Online Appendix

For the article: "Collective Bargaining for Women: How Unions Can Create Female-Friendly Jobs"

Viola Corradini

Lorenzo Lagos

Garima Sharma

CONTENTS

A   APPENDIX TABLES

TABLE A.1

Clauses in the Intuitive Definition of Female-Centric Amenities

Group	Clause Type	Description
Leaves		
	Abortion leave	Leave in cases of miscarriage/abortion
	Adoption leave	Leave following the adoption of a child
	Maternity leave	Leave concerning the birth of a child
	Paid leave	Leave during which worker receives normal pay
	Unpaid leave	Leave during which worker does not receive normal pay
	Other: holidays and leaves	Provisions on holidays/leaves outside predefined clause types
	Female workforce	General provisions concerning female workers
Maternity and childcare		
	Childcare assistance	Payments to assist with childcare support
	Maternity assistance	Payments to assist with becoming a mother
	Abortion protections	Employment protections concerning miscarriage/abortion
	Maternity protections	Employment protections for mothers
	Paternity protections	Employment protections for fathers
	Policy for dependents	Workplace benefits that apply to dependents
Workplace harassment and discrimination		
	Sexual harassment	Rules/penalties pertaining to harassment in the workplace
	Equal opportunities	Initiatives/statements on equality of opportunity for workers
Flexibility and part-time work		
	Workday controls	Rules restricting the duration of the workday
	Special shifts	Work shifts for subgroups of workers, e.g., women, minors, students
	On-call	Rules on workers' availability outside of the normal workday
	Uninterrupted shifts	Rules concerning back-to-back shifts
	Part-time contracts	Directives on temporary/part-time employment contracts

Notes: Table lists the *Sistema Mediador* clause types used in our intuitive definition of female-centric amenities. The descriptions provided in this table are purposefully vague—clauses of a given type can vary to some degree. The clauses were chosen based on the content of CUT's fight plan and the existing literature on workplace amenities valued by women, restricting ourselves to only 20 clause types.

TABLE A.2

Examples of Female-Centric Amenities

Childcare assistance	The company will reimburse all female employees, the monthly amount of R$ 110, as a "day care allowance", per child up to 6 years old. This benefit applies to any employee with custody of the child(ren).
Absences	The employee will receive full pay for absences upon proof of the following cases: a) bereavement (5 consecutive days); b) hospitalization of direct family or legal dependents; c) medical and dental consultations; d) marriage (5 working days)
Adoption leave	The employee who adopts or obtains legal custody for adoption will be granted maternity leave as follows: a) 120 days for children up to 1 year old; b) 60 days, for children from 1 to 4 years old; c) 30 days for children from 4 to 8 years old.
Other: holidays and leaves	The start of vacations cannot coincide with Saturdays, Sundays, holidays, or days already compensated. Vacations will start on the first working day of the week, communicated to the union within 10 working days by the company.
Seniority pay	The company will pay the employee who completes 5 uninterrupted years of work an additional 5% per length of service payable monthly, calculated on the monthly fixed base salary.

Notes: Table lists examples of CBA clauses from the top 5 clause types selected as "female-centric" or "male-centric" based on our data-driven approach—refer to Section III.B for details on the data-driven approach. The clauses were selected based on the number of unique tokens appearing in the clause that are within the top 20 TF-IDF tokens of each specific clause type.

TABLE A.3
Examples of Male-Centric Amenities

On-call pay	The company will pay an additional 35% of the normal hours to employees, when scheduled to be on-call. This additional pay will not apply when the on-call becomes a service actually provided, in which case overtime will be due.
Life insurance	The company will maintain group life insurance, guaranteeing a single and total indemnity of at least R$ 10,000 in the event of death or permanent disability of the employee resulting from an accident at work.
Strike procedures	The union assumes formal commitment not to promote or encourage stoppages, except in cases of non-compliance with clauses of this agreement or current laws, and even so, only after communicating the transgressions in writing to the employers.
Other: protections for injured workers	The company will communicate to Social Security, and subsequently to the union, injuries incurred by employees at the company or while commuting to/from work.
Profit sharing	The company will maintain a Profit Sharing Program with the amount made available for payment may be up to 1 nominal salary per employee. The payment period after the calculation of the results will be the month of February.

Notes: Table lists examples of CBA clauses from the top 5 clause types selected as "female-centric" or "male-centric" based on our data-driven approach—refer to Section III.B for details on the data-driven approach. The clauses were selected based on the number of unique tokens appearing in the clause that are within the top 20 TF-IDF tokens of each specific clause type.

TABLE A.4
Robustness of Data-Driven Female-Centric Amenities

Clause type	Times selected: data-driven (out of 6 methods)	Selected in baseline data-driven approach: (no state and industry FEs)	(state and industry FEs)
Childcare assistance	6	1	1
Absences	6	1	1
Adoption leave	6	1	0
Other: holidays and leaves	6	1	1
Seniority pay	6	1	1
Maternity protections	6	1	0
Paid leave	6	1	1
Night pay	6	1	0
Abortion leave	6	1	0
Policy for dependents	6	1	0
Waiving union fees	6	1	1
Salary adjustments/corrections	6	1	1
Nonwork-related injury protections	5	1	0
Extension/reduction of workday	5	1	1
Renewal/termination of the CBA	5	1	0
Medical exams	5	1	0
Abortion protections	4	1	0
Unionization campaigns	4	1	0
Adoption protections	4	0	0
Guarantees to union officers	3	1	1
Health education campaigns	3	1	0
Military service protections	3	0	1
Separation/dismissal	2	0	1
Other employment protections	2	0	0
Awards	1	0	0
Moral harassment	1	0	1
Maternity leave	1	0	0

Notes: Table lists all of the clauses identified as female-centric in any of the 6 methods implemented based on the estimation of Equation (2). Methods vary in 1) the sample of establishments covered by sectoral CBAs used, i.e., a random sample or the full sample; and 2) the measure of PageRank values used to determine gender gaps, i.e., normalized, non-normalized, or rankings. The initial column simply shows the number of times the clause is picked as female-centric by one of these 6 methods (clauses in the table are sorted in descending order as per the values of this column). The next column is an indicator for whether the clauses is selected as a female-centric by the baseline method, i.e., using a random sample and normalized PageRanks. The final column is an indicator for whether the clause is selected as female-centric by the baseline method but where the lasso includes state and industry fixed effects. Note that the Spearman correlation of the coefficients on clauses using the data-driven lasso approach versus an OLS using these same clauses but adding state and industry fixed effects is 0.56 with p-value below 0.01.

TABLE A.5

Robustness of Data-Driven Male-Centric Amenities

Clause type	Times selected: data-driven (out of 6 methods)	Selected in baseline data-driven approach: (no state and industry FEs)	(state and industry FEs)
On-call pay	6	1	1
Life insurance	6	1	1
Strike procedures	6	1	1
Other: protections for injured workers	6	1	1
Female workforce	6	1	1
Machine and equipment maintenance	6	1	1
Duration and schedule	6	1	0
Working environment conditions	6	1	0
Salary payment - means and timeframes	6	1	0
Hazard pay (danger risk)	6	1	0
Workday compensation	6	1	0
Tools and equipment	6	1	0
Profit sharing	5	1	1
Transfers	5	1	0
Safety equipment	5	1	0
Other assistances	5	1	0
Death/funeral assistance	5	1	0
Salary deductions	4	1	0
Equal opportunities	4	0	0
Collective vacations	3	1	0
Union fees	3	0	0
CIPA: accident prevention committee	2	1	1
Unpaid leave	2	0	0
Part-time contracts	2	0	0
Food assistance	1	0	0
Performance evaluation	1	0	0
Employment/hiring rules	1	0	0

Notes: Table lists all of the clauses identified as male-centric in any of the 6 methods implemented based on the estimation of Equation (2). Methods vary in 1) the sample of establishments covered by sectoral CBAs used, i.e., a random sample or the full sample; and 2) the measure of PageRank values used to determine gender gaps, i.e., normalized, non-normalized, or rankings. The initial column simply shows the number of times the clause is picked as male-centric by one of these 6 methods (clauses in the table are sorted in descending order as per the values of this column). The next column is an indicator for whether the clauses is selected as a male-centric by the baseline method, i.e., using a random sample and normalized PageRanks. The final column is an indicator for whether the clause is selected as male-centric by the baseline method but where the lasso includes state and industry fixed effects. Note that the Spearman correlation of the coefficients on clauses using the data-driven lasso approach versus an OLS using these same clauses but adding state and industry fixed effects is 0.56 with p-value below 0.01.

TABLE A.6
Establishment Descriptives: RAIS vs. Analysis Samples

	All RAIS (1)	Amenities sample (2)	Difference p-value (3)	RAIS: employ men and women (4)	Establishment sample (5)	Difference p-value (6)
Employment and firm characteristics						
Size	16.19	143.11	0.00	31.87	150.22	0.00
Share women	0.45	0.38	0.00	0.45	0.40	0.00
Employs both men and women	0.46	0.82	0.00	1.00	1.00	1.00
Single person firm	0.27	0.04	0.00	0.00	0.00	1.00
Single establishment firm	0.77	0.65	0.00	0.77	0.63	0.00
Sector						
Agriculture & extraction	0.09	0.04	0.00	0.05	0.03	0.00
Manufacturing	0.09	0.28	0.00	0.11	0.30	0.00
Construction & utilities	0.05	0.06	0.00	0.04	0.05	0.00
Commerce	0.39	0.23	0.00	0.41	0.24	0.00
Services	0.38	0.39	0.00	0.38	0.39	0.00
Region						
North	0.04	0.05	0.00	0.05	0.05	0.00
Northeast	0.16	0.12	0.00	0.16	0.12	0.00
Central	0.10	0.07	0.00	0.10	0.08	0.00
South	0.21	0.21	0.00	0.20	0.21	0.00
Southeast	0.49	0.56	0.00	0.49	0.54	0.00
N establishments	3,798,207	80,131		1,739,255	61,752	
N workers	61,492,768	11,467,760		48,564,436	9,276,475	
% workforce	100%	19%		79%	15%	

Notes: Table compares descriptive statistics of establishments in Brazil's formal sector (Columns 1 and 4) and our analysis samples, i.e., the amenity sample (Column 2) and the establishment sample (Column 5). The p-values of the differences between these samples are reported in Columns 3 and 6. The bottom of the table includes the number of unique establishments and workers in each sample, as well as the percentage of the formal workforce present in the corresponding sample.

TABLE A.7
Establishments Descriptives: Treated vs. Control

	Amenities sample			Establishment sample		
	Treatment	Control	Diff. p-value	Treatment	Control	Diff. p-value
	(1)	(2)	(3)	(4)	(5)	(6)
Employment and firm characteristics						
Size	198.21	127.03	0.00	200.37	135.95	0.00
Share women	0.36	0.38	0.00	0.38	0.40	0.00
Employs both men and women	0.83	0.82	0.00	1.00	1.00	1.00
Single person firm	0.03	0.04	0.00	0.00	0.00	1.00
Single establishment firm	0.66	0.65	0.11	0.64	0.63	0.06
Sector						
Agriculture & extraction	0.03	0.04	0.00	0.02	0.03	0.00
Manufacturing	0.32	0.27	0.00	0.33	0.29	0.00
Construction & utilities	0.08	0.06	0.00	0.06	0.04	0.00
Commerce	0.21	0.24	0.00	0.19	0.25	0.00
Services	0.37	0.39	0.00	0.39	0.38	0.04
Region						
North	0.04	0.05	0.00	0.05	0.06	0.00
Northeast	0.15	0.11	0.00	0.16	0.11	0.00
Central	0.09	0.06	0.00	0.11	0.07	0.00
South	0.22	0.20	0.00	0.22	0.20	0.00
Southeast	0.50	0.58	0.00	0.46	0.56	0.00
N establishments	18,103	62,028		13,677	48,075	
N workers	3,588,153	7,879,607		2,740,517	6,535,958	

Notes: Table compares descriptive statistics of establishments between the treated (Columns 1 and 4) and comparison groups (Columns 2 and 5) in our analysis samples, i.e., the amenity sample and the establishment sample. The p-values of the differences between the treated and comparison groups are reported in Columns 3 and 6. The bottom of the table includes the number of unique establishments and workers in each group.

