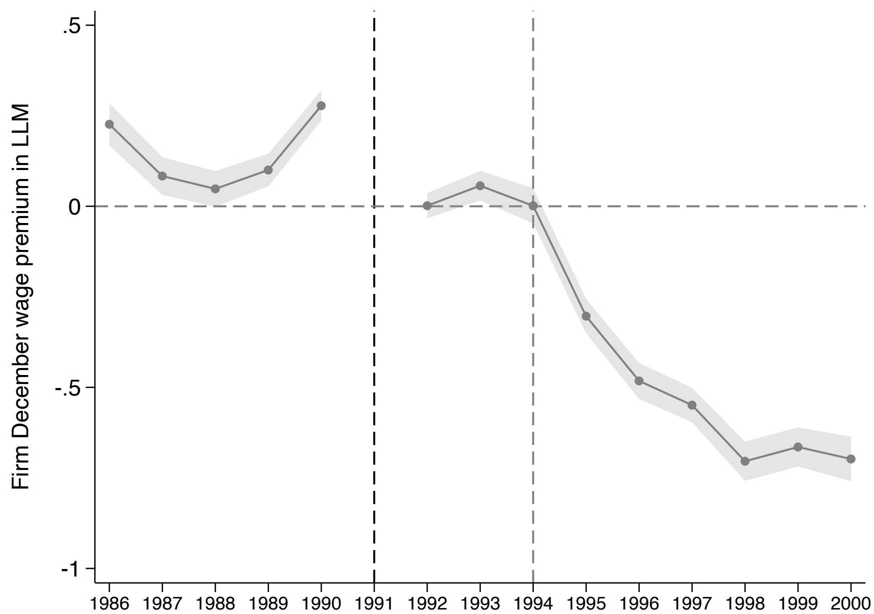
(b) Panel B: Reduced form



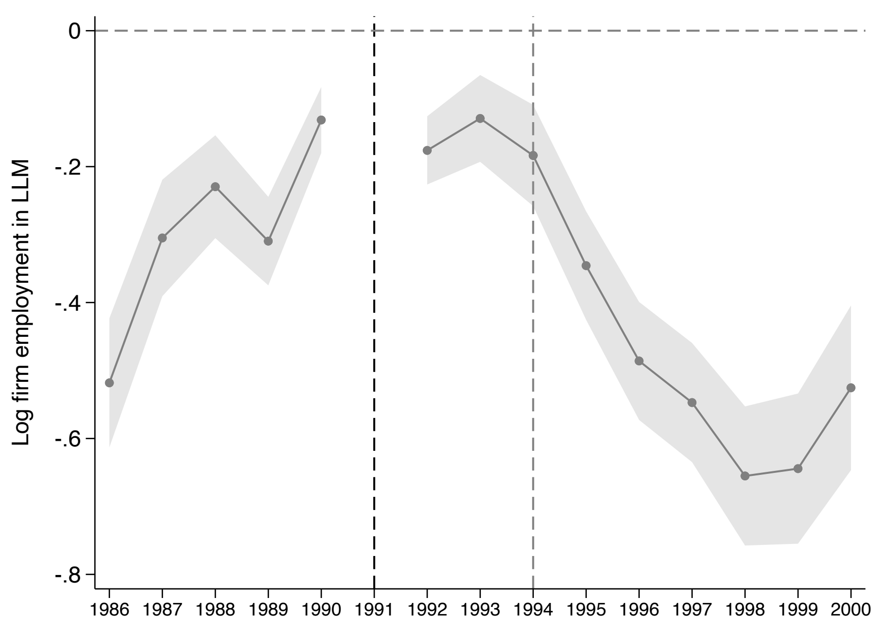
Notes: This figure shows the identifying variation underlying the IV estimate of the within-market cross-firm inverse elasticity of substitution $\frac{1}{\eta}$. Panel A plots the first stage (firm-level employment against the tariff instrument). Panel B plots the reduced form (firm-level wage premia against the tariff instrument). Both panels show binned scatter plots after partialling out local labor market fixed effects.

Figure A.13: Year by year effects of tariff reductions on firm-market-level wage premia and employment

(a) Wage premia

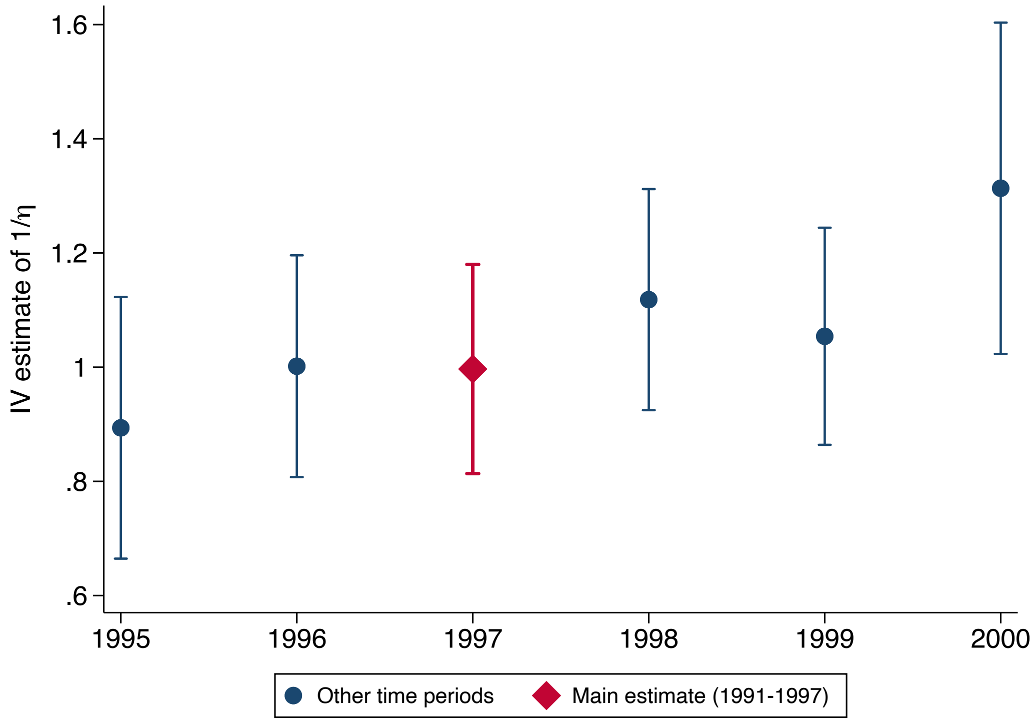


(b) Employment



Note: This figure plots coefficients of regressions of firm-level changes in log employment (from each year to the base year of 1991) on $-\ln\left(\frac{1+\tau_{1994}}{1+\tau_{1990}}\right)$, which is the firm-level change in import competition exposure, separately estimated for each year. Dotted lines indicate the beginning and end of the tariff reductions reform. So that all differences reflect a change from a future year to a past year, for the pre-liberalization years the outcome is the 1991 log employment minus each respective year's log employment, whereas for the post-reform years the outcome is each respective year's log employment minus the 1991 log employment. All regressions are weighted by 1991 firm employment. Standard errors are clustered at the sector level.

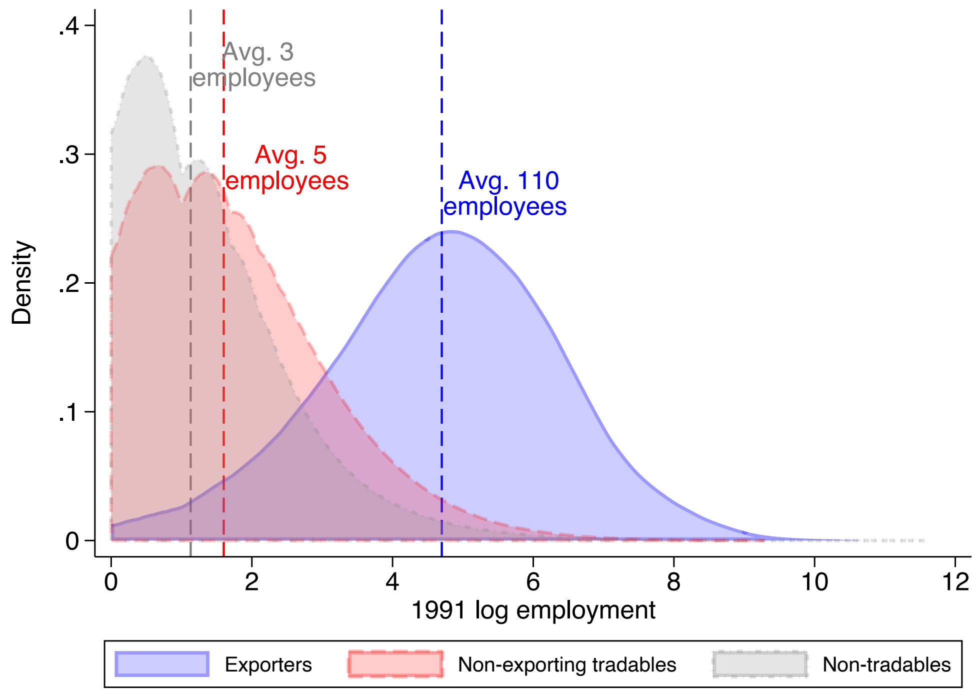
Figure A.14: Year by year estimates of within-market cross-firm inverse elasticity of substitution



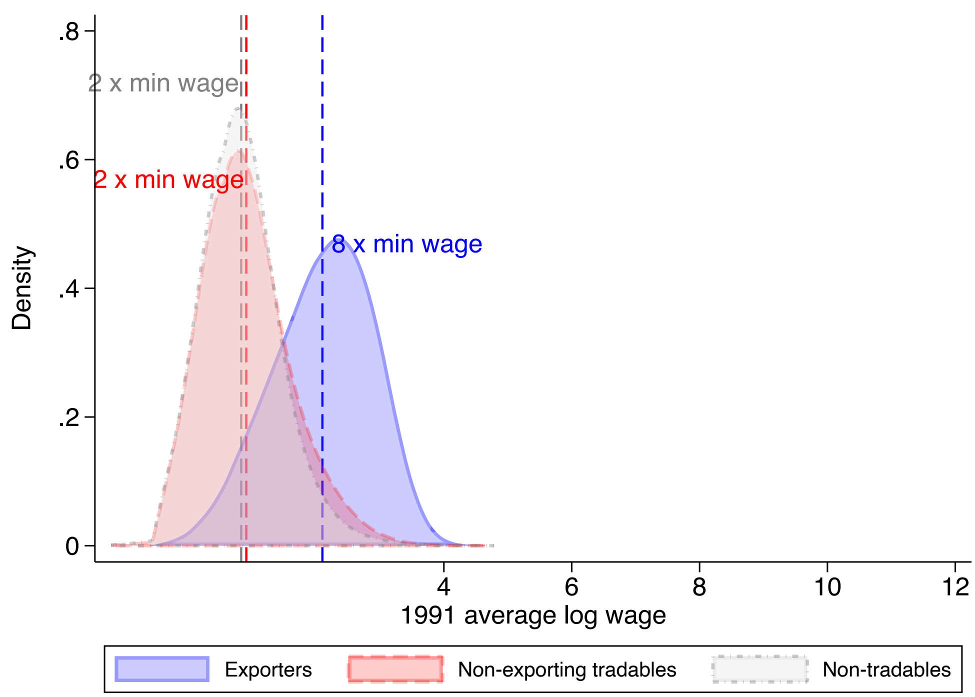
Notes: This figure presents estimates of the within-market cross-firm inverse elasticity of substitution $\frac{1}{\eta}$ estimated separately for each post-liberalization year, using the same specification as in Table 2. The dashed horizontal line indicates the pooled baseline estimate.

Figure A.15: Pre-liberalization distribution of firm size and wages

(a) Distributions of firm log employment



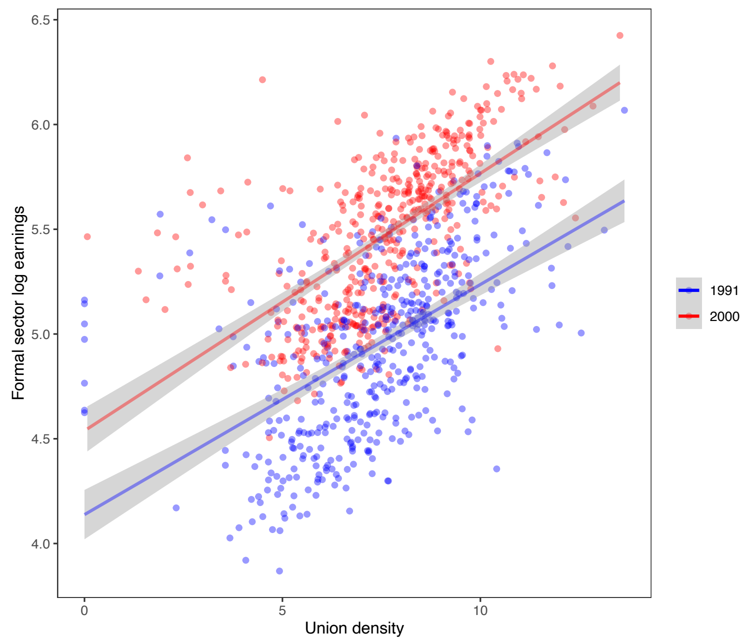
(b) Distributions of firm average log wage



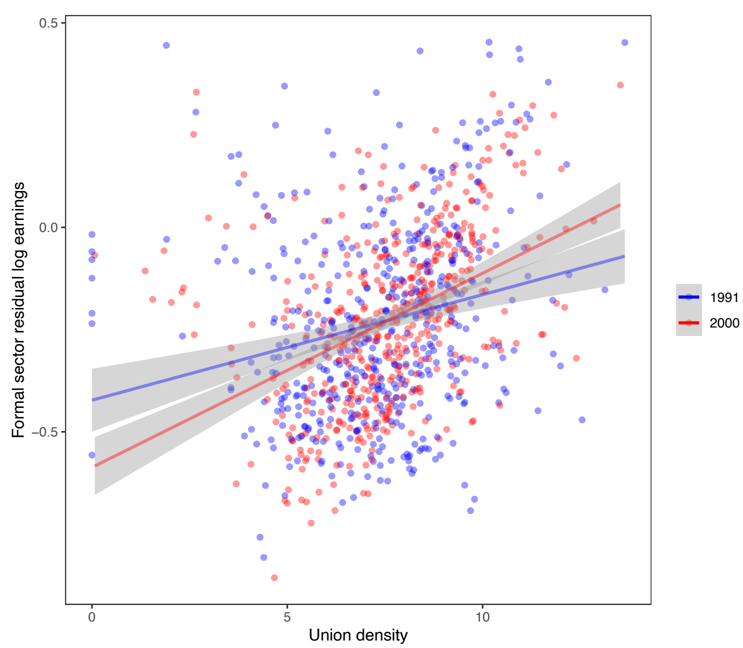
Note: This figure plots pre-liberalization distributions of firm log employment and log December monthly wages for exporters, non-exporters, and non-tradables. Wages are reported as multiples of the national minimum wage.

Figure A.16: Formal sector wage compression and union density

(a) Panel A: Formal sector log earnings and union density



(b) Panel B: Formal sector residual log earnings and union density



Note: This figure examines the relationship between union density and formal-sector wages. Panel A plots average formal-sector log earnings against union density, while Panel B plots residual formal-sector log earnings after netting out observable worker characteristics. Each panel shows two cross-sections, corresponding to 1991 and 2000, with separate linear fits for each year. The horizontal axis reports union density transformed as $\log(1+\text{union density})$ to improve visualization in the presence of skewness. Union density data comes from https://academic.oup.com/jeea/article/23/1/236/7628306 and was collected for the years 1992 and 2001 to match with wage outcomes in 1991 and 2000.

Table A.1: Brazilian workforce during 1990s trade liberalization

	Population Census Year	
	1991	2000
N	51,318,478 (42.6%)	69,279,182 (57.4%)
Formal wage work	0.403 (0.491)	0.313 (0.464)
Informal wage work	0.324 (0.468)	0.344 (0.475)
Self-employed	0.239 (0.427)	0.205 (0.404)
Unemployed	0.044 (0.206)	0.138 (0.345)
Hours worked	43.534 (12.006)	43.950 (14.655)
Earnings (R2000)	331.153 (412.063)	553.811 (704.593)
Earnings (R2000) $- Formal$	383.860 (413.424)	596.340 (653.361)
Earnings (R2000) $- Informal$	279.794 (418.736)	497.524 (725.634)
Earnings (R2000) $- Self-employed$	309.456 (388.220)	568.712 (744.037)
Minimum wage (R2000)	126.514 (0.000)	151.000 (0.000)
Female	0.327 (0.469)	0.404 (0.491)
Age	34.262 (11.490)	34.717 (11.500)
Years of education	5.900 (4.522)	7.033 (4.392)

Note: This table shows descriptive statistics from the Brazilian 1991 and 2000 population censuses, calculated using the 1991 and 2000 census sample extract files from the supplementary materials for Dix-Carneiro and Kovak (2017). Earnings and hours are for main job. All statistics are calculated using individual-level sample extract weights. Earnings statistics exclude top 1% and bottom 1% of earnings distribution. The sample includes all workers employed in the private sector (that is, excludes public administration). Formal wage work follows the standard definition for Brazilian labor markets, which is to have a signed *Carteira de Trabalho* (variables v350 and v0447 in 1991 and 2000 censuses, respectively). Real monthly earnings are based on the IPCA deflator and are expressed in 2000 reais. Minimum wage reports the federal monthly minimum wage for July 1991 and for 2000. Brazil's minimum wage is regulated as minimum monthly earnings for a 44-hour work week.

Table A.2: 1991 descriptive statistics of formal local labor markets (microregion x occupation group pairs)

	Mean	Std. Dev.	p10	p25	Median	p75	p90	N
Number of formal workers	698	5,266	6	16	61	262	1,006	19,514
Number of formal tradable sector workers	293	2,176	0	3	20	101	416	19,514
Number of exporter workers	255	2,076	0	1	10	69	333	19,514
Number of formal non-tradable sector workers	405	3,484	1	6	25	118	496	19,514
Number of formal firms	116	795	3	6	16	55	183	19,514
Number of exporter firms	18	117	0	1	2	8	26	19,514
Payroll HHI (based on wage premia)	0.283	0.248	0.042	0.093	0.207	0.399	0.643	19,513
Payroll HHI (based on wage level)	0.288	0.249	0.044	0.097	0.212	0.406	0.652	19,514
Employment HHI	0.234	0.234	0.027	0.064	0.156	0.333	0.556	19,514
Average December wage (× min. wage)	5.860	5.984	1.672	2.346	3.848	6.922	12.354	19,514
Average December wage premium (× min. wage)	2.484	1.526	1.114	1.474	2.067	3.033	4.395	19,286
Import competition exposure (ΔICE)	0.125	0.091	0.000	0.050	0.130	0.178	0.231	19,514

Note: This table presents descriptive statistics across 19,514 Brazilian local labor markets defined as microregion × occupation group pairs, in the baseline year (1991). Means are unweighted. HHI denotes the Herfindahl-Hirschman Index of employment or payroll concentration. ΔICE is the change in import competition exposure.

Table A.3: Average payroll Herfindahl across local labor markets

	1986 (1)	1991 (2)	1994 (3)	1997 (4)	2000 (5)
Panel A: Payroll Herfindahl (based on wage premia)					
Unweighted average	0.289	0.283	0.274	0.228	0.209
Weighted average (by market employment shares)	0.071	0.078	0.074	0.061	0.058
Weighted average (by market payroll shares)	0.069	0.080	0.077	0.064	0.061
Panel B: Payroll Herfindahl (based on wage levels)					
Unweighted average	0.293	0.288	0.280	0.234	0.214
Weighted average (by market employment shares)	0.073	0.082	0.078	0.065	0.061
Weighted average (by market payroll shares)	0.073	0.084	0.080	0.068	0.065
Panel C: Employment Herfindahl					
Unweighted average	0.246	0.234	0.221	0.184	0.173
Weighted average (by market employment shares)	0.054	0.056	0.053	0.044	0.046
Weighted average (by market payroll shares)	0.049	0.052	0.050	0.041	0.044

Note: This table presents country-level weighted average payroll concentration measures for alternative weights.

Table A.4: Formal employment and worker flows by year

	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Panel A: December 31 Employment (millions)															
Total	10.64	20.80	21.04	21.74	20.79	19.65	18.25	18.84	20.37	23.54	23.29	24.09	23.93	23.64	25.25
Exporters	1.69	3.97	3.88	4.17	3.88	3.59	3.27	3.35	3.47	3.53	3.06	2.79	2.46	2.21	2.26
Non-Exp. Tradables	2.18	3.67	3.49	3.64	3.52	3.25	2.91	3.14	3.96	4.56	4.56	4.85	4.84	4.94	5.40
Non-Tradables	6.77	13.16	13.66	13.94	13.39	12.81	12.07	12.35	12.94	15.45	15.67	16.45	16.63	16.49	17.59
Panel B: Flows In/Out of Formal Sector (millions)															
Hired from Outside	7.84	3.60	3.73	4.01	3.47	3.23	2.95	3.47	4.46	4.63	4.04	4.53	4.37	4.33	5.03
Into Exporters	1.20	0.50	0.45	0.55	0.40	0.39	0.32	0.40	0.44	0.41	0.29	0.27	0.21	0.22	0.25
Into Non-Exp. Tradables	1.68	0.70	0.68	0.80	0.69	0.61	0.55	0.70	1.03	1.04	0.90	1.03	0.98	1.01	1.18
Into Non-Tradables	4.97	2.40	2.60	2.66	2.38	2.23	2.08	2.37	2.98	3.17	2.85	3.23	3.18	3.10	3.60
Separated to Outside	3.00	3.17	3.17	3.43	4.11	3.92	3.40	2.86	3.16	4.19	3.79	3.89	4.13	3.66	3.71
From Exporters	0.40	0.48	0.41	0.48	0.64	0.57	0.46	0.38	0.41	0.54	0.40	0.38	0.34	0.23	0.21
From Non-Exp. Tradables	0.59	0.66	0.59	0.64	0.78	0.74	0.62	0.53	0.67	0.92	0.82	0.84	0.90	0.81	0.85
From Non-Tradables	2.01	2.04	2.17	2.31	2.69	2.61	2.33	1.94	2.08	2.72	2.57	2.67	2.89	2.62	2.65
Panel C: Flows Between Firms (millions)															
From Exporters	0.88	0.96	0.79	0.82	0.83	0.63	0.51	0.52	0.65	0.74	0.53	0.49	0.41	0.36	0.35
From Non-Exp. Tradables	1.08	1.11	0.97	0.94	0.90	0.71	0.58	0.65	0.94	1.18	1.05	1.08	1.02	1.05	1.19
From Non-Tradables	3.83	3.73	3.75	3.55	3.27	2.80	2.36	2.59	3.19	3.92	3.57	3.58	3.46	3.55	3.92

Notes: This table summarizes formal private sector employment and worker flows in Brazil from 1986 to 2000 using matched employer-employee data from RAIS. Panel A reports December 31 employment (in millions) by firm trading status: exporters, non-exporting tradable firms, and non-tradable firms. Panel B reports annual flows into and out of the formal sector, separately by destination or origin firm type. “Hired from Outside” counts workers appearing in the formal sector on December 31 who were not employed in any formal firm on December 31 of the previous year. “Separated to Outside” counts workers employed in a formal firm on January 1 who are not found in any formal firm on December 31 of the same year. Panel C reports within-year job-to-job transitions between formal sector firms, by origin firm type. All values are in millions of workers.

Table A.5: Effects of import competition exposure on local wages and employment

	Δ Import Competition Exposure	Effect per 10% increase in ICE
	(1)	(2)
Panel A: Log number of firms and log formal employment		
Δ Log number of firms	-0.549	-5.489
	(0.045)	(0.447)
Mean (unweighted), 1991	3.651	3.651
Mean (weighted), 1991	7.645	7.645
Δ Log formal employment	-0.440	-4.400
	(0.064)	(0.640)
Mean (unweighted), 1991	4.918	4.918
Mean (weighted), 1991	9.659	9.659
Panel B: Log formal wage premium		
Δ Log formal wage premium	0.029	0.293
	(0.031)	(0.307)
Mean (unweighted), 1991	0.761	0.761
Mean (weighted), 1991	1.072	1.072
Δ De-trended log formal wage premium	-0.206	-2.063
	(0.034)	(0.338)
Mean (unweighted), 1991	1.225	1.225
Mean (weighted), 1991	1.485	1.485
Observations	289,680	289,680
Local labor markets	19,759	19,759

Notes: Standard errors in parentheses, clustered at the local labor market level. Each cell reports the 1997 coefficient from a stacked differences-in-differences regression of long-differenced outcomes on import competition exposure (ICE) interacted with year dummies, with local labor market (mmc × cbo942d) fixed effects.

Table A.6: Effects of 10% increase in import competition on formal sector labor market concentration: Heterogeneity by baseline formal sector share of total employment

	Main specification (1)	Above median formal share (2)	Below median formal share (3)
Δ Payroll Herfindahl (based on wage premium)	0.021	0.017	0.021
	(0.002)	(0.002)	(0.003)
Mean (unweighted), 1991	0.287	0.244	0.377
Mean (weighted), 1991	0.079	0.072	0.215
Δ Payroll Herfindahl	0.021	0.017	0.021
	(0.002)	(0.002)	(0.003)
Median, 1991	0.216	0.172	0.314
Mean (weighted), 1991	0.082	0.076	0.221
Δ Employment Herfindahl	0.025	0.018	0.028
	(0.002)	(0.002)	(0.003)
Median, 1991	0.158	0.122	0.244
Mean (weighted), 1991	0.056	0.051	0.171
Observations	296,340	195,998	96,833
Local labor markets	19,759	13,068	6,457

Notes: Standard errors in parentheses, clustered at the local labor market level. Each cell reports the 1997 coefficient from a stacked differences-in-differences regression of long-differenced outcomes on import competition exposure (ICE) interacted with year dummies, with local labor market (mmc × cbo942d) fixed effects. Coefficients are divided by 10 to report the effect of a 10 percentage point increase in ICE (HHI variables are on a 0–1 scale; the raw regression coefficient is per unit of ICE, so dividing by 10 converts to a 10 percentage point increase on the 0–1 HHI scale, i.e., $(\zeta_{1997} \times 10)/100$). Column (1) uses all markets. Columns (2) and (3) split mmc×cbo942d markets by whether the microregion's formal employment share (formal employment / (informal wage work + self-employed)) from the 1991 census is above or below the median across microregions. Weighted means weight by total formal employment in the mmc×cbo942d pair.