TABLE A.8

Effect of CUT Reform on Negotiated Amenities: Cluster at Union-Level

	Intuitive definition (female clauses)					Data-driven		
	All	Leave	Maternity	Harassment	Flexibility	Female	Male	F/(F+M+1)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Intensive margin (number of clauses)								
$D_i \times \delta_{year \geq 2015}$	0.156*	0.078**	0.042*	0.009**	0.028	0.302**	0.130	0.032*
	(0.083)	(0.040)	(0.023)	(0.004)	(0.031)	(0.144)	(0.159)	(0.017)
Mean outcome	0.94	0.25	0.23	0.02	0.45	1.58	2.55	0.15
Panel B: Intensive margin (sum of unique clause types)								
$D_i \times \delta_{year \geq 2015}$	0.123*	0.047	0.042*	0.008**	0.027	0.155*	0.067	
	(0.067)	(0.031)	(0.022)	(0.004)	(0.021)	(0.080)	(0.095)	
Mean outcome	0.69	0.18	0.20	0.02	0.30	1.26	1.58	
Panel C: Extensive margin								
$D_i \times \delta_{year \geq 2015}$	0.017	0.012	0.020*	0.008**	0.022	0.034*	-0.001	
	(0.015)	(0.011)	(0.012)	(0.004)	(0.015)	(0.020)	(0.015)	
Mean outcome	0.31	0.12	0.15	0.02	0.23	0.36	0.46	
Panel D: As a share of all clauses								
$D_i \times \delta_{year \geq 2015}$	0.005	0.001	0.001	0.000	0.003	0.021	-0.003	
	(0.004)	(0.001)	(0.001)	(0.000)	(0.003)	(0.015)	(0.006)	
Mean outcome	0.05	0.01	0.01	0.00	0.03	0.07	0.14	
Observations	600,840	600,840	600,840	600,840	600,840	600,840	600,840	600,840

Notes: Table reports the coefficients for DID regressions—see Equation (3)—estimating the effect of the CUT reform on the female-centric and male-centric amenities included in CBAs. Columns correspond to different clause groupings and each panel provides a different margin. Panel A reports effects on the total number of clauses in the grouping, an intensive margin measure of amenities. Panel B reports effects on the sum of unique clause types in the grouping, capturing changes to the *space* of female- and male-centric clauses, as opposed to their number. Panel C reports effects on an indicator for whether any clause of the corresponding grouping exists in a contract, i.e., an extensive margin measure of amenities. Panel D uses the share of clauses in the grouping among all clauses in a contract. Under each panel we report the mean of the dependent variable among the treated at baseline (2014). The sample is the filled panel of establishment-union pairs by year. All columns control for pair fixed effects, as well as time-varying state and industry fixed effects. Standard errors are clustered at the union level, instead of at the establishment level, which reduces the number of clusters from around 80 thousand to about 4.4 thousand.

TABLE A.9

Effect of CUT Reform on Negotiated Amenities: CBA coverage in 2014

	Intuitive definition (female clauses)					Data-driven		
	All (1)	Leave (2)	Maternity (3)	Harassment (4)	Flexibility (5)	Female (6)	Male (7)	F/(F+M+1) (8)
Panel A: Intensive margin (number of clauses)								
$D_i \times \delta_{year \geq 2015}$	0.099***	0.044***	0.022***	0.005***	0.028***	0.121***	0.111***	0.009***
	(0.015)	(0.006)	(0.004)	(0.001)	(0.010)	(0.023)	(0.031)	(0.002)
Mean outcome	1.62	0.43	0.39	0.03	0.77	2.71	4.38	0.25
Panel B: Intensive margin (sum of unique clause types)								
$D_i \times \delta_{year \geq 2015}$	0.072***	0.023***	0.023***	0.003***	0.022***	0.077***	0.051***	
	(0.010)	(0.004)	(0.004)	(0.001)	(0.005)	(0.014)	(0.016)	
Mean outcome	1.19	0.31	0.35	0.03	0.51	2.17	2.71	
Panel C: Extensive margin								
$D_i \times \delta_{year \geq 2015}$	0.020***	0.012***	0.011***	0.004***	0.021***	0.005*	0.009**	
	(0.003)	(0.002)	(0.002)	(0.001)	(0.003)	(0.003)	(0.003)	
Mean outcome	0.53	0.21	0.25	0.03	0.40	0.62	0.79	
Panel D: As a share of all clauses								
$D_i \times \delta_{year \geq 2015}$	0.005***	0.001***	0.001***	0.000***	0.003***	0.004***	0.001	
	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.002)	
Mean outcome	0.08	0.01	0.01	0.00	0.06	0.11	0.25	
Observations	366,516	366,516	366,516	366,516	366,516	366,516	366,516	366,516

Notes: Table reports the coefficients for DID regressions—see Equation (3)—estimating the effect of the CUT reform on the female-centric and male-centric amenities included in CBAs. Columns correspond to different clause groupings and each panel provides a different margin. Panel A reports effects on the total number of clauses in the grouping, an intensive margin measure of amenities. Panel B reports effects on the sum of unique clause types in the grouping, capturing changes to the *space* of female- and male-centric clauses, as opposed to their number. Panel C reports effects on an indicator for whether any clause of the corresponding grouping exists in a contract, i.e., an extensive margin measure of amenities. Panel D uses the share of clauses in the grouping among all clauses in a contract. Under each panel we report the mean of the dependent variable among the treated at baseline (2014). The sample is the filled panel of establishment-union pairs by year, restricted to establishment-union pairs with CBA coverage in 2014. All columns control for pair fixed effects, as well as time-varying state and industry fixed effects. Standard errors are clustered at the establishment level.

TABLE A.10

Effect of CUT Reform on Negotiated Amenities: Modified Lasso for Data-Driven Clauses

	Intuitive definition (female clauses)					Data-driven		
	All	Leave	Maternity	Harassment	Flexibility	Female	Male	F/(F+M+1)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Intensive margin (number of clauses)								
$D_i \times \delta_{year \geq 2015}$	0.156***	0.078***	0.042***	0.009***	0.028***	0.315***	0.117***	0.030***
	(0.013)	(0.006)	(0.004)	(0.001)	(0.008)	(0.027)	(0.019)	(0.002)
Mean outcome	0.94	0.25	0.23	0.02	0.45	2.11	1.37	0.20
Panel B: Intensive margin (sum of unique clause types)								
$D_i \times \delta_{year \geq 2015}$	0.123***	0.047***	0.042***	0.008***	0.027***	0.128***	0.079***	
	(0.010)	(0.004)	(0.004)	(0.001)	(0.004)	(0.017)	(0.011)	
Mean outcome	0.69	0.18	0.20	0.02	0.30	1.53	0.88	
Panel C: Extensive margin								
$D_i \times \delta_{year \geq 2015}$	0.017***	0.012***	0.019***	0.008***	0.022***	0.036***	-0.006**	
	(0.003)	(0.002)	(0.002)	(0.001)	(0.003)	(0.003)	(0.003)	
Mean outcome	0.31	0.12	0.14	0.02	0.23	0.35	0.32	
Panel D: As a share of all clauses								
$D_i \times \delta_{year \geq 2015}$	0.005***	0.001***	0.001***	0.000***	0.003***	0.020***	0.001	
	(0.001)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	
Mean outcome	0.05	0.01	0.01	0.00	0.03	0.08	0.06	
Observations	600,840	600,840	600,840	600,840	600,840	600,840	600,840	600,840

Notes: Table reports the coefficients for DID regressions—see Equation (3)—estimating the effect of the CUT reform on the female-centric and male-centric amenities included in CBAs. The lasso for selecting clauses from the data-driven approach includes state and industry fixed effects. Columns correspond to different clause groupings and each panel provides a different margin. Panel A reports effects on the total number of clauses in the grouping, an intensive margin measure of amenities. Panel B reports effects on the sum of unique clause types in the grouping, capturing changes to the *space* of female- and male-centric clauses, as opposed to their number. Panel C reports effects on an indicator for whether any clause of the corresponding grouping exists in a contract, i.e., an extensive margin measure of amenities. Panel D uses the share of clauses in the grouping among all clauses in a contract. Under each panel we report the mean of the dependent variable among the treated at baseline (2014). The sample is the filled panel of establishment-union pairs by year. All columns control for pair fixed effects, as well as time-varying state and industry fixed effects. Standard errors are clustered at the establishment level.

TABLE A.11

Effect of CUT Reform on Female Amenities: Intensive Margin Robustness

	Female-Centric Clauses: Intensive Margin					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Intuitive definition						
$D_i \times \delta_{year \geq 2015}$	0.156***	0.156***	0.156***	0.193***	0.297***	0.099***
	(0.013)	(0.013)	(0.013)	(0.014)	(0.019)	(0.015)
Mean outcome	0.94	0.94	0.94	0.94	0.94	1.62
Panel B: Data-driven definition						
$D_i \times \delta_{year \geq 2015}$	0.302***	0.347***	0.262***	0.331***	0.417***	0.121***
	(0.021)	(0.026)	(0.017)	(0.022)	(0.030)	(0.023)
Mean outcome	1.58	2.05	1.17	1.58	1.58	2.71
Data-driven clauses	baseline	any	all	baseline	baseline	baseline
Geography-year FEs	state	state	state	microregion	micro×ind	state
CBA coverage in 2014	no	no	no	no	no	yes
Observations	600,840	600,840	600,840	600,840	600,840	366,516

Notes: Table reports the coefficients for DID regressions—see Equation (3)—estimating the effect of the CUT reform on female amenities included in CBAs. The dependent variable is the total number of clauses per pair-year as an intensive margin measure, with Panel A using the intuitive definition of female-centric clauses and Panel B using the data-driven approach. Columns (1)-(3) modify the dependent variable by changing the clauses that are chosen as female-centric in the data-driven approach: a) *baseline*: top 20 clauses using a random sample and normalized PageRank values for the gender gaps; b) *any*: counts any of the clauses selected across 6 approaches as female-centric; c) *all*: counts only those clauses that are selected in all 6 approaches as female-centric. Refer to Table A.4 for a list of the clauses used in each of these scenarios. Columns 4 adds more granular time-varying fixed effects at the geographic level, i.e., using microregion instead of state. Columns 5 uses a microregion-industry time-varying fixed effect. Column 6 requires that pairs are covered by a CBA at baseline to test whether effects are driven by changes in the amenities among units with active CBAs rather than by gains in coverage. Standard errors are clustered at the establishment level.

Effect of CUT Reform on Female Amenities: Share of Clauses Robustness

	Female-Centric Clauses: As a Share of All Clauses					
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Intuitive definition						
$D_i \times \delta_{year \geq 2015}$	0.005***	0.005***	0.005***	0.005***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Mean outcome	0.05	0.05	0.05	0.05	0.05	0.08
Panel B: Data-driven definition						
$D_i \times \delta_{year \geq 2015}$	0.021***	0.021***	0.022***	0.017***	0.011***	0.004***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Mean outcome	0.07	0.08	0.04	0.07	0.07	0.11
Data-driven clauses	baseline	any	all	baseline	baseline	baseline
Geography-year FEs	state	state	state	microregion	micro×ind	state
CBA coverage in 2014	no	no	no	no	no	yes
Observations	600,840	600,840	600,840	600,840	600,840	366,516

Notes: Table reports the coefficients for DID regressions—see Equation (3)—estimating the effect of the CUT reform on female amenities included in CBAs. The dependent variable is the share of female-centric clauses among all clauses per pair-year, with Panel A using the intuitive definition of female-centric clauses and Panel B using the data-driven approach. Columns (1)-(3) modify the dependent variable by changing the clauses that are chosen as female-centric in the data-driven approach: a) *baseline*: top 20 clauses using a random sample and normalized PageRank values for the gender gaps; b) *any*: counts any of the clauses selected across 6 approaches as female-centric; c) *all*: counts only those clauses that are selected in all 6 approaches as female-centric. Refer to Table A.4 for a list of the clauses used in each of these scenarios. Column 4 adds more granular time-varying fixed effects at the geographic level, i.e., using micro-region instead of state. Columns 5 uses a microregion-industry time-varying fixed effect. Column 6 requires that pairs are covered by a CBA at baseline to test whether effects are driven by changes in the amenities among units with active CBAs rather than by gains in coverage. Standard errors are clustered at the establishment level.

TABLE A.13

Heterogeneity by Low Representation of Women in the Workplace vs. Union

	Baseline (1)	Low % women in workpl and union (2)	Low % women in workpl only (3)	Low % women in union only (4)
Panel A: Female clauses (data-driven)				
Intensive margin	0.302***	0.476***	0.184*	0.198***
	(0.021)	(0.031)	(0.105)	(0.038)
As a share of all clauses	0.021***	0.033***	0.012***	0.012***
	(0.001)	(0.002)	(0.004)	(0.002)
Panel B: Male clauses (data-driven)				
Intensive margin	0.130***	0.212***	0.008	0.050
	(0.029)	(0.043)	(0.149)	(0.059)
As a share of all clauses	-0.003**	-0.002	-0.011*	-0.006*
	(0.002)	(0.002)	(0.006)	(0.003)
Observations	600,840	267,336	35,934	165,138

Notes: Table shows the effect of the CUT reform on female- and male-centric clauses (data-driven approach) according to the baseline representation of women in the workplace versus union boards. Each cell in the table reports the difference-in-differences coefficient for a separate regression, where the dependent variable depends on the panel (female vs. male clauses) and the row (intensive vs. share margin). The columns denote the sample on which the DID specification in Equation (3) is run. Column (1) is the full analysis sample. Column (2) is restricted to establishments where the 2014 share of female workers and union board members are both below 1/3. Column (3) is restricted to establishments where the 2014 share of female workers (and not union board members) is below 1/3. Column (4) is restricted to establishments where the 2014 share of female union board members (and not workers) is below 1/3. All regressions use the filled panel sample and includes establishment-union pair fixed effects as well as time-varying state and industry fixed effects. Standard errors are clustered at the establishment level.

TABLE A.14

Impact of CUT Reform on Composition of the Female Workforce

	Share poached in (1)	Mean years of age (2)	Mean months of tenure (3)	Mean hours in contract (4)	Mean years of schooling (5)
$D_i \times \delta_{year \geq 2015}$	-0.001	-0.012	0.172	-0.033	-0.001
	(0.002)	(0.041)	(0.215)	(0.025)	(0.010)
Mean outcome	0.209	33.5	43.1	42.0	11.3
Observations	342,207	342,207	342,207	342,207	342,207

Notes: Table reports the coefficients for the establishment-level DID regression from Equation (3), comparing treated to comparison establishments on characteristics of their female workforce. An establishment is treated if the union with which it negotiates is affiliated to CUT in 2012. Each regression includes establishment fixed effects, industry-year fixed effects, and microregion-year fixed effects. Standard errors are clustered by establishment and reported in parentheses.