Table A.7: Effect of import competition exposure on concentration: Microregion boundaries

	Main specification (1)	Local labor market is microregion (2)
Panel A: Formal sector labor market concentration		
Δ Payroll Herfindahl (based on wage premium)	0.213	0.102
	(0.017)	(0.046)
Mean (unweighted), 1991	0.287	0.038
Mean (weighted), 1991	0.079	0.006
Δ Payroll Herfindahl	0.213	0.110
	(0.017)	(0.064)
Mean (unweighted), 1991	0.292	0.113
Mean (weighted), 1991	0.082	0.026
Δ Employment Herfindahl	0.247	0.058
	(0.016)	(0.056)
Mean (unweighted), 1991	0.239	0.071
Mean (weighted), 1991	0.056	0.011
Panel B: Log number of firms and log formal employment		
Δ Log number of firms	-0.549	-0.367
	(0.045)	(0.208)
Mean (unweighted), 1991	3.651	6.646
Mean (weighted), 1991	7.645	10.133
Δ Log formal employment	-0.440	-0.338
	(0.064)	(0.335)
Mean (unweighted), 1991	4.918	9.001
Mean (weighted), 1991	9.659	12.961
Panel C: Log formal wage premium		
Δ Log formal wage premium	0.029	0.116
	(0.031)	(0.131)
Mean (unweighted), 1991	0.761	0.507
Mean (weighted), 1991	1.072	0.895
Δ De-trended log formal wage premium	-0.141	0.106
	(0.031)	(0.131)
Observations	289,680	7,124
Local labor markets	19,759	475

Notes: Standard errors in parentheses. Column 1 clusters at the local labor market level; column 2 clusters at the microregion level. All columns report the 1997 coefficient from stacked differences-in-differences regressions. Weighted means weight by total formal employment in the local labor market.

Table A.8: Effect of import competition exposure on concentration: Alternative exposure measures

	Main specification	ICE weights are firms' base year payroll shares	ICE weights are firms' base year employment shares	ICE tariff shocks are firms' effective tariff protection
	(1)	(2)	(3)	(4)
Panel A: Formal sector labor market concentration				
Δ Payroll Herfindahl (based on wage premium)	0.213	0.259	0.278	0.119
	(0.017)	(0.020)	(0.020)	(0.011)
Mean (unweighted), 1991	0.287	0.287	0.287	0.287
Mean (weighted), 1991	0.079	0.079	0.079	0.079
Δ Payroll Herfindahl	0.213	0.259	0.277	0.121
	(0.017)	(0.020)	(0.020)	(0.012)
Mean (unweighted), 1991	0.292	0.292	0.292	0.292
Mean (weighted), 1991	0.082	0.082	0.082	0.082
Δ Employment Herfindahl	0.247	0.303	0.329	0.141
	(0.016)	(0.019)	(0.020)	(0.011)
Mean (unweighted), 1991	0.239	0.239	0.239	0.239
Mean (weighted), 1991	0.056	0.056	0.056	0.056
Panel B: Log number of firms and log formal employment				
Δ Log number of firms	-0.549	-0.673	-0.736	-0.309
	(0.045)	(0.050)	(0.052)	(0.030)
Mean (unweighted), 1991	3.651	3.651	3.651	3.651
Mean (weighted), 1991	7.645	7.645	7.645	7.645
Δ Log formal employment	-0.440	-0.527	-0.577	-0.225
	(0.064)	(0.073)	(0.076)	(0.044)
Mean (unweighted), 1991	4.918	4.918	4.918	4.918
Mean (weighted), 1991	9.659	9.659	9.659	9.659
Panel C: Log formal wage premium				
Δ Log formal wage premium	0.029	0.037	0.046	0.059
	(0.031)	(0.035)	(0.037)	(0.021)
Mean (unweighted), 1991	0.761	0.761	0.761	0.761
Mean (weighted), 1991	1.072	1.072	1.072	1.072
Δ De-trended log formal wage premium	-0.141	-0.156	-0.150	-0.090
	(0.031)	(0.035)	(0.037)	(0.021)
Observations	289,680	289,680	289,680	289,680
Local labor markets	19,759	19,759	19,759	19,759

Notes: Standard errors in parentheses, clustered at the local labor market level. All columns report the 1997 coefficient from stacked differences-in-differences regressions. Weighted means weight by total formal employment in the mmc × cbo942d pair.

Table A.9: Effect of import competition exposure on concentration: Alternative clustering

	Main specification	Two-way clustered by microregion and occupational group	AKM (2019) standard errors
	(1)	(2)	(3)
Panel A: Formal sector labor market concentration			
Δ Payroll Herfindahl (based on wage premium)	0.213	0.213	0.213
	(0.017)	(0.029)	(0.006)
Mean (unweighted), 1991	0.287	0.287	0.287
Mean (weighted), 1991	0.079	0.079	0.079
Δ Payroll Herfindahl	0.213	0.213	0.213
	(0.017)	(0.028)	(0.007)
Mean (unweighted), 1991	0.292	0.292	0.292
Mean (weighted), 1991	0.082	0.082	0.082
Δ Employment Herfindahl	0.247	0.247	0.247
	(0.016)	(0.028)	(0.006)
Mean (unweighted), 1991	0.239	0.239	0.239
Mean (weighted), 1991	0.056	0.056	0.056
Panel B: Log number of firms and log formal employment			
Δ Log number of firms	-0.549	-0.549	-0.549
	(0.045)	(0.131)	(0.010)
Mean (unweighted), 1991	3.651	3.651	3.651
Mean (weighted), 1991	7.645	7.645	7.645
Δ Log formal employment	-0.440	-0.440	-0.440
	(0.064)	(0.153)	(0.019)
Mean (unweighted), 1991	4.918	4.918	4.918
Mean (weighted), 1991	9.659	9.659	9.659
Panel C: Log formal wage premium			
Δ Log formal wage premium	0.029	0.029	0.029
	(0.031)	(0.068)	(0.006)
Mean (unweighted), 1991	0.761	0.761	0.761
Mean (weighted), 1991	1.072	1.072	1.072
Δ De-trended log formal wage premium	-0.141	-0.141	-0.141
	(0.031)	(0.068)	(0.006)
Observations	289,680	289,680	289,680
Local labor markets	19,759	19,759	19,759

Notes: Standard errors in parentheses. Point estimates are identical across columns; only standard errors differ. Column 1 clusters at the local labor market level, column 2 two-way clusters by microregion and occupational group, and column 3 uses AKM (2019) shift-share standard errors. All columns report the 1997 coefficient from stacked differences-in-differences regressions. Weighted means weight by total formal employment in the mmc × cbo942d pair.

Table A.10: Effect of import competition exposure on concentration: Weighting markets

	Main specification	Weighted by local labor market 1991 employment
	(1)	(2)
Panel A: Formal sector labor market concentration		
Δ Payroll Herfindahl (based on wage premium)	0.213	0.156
	(0.017)	(0.032)
Mean (unweighted), 1991	0.287	0.287
Mean (weighted), 1991	0.079	0.079
Δ Payroll Herfindahl	0.213	0.162
	(0.017)	(0.034)
Mean (unweighted), 1991	0.292	0.292
Mean (weighted), 1991	0.082	0.082
Δ Employment Herfindahl	0.247	0.098
	(0.016)	(0.018)
Mean (unweighted), 1991	0.239	0.239
Mean (weighted), 1991	0.056	0.056
Panel B: Log number of firms and log formal employment		
Δ Log number of firms	-0.549	-0.657
	(0.045)	(0.159)
Mean (unweighted), 1991	3.651	3.651
Mean (weighted), 1991	7.645	7.645
Δ Log formal employment	-0.440	-0.187
	(0.064)	(0.142)
Mean (unweighted), 1991	4.918	4.918
Mean (weighted), 1991	9.659	9.659
Panel C: Log formal wage premium		
Δ Log formal wage premium	0.029	-0.004
	(0.031)	(0.071)
Mean (unweighted), 1991	0.761	0.761
Mean (weighted), 1991	1.072	1.072
Δ De-trended log formal wage premium	-0.141	-0.332
	(0.031)	(0.071)
Observations	289,680	289,680
Local labor markets	19,759	19,759

Notes: Standard errors in parentheses, clustered at the local labor market level. All columns report the 1997 coefficient from stacked differences-in-differences regressions. Weighted means weight by total formal employment in the mmc×cbo942d pair.

Table A.11: Estimates of within-market cross-firm inverse elasticity of substitution $1/\eta$

	Using December wage conditional on observables (Main specification) (1)	Using December wage conditional on worker FE and demo-by-year controls (2)	Using (2) and further conditioning on stayers in firm-market pair (3)	Using December average wage (4)	Using effective rate of protection (5)
Panel A: First stage					
Δ Firm log employment in LLM	-0.556	-0.604	-0.604	-0.556	-0.359
	(0.044)	(0.053)	(0.075)	(0.044)	(0.035)
First stage F	156.771	130.168	65.315	156.771	106.426
Panel B: Reduced form					
Δ Firm wage premium in LLM	-0.550	-0.489	-0.470	-0.531	-0.354
	(0.024)	(0.030)	(0.042)	(0.025)	(0.019)
Panel C: 2SLS					
Labor supply within-market cross-firm inverse elasticity of substitution	0.990	0.811	0.778	0.956	0.984
	(0.089)	(0.080)	(0.110)	(0.089)	(0.109)
Implied upper bound on wage take-home share	50%	55%	56%	51%	50%
Local labor market (LLM) FE	Yes	Yes	Yes	Yes	Yes
Observations	855,104	463,138	182,757	855,104	852,702
Firms	344,534	213,704	89,252	344,534	344,027
Local labor markets	15,730	13,627	9,505	15,730	15,679

Notes: Standard errors clustered by firm in parentheses. All regressions include LLM fixed effects and are weighted by baseline market employment. Column (1): main specification using December wage premium conditional on observables. Column (2): wage premium conditional on worker fixed effects and demographic-by-year controls (strict stayers). Column (3): further restricts column (2) to workers who stay in the same firm-market pair. Column (4): uses December average wage instead of residualized wage premium. Column (5): uses effective rate of protection as instrument instead of TRAINS tariff.

Table A.12: Estimates of within-market cross-firm inverse elasticity of substitution: Alternative samples

	Main specification	Unique producers	Local labor market defined as microregion	Including exiting firms, coding employment and wages at exit as zero
	(1)	(2)	(3)	(4)
Panel A: First stage				
Δ Firm log employment in LLM	-0.556	-0.661	-0.417	-0.556
	(0.044)	(0.151)	(0.037)	(0.044)
First stage F	156.771	19.164	124.666	158.127
Panel B: Reduced form				
Δ Firm's wage premium in LLM	-0.550	-0.347	-0.404	-0.551
	(0.024)	(0.082)	(0.017)	(0.024)
Panel C: 2SLS				
Labor supply within-market cross-firm inverse elasticity of substitution	0.990	0.525	0.969	0.990
	(0.089)	(0.190)	(0.092)	(0.089)
Implied upper bound on wage take-home share	50%	66%	51%	50%
Observations	855,104	696,197	440,966	1,617,416
Firms	344,534	302,514	420,246	720,085
Local labor markets	15,730	13,146	474	18,602

Notes: Standard errors clustered by firm in parentheses. All regressions include local labor market fixed effects and are weighted by baseline market employment. Column (1): main specification with LLM defined as mmc × occupation. Column (2): restricts to markets with a unique producer of each sector. Column (3): defines LLM as microregion (mmc only, no occupation dimension). Column (4): includes firms that exited between 1991 and 1997, coding their exit-year employment and wages as zero (IHS transform). Column (5): restricts to firms in tradable sectors (IBGE subsectors 1–13 and 25).

Table A.13: Estimates of within-market cross-firm inverse elasticity of substitution: Alternative clustering

	Main specification (Clustered by firm) (1)	Clustered by local labor market (2)	Clustered by sector (3)
Panel A: First stage			
Δ Firm log employment in LLM	-0.556	-0.556	-0.556
	(0.044)	(0.067)	(0.107)
First stage F	156.771	69.608	26.984
Panel B: Reduced form			
Δ Firm wage premium in LLM	-0.550	-0.550	-0.550
	(0.024)	(0.103)	(0.103)
Panel C: 2SLS			
Labor supply within-market cross-firm inverse elasticity of substitution	0.990	0.990	0.990
	(0.089)	(0.201)	(0.148)
Observations	855,104	855,104	855,104
Firms	344,534	344,534	344,534
Local labor markets	15,730	15,730	15,730

Notes: Standard errors in parentheses. Column (1) clusters by firm, column (2) by local labor market (mmc × occupation), column (3) by sector (cnae95). All regressions include LLM fixed effects and are weighted by baseline market employment.

Table A.14: Correlates of firm-level wage markdowns

	Dependent variable: $\ln(\mu_{zm})$			
	(1)	(2)	(3)	(4)
Panel A: Employment				
$\ln(\text{Employment}_{zm})$	0.002681	0.002529	0.002425	0.002724
	(0.000623)	(0.000591)	(0.000580)	(0.000636)
Observations	9,683,946	9,683,945	9,578,928	9,683,945
Panel B: Residual wage				
$\ln(\text{Residual wage}_{zm})$	0.000489	0.000410	0.000203	0.002292
	(0.000485)	(0.000384)	(0.000437)	(0.000610)
Observations	9,683,946	9,683,945	9,578,928	9,683,945
Panel C: Average wage				
$\ln(\text{Average wage}_{zm})$	0.000459	0.000336	0.000007	0.001929
	(0.000426)	(0.000345)	(0.000403)	(0.000544)
Observations	9,683,946	9,683,945	9,578,928	9,683,945
Year FE	Yes	Yes	Yes	Yes
Sector FE		Yes		
Firm FE			Yes	
Local labor market FE				Yes
Firms	589,484	589,483	484,466	589,484
Local labor markets	19,054	19,054	19,038	19,053

Notes: This table presents regressions of log firm-level wage markdowns on firm-level employment, residual wages, and average wages. All regressions include year fixed effects. Standard errors in parentheses.

A Data and Methods Appendix

Data on workers and firms: RAIS

Overview. I use Brazil's Relação Anual de Informações Sociais (RAIS) for years 1986 to 2000 as my source of information on workers and firms. RAIS is an administrative employer-employee linked dataset collected by the federal government for the purposes of administering workers' social security. Thus, RAIS covers all workers with signed worker cards (Carteira do Trabalho), namely the entirety of formal sector employment. Firms report RAIS once a year, reporting all workers who ever worked for the firm in the prior calendar year. Firms are required to report a rich set of information about each employment contract (e.g., occupation, admission date, separation date, etc.), as well as worker demographics (i.e., education, date of birth, and gender), separately for each establishment. The municipality of each establishment as well as the economic sector of the firm are also reported.

Wages. RAIS includes two wage variables for years 1986-2000: average monthly earnings and December monthly earnings. Both variables are reported as multiples of the national minimum wage.

Occupation codes. RAIS' occupation codes are 5-digit variables "CBO" (prior to 1994) and "CBO94" (1994 onwards). I focus on the first 2 digits to group workers into occupation groups. Both variables share the same data dictionary, with the only difference between them being phased-out and phased-in occupation codes. I have compiled a complete list of all raw occupation codes, along with the total number of workers in each of them, labels, and flags for which codes were either "phased-out" or phased-in, which I identified based on whether the number of workers changing by more than 100 times between any two years. I then re-classified the first two digits of all phased-out and phased-in codes as "99 - Other occupations," a reclassification that affects roughly 10% of all workers.

Sector codes. RAIS' finest sector codes for 1986-2000 are 4-digit "IBGESUBATIVIDADE" (prior to 1995) and 5-digit "CNAE95" (1995 onwards). I focus on the 5-digit CNAE95 codes to map tariff shocks to firms in RAIS. For firms that exit the data prior to reporting any CNAE95 codes, I assign a CNAE95 code using a correspondence table I constructed using the pre-1995 and post-1995 codes of firms in business in both periods. To each IBGESUBATIVIDADE code I assign the most commonly reported CNAE95 code. Finally, throughout all years I use the first CNAE95 code ever reported by a firm as its official CNAE95 code.

Sample restrictions. I focus on workers employed as of December 31 of each year, and aged 18-65, and with positive December earnings. I exclude all workers in the public sector or with unknown sector. To make sure all public sector workers are excluded, I further exclude workers whose employer's economic activity was not marked as government, but which exert public sector occupations (i.e., Diplomats, Civil servants, and Post office). Finally, following Dix-Carneiro and Kovak (2017) I exclude from all analyses the free trade zone of Manaus.

Data on tariff shocks: TRAINS

I use tariff data from UNCTAD's Trade Analysis Information System (TRAINS), which I download from the World Integrated Trade Solution (WITS)'s website (https://wits.worldbank.org/). The raw tariff data are available for Brazil at the 8-digit HS product level for years 1988 (the first year the data are available) through 2000. As outlined in Section 4, I compute a firm's tariff reduction shock as the change in log one plus the firm's CNAE95 sector code's nominal tariff between years 1990 and 1994.

To map product-level tariffs to CNAE95, I use Brazil's NCM (Nomenclatura Comum do Mercosul) classification, whose first six digits correspond to the HS. NCM codes are mapped to CNAE95 using official IBGE correspondence tables.[47] Over 90% of products map to a single activity code; the remainder are randomly assigned across the handful of activity codes to which they correspond per the concordance tables. CNAE95-level tariffs are weighted averages of product-level nominal tariffs, weighted by the number of tariff lines. For robustness, I compute effective rates of protection (ERP) using the Corden formula: $ERP_j = (t_j - \sum_i a_{ij} t_i)/(1 - \sum_i a_{ij})$, where t_j is sector j's nominal tariff, t_i are input tariffs, and a_{ij} are Leontief coefficients from Brazil's 1985 intersectoral coefficients matrix ("Tabela 20") at Nível 50.[48] Nível 50 sectors are mapped to CNAE95 using CONCLA correspondence tables.

Other data

List of exporters. I classify firms as exporters during the reform period (1990-1994) by matching the list of exporters during that period to RAIS using firms' unique identifiers (CNPJ). The list of exporters was provided by the (extinct as of 2019) Ministry of Development, Industry, and Foreign Trade, currently a part of the Ministry of the Economy, in October 2018.

Census. I use data from the 1991 and 2000 census, available in the supplementary materials from Dix-Carneiro and Kovak (2017).

Methods: wage premia regressions

For each year, I estimate each firm's wage premium in its local labor market as firm ×market fixed effects in a regression of worker log December earnings on the firm × market fixed effects and the same worker observable controls as Dix-Carneiro and Kovak (2017), namely: a dummy for female; 4 age group dummies (25-29; 30-39; 40-49, 50-64); 8 education group dummies (primary school, incomplete primary school, middle school, incomplete middle school, high school, incomplete high school, college, incomplete college). The omitted category is therefore males aged 18-24 with no formal education. Similarly, for each year, I estimate each market's wage premia as a regression of worker log December earnings on the market fixed effects and the previously mentioned worker observable controls.

For the robustness exercise of the within-market cross firm elasticity using wage premia that condition on

[47] https://concla.ibge.gov.br/classificacoes/correspondencias/atividades-economicas.

[48] https://www.ibge.gov.br/estatisticas-novoportal/economicas/contas-nacionais/9085-matriz-de-insumo-produto.html?&t=downloads.

worker fixed effects (e.g., columns (2) and (3) of Appendix Table A.11, I estimate each firm's wage premia in 1991 and 1997 as firm ×market × year fixed effects in a regression—containing years 1991 and 1997—of worker log December earnings on worker fixed effects, the firm ×market × year fixed effects, and worker observable-characteristics-by-year controls.

Methods: effects relative to trend

For wage premia, where positive pre-trends are observed, I also report treatment effects of import competition exposure relative to trend. These effects are estimated as the $\tilde{\beta}$ coefficients from the following regression:[49]

$$\Delta\tilde{Y}_{mt} = \sum_{k\neq 1991} \tilde{\zeta}_k \left(\Delta ICE_m \times 1_{t=k}\right) + \tilde{\delta}_m + \tilde{\delta}_t + \tilde{\epsilon}_{mt} \quad \text{(A.1)}$$

where $\Delta\tilde{Y}_{mt} = \Delta Y_{mt} - \hat{\varphi}\left(\Delta ICE_m \times t\right)$ is the predicted outcome from the following regression, which I estimate using the pre-treatment years 1986-1990 only:

$$\Delta Y_{mt} = \varphi\left(\Delta ICE_m \times t\right) + \nu_m + \nu_t + \nu_{mt} \quad \text{(A.2)}$$

in which ν_m and ν_t are local labor market and year fixed effects, respectively. Causal interpretation of the $\tilde{\beta}_k$ coefficients rely on the identification assumption that more affected markets would have continued to follow the same pre-liberalization growth trend relative to least affected markets.

[49] For didactic purposes, I express the fixed effects in regression equation A.1 and in its non-detrended counterpart (e.g., equation 12) as simply δ_m and δ_t, which makes it easier for the reader to see how this regression is a stacked difference-in-differences specification. In practice, the (equivalent for ζ_k) regressions I actually estimate are of the form:

$$\Delta Y_{mt} = \alpha + \sum_{k\neq 1991} \zeta_k \left(\Delta ICE_m \times 1_{t=k}\right) + \sum_{k\neq 1991} \delta_k \left(1_{t=k}\right) + \sum_{m\neq b} \delta_m + \epsilon_{mt}$$

where the constant α is included in the estimation, the base year fixed effect δ_{1991} is omitted, and one market fixed effect δ_b is also omitted. I implement this using the command reghdfe in Stata, absorbing market fixed effects only (i.e., no standard errors are estimated for those and one is automatically omitted), and manually add regressors for all year fixed effects except for the base year. While producing identical point estimates for ζ_k as equation 12, this approach has the advantage of giving, via estimates for the constant and year fixed effects relative to base year, a descriptive account of what is happening to the least intensively treated markets over time relative to the base year, which is helpful for interpretation.

B Model Appendix

This Appendix provides detailed derivations for various expressions in Section 2.

B.1 Derivation of labor supply equation

Consider an economy consisting of a continuum of homogenous workers j, a large but finite number of local labor markets m, and a finite number of firms z within each local labor market. Each worker chooses to which firm-market pair zm they provide h_{zm}^j units of labor by minimizing their indirect disutility of work V_{zm} subject to making reservation earnings $y^j \sim F(y)$:

$$\min_{zm} V_{zm}^j = \ln l_{zm}^j + \ln \xi_m + \ln \xi_{zm} - \xi_{zm}^j$$
$$\text{s.t. } l_{zm}^j w_{zm} \geq y_j$$

where ξ_{zm}^j is an idiosyncratic taste for working at firm z in market m, and ξ_m and ξ_{zm} are taste shifters common to all workers. This is equivalent to

$$\max_{zm} \ln w_{zm} - \ln y_j - \ln \xi_m - \ln \xi_{zm} + \xi_{zm}^j$$

Now suppose ξ_{zm}^j follows the following Gumbel distribution, a member of the General Extreme Value (GEV) family:

$$G\left(\left\{\xi_{zm}^j\right\}\right) = \exp\left[-\sum_m \left(\sum_{z \in B_m} e^{-\left(1+\frac{\sigma}{1-\sigma}\right)\xi_{zm}^j}\right)^{\frac{1+\frac{\varphi}{1-\varphi}}{1+\frac{\sigma}{1-\sigma}}}\right]$$

where $0 \leq \sigma < 1$ is the index of similarity across firms within a market, $0 \leq \varphi < 1$ is the index of similarity across markets, and B_m is the set of firms in market m.