Downstream Effects by Representation of Women in the Union

	Fem clauses [intensive] (1)	Fem clauses [share] (2)	Share women [managers] (3)	Share extend [mat leave] (4)	Share women [workforce] (5)	Share women [probation] (6)	Retention [women; 20-35] (7)	Retention [men; 20-35] (8)	Absences (9)
High % women in union	0.001	-0.004**	0.005	0.009**	-0.000	0.009*	0.003	0.005	-0.004
	(0.038)	(0.002)	(0.006)	(0.004)	(0.002)	(0.006)	(0.006)	(0.007)	(0.208)
Low % women in union	0.364***	0.025***	0.005*	0.009***	0.003**	0.006*	0.020***	0.014***	-0.219*
	[0.000]	[0.000]	[0.093]	[0.009]	[0.020]	[0.064]	[0.000]	[0.000]	[0.098]
Difference: low vs. high	0.364***	0.030***	-0.001	-0.001	0.003	-0.004	0.017*	0.008	-0.215
	(0.041)	(0.002)	(0.006)	(0.005)	(0.003)	(0.006)	(0.007)	(0.007)	(0.241)
Mean outcome	1.58	0.07	0.33	0.06	0.37	0.36	1.00	1.00	4.76
Observations	592,224	592,224	256,468	138,142	349,131	275,459	10,794,708	18,696,126	331,985

Notes: Table shows heterogeneity in the effect of the CUT reform on female-centric clauses (data-driven approach) and downstream outcomes according to the baseline representation of women in the union boards. The dummy to test for heterogeneity in the effects (H_i) is fully interacted with the treatment dummy (D_i) and the post-period dummy ($\delta_{year \geq 2015}$). In this table, H_i is an indicator for whether the share of women in union boards is below 1/3 in 2014. The table reports the coefficients on the effects that determine the treatment effect for the baseline group (high % women in union) and the differential effect relative to the baseline group (i.e., difference: low vs. high)—with the sum of both coefficients representing the treatment effect for the group of interest (low % women in union). Columns denote the dependent variable in the regression: (1)-(2) use the filled panel sample, (3)-(6) and (9) use the establishment sample, and (7)-(8) use the incumbent sample weighting observations by the inverse of (own-gender) employment at the baseline employer. Standard errors are clustered at the establishment level. Standard errors are shown in parentheses; p-values are reported in brackets.

B APPENDIX FIGURES

FIGURE B1
Impact on Gender Representation in Local Union Boards

(a) Share of women in union board

(b) Woman president or vice-president

Notes: Figures show estimates of the δ_t coefficients for $t \in [2012, 2019]$ (with 2014 omitted) from an event-study specification similar to the one in Equation (3) on measures of women representation within local union boards. The sample is restricted to unions in our analysis sample (results are similar without imposing this restriction). The equation we estimate is slightly different from Equation (3) as the unit of observation here is the union-year, so we include union fixed effects instead of establishment-union pair fixed effects. Figure B1a uses the share of women in the union board as a dependent variable, while Figure B1b uses a dummy indicating whether the union's president (or vicepresident) is a woman. Confidence intervals at a 95% level are reported. Standard errors are clustered by union.

FIGURE B2

Gender Parity in National Leadership by Union Central
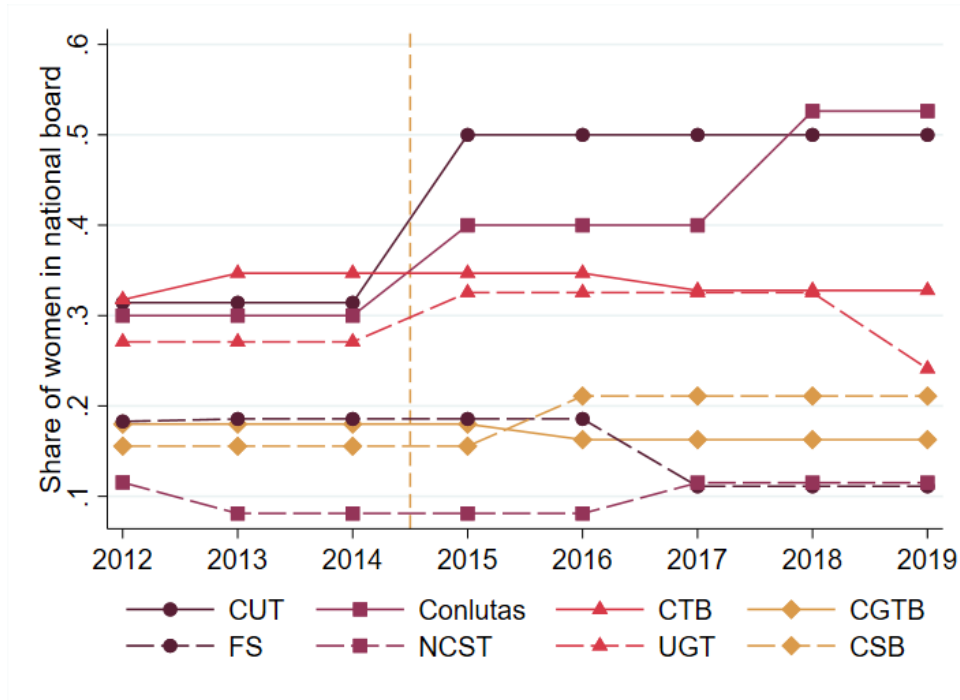
Notes: Figure plots the annual share of women on each union central's national executive committee (*Inter-sindical* is dropped due to missing information on its board). The line for CUT is the same as in Figure IIb, while the unweighted average of all other union centrals make up the other line reported in Figure IIb. Solid lines refer to "combative" union centrals, while dashed lines represent "cooperative" union centrals. The second largest union central and main competitor to CUT is *Força Sindical* (FS). The only other union central that achieves gender parity is *Conlutas*, which was established as an offshoot of the CUT in 2004. As such, *Conlutas* often emulates the CUT but has a very small number of affiliated unions, all of which represent workers in the public sector

FIGURE B3
Example of a Maternity Leave Clause

Notes: Figure shows an example of a maternity leave clause in a CBA. The clause is classified under the "Holidays and Licenses" broad group (9 in total) and the "Maternity Leave" clause types (137 in total). This particular clause extends maternity leave duration from the state-mandated 120 days to 180 days—inclusive to adopting mothers. It also extends post-maternity job protection by 6 months. The paper relies on the clause type classification of the different clauses, ignoring the variation in the text that may exist within each individual clause belonging to a specific type.

FIGURE B4
Sense Checks for Female- and Male-Centric Amenities

(a) Intuitive female clauses and share of women   (b) Data-driven clauses and share of women
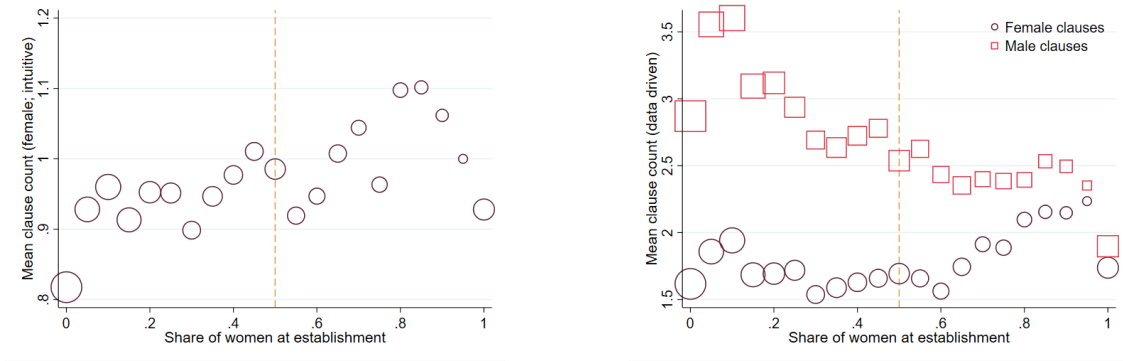
Notes: Figures depict binned scatterplots of the number of female-centric (and male-centric) clauses contained in firm-level CBAs signed at baseline (2014) by the share of women in the workforce of the establishment. The bins in the bottom figures are set to rounded values (in 0.05 increments) of the share of women at the establishment, with the size of the markers scaled to represent the number of pairs observed in a given bin. Figure B4a uses the intuitive definition of female-centric amenities, while Figure B4b uses the data-driven approach for both female- and male-centric amenities. The vertical line indicates 50% of women in the workforce. The sample consists of the establishments in our new contracts panel at baseline (2014). Regressing the y-axis variables in the bottom figures on the share of women at establishments reveals a positive (negative) and statistical significant relation between female (male) centric clauses and the share of women at the establishment. For the intuitive definition of female-centric clauses, the slope is 0.141 (SE 0.048). For the data-driven definition of female-centric clauses, the slope is 0.381 (SE 0.148). For the data-driven definition of male-centric clauses, the slope is -1.121 (SE 0.204).

FIGURE B5
Gender Gaps in PageRank Values and Female-Centric Amenities

(a) Intuitive female clauses

(b) Data-driven female clauses
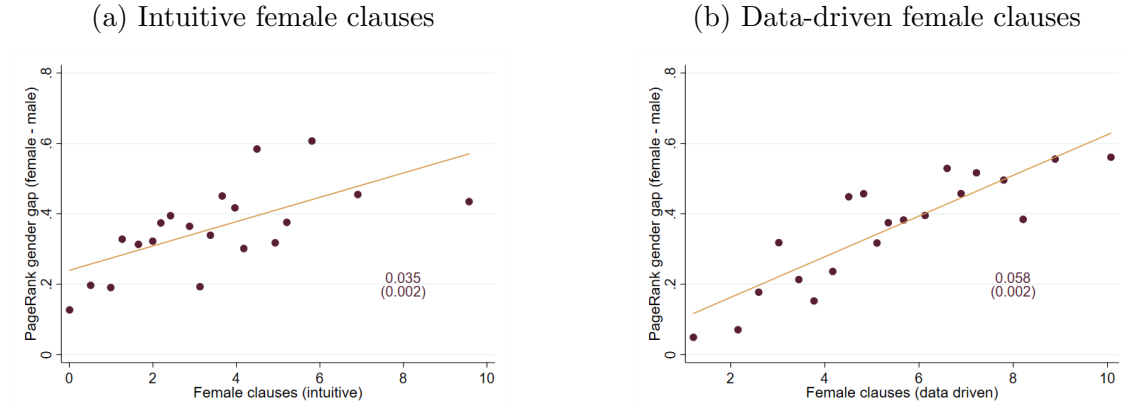
Notes: Figures depict binned scatterplots of the establishment-level gender gaps in PageRank values by the average female-centric clauses from sectoral CBAs applying to the establishment. Figure B5a uses the intuitive definition of female-centric amenities, while Figure B5b uses the data-driven approach. The sample used is the one used to estimate Equation (2), i.e., establishments in the intersection of the gender-specific super-connected sets covered by sectoral CBAs in at least 4 different years between 2009-2016.

FIGURE B6
Union Affiliation to CUT Over Time
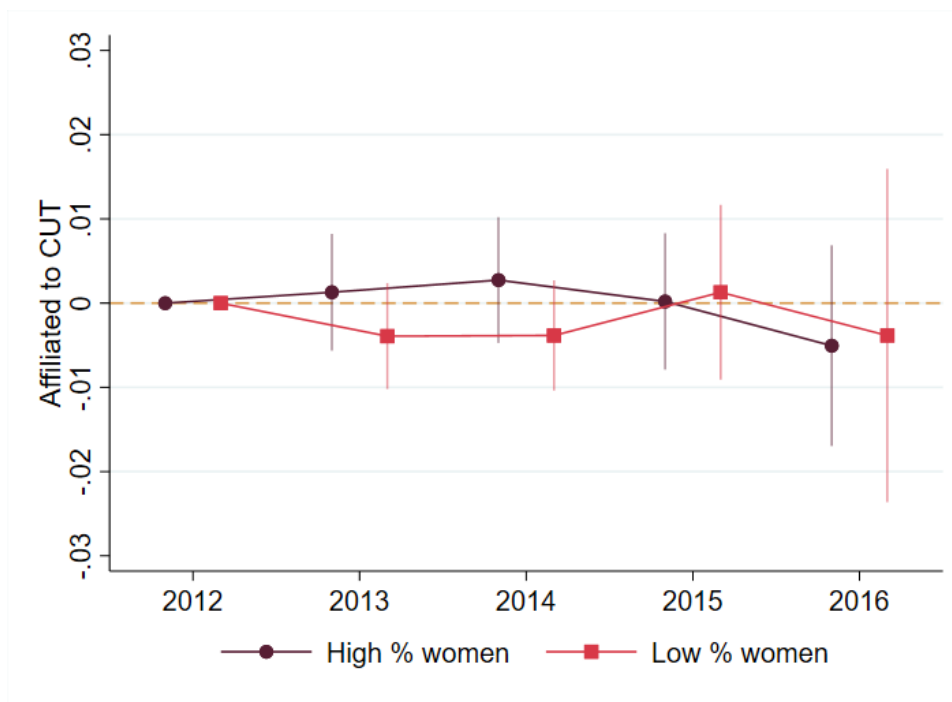
Notes: Figure plots changes in the probability of being affiliated to CUT between 2012 and 2016 separately for unions having either a high or a low share of women among the workers they represent (above or below the mean, i.e., 33% women). Coefficients represent the change with respect to 2012, in which the probability of being a CUT-affiliate is normalized to zero. Unions are weighted by the size of the workforce that they represent, computed by summing the 2012 worker count across establishments negotiating firm-level CBAs with the union. That is, if an establishment negotiates with n unions, we split the workforce count evenly to those n unions (results are robust to removing these weights). The sample is restricted to the unions in the filled panel. Only 3% of unions ever switch affiliation to or from CUT in this sample. Standard errors are clustered at the union level.

Baseline Characteristics of Treated and Control Establishments

(a) Establishment size

(b) Industry

(c) Region

(d) Share of women employees
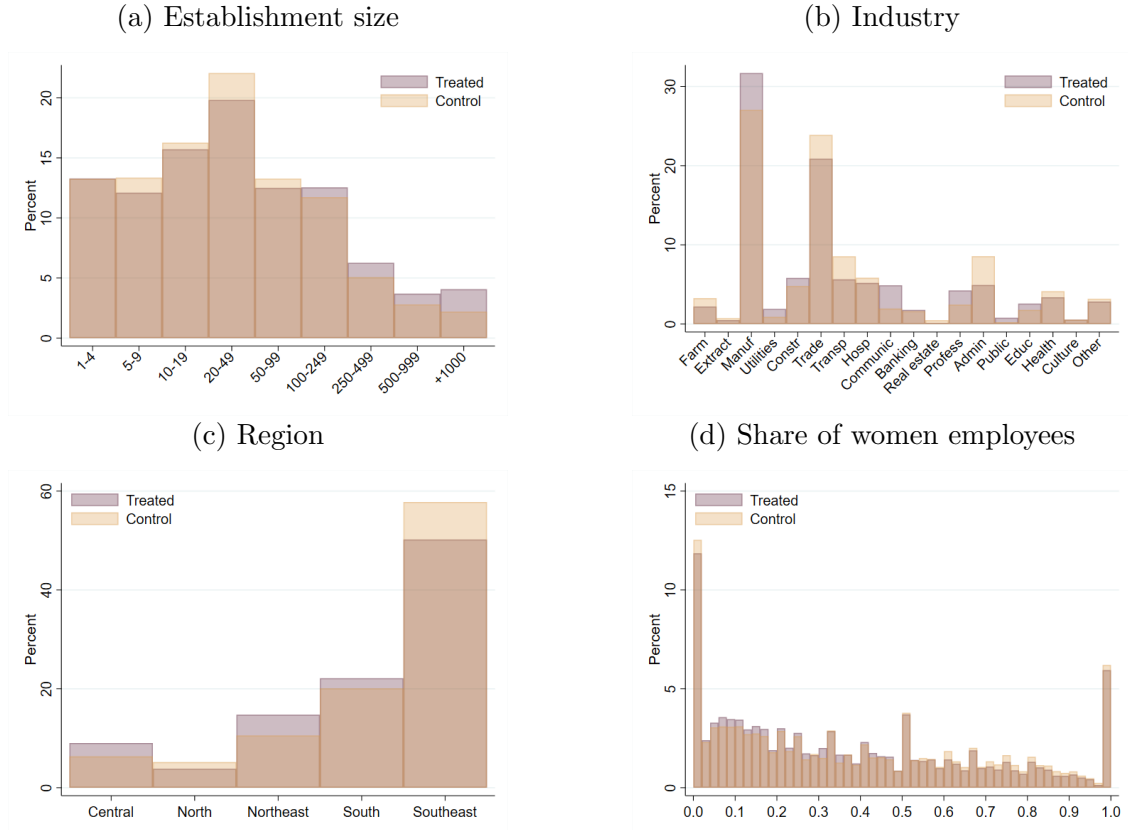
Notes: Figures show the treated and control establishments distributions of size, industry, regional location, and female share of employment at baseline. The establishments come from the starting sample detailed in Table II.

FIGURE B8
Trends in Female-Centric Clauses

(a) Filled panel

(b) New contracts
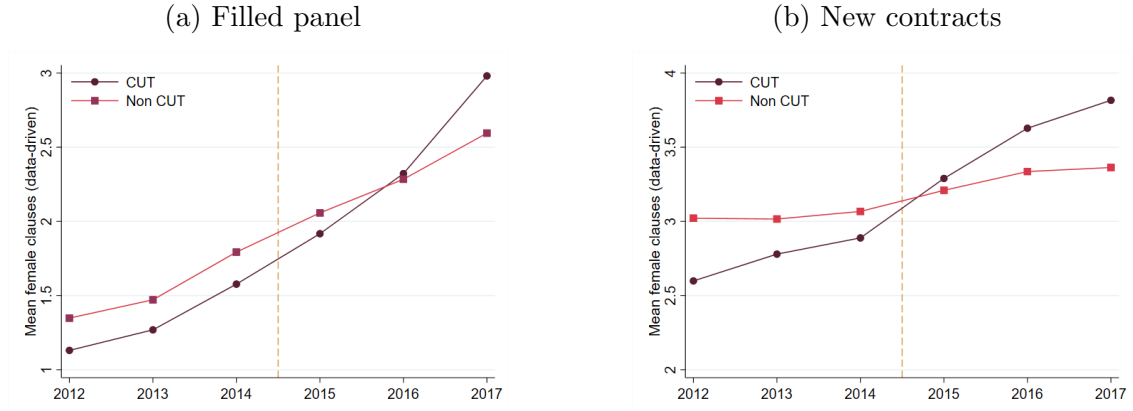
Notes: Figures plot the raw average number of female-centric clauses for treated (CUT) and control (non-CUT) establishment-union pairs over the years. Female-centric clauses are based on the data-driven classification. Figure B8a plots the average number of female-centric clauses for the filled panel, while Figure B8b plots the average number of female-centric clauses in newly signed contracts of the given year. Mean female clauses are lower in the filled panel and react slowly to changes in new contracts because of pairs that do not have CBA coverage in a given year.

FIGURE B9

Trends in Male-Centric Clauses and the Gender Ratio of Clauses

(a) Male-centric clauses

(b) Female-to-male ratio
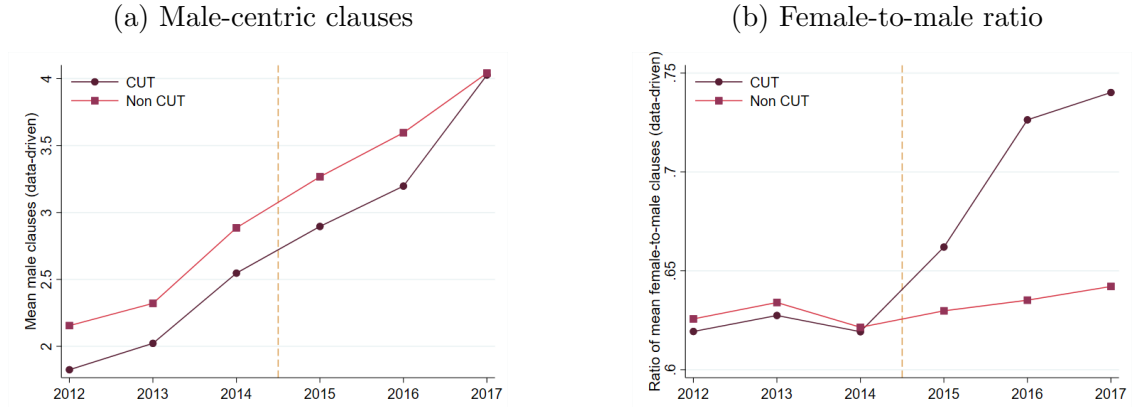
Notes: Figures plot the raw trends in clauses for treated (CUT) and control (non-CUT) establishment-union pairs over the years in the filled panel. Both female- and male-centric clauses are based on the data-driven classification. Figure B9a plots the average number of male-centric clauses, while Figure B9b plots the ratio of the mean female-to-male clauses.

FIGURE B10
Effect of the CUT Reform on Female-Centric Amenities

(a) Intensive margin: number of clauses
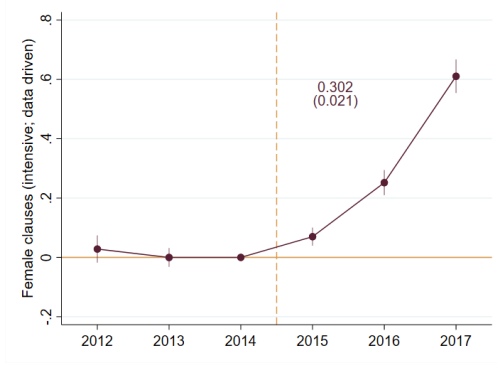
(b) Intensive margin: unique clause types
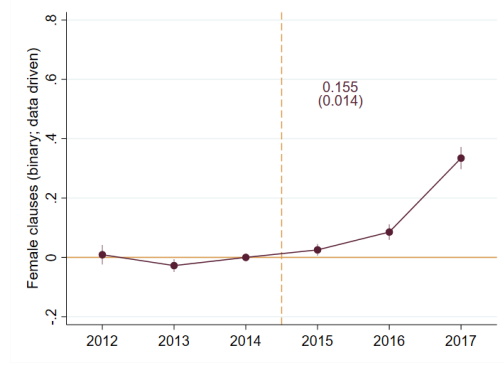
(c) Extensive margin
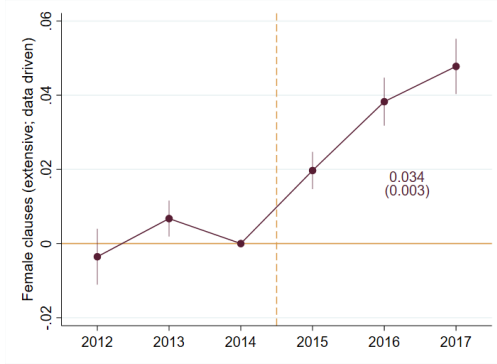
(d) As a share of clauses
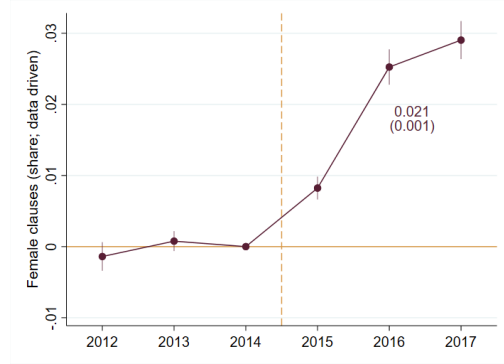
Notes: Figures show estimates of the δ_t coefficients for $t \in [2012, 2017]$ (with 2014 omitted) from the DID specification in Equation (3) on all margins considered for female-centric clauses, defined using the data-driven method. Confidence intervals at a 95% level are reported. Standard errors are clustered at the establishment level. All figures use the filled panel.

FIGURE B11
Effect on Amenities by Low Representation of Women in the Workplace vs. Union
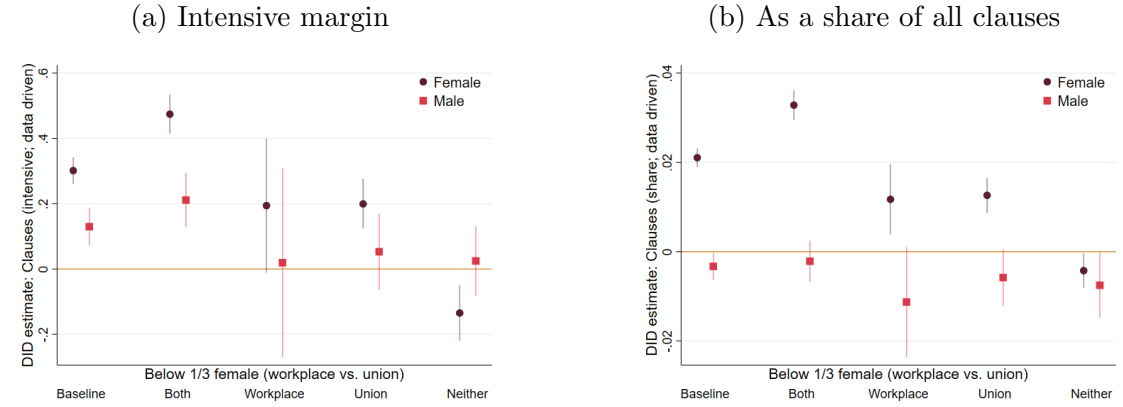
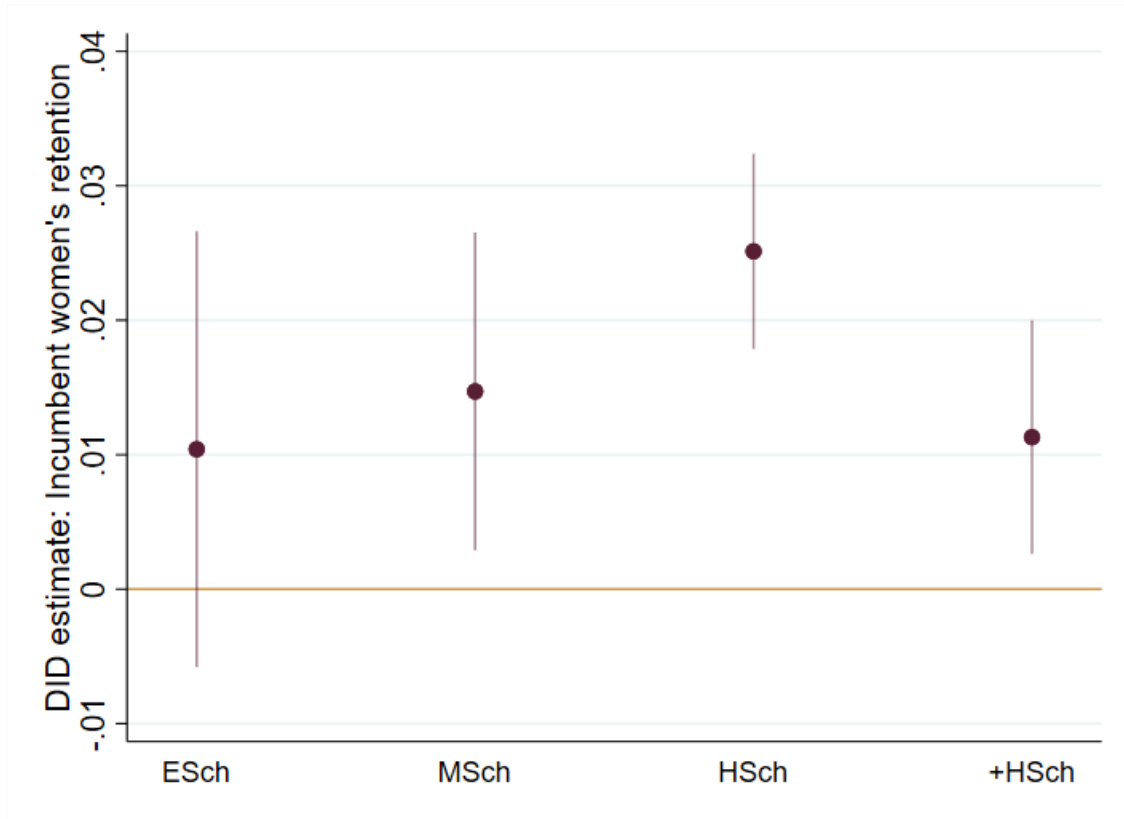| (a) Intensive margin	(b) As a share of all clauses