The probability that worker j chooses firm z in market m is $P\left(\xi_{zm}^j > \ln w_{zm} - \ln y_j - \ln \xi_m - \ln \xi_{zm}\right)$, which can be decomposed as:

$$P_{zm}^j = P(z|B_m)\, P(B_m) \quad \forall j$$

where $P(z|B_m)$ is the probability of choosing firm z conditional on choosing market m with set B_m of firms, and $P(B_m)$ is the probability of choosing market m. By the results in McFadden (1978), P_{zm} can be computed as:

$$
\begin{aligned}
P\left(z|B_m\right) =& \frac{\exp\left[\left(\ln w_{zm} - \ln y^j - \ln \xi_m - \ln \xi_{zm}\right)/(1-\sigma)\right]}{\sum_{k\in B_n} \exp\left[\left(\ln w_{kn} - \ln y^j - \ln \xi_m - \ln \xi_{km}\right)/(1-\sigma)\right]}\\
=& \frac{\exp\left[\left(\ln w_{zm}^{\frac{1}{1-\sigma}} - \ln y^{j\frac{1}{1-\sigma}} - \ln \xi_m^{\frac{1}{1-\sigma}} - \ln \xi_{zm}^{\frac{1}{1-\sigma}}\right)\right]}{\sum_{k\in B_m} \exp\left[\left(\ln w_{km}^{\frac{1}{1-\sigma}} - \ln y^{j\frac{1}{1-\sigma}} - \ln \xi_m^{\frac{1}{1-\sigma}} - \ln \xi_{km}^{\frac{1}{1-\sigma}}\right)\right]}\\
=& \frac{\left(\frac{w_{zm}}{y^j \xi_m \xi_{zm}}\right)^{\frac{1}{1-\sigma}}}{\sum_{k\in B_n}\left(\frac{w_{km}}{y^j \xi_m \xi_{km}}\right)^{\frac{1}{1-\sigma}}}\\
=& \frac{\left(\frac{w_{zm}}{\xi_{zm}}\right)^{\frac{1}{1-\sigma}}}{\sum_{k\in B_n}\left(\frac{w_{km}}{\xi_{km}}\right)^{\frac{1}{1-\sigma}}}
\end{aligned}
$$

and

$$
\begin{aligned}
P\left(B_m\right) =& \frac{\left\{\sum_{z\in B_m} \exp\left[\left(\ln w_{zm} - \ln y^j - \ln \xi_m - \ln \xi_{zm}\right)/(1-\sigma)\right]\right\}^{\frac{1-\sigma}{1-\varphi}}}{\sum_l \left\{\sum_{k\in B_l} \exp\left[\left(\ln w_{kl} - \ln y^j - \ln \xi_l - \ln \xi_{kl}\right)/(1-\sigma)\right]\right\}^{\frac{1-\sigma}{1-\varphi}}}\\
=& \frac{\left[\sum_{z\in B_m}\left(\frac{w_{zm}}{y^i \xi_m \xi_{zm}}\right)\frac{1}{1-\sigma}\right]^{\frac{1-\sigma}{1-\varphi}}}{\sum_l\left[\sum_{k\in B_l}\left(\frac{w_{km}}{y^i \xi_m \xi_{km}}\right)\frac{1}{1-\sigma}\right]^{\frac{1-\sigma}{1-\varphi}}}\\
=& \frac{\left[\left(\frac{1}{\xi_m}\right)^{\frac{1}{1-\sigma}}\sum_{z\in B_m}\left(\frac{w_{zm}}{\xi_{zm}}\right)\frac{1}{1-\sigma}\right]^{\frac{1-\sigma}{1-\varphi}}}{\sum_l\left[\left(\frac{1}{\xi_l}\right)^{\frac{1}{1-\sigma}}\sum_{k\in B_l}\left(\frac{w_{kl}}{\xi_{kl}}\right)\frac{1}{1-\sigma}\right]^{\frac{1-\sigma}{1-\varphi}}}
\end{aligned}
$$

Putting them together

$$
P_{zm}^j = \frac{\left(\frac{w_{zm}}{\xi_{zm}}\right)^{\frac{1}{1-\sigma}}}{\sum_{k\in B_n}\left(\frac{w_{km}}{\xi_{km}}\right)^{\frac{1}{1-\sigma}}} \times \frac{\left[\left(\frac{1}{\xi_m}\right)^{\frac{1}{1-\sigma}}\sum_{z\in B_m}\left(\frac{w_{zm}}{\xi_{zm}}\right)\frac{1}{1-\sigma}\right]^{\frac{1-\sigma}{1-\varphi}}}{\sum_l\left[\left(\frac{1}{\xi_l}\right)^{\frac{1}{1-\sigma}}\sum_{k\in B_l}\left(\frac{w_{kl}}{\xi_{kl}}\right)\frac{1}{1-\sigma}\right]^{\frac{1-\sigma}{1-\varphi}}} \quad \forall j
$$

Let $\eta \equiv \frac{\sigma}{1-\sigma} > 0$, $\theta \equiv \frac{\varphi}{1-\varphi} > 0$, and denote $P_{zm}^j = P_{zm}$ for simplicity. Then:

$$
P_{zm} = \frac{\left(\frac{w_{zm}}{\xi_{zm}}\right)^{1+\eta}}{\sum_{k\in B_n}\left(\frac{w_{km}}{\xi_{km}}\right)^{1+\eta}} \times \frac{\left[\left(\frac{1}{\xi_m}\right)^{1+\eta}\sum_{z\in B_m}\left(\frac{w_{zm}}{\xi_{zm}}\right)^{1+\eta}\right]^{\frac{1+\theta}{1+\eta}}}{\sum_l\left[\left(\frac{1}{\xi_l}\right)^{1+\eta}\sum_{k\in B_l}\left(\frac{w_{kl}}{\xi_{kl}}\right)^{1+\eta}\right]^{\frac{1+\theta}{1+\eta}}}
$$

Finally, define the following wage indices:

$$W_m \equiv \left[\sum_z \left(\frac{w_{zm}}{\xi_{zm}}\right)^{1+\eta}\right]^{\frac{1}{1+\eta}}, \quad W \equiv \left[\sum_m \left(\frac{W_m}{\xi_m}\right)^{1+\theta}\right]^{\frac{1}{1+\theta}}$$

Then

$$P^j_{zm} = \frac{\left(\frac{w_{zm}}{\xi_{zm}}\right)^{1+\eta}}{W_m^{1+\eta}} \times \frac{\left(\frac{W_m}{\xi_m}\right)^{1+\theta}}{W^{1+\theta}} = \left(\frac{w_{zm}/\xi_{zm}}{W_m}\right)^{1+\eta} \times \left(\frac{W_m/\xi_m}{W}\right)^{1+\theta} \tag{B.1}$$

With equation B.1 at hand, total labor supplied to firm z in market m can be found by integrating probabilities P^j_{zm} (times $h^j_{zm} = y^j/w_{zm}$ supplied by each worker) over the continuum of workers:

$$l_{zm} = \int_0^1 P^j_{zm}\left(\frac{y^j}{w_{zm}}\right) dF(y) = w_{zm}^{-1} P_{zm} Y \tag{B.2}$$

where $Y \equiv \int_0^1 y^j dF(y)$ is the country-level labor income. To obtain an expression for l_{zm} that is a function of w_{zm}, parameters, and market-level aggregates, I define the following employment indices:

$$L_m \equiv \left[\sum_z (\xi_{zm} l_{km})^{\frac{1+\eta}{\eta}}\right]^{\frac{\eta}{1+\eta}}, \quad L \equiv \left[\sum_m (\xi_m L_m)^{\frac{1+\theta}{\theta}}\right]^{\frac{\theta}{1+\theta}}$$

which together with equation B.2 and previously defined wage indices imply $Y = \sum_{zm} w_{zm} l_{zm} = WL$ and

$$\begin{aligned}
l_{zm} =& w_{zm}^{-1} P_{zm} Y \\
=& w_{zm}^{-1}\left[\left(\frac{w_{zm}}{\xi_{zm}W_m}\right)^{1+\eta} \times \left(\frac{W_m}{\xi_m W}\right)^{1+\theta}\right] Y \\
=& w_{zm}^{-1}\left[\left(\frac{w_{zm}}{\xi_{zm}W_m}\right)\left(\frac{w_{zm}}{\xi_{zm}W_m}\right)^{\eta} \times \left(\frac{W_m}{\xi_m W}\right)^{\theta}\left(\frac{W_m}{\xi_m W}\right)\right] WL \\
=& \xi_{zm}\xi_m \left(\frac{w_{zm}/\xi_{zm}}{W_m}\right)^{\eta}\left(\frac{W_m/\xi_m}{W}\right)^{\theta} L
\end{aligned}$$

Rearranging:

$$l_{zm} = L\left(\frac{w_{zm}}{W_m}\right)^{\eta}\left(\frac{W_m}{W}\right)^{\theta}\left(\xi_{zm}^{1+\eta}\xi_m^{1+\theta}\right)^{-1} \tag{B.3}$$

B.2 Other proofs and derivations

B.2.1 Equation 3: $w_{zm} = W\left(\frac{l_{zm}}{L_m}\right)^{\frac{1}{\eta}}\left(\frac{L_m}{L}\right)^{\frac{1}{\theta}} \xi_{zm}^{1+\frac{1}{\eta}} \xi_m^{1+\frac{1}{\theta}}$

The inverse function of the residual labor supply equation 2 (same as Appendix equation B.3) is the wage w_{zm} at which l_{zm} units of labor are supplied to firm z at market m. To check that equation 3 satisfies this criterion, plug it into equation B.2 to obtain the identity $l_{zm} = l_{zm}$. I show this in two steps.

First, plug in the expression for w_{zm} into equation B.2 to get:

$$
\begin{aligned}
l_{zm} =& w_{zm}^{-1} P_{zm} Y \\
=& w_{zm}^{-1} \left(\frac{w_{zm}/\xi_{zm}}{W_m}\right)^{1+\eta} \times \left(\frac{W_m/\xi_m}{W}\right)^{1+\theta} Y \\
=& w_{zm}^{\eta} \left(\frac{1}{W_m \xi_{zm}}\right)^{1+\eta} \times \left(\frac{W_m/\xi_m}{W}\right)^{1+\theta} Y \\
=& \left[W \left(\frac{l_{zm}\xi_{zm}}{L_m}\right)^{\frac{1}{\eta}} \left(\frac{L_m \xi_m}{L}\right)^{\frac{1}{\theta}} \xi_m \xi_{zm}\right]^{\eta} \left(\frac{1/\xi_{zm}}{W_m}\right)^{1+\eta} \times \left(\frac{W_m/\xi_m}{W}\right)^{1+\theta} Y \\
=& \left[W \left(\frac{l_{zm}\xi_{zm}}{L_m}\right)^{\frac{1}{\eta}} \left(\frac{L_m \xi_m}{L}\right)^{\frac{1}{\theta}} \xi_m \xi_{zm}\right]^{\eta} \left(\frac{1/\xi_{zm}}{W_m}\right)^{\eta} \times \left(\frac{W_m/\xi_m}{W}\right)^{\theta} WL \left(\frac{1/\xi_{zm}}{W_m}\right) \times \left(\frac{W_m/\xi_m}{W}\right) \\
=& l_{zm} \xi_m^{\eta} \xi_{zm}^{\eta} \left(\frac{\xi_{zm}}{L_m}\right) \left[W \left(\frac{L_m \xi_m}{L}\right)^{\frac{1}{\theta}}\right]^{\eta} \left(\frac{1/\xi_{zm}}{W_m}\right)^{\eta} \times \left(\frac{W_m/\xi_m}{W}\right)^{\theta} \left(\frac{L}{\xi_{zm}\xi_m}\right) \\
=& l_{zm} \left(\frac{\xi_m^{\eta}}{L_m}\right) \left[W \left(\frac{L_m \xi_m}{L}\right)^{\frac{1}{\theta}}\right]^{\eta} \left(\frac{1}{W_m}\right)^{\eta} \times \left(\frac{W_m/\xi_m}{W}\right)^{\theta} \left(\frac{L}{\xi_m}\right) \\
=& l_{zm} \left(\frac{1}{L_m}\right) \left(\frac{L_m \xi_m}{L}\right)^{\frac{\eta}{\theta}} \left(\frac{W \xi_m}{W_m}\right)^{\eta} \times \left(\frac{W_m/\xi_m}{W}\right)^{\theta} \left(\frac{L}{\xi_m}\right) \\
=& l_{zm} \left(\frac{W \xi_m}{W_m}\right)^{\eta} \times \left(\frac{W_m/\xi_m}{W}\right)^{\theta} \left(\frac{L}{L_m \xi_m}\right) \left(\frac{L_m \xi_m}{L}\right)^{\frac{\eta}{\theta}} \\
=& l_{zm} \left(\frac{W}{W_m/\xi_m}\right)^{\eta} \times \left(\frac{W_m/\xi_m}{W}\right)^{\theta} \left(\frac{L}{L_m \xi_m}\right) \left(\frac{L_m \xi_m}{L}\right)^{\frac{\eta}{\theta}} \\
=& l_{zm} \left(\frac{W_m/\xi_m}{W}\right)^{\theta-\eta} \left(\frac{L_m \xi_m}{L}\right)^{\frac{\eta}{\theta}-1} \\
=& l_{zm} \left(\frac{W}{W_m/\xi_m}\right)^{\eta-\theta} \left(\frac{L_m \xi_m}{L}\right)^{\frac{\eta-\theta}{\theta}} \\
=& l_{zm} \left(\frac{W}{W_m/\xi_m}\right)^{\eta-\theta} \left(\frac{L_m \xi_m}{L}\right)^{\frac{\eta-\theta}{\theta}}
\end{aligned}
$$

Second, I show that $\left(\frac{W}{W_m/\xi_m}\right)^{\eta-\theta} \left(\frac{L_m \xi_m}{L}\right)^{\frac{\eta-\theta}{\theta}} = 1$ by expressing the CES wage index W_m as a function of labor and taste shifters, which can be done by first plugging in the expression for w_{zm} into the definition of

W_m:

$$
\begin{aligned}
W_m &= \left[\sum_{k\in B_n}\left(\frac{w_{zm}}{\xi_{km}}\right)^{1+\eta}\right]^{\frac{1}{1+\eta}} \\
&= \left[\sum_{k\in B_n}\left(\frac{\left[W\left(\frac{l_{km}\xi_{km}}{L_m}\right)^{\frac{1}{\eta}}\left(\frac{L_m\xi_m}{L}\right)^{\frac{1}{\theta}}\xi_m\xi_{km}\right]}{\xi_{km}}\right)^{1+\eta}\right]^{\frac{1}{1+\eta}} \\
&= \left[W^{1+\eta}\left(\frac{1}{L_m}\right)^{\frac{1+\eta}{\eta}}\left[\left(\frac{L_m\xi_m}{L}\right)^{\frac{1}{\theta}}\xi_m\right]^{1+\eta}\sum_{k\in B_n}\left(\frac{\left[(l_{km}\xi_{km})^{\frac{1}{\eta}}\xi_{km}\right]}{\xi_{km}}\right)^{1+\eta}\right]^{\frac{1}{1+\eta}} \\
&= \frac{W}{L_m^{\frac{1}{\eta}}}\left(\frac{L_m\xi_m}{L}\right)^{\frac{1}{\theta}}\xi_m\left[\sum_{k\in B_n}(l_{km}\xi_{km})^{\frac{1+\eta}{\eta}}\right]^{\frac{1}{1+\eta}} \\
&= \frac{W}{L_m^{\frac{1}{\eta}}}\left(\frac{L_m\xi_m}{L}\right)^{\frac{1}{\theta}}\xi_m L_m^{\frac{1}{\eta}} \\
&= W\left(\frac{L_m\xi_m}{L}\right)^{\frac{1}{\theta}}\xi_m
\end{aligned}
$$

Thus, $W_m = W\left(\frac{L_m\xi_m}{L}\right)^{\frac{1}{\theta}}\xi_m$. Recall from the first step that completing the proof requires showing that $\left(\frac{W}{W_m/\xi_m}\right)^{\eta-\theta}\left(\frac{L_m\xi_m}{L}\right)^{\frac{\eta-\theta}{\theta}} = 1$. Plugging in the expression for W_m into this equation gives:

$$
\begin{aligned}
\left(\frac{W}{W_m/\xi_m}\right)^{\eta-\theta}\left(\frac{L_m\xi_m}{L}\right)^{\frac{\eta-\theta}{\theta}} &= \left(\frac{W}{\left[W\left(\frac{L_m\xi_m}{L}\right)^{\frac{1}{\theta}}\xi_m\right]/\xi_m}\right)^{\eta-\theta}\left(\frac{L_m\xi_m}{L}\right)^{\frac{\eta-\theta}{\theta}} \\
&= \left(\frac{L_m\xi_m}{L}\right)^{-\frac{(\eta-\theta)}{\theta}}\left(\frac{L_m\xi_m}{L}\right)^{\frac{\eta-\theta}{\theta}} \\
&= 1
\end{aligned}
$$

which completes the proof that w_{zm} is the inverse function of l_{zm}.

B.2.2 Equation 7: $s_{zm} \equiv \frac{w_{zm} l_{zm}}{\sum_k (w_{km} l_{km})} = \frac{\partial \ln L_m}{\partial \ln l_{zm}}$.

To see why this holds, start from the definition of the labor market index L_m in Section 2 to derive $\partial \ln L_m / \partial \ln l_{zm}$ as

$$\frac{\partial \ln L_m}{\partial \ln l_{zm}} = \frac{(\xi_{km} l_{km})^{\frac{1+\eta}{\eta}}}{\sum_{j=1}^{N_m} \left(\xi_{jm} l_{jm}\right)^{\frac{1+\eta}{\eta}}}$$

Now set this aside. Plug in equation 3 to the definition $s_{zm} \equiv w_{zm} l_{zm} / \sum_k (w_{km} l_{km})$ to obtain

$$s_{zm} = \frac{(\xi_{km} l_{km})^{\frac{1+\eta}{\eta}}}{\sum_{j=1}^{N_m} \left(\xi_{jm} l_{jm}\right)^{\frac{1+\eta}{\eta}}}$$

Therefore, $s_{zm} = \partial \ln L_m / \partial \ln l_{zm}$.

B.2.3 Proposition 1: $\mu_m \equiv \frac{\bar{r}_m}{\bar{w}_m} = 1 + \varepsilon_m^{-1} = 1 + \frac{1}{\theta} HHI_m + \frac{1}{\eta} (1 - HHI_m)$,

In this expression, $\bar{w}_m$ and $\bar{r}_m$ are the (employment-weighted) average wage and average marginal revenue product of labor in market m, respectively.

First, I show that $1 + \varepsilon_m^{-1} = 1 + \frac{1}{\theta} HHI_m + \frac{1}{\eta} (1 - HHI_m)$. To see why this holds, let Θ_m denote the set of firms operating in labor market m, and take the (payroll-share-weighted) average of equation 8:

$$\begin{aligned}
\underbrace{\sum_{z \in \Theta_m} s_{zm} \left(1 + \varepsilon_{zm}^{-1}\right)}_{\equiv 1 + \varepsilon_m^{-1}} &= 1 + \sum_{z \in \Theta_m} s_{zm} \left[\frac{1}{\eta} (1 - s_{zm}) + \frac{1}{\theta} s_{zm}\right] \\
&= 1 + \sum_{z} \left[\frac{1}{\theta} s_{zm}^2 + \frac{1}{\eta} \left(s_{zm} - s_{zm}^2\right)\right] \\
&= 1 + \frac{1}{\theta} HHI_m + \frac{1}{\eta} (1 - HHI_m)
\end{aligned}$$

Second, I show that $\sum_{z \in \Theta_m} s_{zm} \left(1 + \varepsilon_{zm}^{-1}\right) = \frac{\bar{r}_m}{\bar{w}_m}$. To see that this equality holds, aggregate the firm-level markdown equation $\frac{r_{zm}}{w_{zm}} = 1 + \varepsilon_{zm}^{-1}$ using payroll shares as weights to get:

$$
\begin{aligned}
\underbrace{\sum_{z \in \Theta_m} s_{zm} \left(1 + \varepsilon_{zm}^{-1}\right)}_{\equiv 1+\varepsilon_m^{-1}} &= \sum_{z \in \Theta_m} s_{zm} \left(\frac{r_{zm}}{w_{zm}}\right) \\
&= \sum_{z \in \Theta_m} \frac{w_{zm} l_{zm}}{\sum_j w_{jm} l_{jm}} \left(\frac{r_{zm}}{w_{zm}}\right) \\
&= \frac{\sum_{z \in \Theta_m} r_{zm} l_{zm}}{\sum_{j \in \Theta_m} w_{jm} l_{jm}} \\
&= \frac{\left(\sum_{z \in \Theta_m} r_{zm} l_{zm}\right) / \left(\sum_{z \in \Theta_m} l_{zm}\right)}{\left(\sum_{j \in \Theta_m} w_{jm} l_{jm}\right) / \left(\sum_{z \in \Theta_m} l_{zm}\right)} \\
&= \frac{\bar{r}_m}{\bar{w}_m} \equiv \mu_m
\end{aligned}
$$

B.2.4 Country-level average wage markdown

I show that a particular country-level average of the market-level average wage markdown (i.e., equation 9) equals the country-level average (employment-weighted) wage markdown. The reader can then directly verify that the resulting expression is the inverse of Berger, Herkenhoff and Mongey (2022)'s expression for the "labor market power adjustment" component of the country-level labor share (see authors' equation 10). Consider the market-level average wage markdown expression from Proposition 1. Then the country-level (employment-weighted) average wage markdown is given by:

$$
\mu \equiv \frac{\bar{r}}{\bar{w}} = \sum_m s_m \mu_m = 1 + \frac{1}{\theta} H\tilde{H}I + \frac{1}{\eta}\left(1 - H\tilde{H}I\right)
$$

where $s_m = \frac{\bar{w}_m l_m}{\sum_m \bar{w}_m l_m}$ is market m's payroll share, $H\tilde{H}I = \sum_m s_m HHI_m$ is the country-level payroll-share-weighted average payroll Herfindahl, and $\bar{w}$ and $\bar{r}$ are the (employment-weighted) average wage and average marginal revenue product of labor at the country-level, respectively.

Proof. Having provided a more detailed proof for Proposition 1, I use the same steps to show the country-level aggregation result more directly. In particular:

$$\begin{aligned}
\mu \equiv \frac{\bar{r}}{\bar{w}} =& \frac{\left(\sum_m \bar{r}_m l_m\right)/\left(\sum_m l_m\right)}{\left(\sum_m \bar{w}_m l_m\right)/\left(\sum_m l_m\right)} \\
=& \frac{\sum_m \bar{r}_m l_m}{\sum_m \bar{w}_m l_m} \\
=& \frac{\sum_m \left(\frac{\bar{r}_m}{\bar{w}_m}\right)\bar{w}_m l_m}{\sum_m \bar{w}_m l_m} \\
=& \sum_m \left(\frac{\bar{r}_m}{\bar{w}_m}\right)\frac{\bar{w}_m l_m}{\sum_m \bar{w}_m l_m} \\
=& \sum_m \mu_m s_m \\
=& \sum_m s_m \left[1 + \frac{1}{\theta} HHI_m + \frac{1}{\eta}\left(1 - HHI_m\right)\right] \\
=& 1 + \frac{1}{\theta} H\tilde{H}I + \frac{1}{\eta}\left(1 - H\tilde{H}I\right)
\end{aligned}$$

□

B.2.5 Corollary 1: $\gamma_t \equiv \frac{d\mu_{mt}}{dX} = \left(\frac{1}{\theta} - \frac{1}{\eta}\right)\beta_t$

In this equation, β_t is the effect of an exogenous shock on the payroll Herfindahl. To derive the expression, plug in $\mu_{mt} \equiv 1 + \varepsilon_{mt}^{-1}$ and differentiate:

$$\begin{aligned}
\gamma_t \equiv \frac{d\mu_{mt}}{dX} =& \frac{d\left(1 + \varepsilon_{mt}^{-1}\right)}{dX} \\
=& \left[\frac{d\left(1 + \varepsilon_{mt}^{-1}\right)}{dHHI_{mt}} \cdot \frac{dHHI_{mt}}{dX}\right] \\
=& \left[\frac{d\left(1 + \varepsilon_{mt}^{-1}\right)}{dHHI_{mt}} \cdot \beta_t\right] \\
=& \left(\frac{1}{\theta} - \frac{1}{\eta}\right)\beta_t
\end{aligned}$$

I then compute standard errors for γ_t under the assumption that the effect on concentration and the labor supply parameters are independent. It follows that:

$$\begin{aligned}
\text{Var}\left(\gamma_t\right) =& \text{Var}\left[\left(\frac{1}{\theta} - \frac{1}{\eta}\right)\cdot \beta_t\right] \\
=& E\left[\left(\frac{1}{\theta} - \frac{1}{\eta}\right)^2\right] E\left[\beta_t^2\right] - \left[E\left(\frac{1}{\theta} - \frac{1}{\eta}\right)\right]^2 \left[E\left(\beta_t\right)\right]^2 \\
=& \left[\text{Var}\left(\frac{1}{\theta} - \frac{1}{\eta}\right) + \left[E\left(\frac{1}{\theta} - \frac{1}{\eta}\right)\right]^2\right]\left[\text{Var}\left(\beta_t\right) + \left[E\left(\beta_t\right)\right]^2\right] - \left[E\left(\frac{1}{\theta} - \frac{1}{\eta}\right)\right]^2 \left[E\left(\beta_t\right)\right]^2
\end{aligned}$$

whose components can all be plugged-in using sample estimates.

B.2.6 Equation 9 under the setup in BHM

I show that equation 9 holds under the additional assumptions on production function and goods market structure in Berger, Herkenhoff and Mongey (2022), henceforth BHM. In that environment, μ_m should be interpreted as the ratio of the average marginal revenue (net of expenditures in non-labor inputs) to the average wage. I show this in two steps.

To start, consider the environment in BHM. Goods markets are perfectly competitive, with $p_{zm} = 1$ for all firms and markets. Firms compete for labor à la Cournot, solving:

$$\max_{k_{zm}, l_{zm}} \pi_{zm} = \underbrace{A_{zm}\left(k_{zm}^{1-\gamma} l_{zm}^{\gamma}\right)^{\alpha}}_{\equiv y_{zm}} - Rk_{zm} - w_{zm}\left(\{l_{zm}, l_{-zm}\}\right) l_{zm} \tag{B.4}$$

where y_{zm} is firm revenues, k_{zm} is capital, A_{zm} is a general firm-market specific productivity term, R is the rental rate of capital (in perfectly competitive capital markets), and w_{zm} is the wage firm w_{zm} would have to pay to obtain l_{zm} units of labor, given nested CES labor supply preferences that yield the same expression for ε_{zm}^{-1}, the firm-specific inverse elasticity of residual supply, as derived in Section 2.1.

First, I show that equation 9 holds when the firm optimizes labor *holding capital constant*, denoting this corresponding average wage markdown by $\mu_m^{k-fixed}$. To avoid confusion due to differences in notation, let $mrpl_{zm}^{k-fixed}$ denote BHM's expression for the marginal revenue product of labor of firm z in market m *holding capital constant*, and continue to use the greek letter μ to denote the wage markdown.[50] Computing the definition of market m's average wage markdown holding capital constant gives:

$$\mu_m^{k-fixed} \equiv \frac{\bar{mrpl}_m^{k-fixed}}{\bar{w}_m} = \frac{\left(\sum_z mrpl_{zm}^{k-fixed} l_{zm}\right) / \sum_z l_{zm}}{\left(\sum_z w_{zm} l_{zm}\right) / \sum_z l_{zm}} \tag{B.5}$$

$$= \frac{\sum_z \alpha\gamma \left(y_{zm}/l_{zm}\right) l_{zm}}{\sum_z w_{zm} l_{zm}} \tag{B.6}$$

$$= \alpha\gamma \frac{\sum_z y_{zm}}{\sum_z w_{zm} l_{zm}} \tag{B.7}$$

where $mrpl_{zm}^{k-fixed} = \partial y_{zm} / \partial l_{zm}|_k = \alpha\gamma \left(y_{zm}/l_{zm}\right)$.

Simplification of equation B.7 can now be done using the equalities in BHM's Proposition 1.1 (with special

[50] In BHM, the greek letter μ refers to the wage take-home share (i.e., the inverse of the wage markdown) holding capital constant.

care given to note the difference in notation across the two papers). Equation B.7 becomes:

$$\begin{aligned}
\mu_m^{k-fixed} = \alpha\gamma \frac{\sum_z y_{zm}}{\sum_z w_{zm} l_{zm}} =& \alpha\gamma \left[\frac{1}{\alpha\gamma} \sum_z s_{zm} \mu_{zm}^{k-fixed} \right] \\
=& \sum_z s_{zm} \mu_{zm}^{k-fixed} \\
=& \sum_z s_{zm} \left[1 + \frac{1}{\theta} s_{zm}^2 + \frac{1}{\eta} \left(s_{zm} - s_{zm}^2 \right) \right] \\
=& 1 + \frac{1}{\theta} HHI_m + \frac{1}{\eta} \left(1 - HHI_m \right)
\end{aligned}$$

where HHI_m is similarly defined as the payroll Herfindahl of labor market m.

Second, I show that, in equilibrium, $\mu_{zm}^{k-fixed} = \mu_{zm}^{k-adjust}$ for all firms z and markets m. In other words, equation 9 holds whether or not optimization of capital is taken into account, *so long as expenditures on capital are netted out of firm revenues*. In this case, $\mu_{zm} = \mu_{zm}^{k-adjust}$ should be interpreted as the ratio of the marginal revenue (net of expenditures in non-capital inputs) product of labor to the wage.

To show this, note first that since $\mu_{zm} = mrpl_{zm}/w_{zm}$, it suffices to show that in equilibrium $mrpl_{zm}^{k-fixed} = mrpl_{zm}^{k-adjust}$. Letting y_{zm}^{net} denote total firm revenues net of capital expenditures, it follows that:

$$\begin{aligned}
mrpl_{zm}^{k-adjust} \equiv \frac{dy_{zm}^{net}}{dl_{zm}} =& \frac{d\left[f\left(k_{zm}, l_{zm}\right) - Rk_{zm} \right]}{dl_{zm}} \\
=& \frac{\partial f\left(k_{zm}, l_{zm}\right)}{\partial k_{zm}} \frac{dk_{zm}}{dl_{zm}} + \frac{\partial f\left(k_{zm}, l_{zm}\right)}{\partial l_{zm}} - R\frac{dk_{zm}}{dl_{zm}} \\
=& \frac{\partial f\left(k_{zm}, l_{zm}\right)}{\partial l_{zm}} + \frac{dk_{zm}}{dl_{zm}} \underbrace{\left(\frac{\partial f\left(k_{zm}, l_{zm}\right)}{\partial k_{zm}} - R \right)}_{=0 \text{ by firm's FOC for } k_{zm}} \\
=& mrpl_{zm}^{k-fixed}
\end{aligned}$$

This result, $mrpl_{zm}^{k-fixed} = mrpl_{zm}^{k-adjust}$, also follows directly from the envelope theorem, as the firm is optimizing its non-labor inputs.

C Extension: Oligopsony in Dual Labor Markets

C.1 Model

I model the labor supply preferences of *people seeking formal employment*. Firms exploit these preferences to set wages below workers' marginal revenue product when maximizing profits while holding local competitors' labor demand decisions constant. The starting point is the labor supply preference structure in Section C.1, which I extend on three dimensions. First, I split each local labor market into two sectors, wage work and self-employment. Second, I allow workers to take into account the probability of involuntary separation into informal wage work—which varies by firm—when making formal sector labor supply decisions. Finally, when bringing the model to the data, whenever possible I allow the model's key elasticities of substitution to vary by age, education, gender, and region.

C.1.1 Discrete choice formal labor supply in dual labor markets

There is a continuum of homogenous workers j, each choosing where to supply l^j effective units of labor.[51] Each worker chooses a single market m from a continuum of local labor markets. Within each market, workers choose a sector s, which is either wage work $\overline{g}$ or self-employment $\underline{g}$. When considering wage work, workers take into account expected earnings from a finite number of formal sector firms z, taking into account their perceived probability of separation into informal wage work o in market m. When considering self-employment, workers take into account expected earnings from starting a small business in market m.

Let l^j_{zgm} denote the effective units of labor that worker j supplies to employment option zsm.[52] Worker j chooses the employment option zsm that minimizes their dis-utility of labor given their preference parameters ξ, expected earnings $\mathbb{E}[l^j_{zgm} w_{zms} | X_{zgm}]$, and their minimum earnings requirement y^j:

$$\min_{zgm} V^j_{zgm} \equiv \ln l^j_{zgm} + \ln \xi_m + \ln \xi_{zgm} - \xi^i_{zgm} \tag{C.1}$$

$$\text{s.t.} \qquad l^j_{zgm} \bar{w}_{zgm} \geq y^j \tag{C.2}$$

where $\bar{w}_{zgm} = \mathbb{E}[w_{zms} | X_{zgm}]$ are expected wages from employment option zsm given firm, sector, and local labor market characteristics. The terms ξ_m, ξ_{gm}, and ξ_{zgm} are market, sector-market, and option-market taste shifters, respectively, while ξ^j_{zgm} is an idiosyncratic worker taste shock for option zsm, drawn from the

[51] I assume that workers are homogenous in both preferences and productivity to center the model discussion around wage differentials driven by worker preferences—the key underlying source of market power—which facilitates exposition and delivers clearer insights. However, both assumptions are easy to relax empirically. I relax the assumption of homogeneous elasticities of substitution in Section C.1, allowing them to vary by gender, education, and age, after presenting the regression equations that I use to estimate model's key elasticities of substitution. As in Section C.1's original model, the assumption of homogenous productivity is straightforwardly mapped to Brazil's rich microdata by focusing on wages conditional on flexible controls for worker gender, education, age, and worker fixed effects when possible.

[52] For formal sector options (i.e., $s = \overline{g}$), $z\overline{g}m$ represents a formal sector firm z in a market m's wage work sector. For self-employment options (i.e., $s = \underline{g}$), the z subscript refers to the worker (as each worker is their own employer).

following nested Generalized Extreme Value (GEV) distribution:

$$G(\xi) = \exp\Big(-\sum_m \Big\{\sum_s \Big(\sum_z e^{-(1+\tilde{\eta})\,\xi^j_{zgm}}\Big)^{\frac{1+\tilde{\rho}}{1+\tilde{\eta}}}\Big)^{\frac{1+\tilde{\theta}}{1+\tilde{\rho}}}\Big]^{\frac{1}{1+\tilde{\theta}}}\Big). \tag{C.3}$$

Equations C.1-C.3 preserve the structure of the original labor supply decision problem in Section C.1, but differ from it in two ways. First, they add self-employment as an alternative to wage work within each market, with elasticity of substitution $\tilde{\rho}$ between the two sectors, as shown in equation C.3. Second, they center workers' labor supply decisions around *expected earnings*.

Introducing earnings uncertainty is a natural way to incorporate both self-employment and informal wage work into the labor supply decisions of people seeking formal sector employment. For self-employment, workers don't know with certainty the returns to starting a small business ex-ante. For wage work, workers might care about the risk of involuntary separation from any formal sector job into informality (and related expected earnings losses)—whether it is inside or outside firms—when making labor supply decisions. I next derive the implications of these extensions to the labor supply curves facing each formal sector firm and, consequently, markdowns.

C.1.2 Labor supply curves faced by individual formal firms

As in the original model, the discrete choice formulation of workers' labor supply decisions with GEV idiosyncratic tastes yield analytic expressions for the probability that worker j chooses formal sector firm z in the wage work sector $\overline{g}$ of market m. Integrating these probabilities over the continuum of workers gives total labor supply to formal sector firm z:

$$l_{z\overline{g}m} = L\left(\frac{\bar{w}_{z\overline{g}m}}{W_{\overline{g}m}}\right)^{\tilde{\eta}}\left(\frac{W_{\overline{g}m}}{W_m}\right)^{\tilde{\rho}}\left(\frac{W_m}{W}\right)^{\tilde{\theta}} \times \left[\xi^{1+\tilde{\eta}}_{zgm}\xi^{1+\tilde{\rho}}_{gm}\xi^{1+\tilde{\theta}}_{m}\right]^{-1}. \tag{C.4}$$

whose inverse is:

$$\bar{w}_{z\overline{g}m} = W\left(\frac{l_{z\overline{g}m}}{L_{\overline{g}m}}\right)^{\frac{1}{\tilde{\eta}}}\left(\frac{L_{\overline{g}m}}{L_m}\right)^{\frac{1}{\tilde{\rho}}}\left(\frac{L_m}{L}\right)^{\frac{1}{\tilde{\theta}}}\xi^{1+\frac{1}{\tilde{\eta}}}_{z\overline{g}m}\xi^{1+\frac{1}{\tilde{\rho}}}_{\overline{g}m}\xi^{1+\frac{1}{\tilde{\theta}}}_{m}. \tag{C.5}$$

These are the labor supply equations for formal sector firms, and—as in the original model—equation C.5 is the key equation whose derivative with respect to formal sector firms' own employment decisions will determine wage markdowns. Importantly, the relevant wage for attracting labor supply to any one formal sector firm is the *expected* wage $\bar{w}_{z\overline{g}m}$ from taking that formal sector job, which takes into account probabilities of separation into informality and local informal sector conditions:

$$\bar{w}_{z\overline{g}m} = p_{z\overline{g}m}w^o_m + (1 - p_{z\overline{g}m})w_{z\overline{g}m} \tag{C.6}$$

where $p_{z\overline{g}m}$ is workers' prior about probability of separation from firm z into informal wage work, w^o_m are expected local informal sector wages, and $w_{z\overline{g}m}$ is firm z's posted formal wage.

In addition, because the elasticity of substitution between self-employment and wage work must necessarily be estimated using census data—which does not include firm boundaries but does include regions—it is useful to introduce the market-level aggregate of equation C.5, the inverse labor supply curve to wage work in each market:

$$W_{\bar{g}m} = W \left(\frac{L_{\bar{g}m}}{L_m}\right)^{\frac{1}{\tilde{\rho}}} \left(\frac{L_m}{L}\right)^{\frac{1}{\tilde{\theta}}} \xi_{\bar{g}m}^{1+\frac{1}{\tilde{\rho}}} \xi_m^{1+\frac{1}{\tilde{\theta}}}. \tag{C.7}$$

Finally, the wage and labor supply indices in equations C.4-C.7 follow the standard nested CES structure:

$$W_{\bar{g}m} = \left[\sum_z \left(\frac{\bar{w}_{z\bar{g}m}}{\xi_{z\bar{g}m}}\right)^{1+\tilde{\eta}}\right]^{\frac{1}{1+\tilde{\eta}}}, \qquad L_{\bar{g}m} = \left[\sum_z \left(\xi_{z\bar{g}m} l_{z\bar{g}m}\right)^{\frac{1+\tilde{\eta}}{\tilde{\eta}}}\right]^{\frac{\tilde{\eta}}{1+\tilde{\eta}}} \tag{C.8}$$

$$W_m = \left[\left(\frac{W_{\underline{g}m}}{\xi_{\underline{g}m}}\right)^{1+\tilde{\rho}} + \left(\frac{W_{\bar{g}m}}{\xi_{\bar{g}m}}\right)^{1+\tilde{\rho}}\right]^{\frac{1}{1+\tilde{\rho}}}, \qquad L_m = \left[\left(\xi_{\underline{g}m} l_{\underline{g}m}\right)^{\frac{1+\tilde{\rho}}{\tilde{\rho}}} + \left(\xi_{\bar{g}m} L_{\bar{g}m}\right)^{\frac{1+\tilde{\rho}}{\tilde{\rho}}}\right]^{\frac{\tilde{\rho}}{1+\tilde{\rho}}} \tag{C.9}$$

$$W = \left[\sum_m \left(\frac{W_m}{\xi_m}\right)^{1+\tilde{\theta}}\right]^{\frac{1}{1+\tilde{\theta}}}, \qquad L = \left[\sum_m \left(\xi_m L_m\right)^{\frac{1+\tilde{\theta}}{\tilde{\theta}}}\right]^{\frac{\tilde{\theta}}{1+\tilde{\theta}}} \tag{C.10}$$

where $W_{\underline{g}m}$ are expected earnings from self-employment in market m.[53]

C.1.3 Firm-specific wage markdowns

Taking logs of Equation C.5 and differentiating it with respect to $l_{z\bar{g}m}$—holding constant labor demand from competing formal firms, priors $p_{z\bar{g}m}$ about the probability of separation, expected returns to self-employment $W_{\underline{g}}$, and expected informal sector wages $\bar{w}_m^o$—gives:

$$\varepsilon_{z\bar{g}m}^{-1} \equiv \frac{\partial \ln \bar{w}_{z\bar{g}m}}{\partial \ln l_{z\bar{g}m}} = \frac{1}{\tilde{\eta}} + \left(\frac{1}{\tilde{\rho}} - \frac{1}{\tilde{\eta}}\right) \underbrace{\frac{\partial \ln L_{gm}}{\partial \ln l_{z\bar{g}m}}}_{s_{z\bar{g}m}} + \left(\frac{1}{\tilde{\theta}} - \frac{1}{\tilde{\rho}}\right) \underbrace{\frac{\partial \ln L_m}{\partial \ln l_{z\bar{g}m}}}_{s_{zm}} \tag{C.11}$$

where

$$s_{z\bar{g}m} \equiv \frac{\bar{w}_{z\bar{g}m} \cdot l_{z\bar{g}m}}{\sum_k \left(\bar{w}_{k\bar{g}m} \cdot l_{k\bar{g}m}\right)} \quad \text{and} \quad s_{zm} \equiv \frac{\bar{w}_{z\bar{g}m} \cdot l_{z\bar{g}m}}{\sum_s \sum_k \left(\bar{w}_{kgm} \cdot l_{kgm}\right)}$$

are formal sector firm z's *expected* wage bill as a fraction of the formal sector's *expected* wage bill (aka, taking into account separation probabilities into informality and local informal sector wages), and as a fraction of each market's overall *expected* wage bill (including self-employment as an alternative to wage

[53] In a nested CES preference structure, $W_{\underline{g}m}$ equals expected log earnings in the limit where the elasticity of substitution between self-employment options is zero. At the micro level, this means that no one worker can provide self-employment for another worker. At the macro level, this is the CES limiting case of Cobb-Douglas preferences over self-employment options. Formally, let $\eta_{\underline{g}}$ denote the elasticity of substitution between self-employment options and let $\mathbb{E}[\pi_{\underline{g}m}^j]$ denote worker j's expected earnings from self-employment in market m. Then $W_{\underline{g}m} = \left(\int_0^1 \mathbb{E}[\pi_{\underline{g}m}^j]^{1+\eta_{\underline{g}}} dj\right)^{1/(1\eta_{\underline{g}})}$ is the CES aggregator over self-employment options in a given market. When $\eta_{\underline{g}} \to 0$, we have $W_{\underline{g}m} \to \exp\left(\int_0^1 \log \mathbb{E}[\pi_{\underline{g}m}^j]\, dj\right)$, the Cobb-Douglas form. Note that $W_{\underline{g}m} > 0$ because workers must satisfy their minimum earnings requirement in equilibrium.

work), respectively.[54] Rearranging:

$$\varepsilon^{-1}_{z\overline{g}m} = \left[\frac{1}{\tilde{\rho}} s_{z\overline{g}m} + \frac{1}{\tilde{\eta}}(1 - s_{z\overline{g}m})\right] + \left(\frac{1}{\tilde{\theta}} - \frac{1}{\tilde{\rho}}\right) s_{zm}. \tag{C.12}$$

Equation C.12 differs from its counterpart in Section C.1 in three ways. First, it states that the elasticity of labor supply to any one formal sector firm depends not only on the ease with which workers can substitute between firms within markets $\tilde{\eta}$, and between markets $\tilde{\theta}$, but also on the ease with which they can substitute between wage work and self-employment within markets $\tilde{\rho}$. Second, the relevant wage bill shares for the markdowns are based on expected wages—taking into account probabilities of separation into wage work and local informal sector wages.

Local average wage markdown. As in the original model, the average wage markdown is $\mu_{\overline{g}m} \equiv 1 + \varepsilon^{-1}_{\overline{g}m}$ and the wage take-home share is $\mu^{-1}_{\overline{g}m}$. Taking a weighted average of the firm-specific inverse elasticity of substitution in Equation C.12 using $s_{z\overline{g}m}$ as weights gives the expression for the inverse elasticity of residual labor supplied to each firm:

$$\bar{\varepsilon}^{-1}_{\overline{g}m} = \frac{1}{\tilde{\rho}} HHI_{\overline{g}m} + \frac{1}{\tilde{\eta}}(1 - HHI_{\overline{g}m}) + \left(\frac{1}{\tilde{\theta}} - \frac{1}{\tilde{\rho}}\right) HHI_{\overline{g}m} \cdot s_m. \tag{C.13}$$

where $HHI_{\overline{g}m} \equiv \sum_{z \in \Gamma_{\overline{g}m}} s^2_{z\overline{g}m}$ is the Herfindahl-Hirschman Index for wage work sector $\overline{g}$ in market m, measuring the wage bill concentration in sector $\overline{g}$ in market m, and

$$s_m \equiv \frac{\sum_{z \in \Gamma_{\overline{g}m}} s_{z\overline{g}m}}{\sum_{z \in \Gamma_m} s_{z\overline{g}m}}$$

is the wage bill share of wage work sector $\overline{g}$ in market m.

No dynamic gains from influencing workers' priors. Importantly, these theoretical results assume that firms do not *internalize* that their hiring and firing decisions at time t might influence workers' priors about the probability of separation from the firm at $t + 1$ or after. This assumption of *no dynamic gains from formation of priors* simplifies firms' wage setting problem by shutting off $p_{z\overline{g}m}$ as an additional lever for (future) wage setting. As a result, it preserves the original model's static structure while allowing priors to enter labor supply decisions. While workers may form priors based on past firm-level involuntary separations (observed in the administrative data) and act accordingly, neither workers nor firms consider *future* benefits or costs of labor market decisions made at t. The possibility of dynamic gains to labor market power from firms' ability to influence worker priors about the possibility of separation from formation of priors (and about other within-firm dynamic considerations, such as promotion timelines, tenure benefits, etc) merits further research.

[54]The equivalence between the partial derivatives and market shares can be shown by computing the partial derivatives given the definition of labor supply indices, and then contrasting this result with what we obtain when plugging the inverse labor supply equation into the definition of wage bill market shares.

C.2 Estimation strategy

C.2.1 Substitution Within markets, across firms

I next show how $1/\tilde{\eta}$ can be estimated by combining the IV estimate for $1/\eta$ in Section C.1 with an estimate of the bias introduced into that estimation strategy if the extended model is true instead. That is, under the assumption that workers take into account the probability of involuntary separation into informality when making formal sector labor supply decisions. First, re-arrange the expected wage equation we have:

$$\bar{w}_{z\bar{g}m} = p_{z\bar{g}m} w_m^o + (1 - p_{z\bar{g}m}) w_{z\bar{g}m} = w_{z\bar{g}m} \left(1 + p_{z\bar{g}m}\sigma_{z\bar{g}m}\right) \tag{C.14}$$

where $\sigma_{z\bar{g}m} \equiv (w_m^o - w_{z\bar{g}m})/w_{z\bar{g}m}$ is the wage gap between the local informal sector wage w_m^o (for observationally equivalent workers) and firm z's formal wage. Note that this expected wage gap may be negative (informal sector pays less) or positive (informal sector pays more). Taking logs:

$$\ln \bar{w}_{z\bar{g}m} = \underbrace{\ln w_{z\bar{g}m}}_{\text{Formal wage}} + \underbrace{\ln\left(1 + p_{z\bar{g}m}\sigma_{z\bar{g}m}\right)}_{\text{Expected wage gap}} \tag{C.15}$$

Next, take logs of Equation C.5 and let δ_m denote a market fixed effect to write:

$$\ln \bar{w}_{z\bar{g}m} = \frac{1}{\tilde{\eta}} \ln l_{z\bar{g}m} + \delta_m + \ln \xi_{z\bar{g}m}^{1+\tilde{\eta}} \tag{C.16}$$

Bias formula. Given equation C.16, the within-market cross-firm inverse elasticity of substitution $\frac{1}{\tilde{\eta}}$ in the extended model can estimated as the second-stage coefficient from an IV regression where $\ln l_{z\bar{g}m}$ is instrumented by an exogenous labor demand shock $X_{z\bar{g}m}$. Expanding expected wages per equation C.15 gives and letting $(*)$ denote partialled-out variables to apply the Frisch-Waugh-Lovell theorem, we have:

$$\frac{1}{\tilde{\eta}} = \frac{\text{Cov}(\ln \bar{w}^*_{z\bar{g}m}, X_{z\bar{g}m})}{\text{Cov}(\ln l^*_{z\bar{g}m}, X_{z\bar{g}m})} = \underbrace{\frac{\text{Cov}(\ln w^*_{z\bar{g}m}, X_{z\bar{g}m})}{\text{Cov}(\ln l^*_{z\bar{g}m}, X_{z\bar{g}m})}}_{\frac{1}{\eta}\text{: Supply response to formal wage}} + \underbrace{\frac{\text{Cov}[\ln(1 + p_{z\bar{g}m}\sigma_{z\bar{g}m})^*, X_{z\bar{g}m}]}{\text{Cov}(\ln l^*_{z\bar{g}m}, X_{z\bar{g}m})}}_{\Omega\text{: Supply response to expected wage gap}} \tag{C.17}$$

where $\sigma_{z\bar{g}m} \equiv (w_m^o - w_{z\bar{g}m})/w_{z\bar{g}m}$ is the pay gap between the local informal sector and firm z's formal wage for an effective (namely, equally productive) unit of labor.

The first term in equation C.17—the labor supply response to a firm's change in its formal wage—equals the IV estimate for the within-market cross-firm elasticity of substitution $1/\eta$ in the original model in Section C.1. If the original model were true, workers do not take into account the possibility of separation into informality when making labor supply decisions and, as a result, the second term in equation C.17 is zero and the elasticities in the extended and original models are the same.

However, *if the extended model is true*, then the relevant elasticity of substitution for wage setting is $1/\tilde{\eta}$. The term Ω in equation C.17 is thus the misspecification bias in the original model if the extended model is true. This bias is zero if there is no within-market cross-firm variation in perceived probabilities of involuntary separation into the informal sector at the time workers make decisions. If the bias is non-zero, its sign depends not only on local informal sector wages, but also on the joint distribution (across formal sector firms in the same market) between perceived probabilities of separation and formal sector wages.

Interpretation. If Ω is *negative*, *less* labor is supplied to firms with higher expected wage gaps relative to the same market's informal sector wage. That is, workers dislike firms that in expectation pay less than the local informal sector.[55]. Conversely, if Ω is *positive*, *more* labor is supplied to firms with higher expected wage gaps relative to the same market's informal sector wage. That is, workers prefer to supply labor to a formal firm *despite* higher expected wages in the local informal sector. In this case, the inverse elasticity of substitution $1/\eta$ estimated based on the original model is *too small* relative to the true inverse elasticity $1/\tilde{\eta}$, meaning that *firms have more labor market power* than originally estimated.

Regression. I estimate Ω, the labor supply response to changes in expected formal-informal wage gaps *conditional on firm-level changes in formal wages*, by running the following second stage regression:

$$\Delta \ln(1 + p_{z\overline{g}r}\sigma^{c}_{z\overline{g}r}) = \Omega\Delta \ln l^{c}_{z\overline{g}r} + \delta_z + \delta_r + \delta_c + \epsilon_{z\overline{g}r} \tag{C.18}$$

where the left-hand side are 1991-2000 changes in expected wage gaps for each demographic cell, the right-hand side are 1991-2000 changes in employment at firm z in microregion r for demographic cell c, and the δ terms are firm, region-cell fixed effects. Including a firm fixed effect ensures that Ω captures the supply response to changes in expected wage gaps relative to the informal sector conditional on firm-level changes in wage levels, the identifying variation for η in both the original and extended models.

Measurement. Brazil's employer-employee linked dataset RAIS includes information on separation reason for all separations, discerning between firings and quits. I use this information to proxy for workers' prior about the *probability of involuntary separation* from firm z in microregion r using:

$$p_{z\overline{g}r} = \frac{\text{Fired workers from firm } z\overline{g}r}{\text{Fired workers from + Stayers in firm } z\overline{g}r} \tag{C.19}$$

Since firm employment is measured as of December 31 of each calendar year, I measure priors for year t as of December 31 of year t using separations during that elapsed calendar year. For example, priors as of December 31, 1991 are based on separations between January 1, 1991 and December 30, 1991. Finally, to ensure that the formal-informal wage gaps I measure are not confounded by productivity differences across demographics, I calculate $\sigma_{z\overline{g}m}$ separately for each demographic cell, following the definitions as in the self-employment analysis. I merge RAIS' data for 1991 and 2000 with the census those years at the microregion and demographic cell level, matching on gender, education group, and age group dummies.