Notes: Figures show estimates of the treatment effect ($\delta_{year \geq 2015}$) from the DID specification in Equation (3) on the number of female- and male-centric clauses (data-driven approach) computed on the full analysis sample (*Baseline*) and four mutually exclusive and collectively exhaustive subsamples based on the representation of women at the workplace versus the union. Low representation of women in the workplace (union) refers to having a 2014 share of female workers (union board members) below 1/3. The *Both* (*Neither*) subsample denotes when there is low female representation in both (neither) the workplace and (nor) the union. The *Workplace* (*Union*) subsample denotes when there is low female representation in the workplace (union) only. We use the filled panel. Confidence intervals at a 95% level are shown. Standard errors are clustered at the establishment level.

FIGURE B12
Effects on Wages and Employment

(a) Wages: new female workers

(b) Wages: new male workers

(c) Log employment

(d) Log wage bill
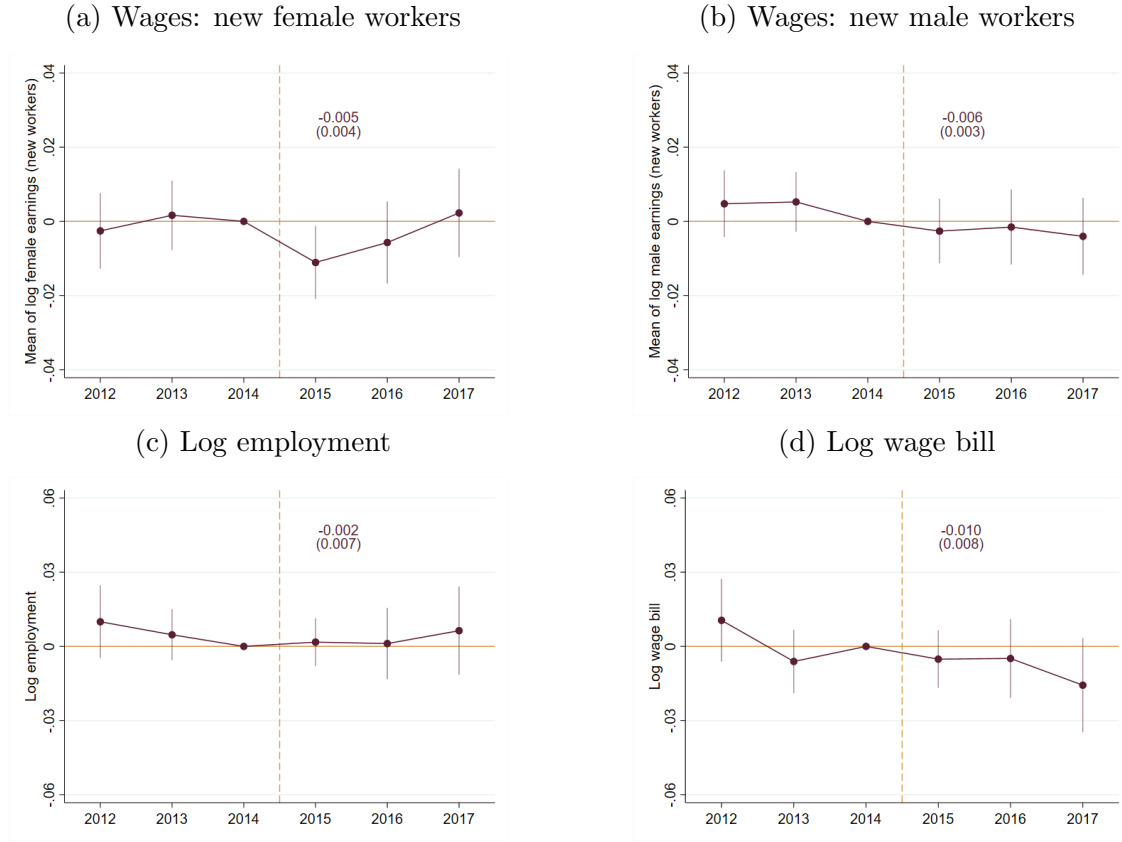
Notes: Figures report the results of the establishment-level DID regression in Equation (3) with outcome variables: mean log wages for new female hires (i.e., tenure below 12 months), mean log wages for new male hires (i.e., tenure below 12 months), log of total employment, and log of the wage bill. Each regression includes establishment fixed effects, industry-year fixed effects, and microregion-year fixed effects. The figure plots estimates of the δ_t coefficients for $t \in [2012, 2017]$ with 2014 omitted. Confidence intervals at a 95% level are reported. Standard errors are clustered by establishment.

FIGURE B13
Heterogeneity in Women's Retention by Education Group

Notes: Figure show estimates of the treatment effect ($\delta_{year \geq 2015}$) from the DID specification in Equation (3) on retention using subsamples of incumbents by education groups. We split the sample by level of education at baseline, i.e., completed elementary or less (ESch), some or completed middle school (MSch), some or completed high school (HSch), and more than high school (+HSch). To make treatment effects in worker-level regressions interpretable as establishment-level averages, we weight each incumbent worker by the inverse of (own-gender) employment at their baseline employer. Confidence intervals at a 95% level are shown. Standard errors are clustered at the establishment level.

FIGURE B14

Spillover Effects at Multi-Establishment Firms
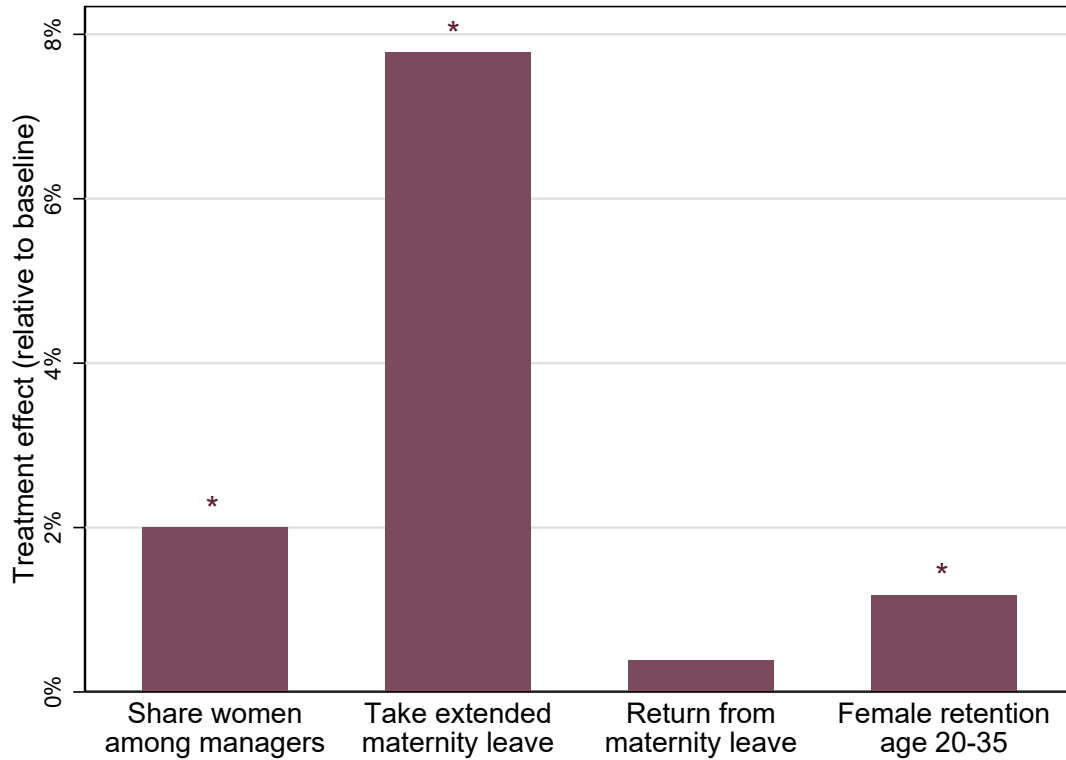
Notes: Figure shows the effects (in percentage terms relative to baseline) from DID regressions—see Equation (3)—estimating the spillover effect of the CUT reform at multi-establishment firms on firm environment outcomes and female retention. In this analysis, for each one of our samples, we use the first 8 digits of the CNPJ identifier to observe whether establishments in the control group are part of a firm that has an establishment in the treated group. Once that has been determined, we drop all treated observations and redefine treatment ($D_i = 1$) as an observation in the original control group whose establishment belongs to a firm that has an originally defined treated establishment. The regressions for the first three bars use the establishment sample and are related to the firm environment (see Figure V). The regression for the last bar uses the incumbent sample and is weighted by the inverse of (own-gender) employment at the baseline employer. Standard errors are clustered at the establishment level.

C DATA APPENDIX

C.A Sample construction

To analyze the CUT reform's impact on various outcomes, we construct three main analysis samples. The first is a sample to study changes in CBA clauses at the establishment-union pair level (henceforth, simply *pair* level). The second is a sample at the establishment level to study changes in the workplace. The third is a sample at the worker level used to track the labor market outcomes of incumbent workers. In addition to these three main samples, we also construct two panel datasets at the local union level and at the union central level to study the gender composition of their boards.

Amenities sample   Amenities (on paper) are captured by negotiated CBA clauses signed by establishment-union pairs. We first construct a yearly panel of the new CBAs signed by a pair in a given year, i.e., new contracts. We then use this sample to construct a balanced panel containing the active clauses applying to a pair over time, i.e., filled panel.

1. New contracts: We construct this sample using the set of CBAs registered on *Sistema Mediador*. We restrict to valid, non-amendment, firm-level CBAs signed between 2012 and 2017 (inclusive). Each CBA contains information on who signs the agreement—the CNPJ identifiers of the employer(s) and union(s) signing it—and, importantly for our analysis, how many clauses it contains classified into clause types.[1]

The union identifier allows us to merge these data with data on union affiliation to union centrals coming from CNES. The employer identifier allows us to merge these data with information in RAIS, e.g., industry, microregion, and employment. We drop CBAs signed by unions with missing information about their 2012 union central affiliation (around 1.5% of contacts).[2] We additionally drop contracts signed by multiple unions with different union central affiliations: this is fewer than 0.33% of CBAs.[3]

Almost all pairs negotiate at most one contract per year: 96% of CBAs are the only agreement signed by a pair that year and 85% of pairs always negotiate at most one CBA per year during our study period. As for the remaining 15%, we take the maximum count of a given clause type across the CBAs negotiated by the pair in a given year.[4] In this way

---

[1]*Sistema Mediador* classifies clauses into 137 categories, e.g., maternity assistance, overtime pay, life insurance, procedures in relation to strikes and strikers, etc.

[2]Unions that decide not to affiliate with any union central—which are registered in CNES as "Not-Affiliated"—are not dropped. The CBAs signed by these unions are part of the control group.

[3]Of the remaining agreements, 89.8% are negotiated between a single establishment and a single union, 7.3% are negotiated by a single union with two or more establishments, 2.5% are signed by one establishment and two or more unions with the same CUT or non-CUT affiliation, and only 0.5% by multiple unions and multiple establishments.

[4]We do this to avoid double-counting clauses as the multiple agreements per pair-year often result from

we obtain a sample of newly negotiated CBAs at the pair-year level, reporting the number of clauses for each clause type.

On the signing establishment's side, we restrict to pairs where the employer has non-missing industry and microregion information, and it employs workers at baseline (2014). These restriction drop 8.5% of observations. This comprises the starting sample with observations at the pair-year level reported in the descriptive statistics of Table II.

2. Filled panel: This sample fills in the amenities information for pairs in the *new contracts* sample for years when a new firm-level CBA was not signed. In filling the panel, we consider the institutional context regarding the automatic extension of CBAs into the future. That is, for a given pair, contracts expiring after September 2012 are automatically extended into the future until a new CBA is signed (Lagos, 2024). Although CBAs expiring before that date were not extended, we observe contracts starting 3 years prior to our study period, i.e., starting in 2009. Since the maximum duration of a CBA is 24 months, by the start of our study period (i.e., 2012) we can already be certain whether any CBA applies to a given pair-year. As such, these institutional features allow us to generate a balanced panel at the pair-year level.