Panel A of Figure C.4 plots a binned scatter of firm-level probabilities of involuntary separation distribution against informal-formal wage gaps. The probability is on average 3%. It is highest in local labor markets

[55] The possibility that informal sector options might pay more than formal sector options should not be discarded. Appendix Figure C.4 shows that a non-trivial share of regions exhibit larger monthly earnings for workers of the same demographic composition than monthly earnings at formal sector firms

where the informal sector pays less than the formal sector (for the same demographic cell). This means that workers who are at higher risk of separation would also face lower-paying informal sector jobs. Panel B of Figure C.4 shows that a substantial number of jobs are in local labor markets where the informal sector pays less, but there is a sizable number of jobs for which informal wage work pays higher wages than the formal sector. This is consistent with the residual wage distributions plotted in Figure C.3.

Instruments. Since Ω is a labor supply parameter governing within-market cross-firm substitution in response to formal-informal wage gaps, its identification requires labor demand shocks that vary: (a) across firms within markets, as in the identification of η in the original model; and (b) *within* firm relative to the informal sector, since Ω captures labor supply responses to changes in wage gaps conditional on changes in formal wage levels, captured by η instead. I leverage demographic heterogeneity within firm to construct such an instrument as the interaction between firm-level changes in import tariff reductions $\Delta \ln(1 + \tau_{i(z)})$—the labor demand used to identify η in the original model—and demographic-specific regional tariff reductions ΔRTR_m^c, the labor demand shock used to identify the elasticities of substitution between wage work and self-employment in the extended model.

The key idea is that, conditional on firm-level changes in formal wages, which identify η, the remaining within-market cross-firm labor supply reallocation is driven by cross-firm differences in expected wage gaps relative to the informal sector. Expected wage gaps are, in turn, affected both by the direct import tariff reduction shock to the firm—which change workers' perceived probability of separation from any one firm into informality over the decade—and by regional tariff reductions, which change equilibrium informal sector wages for each demographic group. The exclusion restriction assumes that, conditional on firm, region, and demographic group fixed effects, the only way labor supply is affected by tariff reductions is by changing separation probabilities and equilibrium wages, as opposed to workers' distaste for individual firms or for formal versus informal wage work.

Unemployment insurance. Finally, to test how the prospect of receiving unemployment insurance in the event of involuntary separations affect workers' formal sector labor supply decisions, I consider an alternative measure of the expected wage. I do this by measuring the expected wage after separation as 4/12 the same wage as paid by the formal firm (i.e., the typical unemployment insurance benefit at the time) and 8/12 the local informal sector wage. [56]

C.2.2 Substitution to self-employment and across markets

To arrive at the regression equation to estimate the elasticity of substitution between wage work and self-employment, take logs of equation C.7, express it in long-differences, add a constant to absorb country-level

[56] While Brazil's regulation of unemployment insurance benefits in the 1990s were quite complicated—sometimes featuring earnings brackets to limit benefit amounts, and being contingent on proof of time of service—the typical benefit amount was four-months salary, paid monthly. See Brasil (1990) for the 1990 law that instituted the unemployment insurance system. Details on benefit brackets and how the rules have changed over time are available at https://www.debit.com.br/tabelas/seguro-desemprego?utm_source=chatgpt.com.

changes, and group taste shifters into an error term $\epsilon_{\overline{g}m}$ to get:

$$\Delta \ln W_{\overline{g}m} = \alpha + \frac{1}{\tilde{\rho}} \Delta \ln\left(\frac{L_{\overline{g}m}}{L_m}\right) + \frac{1}{\tilde{\theta}} \Delta \ln\left(\frac{L_m}{L}\right) + \epsilon_{\overline{g}m} \tag{C.20}$$

A challenge in estimating equation C.20 in the context of trade liberalization is that each labor market has a single wage work sector, and import tariff reductions were primarily a labor demand shock to wage work. As a result, $\tilde{\rho}$ cannot be separately identified from $\tilde{\theta}$ unless markets can be partitioned into multiple sub-markets,[57] each with their own wage work and self-employment sectors and their own differential shock to wage work labor demand (the key within-market, cross-sector variation needed to identify $\tilde{\rho}$),[58] but all subject to the same market-level general equilibrium effects of trade liberalization (the key cross-market variation needed to identify $\tilde{\theta}$).

Regression. I overcome this challenge by partitioning labor markets into worker demographic cells, each with their own wage work and self-employment sectors. Let g denote demographic groups defined by gender (men; women), education (primary; secondary; tertiary), and age (young: ages 18-29; middle: ages 30-49; old: ages 50-64) and c denote the fully saturated demographic cells defined by these three groups (e.g., young men with at most primary education). I then estimate:

$$\Delta \ln W^c_{\overline{g}m} = \frac{1}{\tilde{\rho}} \Delta \ln\left(\frac{L^c_{\overline{g}m}}{L^c_m}\right) + \frac{1}{\tilde{\theta}} \Delta \ln\left(\frac{L^c_m}{L^c}\right) + \delta_r + \epsilon^c_{\overline{g}m} \tag{C.21}$$

where Δ denotes long-differences within cell by microregion pairs and δ_r is a region fixed effect (denoting one of Brazil's major 5 regions) to absorb regional general equilibrium effects of trade liberalization.

Measurement. I estimate regression equation C.21 using data from Brazil's 1991 and 2000 censuses and defining markets as microregions.[59] While $W^c_{\overline{g}m}$, $L^c_{\overline{g}m}$, and L^c_m are technically wage and labor supply indices that reflect lower-nest elasticities and taste shifters, to keep the analysis entirely executable with census data I use the corresponding directly observable measures of these objects. Specifically, I measure $L^c_{\overline{g}m}/L^c_m$ as demographic cell c's wage work employment share in microregion m, and $\ln W^c_{\overline{g}m}$ as demographic cell's c's log residual real monthly earnings in microregion m . Log real monthly earnings are residualized conditional on a fully saturated vector of the gender, education, and age variables, so that the variation in wages used to estimate the model's preference parameters are not driven by productivity differences across demographic cells.

[57] Note that identification of lower-nest elasticities also requires that each partition of the wage work sector has within-sector cross-firm variation in tariff reductions, which precludes partitioning the wage work sector by industry.

[58] An alternative approach is to introduce additional cross-market variation in import tariff reductions, such as those generated by interactions between regional tariff reductions and baseline market characteristics. I explore this approach in Appendix Table C.7, where I instrument equation C.20 with regional tariff reductions and its interaction with: (a) each microregion's log maximum distance to the nearest labor law enforcement office, borrowing the estimation strategy from Ponczek and Ulyssea (2022); and (b) each microregion's 1991 share of employment that was formal. The approach lacks power but yields similar point estimates, especially for $\tilde{\theta}$, relative to my preferred specification, which adds within-market cross-group variation to better identify $\tilde{\rho}$.

[59] While the 1991 and 2000 censuses contain occupation codes, they are missing for a large share of observations, and the two years also follow different definitions from each other (and from RAIS). As a result, I define labor markets more broadly as microregions only for the purposes of estimating the extended model's elasticities of substitution between self-employment and wage work and across markets.

Instruments. I instrument the two endogenous variables in equation C.21 with cell-specific regional tariff reductions ΔRTR_m^c and their interaction with: (a) each cell's 1991 formal sector employment share, as tariffs hit formal firms most directly; and (b) demographic group dummies, to soak up variation and allow for heterogeneous effects by groups. The primary instrument ΔRTR_m^c is calculated following the shift-share methodology in Dix-Carneiro and Kovak (2017), featuring exposure shares that are cell-industry-microregion-specific and include *all* workers—formally or informally employed—in tradable sector industries. The exclusion restriction assumption is all cell-specific regional tariff reductions and their interactions with baseline and demographic characteristics are orthogonal to changes in workers' tastes-shifters ξ for wage work and for the market.

Figure C.7 shows that regional tariff reductions reduced *both* employment and earnings in wage work. This same-sign effect is consistent with tariff reductions tracing out labor supply to wage work. The same shock had effects of opposite signs for self-employment: regional tariff reductions *increased* the number of workers in self-employment in more affected regions but, on average, *reduced* their earnings. These opposite-sign effects are consistent with the release of workers from wage work constituting a labor supply shock to the self-employment sector.

Appendix Table C.7 shows a similar pattern for informal wage work as for self-employment, accompanied by a near-zero effect on formal sector residual real monthly earnings. The latter finding differs from the significantly *negative* effects of regional tariff exposure documented based on RAIS-reported earnings in prior literature (e.g., Dix-Carneiro and Kovak (2017), Section C.1). Informality within firms, as studied by Ulyssea (2018) and, more recently, Derenoncourt et al. (2025), is a strong candidate explanation for this near-zero effect. If higher-wage formal sector workers are moved "off the books," they disappear from RAIS,[60] but they might still report that they are formally employed in the Census, as their job is the same as before. This induces a positive selection into the set of workers marked as formally employed in the Census, even if they are in fact informal. In the absence of data that can discern the intensive margin and extensive margin of informality during this period, it is hard to make empirical progress on elasticities of substitution between informal and formal wage work.

C.2.3 Heterogeneous preferences by demographics

An advantage of partitioning markets by demographic cells—and in constructing cell-specific regional tariff reductions— is that it is the first step in estimating demographics-based heterogeneity in elasticities of substitution. I estimate heterogeneity in the elasticity of substitution between wage work and self-employment by demographic groups as follows:

$$\Delta \ln W_{gm}^c = \frac{1}{\tilde{\rho}^g} \Delta \ln \left(\frac{L_{gm}^c}{L_m^c} \right) + \delta_m + \epsilon_{gm}^c \tag{C.22}$$

[60]Moving a worker off the books–potentially tagging their leaving as a quit so that no firing costs are paid—saves the firm nearly 100% the worker's wage in the labor costs that are now evaded, even if the worker keeps the same take-home earnings. See Debaere (2003) for a table listing non-wage labor costs to the firm.

where δ_m is a market fixed effect, absorbing market-level general equilibrium effects of trade common to all demographic cells. I use the same cell-level instruments as for equation $C.21$. The exclusion restriction assumption is that the shocks to wage work labor demand are orthogonal to workers' tastes-shifters ξ for wage work in that market. Unfortunately, there is not enough within-group cross-cell variation in tariff reductions to estimate equation $C.22$ separately by demographic groups, which would yield estimates of heterogeneity in the cross-market elasticity of substitution $\tilde{\theta}$.

If exploited by firms when setting wages, such heterogeneity could generate differential wage markdowns by demographics. Thus, partitioning labor markets by demographic groups effectively extends the original model in Section C.1 in a third direction, contributing the first labor market oligopsony framework inclusive of self-employment and informal wage work to a vast literature on wage differentials by demographics, where evidence on the link between wage markdowns and gender wage gaps has begun to emerge (e.g., see Sharma (2023), Hoang, Mitra and Pham (2024)).

C.2.4 Wage markdowns

I combine estimates of the extended model's parameters with direct measures of labor market equilibrium wage bill shares to compute local labor market m's average wage markdown according to the extended model. The markdown is given by:

$$\mu_{\overline{g}m} \equiv 1 + \varepsilon_{\overline{g}m}^{-1} = \left[\frac{1}{\tilde{\rho}} HHI_{\overline{g}m} + \frac{1}{\tilde{\eta}}(1 - HHI_{\overline{g}m})\right] + \left(\frac{1}{\tilde{\theta}} - \frac{1}{\tilde{\rho}}\right) HHI_{\overline{g}m} \times s_m. \tag{C.23}$$

where $1/\tilde{\eta}$ is calculated as $1/\tilde{\eta} = 1/\eta + \Omega$ using estimates of the bias Ω introduced by expected involuntary separation into informality, $s_{z\overline{g}m}$ and s_{zm} are formal sector firm z's *expected* wage bill as a fraction of the formal sector's *expected* wage bill, and as a fraction of each market's overall *expected* wage bill (including self-employment as an alternative to wage work), respectively. Table **??** summarizes the model parameters and key endogenous objects—which act as weights of these preference parameters in computing workers inverse elasticity of residual labor supply—along with their main estimates and ranges based on detected heterogeneity.

C.2.5 Policy effects on wage markdowns

Let X_m denote an exogenous shock to labor demand in local labor market m. Then this shock's effect on local markdowns is given by its effect on average local inverse elasticities of labor supply, which follows:

$$\frac{\partial \bar{\varepsilon}_{\overline{g}m}^{-1}}{\partial X_m} = \left(\frac{1}{\tilde{\rho}} - \frac{1}{\tilde{\eta}}\right) \times \beta_{\text{HHI}} + \left(\frac{1}{\tilde{\theta}} - \frac{1}{\tilde{\rho}}\right)\left[s_m \times \beta_{\text{HHI}} + HHI \times \beta_{s_m}\right] \tag{C.24}$$

where HHI (short for $HHI_{\overline{g}m}$) market's formal sector's expected wage bill Herfindahl Index, s_m is market m's wage work wage bill share (that is, formal plus informal wage work divided by total work, which includes self-employment), and the β coefficients are the partial (aka, first derivative, reduced form) effects of the labor demand shock on Herfindahl and wage work shares, respectively.

C.3 Results

C.3.1 Elasticities of substitution

Substitution within markets, across firms. Table C.2 reports IV estimates of the within-market cross-firm inverse elasticity of substitution, $1/\eta$, and its regional heterogeneity. Column (1) replicates the 0.990 estimate in Section C.1, which is the relevant within-market cross-firm elasticity of substitution in Section C.1 if workers are indifferent to the threat of involuntary separation to informal wage work (namely, if $\Omega = 0$). The near-unit value of this elasticity is nearly 7 times higher than its corresponding value for the United States, and remains—as in Section C.1—the leading driver of Brazil's high markdown levels.

This near-unit and relatively inelastic value is not homogeneous across Brazil, however. Columns (2)-(5) report heterogeneity by Brazil's major regions. Column (2) shoes that the pooled elasticity is driven by the Southeast— the largest region by employment and where the formal sector made up more than half of total employment share in 1991, as shown in the microregions to the left of Figure 1—where the within-market cross-firm elasticity of substitution is estimated at 1.029.

The picture is very different in the Northeast, Brazil's second largest region by population and employment. Column (2) shows that the within-market cross-firm elasticity of substitution in the Northeast is nearly half of that in the Southeast, at 0.458. Since nearly all employment in the Northeast is informal—these are primarily the microregions to the right of Figure 1, with the exception of a few metropolitan areas (e.g., PE-3 and AL-3), where formality is high—more elastic within-market cross-firm elasticity of substitution in this region suggests raises the interesting possibility of congestion—or larger employer-employee matching frictions within the formal sector—in markets where many more firms are formal and small. Columns (4) and (5) report coefficients pooling with other regions where the first stage of Brazil's firm-level tariff shocks is too weak to yield estimates restricted to the smaller regions of the North, South, and Center West. I pool the North with the Northeast given their similarity in levels of development, and the South and Center West with the Southeast, for similar reasons. Both estimates are attenuated when pooled.

Bias from separation into informal wage work. Panel A of Figure C.8 plots the reduced form and first stage relationships that identify Ω. They are both negative: employment and expected informal-formal wage gaps declined in firms and markets more exposed to import competition. Combined these imply that Ω is positive, such that the threat of involuntary separation from a formal job increases firms' market power. For ease of interpretation, Panel B of Figure shows the OLS relationship between the two variables. The OLS says that firms that grow are also those that distance themselves–in terms of earnings–from local labor market conditions. Some of these firms pay more than local informal jobs, but many pay less, as shown in Panel B of Figure C.3. Table C.4 reports the implied IV coefficients and its heterogeneity by demographic groups.

Table C.4 examines the bias term Ω in the within-market cross-firm elasticity of substitution presented in Table C.2 if the extended model is true, namely, if workers take into account the probability of involuntary separation into informal wage work when making labor supply decisions. The most important take-away is that the magnitude of the bias is very small. At 0.0331 on average per Column (2), Ω shows that the within-market cross-firm inverse elasticity of substitution is roughly 3% *smaller* in Section C.1's model than in the extended model where involuntary separations are considered. In other words, the threat of involuntary separation into informal wage work slightly increases formal firms' labor market power. Columns (2)-(4) then show some heterogeneity in this estimate by demographics, though the magnitudes are not significantly different from each other.

Unemployment insurance. An interesting question is whether the value of unemployment insurance alters workers' attitude towards being fired from a formal sector job. Appendix Table C.4 tests this hypothesis. Consistent with the cross-country evidence Amodio et al. (2025), I find that unemployment insurance makes workers less responsive to the threat of involuntary separation, curbing labor market power. This results from contrasting the magnitude of the Ω bias introduced by involuntary separations if it is estimated with or without the value of unemployment insurance benefits, available in Appendix Table C.5. When Ω is positive, the threat of involuntary separation to the informal sector increases firms' labor market power. When it is negative, the informal sector curbs that market power. While all estimates I find for Ω are positive, adding unemployment insurance to the estimation reduces Ω from 0.0331 to 0.0221. This suggests that unemployment insurance, on average, operates like a substitute to self-employment at Brazil's level of economic development. This is consistent with the findings in Amodio and de Roux (2021). The elasticity of substitution to self-employment is however a much stronger force in curbing labor market power than the contribution of unemployment insurance to its relevant elasticity ($1/\eta$).

Substitution to self-employment. Tables C.6 and C.8 report estimates of inverse elasticities governing substitution between wage work and self-employment, $1/\tilde{\rho}$, as well as cross-market substitution, $1/\tilde{\theta}$. The magnitudes are generally small, indicating highly elastic reallocation between wage work and self-employment and across markets, which limits equilibrium labor market power. Heterogeneity results show that substitution is particularly elastic for demographic groups with stronger outside options, reinforcing the role of self-employment and informal work as key competitive constraints on formal-sector wage setting.

Appendix Figure C.6 plots the first stage of equation C.21—namely, the effect of demographic-specific regional tariff reductions on log share of wage work employment—separately by demographic groups. These figures are demographic-specific versions of Figure C.5, which plots the pooled, across groups, relationship between wage work and regional tariff reductions. The variation in Appendix Figure C.6 is the one I use to estimate the elasticities of substitution, reported in Tables C.6 and C.8.

C.3.2 Wage markdowns and key take-aways

On average, Brazilian formal sector workers were paid 51 cents of their marginal revenue product of labor in 1991. However, this average masks substantial heterogeneity across regions and across demographics. Take-home shares are highest for less educated workers, for whom the minimum wage is a substantial increase relative to their earnings in either informal wage work or self-employment, and for women, for whom self-employment earnings are higher than either formal sector earnings or informal sector earnings at any point of the formal wage take-home share distribution (see, e.g., Figure C.15).

Figure C.9 shows the distributions of average wage take-home shares—the inverse of wage markdowns—

across microregions and across occupations, since a local labor market is defined as a microregion x 2-digit occupation pair. The dispersion is driven by regional heterogeneity in the within-market cross-firm inverse elasticity of susbtitution $1/\eta$. Most microregions have wage markdowns around 50 cents on the dollar, with that mass increasing between 1991 and 2000. Microregions in the North and Northeast of Brazil, for which within-market cross-firm elasticities of substitution are almost half those of the rest of the country, feature higher wage take-home shares. The dispersion across occupations is centered between 50 and 55 cents on the dollar.

Figure 9 displays the microregion-level average markdowns by microregion, along with each microregion's share of employment in formal wage work, informal wage work, and self-employment. A clear pattern is that wage take-home shares are typically higher, though not always, in places where a larger share of the workforce is self-employed. Regions that are primarily formal, such as microregions in the states of São Paulo or Santa Catarina, show little dispersion of average wage take-home shares, though variation still exists as places differ in demographic composition, and demographic groups have different elasticities of substitution to self-employment and respond differently to the threat of separation into informal wage work. Appendix Figures C.13 through C.15 replicate these figures but showing demographic-specific local wage markdowns.

It should also be noted that a substantial share of the earnings gaps between the informal, informal, and self-employment sector is accounted for by the demographics of the workers, as shown in Figure C.3. In fact, the 1991 residual earnings distribution shows that observationally-equivalent workers were paid on average higher earnings in either the informal wage work sector or in self-employment. By 2000, the residual earnings gap between the formal and informal wage work sectors closes, whereas self-employment–primarily composed of women—remained a higher return activity on average.

Wage markdowns and concentration. Since the model in Section C.1 features ample heterogeneity in preference parameters and implies that local labor market conditions outside of the formal sector affect firms' wage setting, the relationship between local labor market concentration and wage markdowns is no longer a linear function of formal sector concentration, as in Section C.1's framework. The relationship is now tied to many more variables, including cross-market variation in demographic composition.

Figure C.11 and its demographic-specific versions (Appendix Figures C.12) plot the relationship between the formal sector's local Herfindahl-Hirschmann Index (HHI)–measured with respect to expected wages—and wage take-home share estimates. It shows a non-linear relationship likely driven by cross-regional differences in elasticities of substitution. Wage take-home shares increase in labor market concentration for markets with numerous similarly sized firms (i.e., HHI below 0.2). Past that threshold, labor market concentration among formal sector firms typically reduces formal sector wage take-home shares.

C.3.3 Policy effect of trade liberalization on wage markdowns

I apply the policy effect formula to compute the effect of a 10% increase in Regional Tariff Reductions (RTR) on average wage markdowns. Table 8 reports the country-average effect and heterogeneity across regions and demographics. All effects are in percentage points (pp) of the take-home share $\mu = w/MRP$. A positive

$\Delta\mu$ means wages moved closer to MRP—reducing employer market power—while a negative $\Delta\mu$ means the wedge widened.

Reduced-form effects of RTR on market structure. The policy effect formula combines the structural elasticities with reduced-form estimates of how trade changed market structure. A unit RTR increase raises formal employer concentration by $\beta_{HHI} = 0.722$ (SE = 0.075), so a 10% RTR increase raises HHI by 7.2 percentage points—large relative to the baseline mean of 16.7%. A unit RTR increase also decreases the formal sector's expected wage bill share by $\beta_{s_r} = -0.477$ (SE = 0.069), so a 10% increase reduces s_r by 4.8 percentage points from a baseline of 78.5%. Both coefficients come from regressions $\Delta Y_m = \alpha + \beta \cdot RTR_m + \gamma_r + u_m$, where ΔY_m is the 1991–2000 change in HHI or s_r, γ_r are region fixed effects, and standard errors are clustered at the microregion level. The HHI regression is estimated at the market (mmc × cbo942d) level with formal wagebill weights; the s_r regression at the mmc level since s_r varies only across microregions. For informality and region subgroups, both coefficients are re-estimated within each subgroup; for demographic subgroups, pooled coefficients are used.

Average effect: two opposing channels. These reduced-form effects enter the markdown formula through two opposing channels. The *within-market net substitution channel* (−0.0017) reduces markdowns: as trade increases concentration, the formula shifts weight from the less elastic cross-firm margin ($1/\tilde{\eta}$) toward the more elastic wage work–self-employment margin ($1/\tilde{\rho}$), so self-employment disciplines employers even as formal markets become more concentrated. The *cross-market substitution channel* (+0.0058) increases markdowns: as the formal sector shrinks and concentrates, the HHI–s_r interaction shifts weight toward the least elastic cross-market margin ($1/\tilde{\theta}$), so remaining formal employers face workers with fewer cross-market options. At the country average, these channels nearly offset: $\partial\varepsilon^{-1}/\partial RTR = +0.004$, translating into an essentially zero change in the take-home share from a baseline of 50.9%.

Heterogeneity. The near-zero country average masks modest heterogeneity driven by variation in local informality and labor supply elasticities. In high-informality regions—where self-employment is prevalent and baseline concentration is high ($HHI = 0.43$ vs. 0.16 in low-informality regions)—the take-home share falls by −0.54 pp. In low-informality regions, the effect is a small increase (+0.04 pp). Regional heterogeneity is also modest: the take-home share rises by +1.18 pp in the North and Northeast, while in the Southeast, South, and Center-West the effect is −0.06 pp.

Across demographics, all effects are close to zero. Tertiary-educated workers ($1/\tilde{\rho} = 0.93$, $\Delta\mu = +0.2$ pp) and young workers (0.85, +0.1 pp) see the largest—though still small—positive effects. Middle-aged workers ($1/\tilde{\rho} = 0.35$, $\Delta\mu = -0.1$ pp), older workers (0.30, −0.1 pp), and women (0.39, −0.1 pp) see small decreases, while men (0.66, 0.0 pp) and primary- and secondary-educated workers (−0.0 pp each) are essentially unaffected.

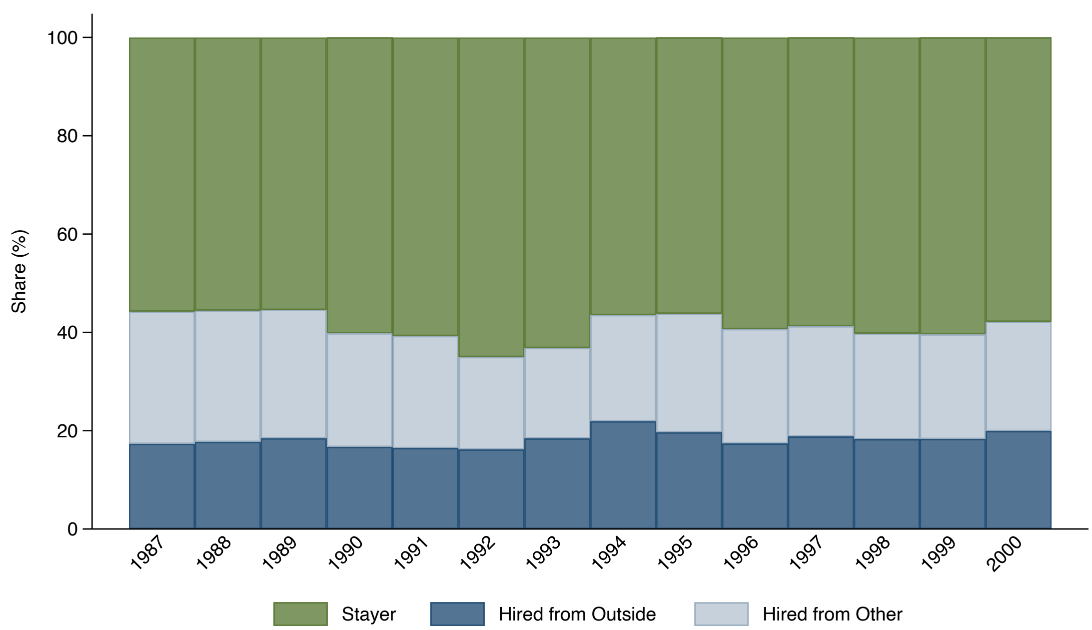
C.4 Summary

This appendix extends the baseline model to incorporate self-employment, the risk of involuntary separation into informal wage work, and preference heterogeneity by demographics and region. The key findings are: (i) labor supply to formal firms is highly inelastic at the within-market cross-firm margin ($1/\tilde{\eta} \approx 1.01$), with the bias from ignoring informal wage work small ($\Omega \approx 0.02$) and partially offset by unemployment insurance; (ii) substitution between wage work and self-employment is substantially more elastic ($1/\tilde{\rho} \approx 0.47$), confirming that self-employment curbs formal firms' power, especially for women, older workers, and less educated workers; (iii) cross-market mobility is the least elastic margin ($1/\tilde{\theta} \approx 1.19$), indicating large geographic and occupational frictions; (iv) average take-home shares are 51 cents on the dollar—close to the baseline's 50 cents—but range from 46 to 73 cents depending on local informality and workforce composition (Table 7); (v) the effect of trade on markdowns is essentially zero at the country average ($\Delta\mu \approx 0$ pp for a 10% RTR increase), as within-market and cross-market substitution channels nearly offset. Heterogeneity across informality and region is modest: the take-home share falls by -0.54 pp in high-informality regions and rises by +0.04 pp in low-informality regions, while demographic effects are uniformly close to zero.

Figure C.1: Formal sector worker flows

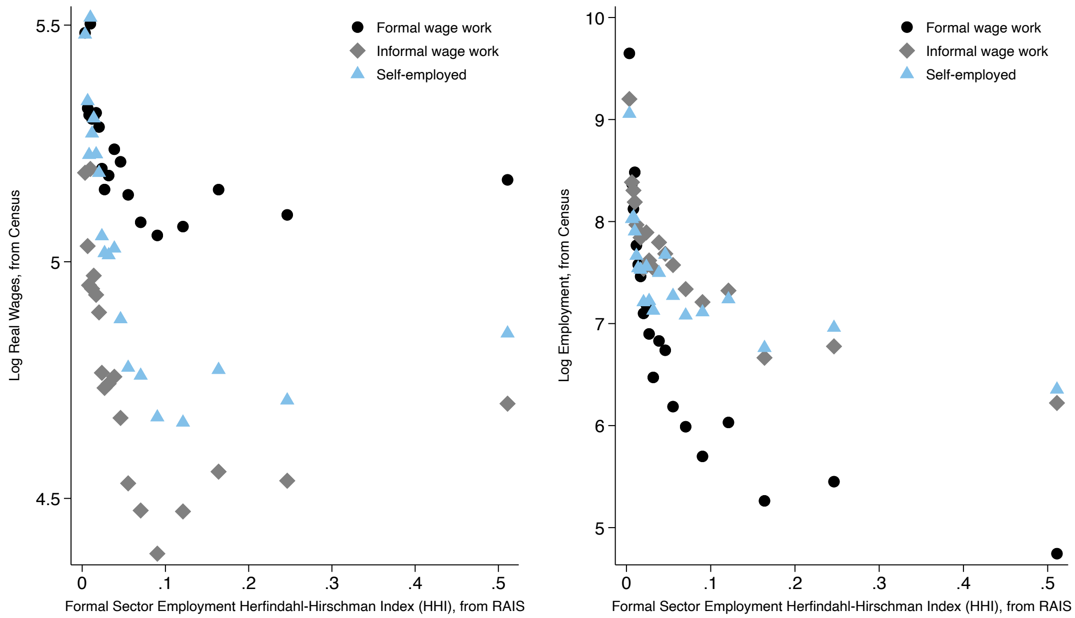
(a) Total firm-to-firm and firm-to-outside flows

(b) Flow shares



Notes: This figure plots the total number of new hires directly from formal sector firms ("Hired from Other"), workers separated from a formal firm and are matched within one year to another formal firm ("Sep to Other"), new hires from outside the formal sector ("Hired from Outside"), separations to outside the formal sector ("Sep to Outside"), and workers who do not switch employers within the year ("Stayer") as reported by formal sector firms in Brazil's RAIS employer-employee linked dataset. Note that "Hired from Other" and "Separated to Other" match, but are measured separately, the former based on records from the receiving firm, the latter based on records from the origin firm. Flows include within-year transitions and exclude flows to/from public administration, retirements, and deaths.

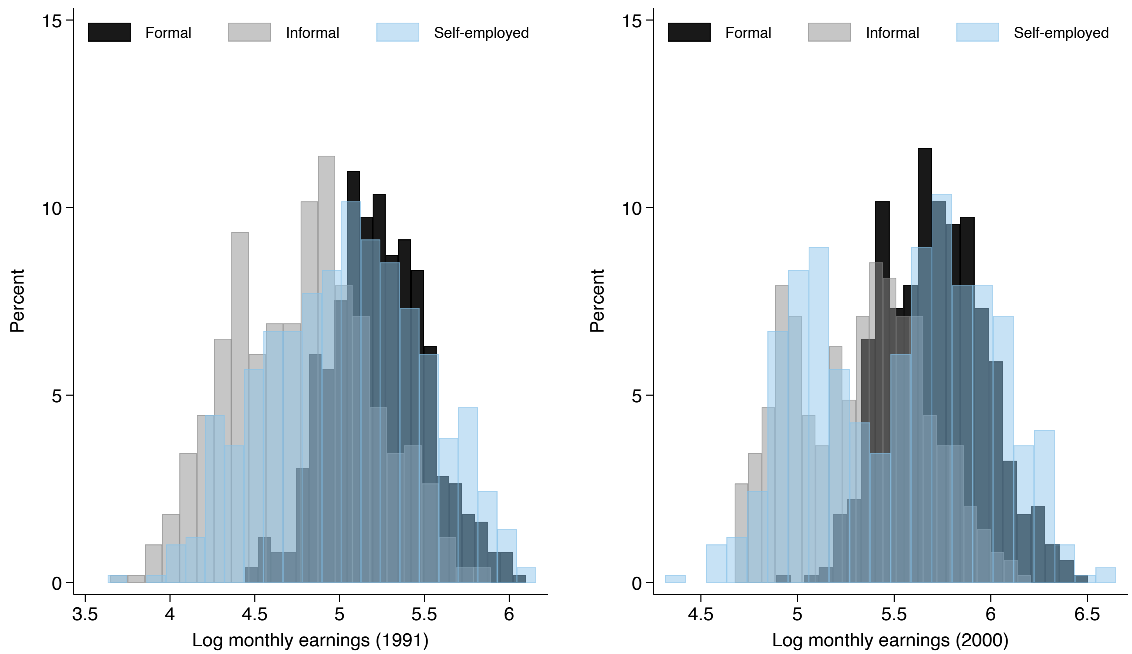
Figure C.2: 1991 wages, employment, and formal sector concentration (Census)



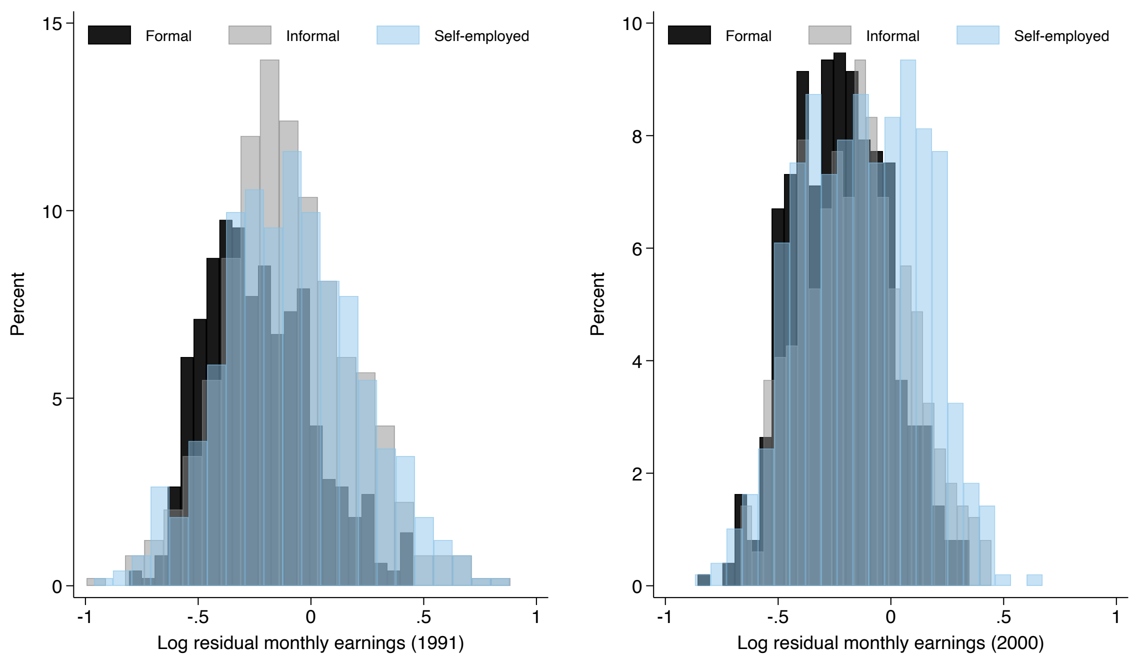
Notes: This figure plots wages and employment as a function of formal sector HHI in microregions using Census data for 1991.

Figure C.3: Formal, informal, and self-employment earnings

Panel A: Raw earnings



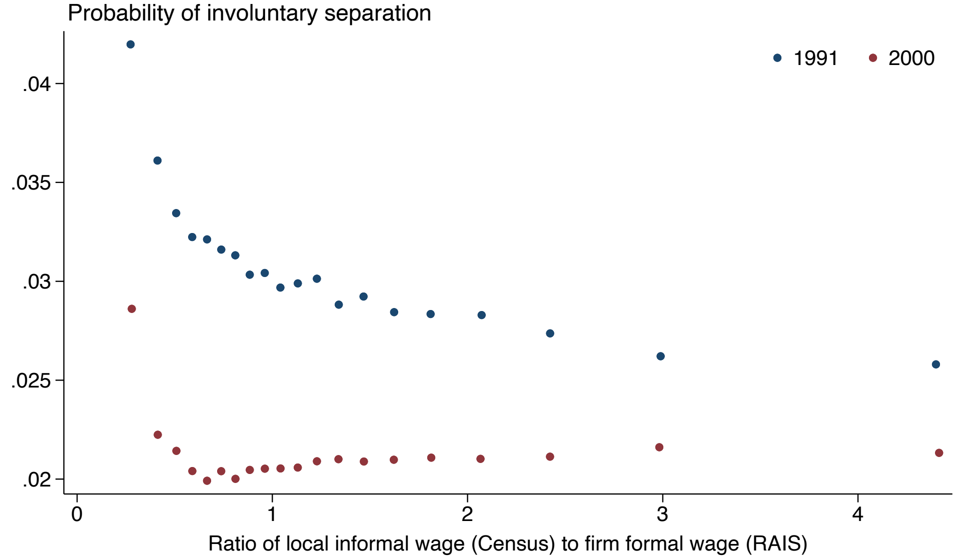
Panel B: Residual earnings



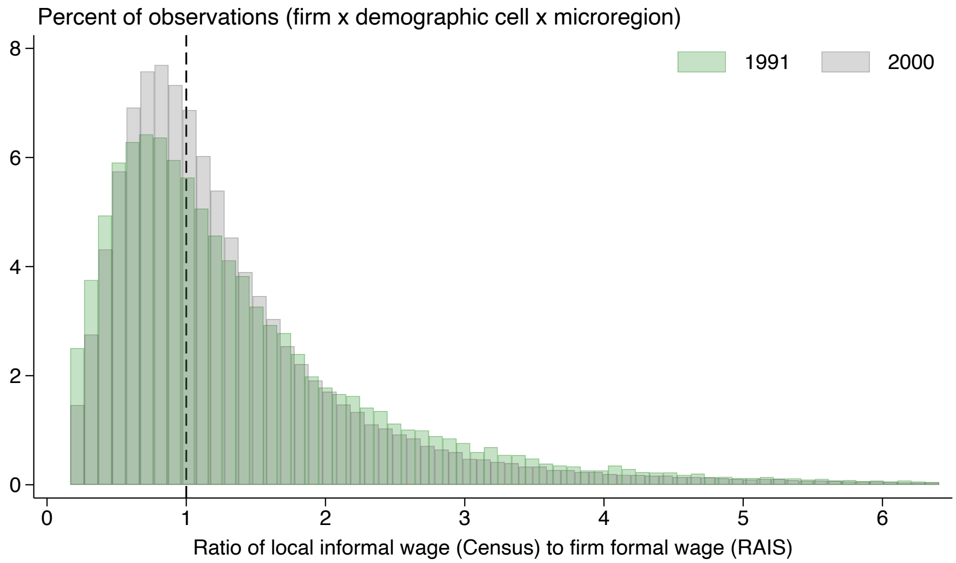
Notes: This figure plots the distribution of raw and residual formal real monthly earnings, informal real monthly earnings, and self-employment real monthly earnings, across 486 Brazilian microregions. Residual real monthly earnings are conditional on a fully saturated vector of dummies for gender, age groups, and education groups. Real earnings are expressed in 2000 reais.

Figure C.4: Probability of involuntary separation from formal firm and informal-formal wage gaps

Panel A: Probability of involuntary separation and wage gap



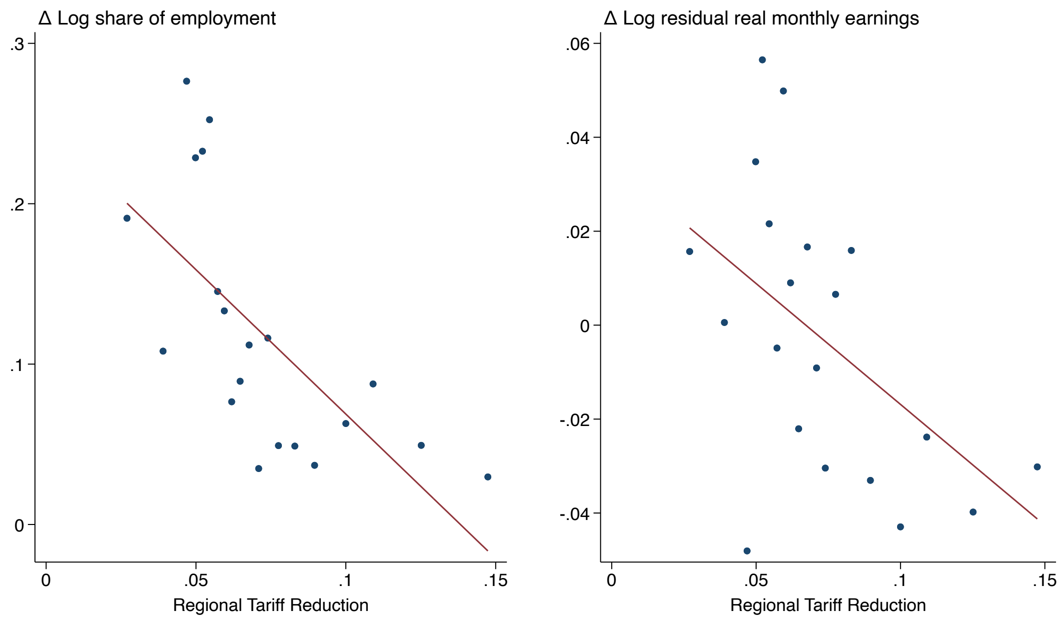
Panel B: Wage gap distribution



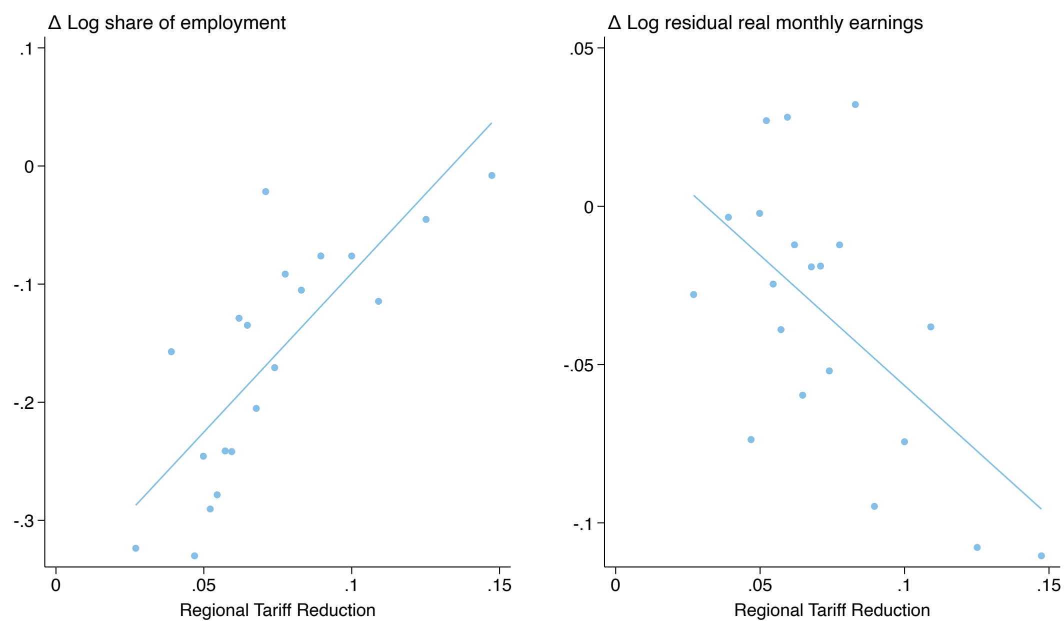
Notes: This figure shows the relationship between probabilities of involuntary separation from a formal sector firm in RAIS and the gap between local informal sector monthly earnings and the firm's formal monthly earnings, separately computed for each demographic cell. The sample includes all formal sector firms in RAIS in 1991 and 2000, merged at the microregion X demographic cell X year level to the corresponding census data. Each observation is a firm X microregion X demographic cell (gender X education group X age group dummies). Monthly earnings gaps are calculated by first converging all reported earnings to 2000 constant Reais. The probability of involuntary separation is calculated using separations between January 1 and December 30 of each year. Firm total formal employment is measured as of December 31 of each year.

Figure C.5: Effect of Regional Tariff Reductions on wage work and self-employment

Panel A: Wage work



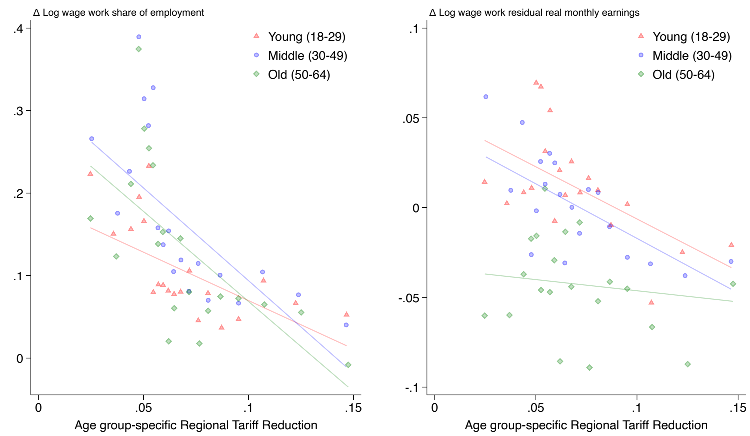
Panel B: Self-employment



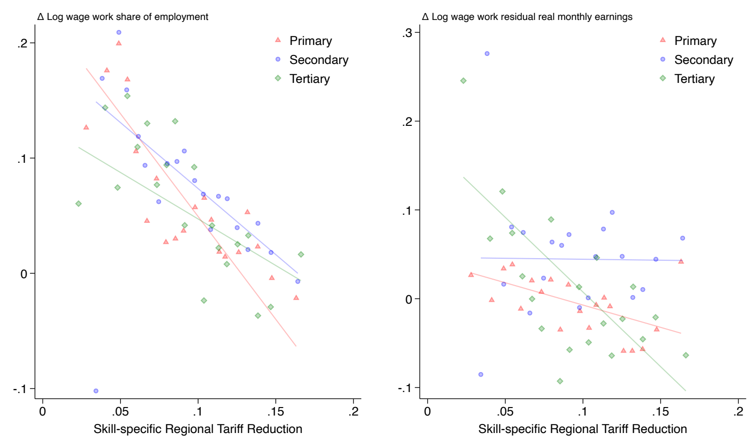
Notes: This figure shows the correlation between Regional Tariff Reductions (i.e., microregion-level exposure to 1990-1994 import tariff reductions from Dix-Carneiro and Kovak (2017)), changes in employment shares (left panel), and changes in log residual real monthly earnings (right panel) between the 1991 and 2000 census, for individuals employed in wage work (formal or informal) (Panel A) versus in self-employment (Panel B) across 486 Brazilian microregions. Log residual real monthly earnings are conditional on flexible controls for gender, age, and education. Real earnings are computed from nominal earnings by expressing all values in 2000 reais.

Figure C.6: Effect of Regional Tariff Reductions on wage work by demographics

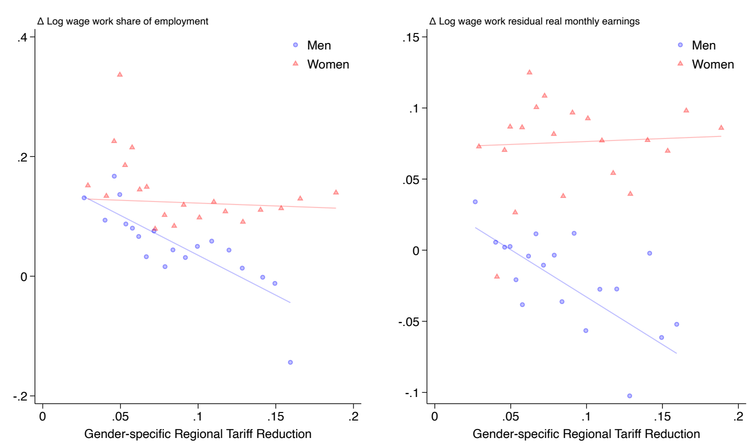
Panel A: By age



Panel B: By education



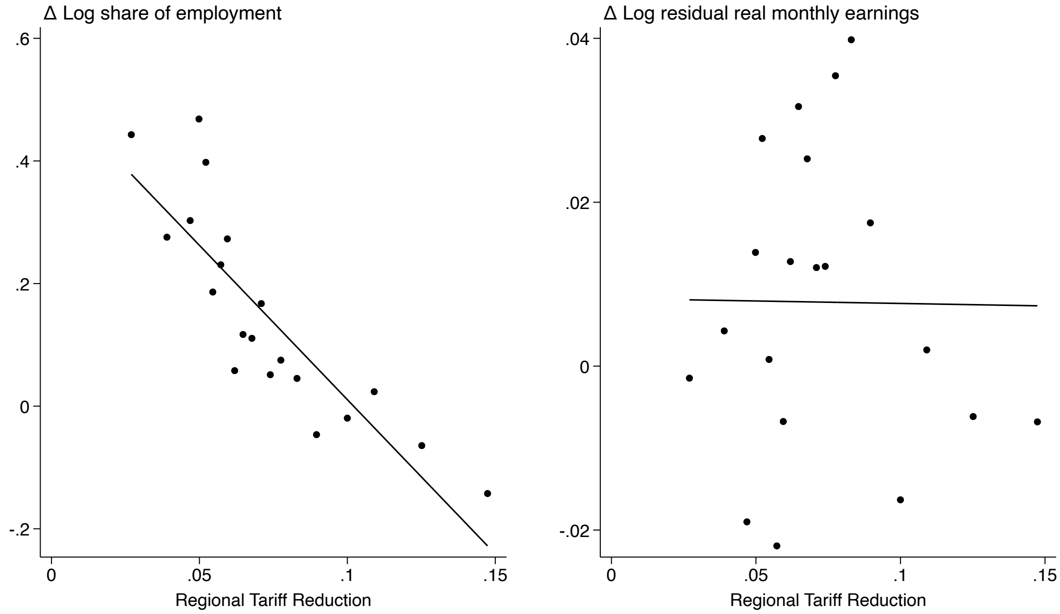
Panel C: By gender



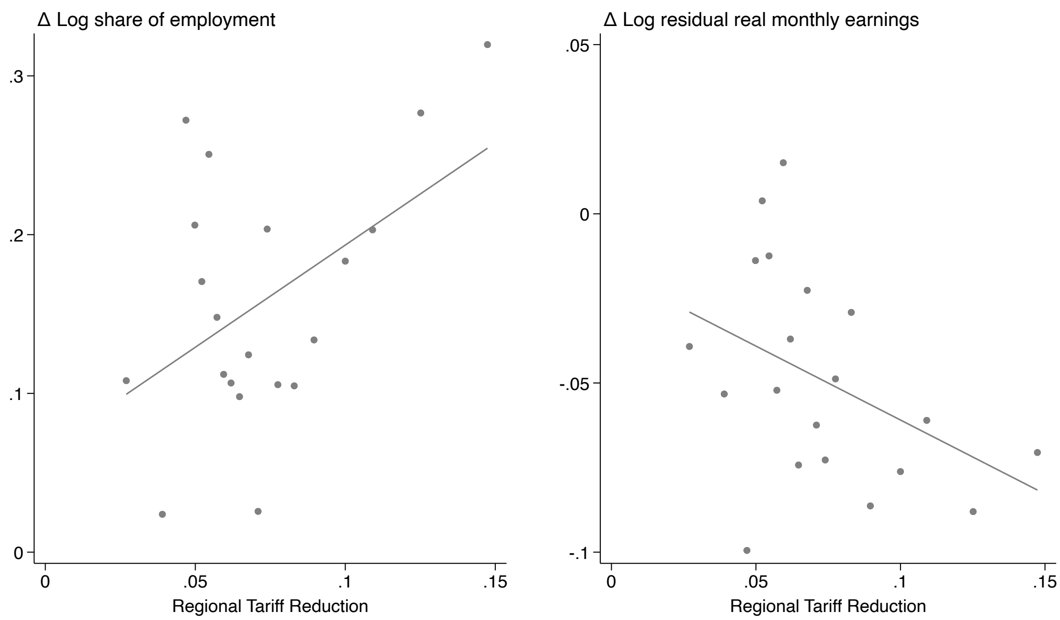
Notes: is figures shows the correlation between Regional Tariff Reductions (i.e., microregion-level exposure to 1990-1994 import tariff reductions from Dix-Carneiro and Kovak (2017)), changes in employment shares (left panel), and changes log residual real monthly earnings (right panel) between the 1991 and 2000 census, for self-employed individuals across 486 Brazilian microregions, separately for different demographic groups. Log residual real monthly earnings are conditional on a fully saturated vector of dummies for gender, age groups, and education groups. Real earnings are computed from nominal earnings by expressing all values in 2000 reais.