To aggregate amenities at the pair-year level, for each year we only consider the contract(s) covering at least 6 months of the year.[5] If more than one contract per pair-year remains, we take the maximum count of a given clause type across CBAs—similarly to what done for the *new contracts* sample. If a pair is not covered by a firm-level CBA in a given year (even after filling the panel), we set the clause count for each clause type to zero. As such, this procedure produces a yearly balanced panel at the establishment-union pair level.

Establishment sample To study changes in the workplace, we match the contracts in our *amenities sample* to the signing establishments in RAIS. Establishments covered by contracts negotiated by unions affiliated to CUT in 2012 form our treatment group, while establishments covered by CBAs signed by unions not affiliated to CUT in 2012 make up our comparison group.

We start with the list of establishments that are part of the pairs in our *new contracts* sample. We restrict to establishments employing both men and women at baseline, dropping 15,550 establishments. We further restrict this list to establishments in the geographic coverage of their "baseline CBA", defined as the firm-level agreement closest to the 2015 CUT reform among those signed by the establishment. The reason for this restriction is that, for multiple-establishment firms, the CNPJ listed as the employer counterpart in the CBA need

---

misclassified CBA amendments or single-issue CBAs that are renegotiated more frequently than a year.

[5]All other restrictions used in the *new contracts* sample apply.

not be covered.[6] Restricting to signing establishments in the geographic coverage of their baseline CBA further drops 8,684 establishments, leaving us with 61,752 establishments.

For each establishment in this list we compute outcomes at the establishment-year level, such as mean log wages or total female employment, either using all job spells registered at that establishment in the year or using workers' "main job spell" in each year. We define the "main job spell" as the employment spell at which the worker worked the longest during the year. In case all job spells have the same duration, we break ties by keeping only one spell at random.

Because the same establishment can negotiate CBAs with more than one union, the final step to construct the *establishment sample* involves determining treatment status at the establishment level. We assign establishments to the treatment group as long as they are part of at least one treated pair. In practice, this decision is innocuous. Because the great majority (93.5%) of establishments always bargain with the same union, treatment assignment is trivially defined for most establishments. There are 4.4% of establishments that sign CBAs with more than one union over the time frame we consider, and all the unions they negotiate with have the same treatment status, e.g. they are all affiliated to CUT (or they are not) in 2012. The remaining 2.1% of establishments negotiate with more than one union over time and these unions have different treatment status. We conservatively assign this last group of establishment to the treatment group, which should run counter to finding effects if some of these establishments are not affected by the CUT reform.

Incumbent workers sample Incumbent workers are defined as those employed at a treated or comparison establishment as of 2014 (based on the *establishment sample*). Their treatment status depends on the treatment status of their baseline (2014) employer, as explained above in the description of the *establishment sample* construction. Leveraging the linked employer-employee feature of RAIS, incumbent workers are tracked across jobs from 2012 to 2017—that is, we are not restricting to job spells at employers in the *establishment sample*. In constructing this sample, we only consider the "main job spell" for each worker in each year.

Union and union central boards For each Brazilian union central, we construct a yearly panel with information on the gender composition of their national board between 2012 and 2019. The raw data contains the full name of all the board members, which allows us to infer their gender. We do so using the R package *genderBR*, which codes a name as female

---

[6]Firm-level CBAs apply to workers at all establishments of the signing firm that are in the geographic coverage specified in the contract. In case of multi-establishment firms, the establishment signing a CBA could be the firm headquarter but the contract might cover only subsidiaries located in other municipalities.

if most people with that name are women in the Brazilian census (and similarly for men).[7] We use this data to check that the introduction of the CUT gender quota had bite.

We similarly construct a yearly panel with information on the gender composition of local union boards, the gender of their presidents and vice-presidents, and their affiliation to union centrals between 2012 and 2019. We use these data 1) to assign treatment status to unions; 2) to understand whether the reform had spillovers on local union boards; and 3) conduct heterogeneity analyses concerning women's representation in unions.

C.B   Construction of variables

Amenities   In the analysis we adopt two different ways of classifying clauses as female-centric amenities. The first is guided by intuition to select clause types that are of plausibly of greater value to women than men (intuitive definition). The second definition is data-driven, where we use lasso to pick clauses that are most predictive of women's value of employment (relative to men) at an establishment in the cross-section. An important advantage of the data-driven approach—compared to the intuitive definition—is that it also identifies clauses that are valued relatively more by men, i.e., male-centric amenities.

We also generate four different outcome margins for clauses at the pair-year level. First, the *intensive margin (number of clauses)* measures the sum of the clause counts from the clause types categorized as either female- or male-centric in the corresponding contract. Second, the *intensive margin (sum of unique clause types)* measures the sum of clause type indicators for those categorized as either female- or male-centric in the corresponding contract. Third, the *extensive margin* simply indicates whether any female (or male) clause exists in the CBA of interest. Finally, we calculate the *share* of the intensive margin (count) relative to the total clause count in the CBA.

1. Intuitive definition: Guided by CUT's "fight plan" and previous work documenting the value women place on flexibility (Goldin and Katz, 2011; Mas and Pallais, 2017; Maestas et al., 2023), we identified 4 themes as female-centric: 1) leaves; 2) maternity and childcare; 3) workplace harassment and discrimination; and 4) flexibility and part-time work. From these themes we restricted ourselves to select 20 clause types. These clauses are listed in Table A.1—which includes clauses on maternity leave, childcare assistance, prevention of sexual harassment—all of which are conceivably of greater value to women than men.

2. Data-driven definition: The data-driven definition of amenities selects clauses that are most predictive of gender differences in the value of employment at an establishment, controlling for gender-specific wage premiums.[8] In practice, we estimate the following cross-

---

[7]Developed by Fernando Meireles and posted on GitHub (Meireles, 2023).
[8]Section III.B provides a detailed justification for this approach.

sectional specification using lasso:

$$V_j^F - V_j^M = \alpha + \beta_w^F \psi_j^F - \beta_w^M \psi_j^M + \sum_{z \in Z} \beta_z a(z)_j + \epsilon_j \tag{1}$$

where V_j^G is the PageRank value of establishment j for workers of gender G, ψ_j^G is the establishment fixed-effect for workers of gender G at employer j from an AKM regression on wages, and $a(z)_j$ is the average clause count of amenity z (one among the 137 clause types) offered in the CBAs covering workers. We select the 20 clause types with the highest β_z and label them as "female-centric" amenities. Conversely, the 20 clause types with the lowest β_z comprise our "male-centric" amenities. Results are shown in Table I.

PageRank values. To estimate PageRank values we take job spells of full-time workers, ages 18-54, on open-ended contracts, and earning monthly wages in private sector establishments from RAIS (2009-2016). For each gender, we find the largest strongly connected set of establishments based on worker flows, i.e., a link between two establishments is defined as having at least one inflow and one outflow. We restrict to establishments that have at least 10 hires overall, with at least one of these coming from non-employment. To solve for the vector of PageRank values (see Appendix E), we follow Morchio and Moser (2020) and only consider employment-to-employment flows to be month-to-month job transitions. In addition, we set the damping factor used in finding the fixed point in the linear system of normalized flows to 0.8—one of the standard values in computer science. That is, the "random surfer" moving through the labor market restarts his search at a new establishment with 80% probability. As shown in Sorkin, 2018, PageRank values are unique up to an unknown multiplicative factor. Below we discuss robustness to assumptions about the multiplicative factor applying to women versus men to obtain $V_j^F - V_j^M$.

Wage premiums. To estimate the establishment fixed effect from AKM we take job spells of full-time workers, ages 18-54, on open-ended contracts, and earning monthly wages in private sector establishments from RAIS (2009-2016). For each gender, we find the largest strongly connected set of establishments based on worker flows, i.e., a link between two establishments is defined as having at least one inflow and one outflow. We restrict to establishments that have at least 10 workers (on average across years) and are observed at least 4 years in RAIS. Following Gerard et al. (2021), the model includes dummies for individual workers (α_i) and individual establishments (ψ_j), year dummies interacted with five education dummies, and quadratic and cubic terms in age interacted with the education dummies (X_{it})—see Appendix E. For the baseline year, the worker effects are measured as of age 40 to correspond to the approximate peaks of experience profiles. The establishment fixed effects for each gender—i.e., ψ_j^F and ψ_j^M—are normalized relative to the restaurant

industry, where rents are assumed to be negligible.

Clause counts. To get a measure of $a(z)_j$ for each establishment, we take a yearly average of the number of clauses in each of the 137 clause groups found in sectoral CBAs negotiated between 2009 and 2016. To assign coverage from sectoral CBAs to establishments, we first need to map the signing employer association to the firms being represented. Using the equivalent of a FOIA request, we obtained the universe of establishments paying dues to employer associations. We then take sectoral CBAs and match them to all establishments paying dues to the signing employer association. The next step is to assign coverage only to establishments located in the geographic region specified in the CBA. Finally, to reduce overlap in CBA coverage, we exploit information on negotiated wage floors to assign a "main CBA" to each establishment-year.[9]

Robustness. We check the robustness of our data-driven method on two dimensions: 1) two different ways of selecting the establishment sample used in the regressions: either a 50% random split-sample (used in our baseline approach) or the full estimation sample of establishments; and 2) three definitions of the gender gap in PageRank values, i.e., $V_j^F - V_j^M$. The first definition (used in our baseline approach) chooses the establishment with the smallest wage premium gap as the normalizing establishment, and then adjusts female values relative to the male values by multiplying the former by the ratio of the female-to-male PageRank values of the normalizing establishment. The second definition simply assumes the multiplicative factor is the same for both genders, i.e., no normalization is needed. The third definition uses a (within-gender) normalized index from 0 to 100 of V_j^F and of V_j^M.[10]

Tables A.4 and A.5 show all the clause types selected by any of the combinations above. These tables also show how many of these 6 different combinations choose a given clause type as either female- or male-centric, as well as those selected under the baseline approach but adding state and industry fixed effects.

Labor market outcomes We briefly describe how we define the outcomes used for the establishment-level and incumbent worker-level analyses. While for all worker-level outcomes we use the main job spell, some establishment-level outcomes are constructed with all job spells. We first describe establishment-level outcomes derived with all job spells and then those derived using main job spells. Finally, we describe worker-level outcomes.

Establishment level outcomes - all job spells:

---

[9]Specifically, we first define an establishment's "core union" to be the modal union involved in negotiating wage floors that have bite on the wage distribution. Among the CBAs negotiated by the "core union" in a given year, the "main CBA" is the one with the wage floor that has the largest mass of workers.

[10]The values are linearly rescaled within each gender group, setting the minimum to 0 and the maximum to 100. This normalization makes the variance of PageRank index values more similar across gender groups and ensures an identical range, while still preserving relative differences within each group.

- Total employment. The total number of workers employed at an establishment in a given year.

- Share of women in the workforce. Share of women employed in a given establishment-year among all workers.

- Share of women in the probationary workforce. Share of women employed in a given establishment-year with less than 3 months of tenure among all workers with fewer than 3 months of tenure. Brazil's federal labor code allows for at most 3 months of probation, after which employment terminations imply severance payments.

- Absences. Percent of total workdays in a given establishment-year that are not worked due to leaves of absence. The numerator is the number of days taken as leave by workers (excluding maternity leaves and military service) at the establishment throughout the year. The denominator is the number of workers employed at the establishment for the entire year multiplied by 365.

- New hires. Number of workers recently hired by the establishment, defined as the number of workers employed in a given establishment-year with less than 12 months of tenure.

- Share of women among new hires. Share of women employed in a given establishment-year with less than 12 months of tenure among all workers with fewer than 12 months of tenure.

- Share of women among separating workers. Share of women among workers who separate from the establishment in that year. Separating workers are defined as those who are no longer employed at the establishment by the end of the year.

- Establishment exit. A dummy variable indicating whether the establishment does not appear in RAIS in 2017.

Establishment level outcomes - main job spell:

- Mean log wage. For any given worker subgroup, we take the mean of the wage outcome (defined below) in logs across all workers in the subgroup employed at the establishment in that year. This variable is defined for the following worker subgroups: women and men with more than 12 months of tenure, women and men with less than 12 months of tenure.

- Mean gender wage gap. The difference between the mean log wage for women and the mean log wage for men for a given establishment-year.

- Wage bill. The monthly wage bill for the establishment. That is, we sum the wage outcome (defined below) for all workers employed by the establishment in that year.

- Share of women poached in. Share of new female hires that are poached from another firm among all female workers. New hires are defined as workers with less than 12 months of tenure at that establishment in a given year. Poached hires are defined as workers who in the preceding year worked at another firm in RAIS, as opposed to being unemployed or out of the (formal) workforce.

- Age of female workforce. Mean age of female workers employed at an establishment in a given year.

- Tenure of female workforce. Average months of tenure of female workers employed at an establishment in a given year.

- Hours of female workforce. Average contracted hours of work per week of female workers employed at an establishment in a given year. Weekly contracted hours are those agreed upon hiring, and do not include overtime work.

- Education of female workforce. Average years of schooling of female workers employed at an establishment in a given year.

- Share of women among managers. The share of women among workers with an occupation code corresponding to a managerial role. Occupation codes corresponding to manager positions are those starting with 12, 13 or 14 (as per CBO: *Classificação Brasileira de Ocupações*).