Figure C.7: Effect of Import Tariff Reductions on formal versus informal wage work

Panel A: Formal Wage Work



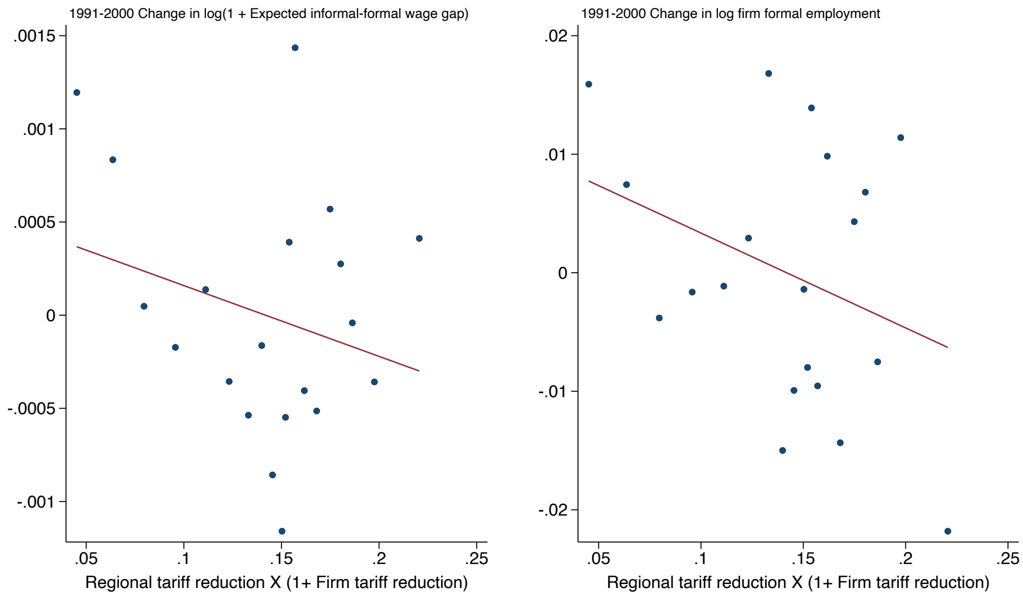
Panel B: Informal Wage Work



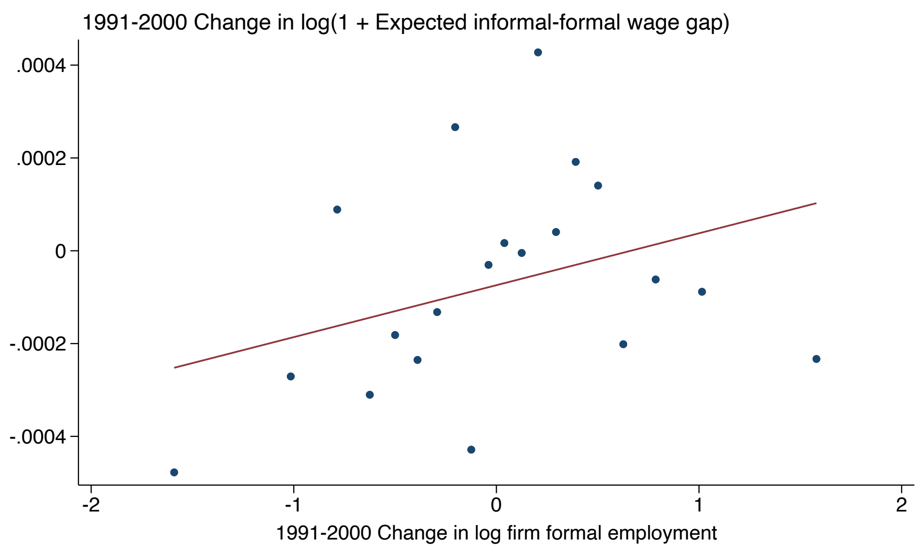
Notes: Each panel shows the correlation between Regional Tariff Reductions (microregion-level exposure to 1990-1994 import tariff reductions from Dix-Carneiro and Kovak (2017)), changes in employment shares (left) and log residual real monthly earnings (right) between 1991 and 2000 across 486 Brazilian microregions. Log residual real monthly earnings are conditional on a fully saturated vector of dummies for gender, age groups, and education groups. Real earnings are expressed in 2000 reais.

Figure C.8: Firm-specific labor supply response Ω to expected informal-formal wage gaps

Panel A: Visual IV: Reduced form (left) and First Stage (right)



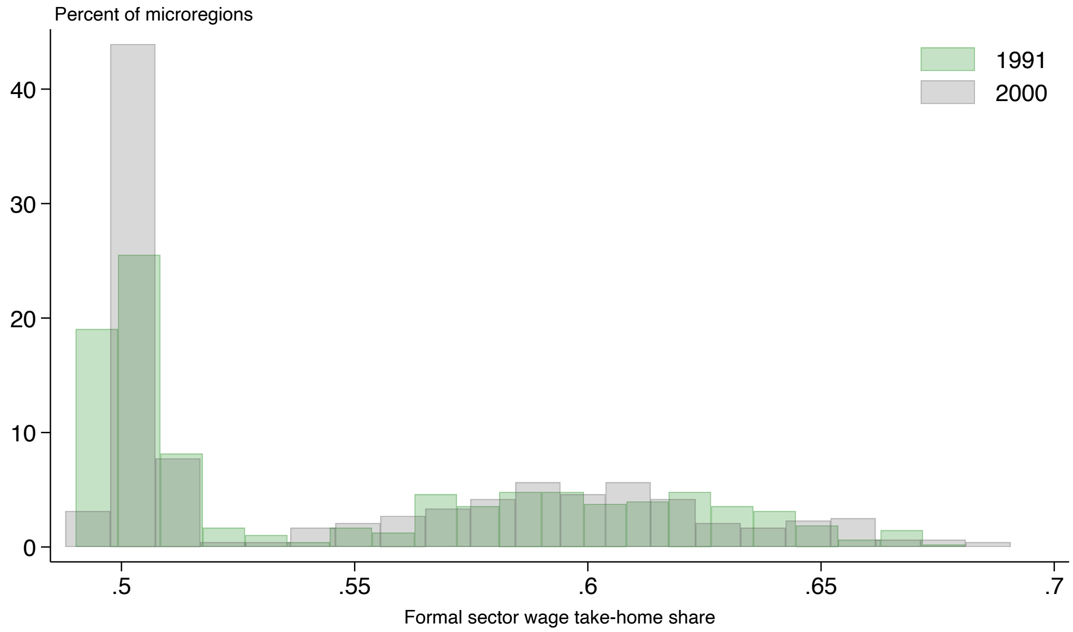
Panel B: Visual OLS



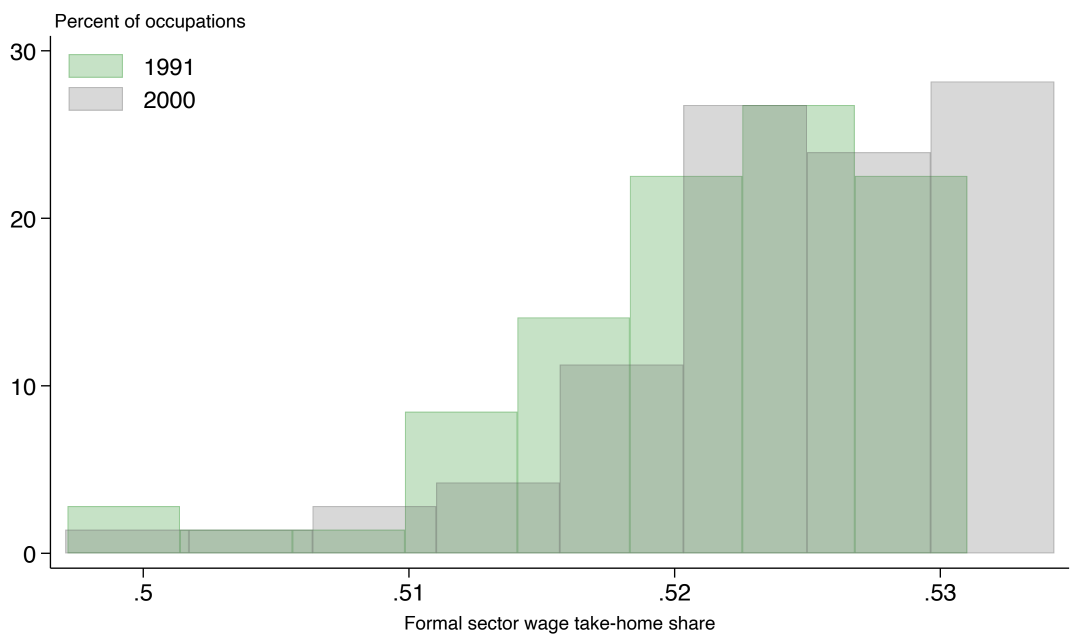
Notes: This figure shows binned scatters of 1991-2000 changes in expected informal-formal wage gaps, firm formal employment, and import tariff reductions. The sample includes all formal sector firms in RAIS in 1991 and 2000, merged at the microregion X demographic cell X year level to the corresponding census data. Each observation is a firm X microregion X demographic cell (gender X education group X age group dummies). All binned scatters show residual variation conditional on firm, microregion, and demographic cell fixed effects. Expected informal-formal wage gap is calculated separately by firm, microregion, and demographic cell as $p_{z\bar{s}r}\sigma^{c}_{z\bar{s}r}$. The first term, $p_{z\bar{s}m}$, is the probability of *involuntary* separation from firm z, measured as the ratio of firings to firings plus stayers in firm z. The second, $\sigma^{c}_{z\bar{s}m} \equiv (w^{o}_{m} - w^{c}_{z\bar{s}m})/w^{c}_{z\bar{s}m}$, is the monthly earnings gap between the local informal sector and firm z's formal wage for that demographic cell. Monthly earnings gaps are calculated by first converging all reported earnings to 2000 constant reais. $p_{z\bar{s}m}$ is calculated using separations between January 1 and December 30 of each year. Firm total formal employment is measured as of December 31 of each year. Regional tariff reductions are microregion-level exposure to 1990-1994 import tariff reductions from Dix-Carneiro and Kovak (2017) and differ by demographic group.

Figure C.9: Formal sector wage take-home share in dual labor markets

Panel A: Distribution across microregions

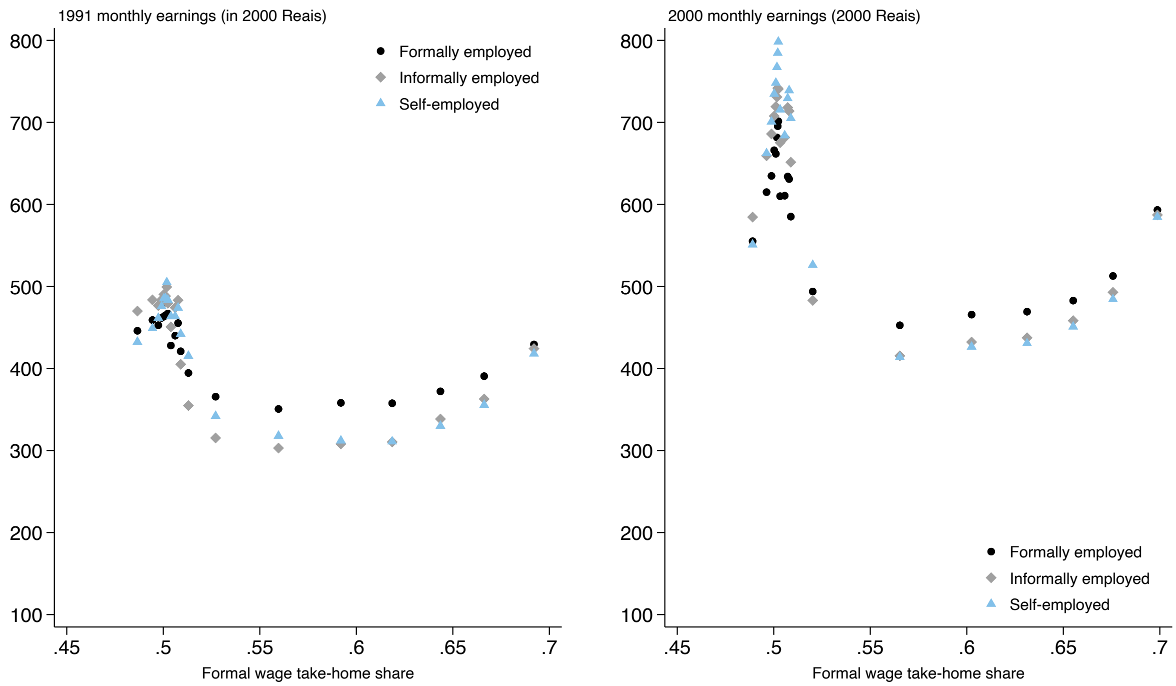


Panel B: Distribution across occupations



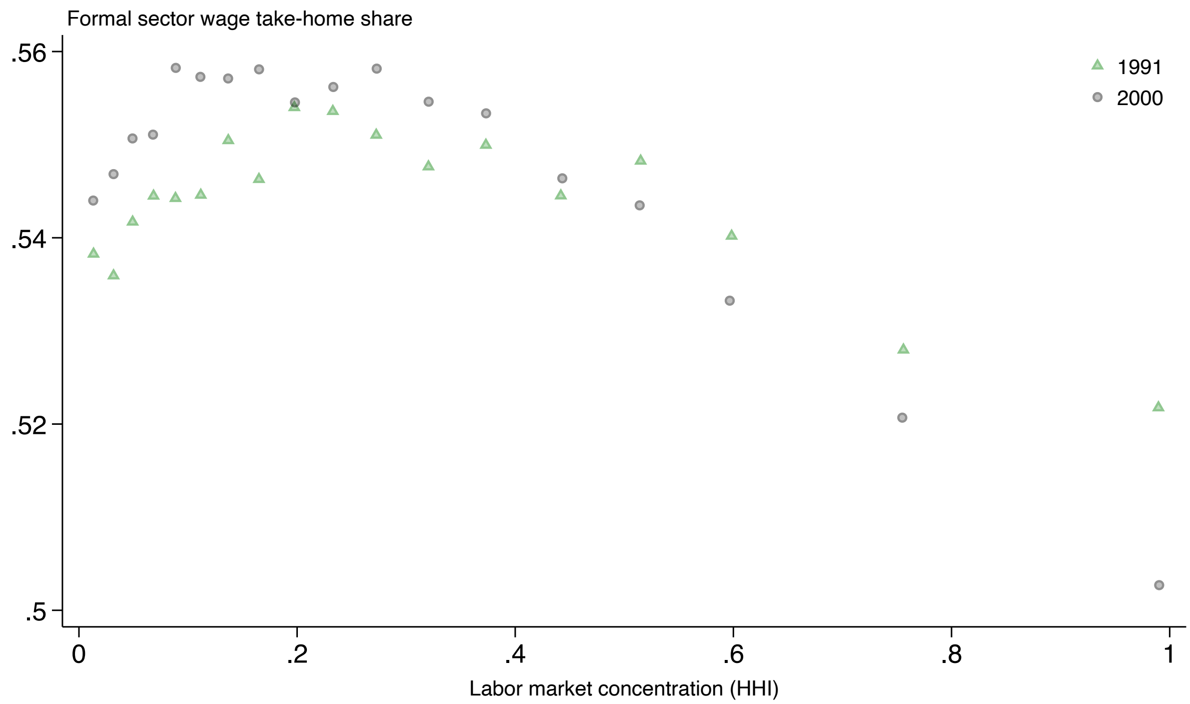
Notes: This figure show the distribution of average wage take-home shares (the inverse of average wage markdown) across microregions (Panel A) and occupations (Panel B), separately by year. Average wage take-home shares are calculated at the local labor market level (microregion x occupation) following equation C.23, using region-specific estimates of $1/\tilde{\eta}$ and pooled estimates of $1/\tilde{\rho}$ and $1/\tilde{\theta}$. Panels A and B aggregate these estimates by microregion and 2-digit occupation, respectively, weighing observations by total formal wage bill.

Figure C.10: Wages and formal sector wage take-home shares



Notes: This figure plots binned scatters between real monthly earnings (separately by employment type) and average formal sector wage take-home shares across microregions. See Figure C.9 for details and Appendix Figures C.12 for demographic-specific binned scatters.

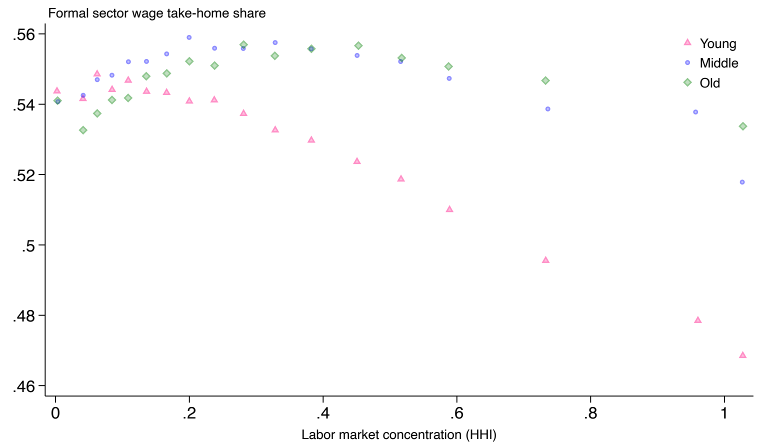
Figure C.11: Concentration and formal sector wage take-home shares



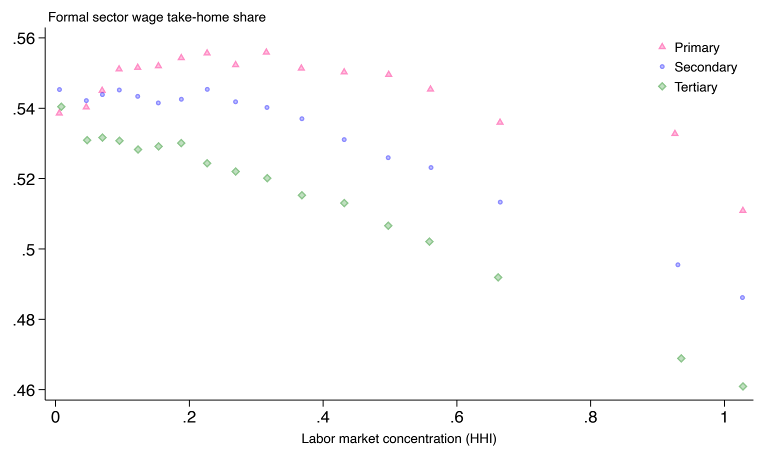
Notes: This figure plots binned scatters between average formal sector wage take-home shares and formal firms' local labor market concentration (HHI), measured in terms of *expected* wage bill shares, which take into account probabilities of separation into informal wage work, across microregion x occupation pairs, separately by year. See Figure C.9 for details and Appendix Figures C.13 through C.15 for demographic-specific distributions.

Figure C.12: Formal sector wage take-home share and labor market concentration

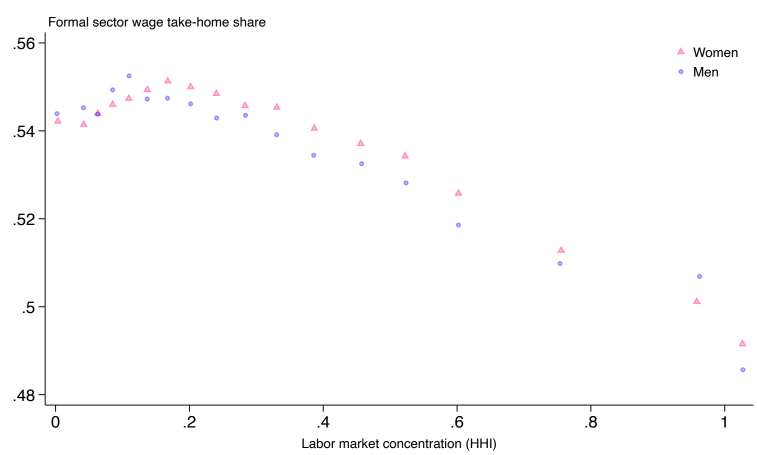
Panel A: By age



Panel B: By education



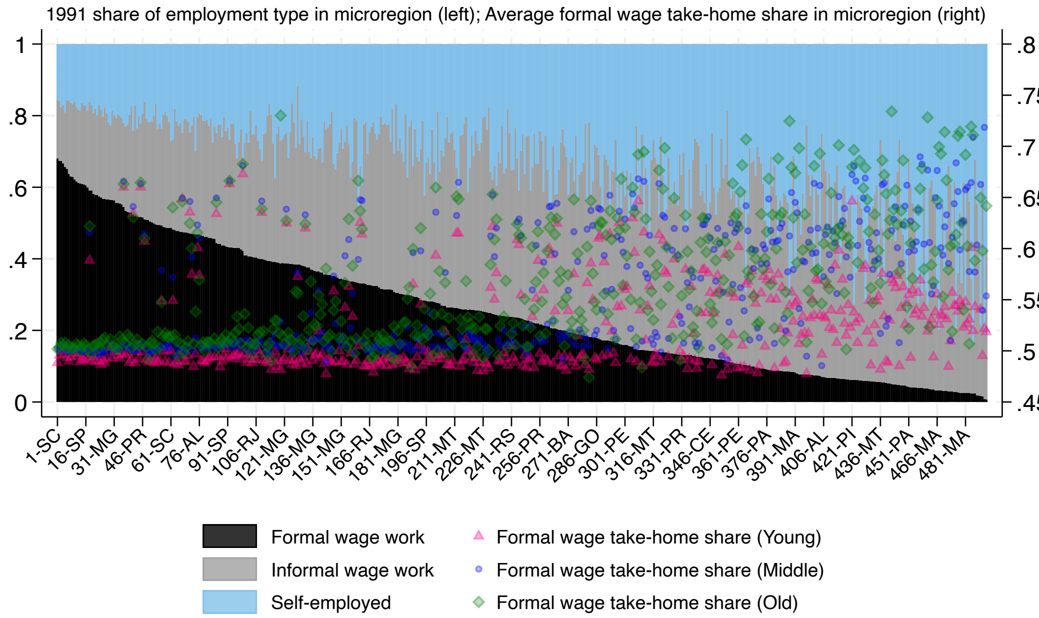
Panel C: By gender



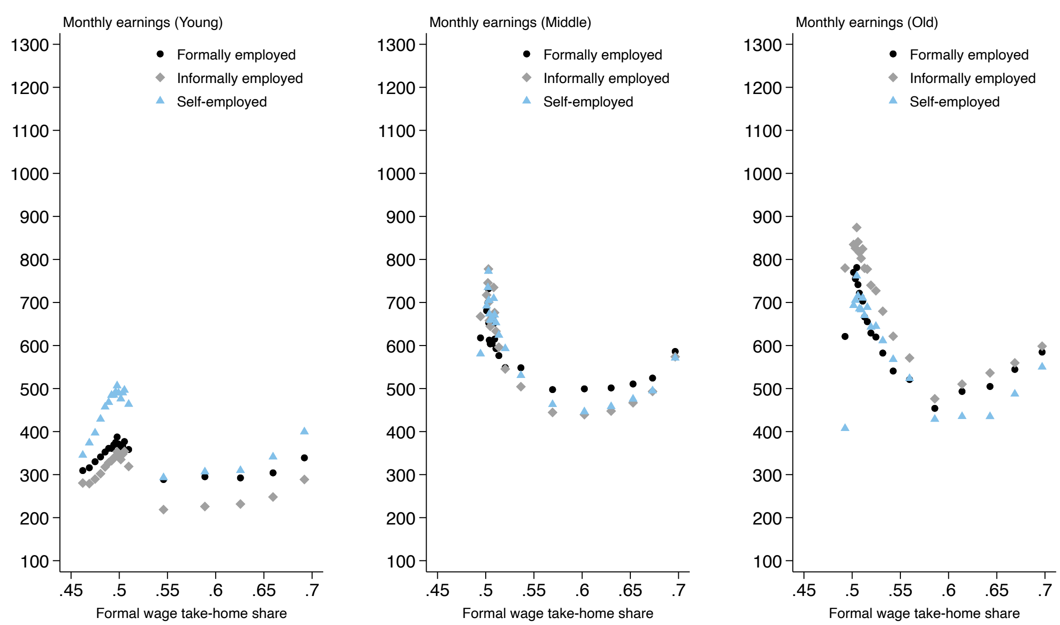
Notes: This figure plots binned scatters between demographic-specific average formal sector wage take-home shares and formal firms' local labor market concentration (HHI). See notes to Figure C.11.