- Maternity leave benefits. The share of women taking maternity leaves longer than 120 days among women employed at an establishment that start their maternity leave in a given year. We are able to identify women taking maternity leave thanks to detailed information on both the length and the reason of the three longest leave spells per job spell. We think that it is very unlikely that maternity leaves are not among the three longest leave spells in a year for a woman on maternity leave. For this reason we are confident that we are observing the near universe of maternity leave spells.

- Job protection after maternity. The share of women working at the same employer where they were working at the start of maternity leave by end-of-year for the year

when their maternity leave ends, among women employed at said establishment who start their maternity leave in the same year.

- Injury leave. The share of workers taking leave due to a workplace injury among all workers employed at an establishment during a given year.

Establishment level outcomes - not in RAIS:

- CBA wage adjustments. The largest percentage wage adjustment negotiated among the firm-level CBAs covering an establishment. For years without a wage adjustment clause or without a negotiated CBA, the assigned wage adjustment is zero.

- Profit margin. The mean profit margin (in percentage terms) over 2012-2014 and 2015-2017. The sample is restricted to establishments reporting profit margin information to Orbis in both the pre- and post-reform periods.

Worker level outcomes - main job spell of incumbent workers

- Wages. The average monthly earnings that a worker makes during a job spell in a given year. We always use earnings in real terms by using the December CPI (i.e., the *Índice Nacional de Preços ao Consumidor* reported by IBGE) with 2015 as the base year.

- Retention. A dummy that indicates whether the worker is observed working at the baseline employer in any given year, where the baseline employer is defined as the (main) establishment of employment in 2014.

- Employed in formal sector. A dummy that indicates whether the worker is observed working in the formal sector in that year, i.e., they have a job spell registered in RAIS in that year.

D   QUALITATIVE INFORMATION ON THE CUT REFORM

We conducted a systematic search of multimedia sources, archival documents, and structured interviews to better understand the precursors and aftermath of the CUT reform. Given the open-ended nature of the search, this appendix details our approach and the rich variety of sources that inform our analysis.

Union Central Websites

Most union centrals have official websites that allow one to search internal news reports and multimedia describing the activities of the organization and its members. We searched for content using the words *mulher* (woman) and *paridade* (parity) in the years surrounding the reform, since these were the most publicized aspects of the 2015 CUT reform. Some of the most interesting items found in the search include (by union central):

- CUT: CUT administered a training course to over 50 female union leaders whose central objective was to "introduce feminism within CUT so that women leaders understand the struggle of women in all spaces, in all areas, and to further empower the theme of parity that we are going to implement this year"

- CSB: the board of directors and other leaders held a discussion on gender issues (we were unable to find any information on follow-up)

- CTB: a female leader of the national board called for a more feminist agenda (we were unable to find any information on follow-up)

- NCST: published a bulletin on gender issues called *Boletim Observa Gênero*

These items illustrate that many union centrals discuss topics related to women. However, there was no evidence of concrete actions taken to advance women's issues, unlike the 2015 CUT reform we describe in Section II.B.

A similar but wider search within CUT's official website allowed us to obtain more detailed information on what leaders highlighted about the reform, as well as subsequent actions that the CUT has taken in more recent years.[11] On the former, for example, we found evidence of frustration with the effective power that women leaders gained at the national board of the CUT, which further supports our evidence that raising the voice of existing

---

[11]In terms of precursors to the reform, the search pointed us to the original opinion piece authored by CUT leaders in response to the PT's adoption of gender parity (link), which led to the CUT reform Godinho Delgado, 2017.

women was the critical feature of the reform in 2015.[12] On the subsequent actions taken by the CUT to strengthen the female-focused priorities, we find for example that the CUT did not allow any of its state or industry bodies to disband their internal women's collectives or secretariats in response to a dramatic drop in union revenue triggered by the removal of mandatory dues in 2017 (link). We also find that the CUT schools have launched training programs committed to having a class size that is at least 50% women (link).

Archives

To gather information on official actions taken by union centrals during the time of the reform, we searched their archives. This search focused on finding out what happened at their congresses and similar gatherings. The fact that CUT has its own library to manage its archives—known as *Centro de Documentação e Memória Sindical da CUT (Cedoc)*— enabled us to perform a very thorough search.[13] For other union centrals, we were restricted to search within their official websites as well as the *Centro de Memoria Sindical* .

The archives in *Cedoc* were a trove of information for our analysis. Not only were we able to look at documents from the CUT's national gathering of women (*8º Encontro Nacional de Mulheres*) and its national congress from 2015 (*12º CONCUT*) that were central to cementing the reform, but we also obtained resolutions by CUT-affiliated confederations as well as materials produced by the CUT training schools. Below is a list of some of the most important documents that informed our analysis:

- The resolutions from all of CUT's national congresses (or *CONCUTs*)

- Supporting documents for the *12º CONCUT*, including a companion booklet for the matters to be debated at the congress, summarizing the proposals suggested by members to be discussed (984 in total) where each proposal—including the female-friendly fight plan—is linked to the confederations that listed it as a priority

- The resolutions and amendments passed during the *8º Encontro Nacional de Mulheres*

- Resolutions passed by CUT-affiliated confederations: CNM, CONTRACS, and CNTSS[14]

---

[12]For example, the head of SNMT in 2023, Junéia Batista, mentioned that the next course of action is gender parity in decision-making positions at CUT (link). Despite this desire for a stronger form of parity, a leader of the teacher's confederation affiliated to CUT notes how discussions on collective agreements without women have become "the exception rather than the rule" (link).

[13]We are indebted to Adalto Carvalho, the librarian who digitized numerous documents for us in the process, and also invited us to virtually join an informative debate on the role of women in the development of CUT unionism between Fátima da Silva and Didice Godinho Delgado (link).

[14]For example, the *pautas* proposed by CNM in 2015 include 1) defend the right to daycare guaranteed for children aged zero to six; 2) reimbursement of daycare assistance of at least 50% of the wage floor; 3) guarantee 180 days of maternity leave; and 4) relaunch the training course for metallurgical women.

- Archives of specific confederations, with information on the campaigns carried out by year, as well as video recordings of their national congresses[15]

- The courses offered at CUT training schools along with some activity handbooks that included a section focused on "gender clauses"

The resolutions, fight plans, and supporting documents from national congresses of other union centrals around the time of the CUT reform do not emphasize the need to raise the voice of women. For example, during Força Sindical's national congress in 2013, plenary sessions for women were held and women made up 30% of the board for the first time in their history. The resolutions from CSB's national congress in 2016 only mention the word *mulher* (woman) 7 times. Lastly, the new bylaws drafted in 2013 by NCST merely commit to combating discrimination of all kinds (including gender) and require that the head of the women's division be a woman, which was *de facto* already the case.

We also searched for evidence on how employers responded to the CUT reform, as well as any news reports on changes to worker satisfaction. Our search focused on news archives like the *Hemeroteca Digital Brasileira* and the *Acervo - Folha de São Paulo* but did not produce any results worthy of note.

Academic works

To hear the voices of people affected by the CUT reform, we searched academic works that conducted structured interviews and surveys with subjects of interest. Our search produced results that were mostly dissertations in fields of social sciences outside economics at universities in Brazil and Portugal.[16] These papers were incredibly valuable since they captured the sentiment regarding women's role in Brazil's unions around the time of the CUT reform. Below are some of the highlights from this search.

Martins (2021) conducts an in-depth study of the banking workers' union in São Paulo from 2000 to 2020, which has been a union at the forefront of women's issues in unionism even prior to the 2015 reform. Most of her interviews are with current female leaders that rose up the ranks of this union. These leaders document in detail the challenges that women face within Brazilian unionism. Some examples include union meetings occurring late in the day that do not provide childcare, leaders disregarding women's demands as not appealing to the base, the importance of the *pautas* in generating change, and the lack of women in discussions surrounding these *pautas*.

---

[15]For example, in 2016 CONTRACS launched one campaign on daycare for employees and another one on fostering respect to female workers (*#Elasmerecemrespeito*).

[16]We relied on Google Scholar as a search engine, as well as theses repositories from Brazil, e.g., *Biblioteca Digital Brasileira de Teses e Dissertações* and *Biblioteca Digital de Teses e Dissertações da Universidade de São Paulo (USP)*.

Munhoz and Silotto (2019) follows the 3 largest unions affiliated to CUT in the state of São Paulo and asks whether the gender quota had any spillovers to the local unions. Aligned with our causal analysis, their descriptive statistics on these unions show that there were negligible spillovers. From the point of view of the union leaders they interview, the gender quota at the CUT's national board is of little consequence to workers without placing women in positions with decision-making power at local unions. This supports our evidence that raising the voice of existing women was the critical feature of the reform in 2015.

Silva (2021) studies women's involvement in the family agricultural workers' federation in Rio Grande do Sul. One of the events she follows is how this federation—which is affiliated to CUT—restarted its women's collective in 2016. Similarly, Franco Oliveira (2017) followed the women's collective at CUT-Bahia during 2016-2017 and conducted interviews with attendees on the subject of women in unionism after the 2015 reform. She documents that this state branch focused on offering more training courses on women's issues—an initiative similar to those made by the CUT confederations that prioritized the female-friendly fight plan. Moreover, based on the interviews conducted, she reaches the conclusion that even male union leaders have become more attuned to women's issues and their demands for power in the union.[17] As pointed out in her study of feminism within CUT, Recoaro (2022) argues that having a woman in power is not necessary for pushing women's demands forward; the necessary condition is that the person in power (man or woman) is convinced that these demands are a priority. This aligns with our analysis, in that the numerous initiatives to raise the voice of women within CUT shifted the bargaining priorities at local unions.

---

[17]An interview with a male union leader in Oliveira (2022) highlights this new progressive stance by male union leaders. In this case, the leader from telemarketers' union invokes an argument made by some feminists in the union movement: that employers are more open to female amenities that reinforce women's role as caregivers rather than workers, e.g., childcare assistance.

E AKM AND PAGERANK MODEL

Our data-driven approach to identify female- and male-centric amenities requires establishment level estimates of gender-specific PageRank values and AKM wage premiums. This appendix presents the model underlying these estimates. For simplicity, we present the model without any reference to gender specificity. We also use establishment and firm interchangeably.

Denote $\tilde{V}_j$ as the common value of employment for any worker i at firm j. Common value means that all workers agree on $\tilde{V}_j$ such that a single job ladder exists ranking firms according to this value. All else equal, workers value higher compensation bundles so that one can write $\tilde{V}_j = h\left(w_j, a_j\right)$, where $h(\cdot)$ in strictly increasing in both the wage w_j and the amenity a_j arguments. The utility of workers from employment at the establishment, however, is heterogeneous and given by $u_{ij} = h(w_j, a_j) + \varepsilon_{ij}$, where ε_{ij} captures an individual's idiosyncratic preferences for working at j.

PageRank values

The starting point here is $u_{ij} = \tilde{V}_j + \varepsilon_{ij}$. In a market with only two firms and independently distributed type I Extreme Value ε_{ij} across workers, the probability that a worker prefers firm j over k is given by $\frac{\exp(\tilde{V}_j)}{\exp(\tilde{V}_j) + \exp(\tilde{V}_k)}$. With N workers and letting M_{jk} denote the number of workers choosing firm j over k, the following relation between employment decisions and valuations of firm-specific employment is simply $M_{kj}/M_{jk} = \exp(\tilde{V}_k)/\exp(\tilde{V}_j)$.

In a labor market with multiple firms $j \in \mathcal{J}$, the above condition imposes a restriction on each pair of firms, i.e.,

$$M_{kj}\exp(\tilde{V}_j) = M_{jk}\exp(\tilde{V}_k), \forall j \in \mathcal{J}. \tag{2}$$

Following Sorkin (2018), one can relax this condition by imposing a single restriction per firm that guarantees a consistent valuation of employers (e.g., no Condorcet cycles), as well as a unique set of firm-level values that best explains worker flows across establishments. Summing equation (2) across all employers and rearranging terms gives

$$\underbrace{\frac{\overbrace{\sum_{j\in\mathcal{J}} M_{kj}\exp(\tilde{V}_j)}^{\text{value-weighted entry}}}{\underbrace{\sum_{j\in\mathcal{J}} M_{jk}}_{\text{exits}}} = \underbrace{\exp(\tilde{V}_k)}_{\text{value},}} \tag{3}$$

45

which implies a single linear restriction per establishment.

The intuition behind equation (3) is that a valuable firm tends to be chosen over other valuable firms and has fewer workers leave it. This recursive definition of $\exp(\tilde{V}_j)$ is closely linked to Google's PageRank algorithm for ranking web-pages in a search. Along these lines, one can solve for $\exp(\tilde{V}_j)$ as a fixed point in a linear system. Moreover, a unique solution exists if the set of employers are strongly connected, i.e., an establishment has to both hire a worker from and have a worker hired by another establishment in the set.