Figure C.13: Formal wage take-home share in dual labor markets: Heterogeneity by age

Panel A: 1991 distribution across microregions



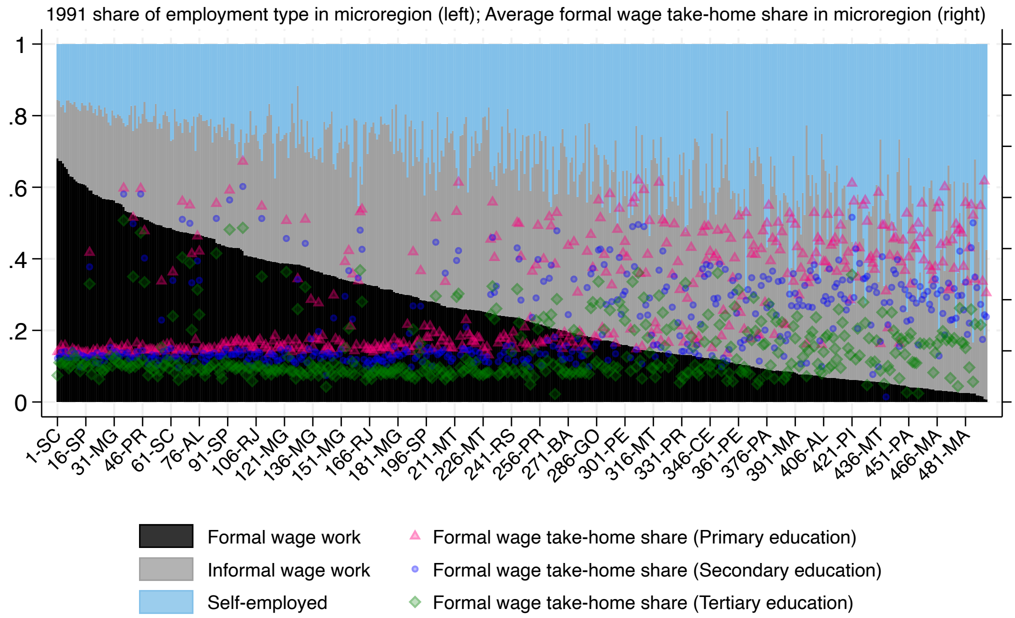
Panel B: Wages and wage take home-shares



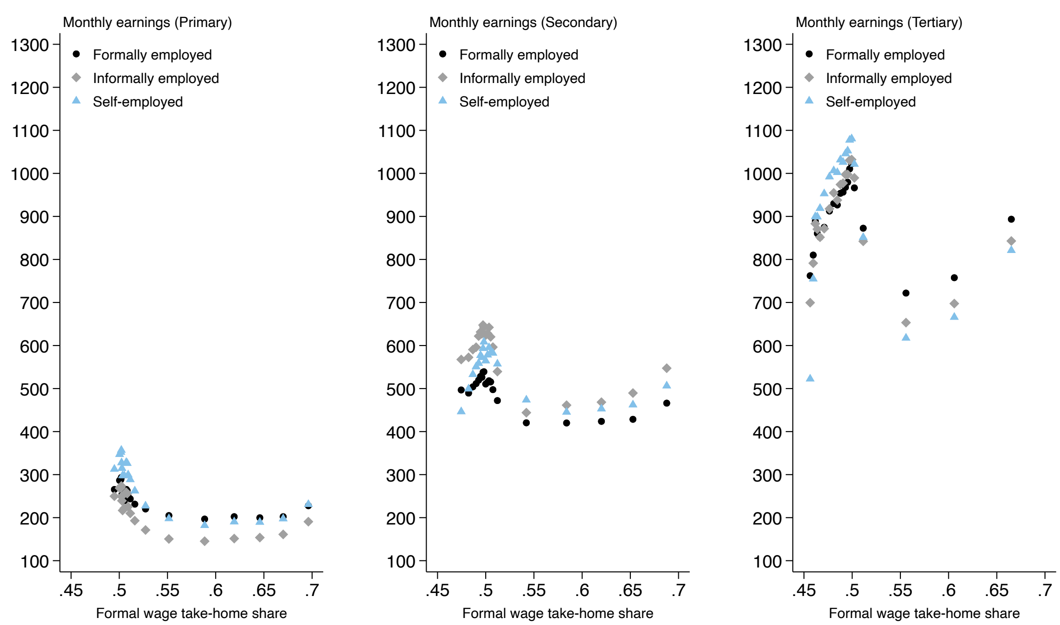
Notes: This figure presents age-specific versions of Figures 9 and C.10. Average wage take-home shares are calculated at the local labor market level (microregion x occupation) following equation C.23, using region-specific estimates of $1/\tilde{\eta}$, demographic-specific estimates of $1/\tilde{\rho}$, and the pooled estimate for $1/\tilde{\theta}$. For Panel A, average take-home shares are aggregated at the microregion level using formal sector wage bill as weights.

Figure C.14: Formal wage take-home share in dual labor markets: Heterogeneity by education

Panel A: 1991 distribution across microregions



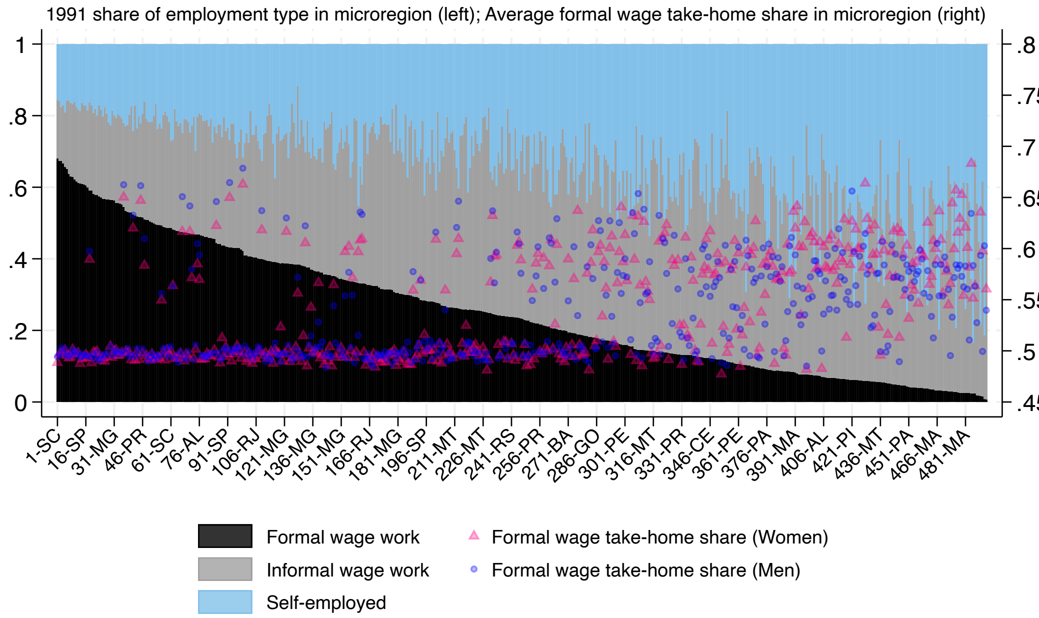
Panel B: Wages and wage take home-shares



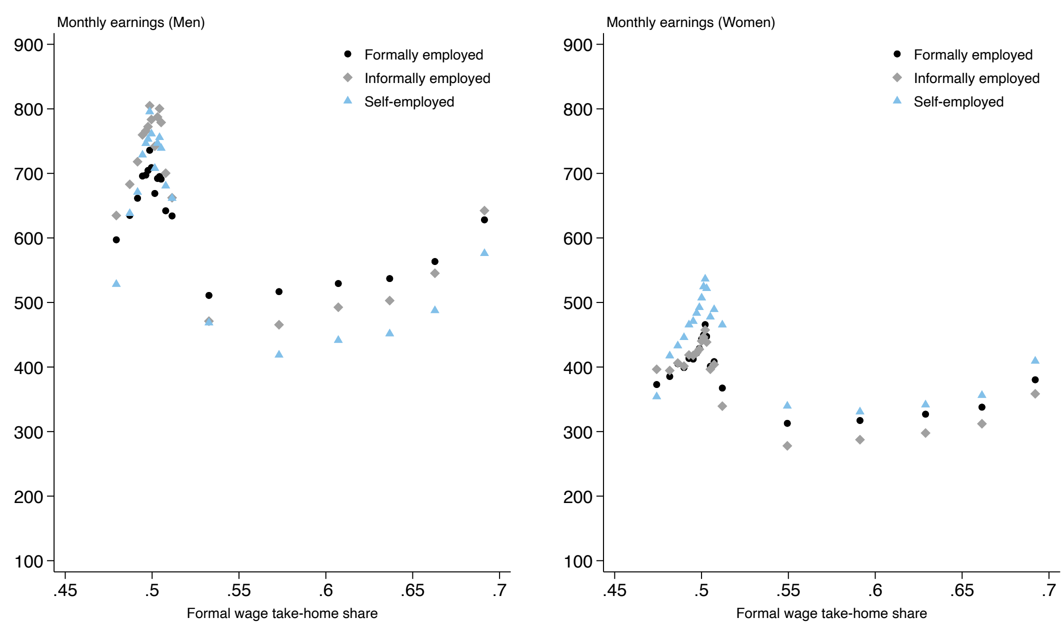
Notes: This figure presents education-specific versions of Figures 9 and C.10. Average wage take-home shares are calculated at the local labor market level (microregion x occupation) following equation C.23, using region-specific estimates of $1/\tilde{\eta}$, demographic-specific estimates of $1/\tilde{\rho}$, and the pooled estimate for $1/\tilde{\theta}$. For Panel A, average take-home shares are aggregated at the microregion level using formal sector wage bill as weights.

Figure C.15: Formal wage take-home share in dual labor markets: Heterogeneity by gender

Panel A: 1991 distribution across microregions



Panel B: Wages and wage take home-shares



Notes: This figure presents gender-specific versions of Figures 9 and C.10. Average wage take-home shares are calculated at the local labor market level (microregion x occupation) following equation C.23, using region-specific estimates of $1/\tilde{\eta}$, demographic-specific estimates of $1/\tilde{\rho}$, and the pooled estimate for $1/\tilde{\theta}$. For Panel A, average take-home shares are aggregated at the microregion level using formal sector wage bill as weights.

Table C.1: Wage markdowns in dual labor markets: Worker preference parameters and weights

		Estimation				
					Statistically signif. heterogeneity	
Parameter (1)	Definition (2)	Data (3)	Tables (4)	IV estimate (5)	Range (6)	Detected by (7)
$1/\eta$	Within-market cross-firm inverse elasticity of substitution	RAIS + Tariffs	C.2	0.990	0.458 - 1.029	Region
Ω	Bias in $1/\eta$ due to expected involuntary separation into informal wage work	RAIS + Census + Tariffs	C.4	0.0221	0.0279 - 0.0354	Age, gender
$1/\tilde{\rho}$	Inverse elasticity of substitution between wage work and self-employment	Census + Tariffs	C.6 and C.8	0.469	0.303 - 0.935	Age, educ, gender
$1/\tilde{\theta}$	Cross-market inverse elasticity of substitution	Census + Tariffs	C.6	1.191	-	-
Directly measured labor market equilibrium objects ("weights")						
$HHI_{\bar{g}m}$	*Expected* wage bill concentration among formal firms in market m.	RAIS + Census				
s_m	Wage work sector wage bill share as fraction of overall wage bill in market m.	RAIS + Census				

Notes: Column (3) lists the datasets used; column (6) lists the heterogeneity ranges; column (7) lists the heterogeneity dimensions. A market is a microregion × 2-digit occupation pair.

Table C.2: IV estimates of $1/\eta$ and its heterogeneity across major regions

	$\Delta \ln w_{z\bar{g}m}$				
	(1) Main	(2) NE	(3) SE	(4) NE + N	(5) SE + S + CW
$\Delta \ln l_{z\bar{g}m}$	0.990***	0.458*	1.029***	0.403*	1.005***
	(0.181)	(0.233)	(0.150)	(0.207)	(0.170)
First-stage F	82.10	18.10	112.6	19.84	87.93
Observations	847391	99885	530079	114961	732430

Notes: This table shows IV estimates of $1/\eta$, the within-market cross-firm inverse elasticity of residual labor supply in Section C.1 (and in Section C.1 if $\Omega = 0$) and its heterogeneity by Brazil's major regions. The sample includes all formal sector firms in RAIS in 1991 and 1997. Each observation is a firm X local labor market cell. A local labor market is a microregion x 2-digit occupation pair. All regressions include local labor market fixed effects. The instruments are firm-level import reductions from Section C.1. Column (1) replicates the average elasticity from Section C.1. Column (2) re-estimates Column (1) with a microregion fixed effect instead of a microregion x 2-digit occupation pair fixed effect. Column (3) re-estimates Column (1) with a 2-digit occupation fixed effect instead of a microregion fixed effect. Columns (4)-(7) re-estimates Column (1) limiting the sample to microregions within major regions: NE (Northeast), SE (Southeast), N (North), S (South), and CW (Center West). Limiting the sample to the N, S, or CW alone yields weak first stages and statistically insignificant results, hence why Columns (4)-(5) present results pooling them with Brazil's largest regions in terms of population (Northeast and Southeast). Standard errors are clustered two-ways by microregion and 2-digit occupation. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.3: Heterogeneous inverse elasticity by worker demographics

	Δ firm log wage premium in local market				
	(1)	(2)	(3)	(4)	(5)
Δ Firm log employment in local market	0.990	1.083	1.104	1.008	1.212
	(0.089)	(0.145)	(0.088)	(0.088)	(0.150)
Δ Firm log employment in local market		-0.0023			-0.00263
x (Baseline female share of employment)		(0.003)			(0.003)
Δ Firm log employment in local market			-0.0176		-0.0182
x (Baseline college-educated share of employment)			(0.003)		(0.003)
Δ Firm log employment in local market				-0.000814	0.000129
x (Baseline over-40-years-old share of employment)				(0.001)	(0.002)
Observations	855,104	855,104	855,104	855,104	855,104

Notes: Standard errors clustered by firm in parentheses. All specifications estimated by 2SLS using the change in log tariff and its interaction with the baseline demographic share as instruments. Baseline demographic shares are measured at the firm-market level in 1991 and scaled by 100. All regressions include LLM fixed effects and are weighted by baseline market employment.

Table C.4: IV estimates of Ω and its heterogeneity by demographic groups

	$\Delta \ln(1 + p_{z\bar{s}m}\sigma^{c}_{z\bar{s}m})$			
	(1)	(2)	(3)	(4)
	All	By gender	By education	By age
$\Delta \ln l^{c}_{z\bar{s}m}$	0.0331**			
	(0.0130)			
Men		0.0332**		
		(0.0132)		
Women		0.0354**		
		(0.0162)		
Primary education			0.0119	
			(0.0260)	
Secondary education			-0.0409	
			(0.0574)	
Tertiary education			0.0416	
			(0.0383)	
Young (18-29)				0.0344**
				(0.0144)
Middle (30-49)				0.0351***
				(0.0134)
Old (50-64)				0.0279*
				(0.0166)
First-stage F	1.608	3.250	2.757	6.155
Anderson-Rubin F		4.372	24.68	4.124
Observations	395305	395305	395305	395305

Notes: This table shows IV estimates of Ω, the bias term in IV estimates of the within-market cross firm elasticity of substitution in the model in Section C.1 if the extended model is true, and its heterogeneity by worker characteristics. The sample includes all formal sector firms in RAIS in 1991 and 2000, merged at the microregion X demographic cell X year level to the corresponding census data. Each observation is a firm X microregion X demographic cell (gender X education group X age group dummies). All regressions include firm, microregion, and demographic cell fixed effects. Expected informal-formal wage gap is calculated separately by firm, microregion, and demographic cell as $p_{z\bar{s}r}\sigma^{c}_{z\bar{s}r}$. The first term, $p_{z\bar{s}m}$, is the within-year probability of *involuntary* separation from firm z. The second, $\sigma^{c}_{z\bar{s}m} \equiv (w^{o}_{m} - w_{z\bar{s}m})/w_{z\bar{s}m}$, is the monthly earnings gap between the local informal sector and firm z's formal wage for that demographic cell. Monthly earnings gaps are calculated by first converging all reported earnings to 2000 constant Reais. $p^{c}_{z\bar{s}m}$ is calculated using separations between January 1 and December 30 of each year. Firm total formal employment is measured as of December 31 of each year. The instrument for the IV estimate is the interaction between firm-level import reductions and regional tariff reductions, interacted with demographic group dummies and firm baseline formal employment. Regional tariff reductions are microregion-level exposure to 1990-1994 import tariff reductions from Dix-Carneiro and Kovak (2017) and differ by demographic group. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.5: Sensitivity of Ω and to the value of unemployment insurance

	$\Delta \ln(1 + p_{z\bar{s}m}\sigma^c_{z\bar{s}m})$		
	(1)	(2)	(3)
	OLS	Without UI	With UI
$\Delta \ln l^c_{z\bar{s}m}$	0.000443***	0.0331**	0.0221**
	(0.000135)	(0.0130)	(0.00856)
First-stage F		1.608	1.608
Observations	407862	395305	395305

Notes: This table shows OLS and IV estimates of Ω, the bias term in IV estimates of the within-market cross firm elasticity of substitution in the model in Section C.1 if the extended model is true, and its heterogeneity by worker characteristics. The sample includes all formal sector firms in RAIS in 1991 and 2000, merged at the microregion X demographic cell X year level to the corresponding census data. Each observation is a firm X microregion X demographic cell (gender X education group X age group dummies). All regressions include firm, microregion, and demographic cell fixed effects. Expected informal-formal wage gap is calculated separately by firm, microregion, and demographic cell as $p_{z\bar{s}r}\,\sigma^c_{z\bar{s}r}$. The first term, $p_{z\bar{s}m}$, is the within-year probability of *involuntary* separation from firm z. The second, $\sigma^c_{z\bar{s}m} \equiv (w^o_r - w_{z\bar{s}m})/w_{z\bar{s}m}$, is the monthly earnings gap between the local informal sector w^o_r and firm z's formal wage for that market and demographic cell. Column (3) substitutes $w^o_{r,UI} = (4 * w_{z\bar{s}m} + 8 * w^o_r)/12$ for the informal sector wage to account for unemployment insurance benefits of 4-months salary. Monthly earnings gaps are calculated by first converging all reported earnings to 2000 constant Reais. $p_{z\bar{s}m}$ is calculated using separations between January 1 and December 30 of each year. Firm total formal employment is measured as of December 31 of each year. The instrument for the IV estimate is the interaction between firm-level import reductions and regional tariff reductions, interacted with demographic group dummies and firm baseline formal employment. Regional tariff reductions are microregion-level exposure to 1990-1994 import tariff reductions from Dix-Carneiro and Kovak (2017) and differ by demographic group. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.6: IV Estimates of $1/\tilde{\rho}$ and $1/\tilde{\theta}$

	(1)
	$\Delta \ln W^c_{\bar{s}m}$
$\Delta \ln(L^c_{\bar{s}m}/L^c_m)$	0.469***
	(0.0805)
$\Delta \ln(L^c_m/L^c)$	1.191*
	(0.619)
First-stage F ($1/\rho$)	66.57
Anderson-Rubin F ($1/\rho$)	60.88
First-stage F ($1/\theta$)	3.207
Anderson-Rubin F ($1/\theta$)	35.33
Observations	8055

Notes: This table shows second stage estimates from an instrumental variables regression that estimates the inverse elasticity of substitution between wage work and self-employment $1/\tilde{\rho}$ as the coefficient on $\Delta \ln(L^c_{\bar{s}m}/L^c_m)$, and the cross-market elasticity of substitution $1/\tilde{\theta}$ as the coefficient on $\Delta \ln(L^c_{\bar{s}m}/L^c_m)$. The sample includes 486 microregions and 18 demographic cells defined by 8 major demographic groups (2 by gender) x (3 by education) x (3 by age). The outcome variable is the 1991-2000 change in log residual real monthly earnings among individuals employed in wage work, either formally or informally, for each demographic group. The dependent variables are (1) the 1991-2000 change in log share of wage work employment in a microregion for each demographic cell; and (2) the 1991-2000 change in the log share of microregion total employment relative to national employment for each demographic cell. The instruments are group-specific Regional Tariff Reductions interacted with demographic group dummies and with the 1991 formal share of employment in the microregion. The regression includes region fixed effects for each of Brazil's five major regions and is weighted by each cell's 1991 microregion population. Log residual real monthly earnings are conditional on flexible controls for gender, education, and age. Real monthly earnings are based on the IPCA deflator and are expressed in 2000 reais. Standard errors in parentheses are clustered by microregion. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.7: Estimates of $1/\tilde{\rho}$ and $1/\tilde{\theta}$: Alternative instruments

	(1)
	$\Delta \ln W_{\bar{s}m}$
$\Delta \ln(L_{\bar{s}m}/L_m)$	0.120
	(0.133)
$\Delta \ln(L_m/L)$	1.111***
	(0.405)
First-stage F for $\Delta \ln(L_{\bar{s}m}/L_m)$	38.19
Anderson-Rubin F for $\Delta \ln(L_{\bar{s}m}/L_m)$	49.92
First-stage F for $\Delta \ln(L_m/L)$	1.965
Anderson-Rubin F for $\Delta \ln(L_m/L)$	2.454
Observations	478

Notes: This table shows second stage estimates from an instrumental variables regression that estimates the inverse elasticity of substitution between wage work and self-employment $1/\tilde{\rho}$ as the coefficient on $\Delta \ln(L_{\bar{s}m}/L_m)$, and the cross-market elasticity of substitution $1/\tilde{\theta}$ as the coefficient on $\Delta \ln(L_{\bar{s}m}/L_m)$. The sample includes 486 microregions. The outcome variable is the 1991-2000 change in log residual real monthly earnings among individuals employed in wage work, either formally or informally, in each microregion. The dependent variables are (1) the 1991-2000 change in log share of wage work employment in a microregion; and (2) the 1991-2000 change in the log share of microregion total employment relative to national employment. The instruments are Regional Tariff Reductions interacted with (a) each microregion's log maximum distance to the nearest labor law enforcement office, borrowing the estimation strategy from Ponczek and Ulyssea (2022); and (b) each microregion's 1991 share of employment that was formal. The regression includes region fixed effects for each of Brazil's five major regions and is weighted by each cell's 1991 microregion population. Log residual real monthly earnings are conditional on flexible controls for gender, education, and age. Real monthly earnings are based on the IPCA deflator and are expressed in 2000 reais. Standard errors in parentheses are clustered by microregion. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C.8: IV Estimates of $1/\tilde{\rho}$ and its heterogeneity by demographic groups

	$\Delta \ln W^c_{\bar{s}m}$			
	(1)	(2)	(3)	(4)
	All	By gender	By education	By age
$\Delta \ln(L^c_{\bar{s}m}/L^c_m)$	0.482***			
	(0.0594)			
Men		0.660***		
		(0.186)		
Women		0.390***		
		(0.122)		
Primary education			0.0686	
			(0.0668)	
Secondary education			0.460***	
			(0.177)	
Tertiary education			0.935***	
			(0.307)	
Young (18-29)				0.850***
				(0.157)
Middle (30-49)				0.355***
				(0.100)
Old (50-64)				0.303***
				(0.104)
First-stage F	52.64	90.64	33.53	30.18
Anderson-Rubin F	52.64	105.1	44.43	44.89
Observations	8055	8055	8055	8055

Notes: This table shows second stage estimates from instrumental variables regressions that estimate the inverse elasticity of substitution between wage work and self-employment $1/\tilde{\rho}$ jointly or separately by demographic groups. The sample includes 486 microregions and 18 demographic cells defined by 8 demographic groups (2 by gender) x (3 by education) x (3 by age). The outcome variable is the 1991-2000 change in log residual real monthly earnings among individuals employed in wage work, either formally or informally, for each demographic cell. The dependent variable is the 1991-2000 change in log share of wage work employment in a microregion for each cell. The instruments are cell-specific Regional Tariff Reductions interacted with group dummies and with each cell's 1991 formal share of employment in the microregion. All regressions include microregion fixed effects. Column (1) reports a pooled regression coefficient. Columns (2)-(4) report coefficients on the dependent variable interacted with dummies for each of the 8 major demographic groups. Columns (1) and (2) are weighted by each cell's 1991 microregion population. Columns (3) and (4) are unweighted due to substantial heterogeneity in workforce composition by education and age across microregions, resulting in larger and noisier estimates. Log residual real monthly earnings are conditional on flexible controls for gender, education, and age. Real monthly earnings are based on the IPCA deflator and are expressed in 2000 reais. Standard errors in parentheses are clustered by microregion. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.