AKM premiums

The starting point again is $u_{ij} = \tilde{V}_j + \varepsilon_{ij}$ but with the assumption that $\tilde{V}_j = \beta \log(w_j - b) + \eta \log(a_j - q)$. The parameters b and q are the workers' reference wage and amenity levels, and $\epsilon_{i,j}$ refers to the idiosyncratic preferences from working at establishment j. Assuming that the $\{\epsilon_{i,j}\}$ are independent draws from a Type I Extreme Value distribution and the number of establishments $\mathcal{J}$ is very large, workers' choice probabilities are closely approximated by exponential probabilities.[18] Hence, the establishment-specific labor supply functions are approximated by:

$$\log(L_j) = \log(\lambda) + \beta \log(w_j - b) + \eta \log(a_j - q). \tag{4}$$

The employer's problem is to post the wages and amenities that minimize production costs given labor supply in (4). The posted wages and amenities are common to all workers since employers cannot discriminate on the basis of their idiosyncratic preferences $\{\epsilon_{i,j}\}$.[19] The optimal choice is the solution to the following cost-minimization problem:

$$\min_{w,a}(w_j + \xi_j a_j)L(w_j, a_j) \quad \text{s.t. } T_j f(L(w_j, a_j)) \geq \bar{Y}, \tag{5}$$

where ξ_j captures heterogeneity in the marginal cost of amenity provision across employers.

The first order conditions imply that the optimal compensation package is given by:

$$w_j = T_j f'(L_j)\mu_j \left(\frac{e_{wj}^L}{1 + e_{wj}^L + e_{aj}^L} \right) \tag{6}$$

$$a_j = T_j f'(L_j)\mu_j \left(\frac{e_{aj}^L}{\xi_j(1 + e_{wj}^L + e_{aj}^L)} \right). \tag{7}$$

---

[18]The exponential probabilities are $p_j \approx \lambda \exp(\beta \log(w_j - b) + \eta \log(a_j - q))$, where λ is a constant common across all establishments in the market.

[19]This asymmetry in information, rather than labor market concentration, is the source of monopsony power. Recall that $\mathcal{J}$ is large so as to ignore strategic interactions in posting.

Rearranging equations (6) and (7), one can write wages and amenities as weighted averages of the marginal revenue product of labor and their respective reference values, i.e.,

$$w_j = \left(\frac{\beta}{1+\beta+e_{aj}^L}\right) T_j f'(L_j)\mu_j + \left(\frac{1+e_{aj}^L}{1+\beta+e_{aj}^L}\right) b \tag{8}$$

$$a_j = \left(\frac{\eta}{\xi_j(1+\eta+e_{wj}^L)}\right) T_j f'(L_j)\mu_j + \left(\frac{1+e_{wj}^L}{1+\eta+e_{wj}^L}\right) q. \tag{9}$$

Assume a linear technology $f(L_j) = \theta L_j$ and price-taking employers in the output market to specify the marginal revenue product of labor: $T_j f'(L_j)\mu_j = T_j P_j \theta$. To simplify further, assume that reference wages and amenities are proportional to productivity ($b = \bar{b}\theta$ and $q = \bar{q}\theta$). Rearranging terms and taking logs results in

$$\log(w_j) = \log\left(\frac{\theta\bar{b}(1+e_{aj}^L)}{1+\beta+e_{aj}^L}\right) + \log\left(1+\beta R_j^w\right) \tag{10}$$

$$\log(a_j) = \log\left(\frac{\theta\bar{q}(1+e_{wj}^L)}{1+\eta+e_{wj}^L}\right) + \log\left(1+\eta R_j^a\right), \tag{11}$$

where $R_j^w = T_j P_j / [(1+e_{aj}^L)\bar{b}]$ and $R_j^a = T_j P_j / [\xi_j(1+e_{wj}^L)\bar{q}]$. With relatively small values of βR_j^w and ηR_j^a, log wages and log amenities are functions of a fixed worker component and a fixed establishment component as in Abowd, Kramarz, and Margolis (1999)—henceforth AKM. Specifically,

$$\log(w_j) = \log\left(\frac{\bar{b}(1+e_{aj}^L)}{1+\beta+e_{aj}^L}\theta\right) + \beta R_j^w \tag{12}$$

$$\log(a_j) = \log\left(\frac{\bar{q}(1+e_{wj}^L)}{1+\eta+e_{wj}^L}\theta\right) + \eta R_j^a. \tag{13}$$

In short, equations (12) and (13) imply that the wages and amenities of workers can be written in the form $\log(w_j) = \alpha^w + \psi_j^w$ and $\log(a_j) = \alpha^a + \psi_j^a$, where $\psi^w = \beta R_j^w$ is an establishment-specific wage premium and $\psi^a = \eta R_j^a$ is an establishment-specific amenity premium. To separately identify these premiums from the worker fixed effects, one must focus on a set of firms that are connected through worker flows.

F   WELFARE EFFECTS OF THE CUT REFORM

The CUT reform increased female-centric amenities and made CUT establishments more valuable to women, as signaled by higher retention rates and job queues. By how much did women's welfare change? What about the reform's impact on men's welfare? We describe our approach to quantify changes in worker welfare here.

F.A   Approach and Intuition

We quantify the CUT reform's effect on worker welfare through a revealed preference approach that (i) relies on a few sufficient statistics that are easily computable in the data; and, thus, (ii) takes no stance on the precise functional form linking amenities to worker utility. In particular, we adapt a framework used to evaluate changes in consumer welfare from introducing new or improved product varieties (Feenstra, 1994; Redding and Weinstein, 2016) to our labor market setting.

For tractability, we assume that workers possess CES preferences over employers, as is common in the consumer setting (Feenstra, 1994; Atkin et al., 2015). As shown in Anderson, De Palma, and Thisse (1992), a key advantage of CES is that it generates the same labor supply to firms as obtained by aggregating workers' discrete choices over where to work based on where they obtain the highest utility. This is a common way of modeling the labor market (Card et al., 2018; Sorkin, 2018; Berger, Herkenhoff, and Mongey, 2022; Lamadon, Mogstad, and Setzler, 2022). In the next subsection we microfounds CES demand using such discrete choices.

Because of the reform, the quality of CUT-affiliated firms is changing. From a modeling perspective, this is analogous to a situation in which the quality of certain goods is improving or when new, improved, good varieties are introduced in the market. Then, just as gains to consumer welfare from improving product varieties can be measured though changes to the price index—i.e., the change in cost of purchasing one util worth of utility—the gains to worker welfare from improving workplace amenities can be measured through changes to the wage index—i.e., how much more (or less) the representative worker earns to work one disutility-weighted hour.

Under CES preferences, only four sufficient statistics quantify the change in worker welfare, i.e., measure the change in the wage index. First, welfare increases with the share of total labor income found at treated establishments, which captures workers choosing to move to these employers after they improve amenities. Second, the same change in labor income at treated establishments corresponds with a higher increase in welfare if workers are less elastic to begin with, since it takes a larger improvement in amenities to draw them away. Third, welfare is higher if workers are drawn away from non-CUT firms with initially

low value, capturing a bigger upgrade in employer quality across regimes. Finally, welfare increases with wages at non-CUT establishments, potentially capturing the pro-competitive spillover effects of the reform.

F.B   Model

In each period, a representative household with CES preferences over employers is willing to work a fixed number of (dis)utility-weighted hours. It chooses labor supply to each firm to maximize total income, subject to this hours constraint:

$$\max_{\{n_{jt}\}} \sum_{j \in \mathcal{J}_t} w_{jt} n_{jt} \quad s.t. \quad \left[\sum_j (b_{jt} n_{jt})^{\frac{1+\eta}{\eta}} \right]^{\frac{\eta}{\eta+1}} = N, \tag{14}$$

where $\mathcal{J}_t$ denotes the set of firms operating at time t, n_j is the number of hours supplied to firm j, w_j is the wage at j, η is the elasticity of substitution across firms, and b_j represents the "taste-shifter" for firm j. b_j captures all non-wage attributes that commonly affect each worker's utility at j. Worse amenities increase this disutility b_j. We assume a utility-posting world without job rationing, where a firm accepts any worker who wishes to work there. For simplicity, since worker welfare only depends on firms' final wage and amenity offers, regardless of how firms arrive at them, we do not model the firm side.

Optimal labor supply to each firm is given by:

$$n_j^* = \left(\frac{w_j}{\tilde{W}} \right)^{\eta} \frac{1}{b_j^{1+\eta}} N \tag{15}$$

, where $\tilde{W}$ is a book-keeping term called the wage index, defined as:

$$\tilde{W} = \left[\sum_{j \in \mathcal{J}} \left(\frac{w_j}{b_j} \right)^{1+\eta} \right]^{\frac{1}{1+\eta}} \tag{16}$$

.

The wage index measures how much the representative worker is paid to work a disutility-weighted hour, and serves as a measure of welfare.[20]

---

[20]This can be seen by taking the envelope condition around the optimal solution to the worker's problem: $\sum_j w_j n_j = \tilde{W} N$. Formally

$$\frac{\partial}{\partial N} \sum_{j \in \mathcal{J}} w_j n_j^*(w_j, w_{-j}) = \tilde{W}$$

. The wage index is to welfare in the labor setting like the price index is to welfare in consumer theory. In consumer theory, the price index captures the cost of purchasing one util of utility. Welfare rises as it gets cheaper to purchase one more util. Here, an increase in $\tilde{W}$ means workers are paid more for providing one

Any change in the wage index across two periods captures changes to worker welfare, measured by the ratio:

$$\phi_{t-1,t} = \frac{\tilde{W}_t}{\tilde{W}_{t-1}}$$

The CUT reform changes amenities, or taste shifters b_{jt}, at treated establishments. The key challenge in estimating welfare changes is that these $\{b_{jt}\}_{j \in \mathcal{J}_t}$ are unobserved. However, as first shown in (Feenstra, 1994), assuming CES preferences allows us to overcome this challenge. Under CES, any welfare change depends only on the *observed* pre- and post-reform wages and employment at CUT and non-CUT employers.[21] Formally:

$$\ln \phi_{t-1,t} = -\frac{1}{1+\eta} \ln\left(\frac{\lambda_t}{\lambda_{t-1}}\right) - \frac{1}{1+\eta} \ln\left(\frac{\overline{S^*}_t}{\overline{S^*}_{t-1}}\right) + \ln\left(\frac{\overline{w^*}_t}{\overline{w^*}_{t-1}}\right) \qquad (17)$$

where λ_t is the share of total labor income in t at non-CUT firms, $\overline{S^*}_t$ is a geometric average of the share of labor income at each non-CUT firm in t, and $\overline{w^*}_t$ is a geometric average of period t wages at non-CUT firms. The asterisk * denotes that operations are taken over non-CUT firms.

Changes in welfare depend on three terms, as per Equation (17). The first, the "variety-adjustment" term, depends on the ratio of the share of total labor income at non-CUT firms after relative to before the reform $\frac{\lambda_t}{\lambda_{t-1}}$. This ratio captures welfare changes through a revealed preference logic: workers substitute toward CUT firms once their amenities improve, lowering the share of the labor income at non-CUT firms and increasing welfare. The magnitude of this change depends on the elasticity of substitution across firms. If workers are inelastic (η is low), the same move toward amenity-improving CUT-firms implies a larger welfare increase because it takes a bigger improvement in amenities to draw workers away.

The term $\frac{\overline{S^*}_t}{\overline{S^*}_{t-1}}$ captures the heterogeneity in labor income at non-CUT firms: welfare increases by more if CUT firms draw workers away from less valued non-CUT firms, thereby increasing dispersion in and lowering the geometric mean of their wage bill share. As in the "variety-adjustment" term, the implied effects are larger as workers become more inelastic. The final term $\frac{\overline{w^*}_t}{\overline{w^*}_{t-1}}$ represents a change in wages at non-CUT firms, possibly as a pro-competitive response to the reform. As these "outside" wages increase, so too does welfare.

Derivation of the Welfare Equation To derive equation (17), note that

$$\lambda_t = \frac{\sum_{j \in \Omega_{t,t-1}} w_j n_j}{\sum_{k \in \mathcal{J}_t} w_k n_k} = \frac{\sum_{j \in \Omega} \left(\frac{w_j}{b_j}\right)^{1+\eta}}{\sum_k \left(\frac{w_k}{b_k}\right)^{1+\eta}} = \left[\frac{W_t^*}{W_t}\right]^{-(1+\eta)}$$

---

additional unit of disutility-weighted labor supply, thereby increasing their welfare.

[21]Under CES, the relative (dis)utility of working at an employer is captured by its expenditure share, which depends exclusively on prices and quantities.

where $\Omega_{t,t-1} = \mathcal{J}_t \cap \mathcal{J}_{t-1}$ are firms common to both periods—in our case, non-CUT firms. The asterisk * in W_t^* and W_{t-1}^* denotes that these are wage indices over the common set of firms.

We can thus re-write equation (17) as:

$$\phi_{t-1,t} = \left[\frac{\lambda_t}{\lambda_{t-1}} \right]^{-\frac{1}{1+\eta}} \left[\frac{\sum_{j \in \Omega_{t,t-1}} (\frac{w_{jt}}{b_{jt}})^{1+\eta}}{\sum_{j \in \Omega_{t,t-1}} (\frac{w_{jt-1}}{b_{jt-1}})^{1+\eta}} \right]^{\frac{1}{1+\eta}} = \left[\frac{\lambda_t}{\lambda_{t-1}} \right]^{-\frac{1}{1+\eta}} \frac{\tilde{W}_t^*}{\tilde{W}_{t-1}^*} \tag{18}$$

Finally, λ_t is the share of the wage bill at common firms in t (using wages at t).

To further decompose $\tilde{W}_t^* / \tilde{W}_{t-1}^*$, denote with S_j the share of labor income (the consumer expenditure in our consumer problem analogy) at any given firm:

$$S_j = \frac{w_j n_j}{\sum_k w_k n_k} = \frac{(\frac{w_j}{b_j})^{1+\eta}}{\sum_k (\frac{w_k}{b_k})^{1+\eta}} \tag{19}$$

. We then can write:

$$\tilde{W}_t^* = (S_{jt}^*)^{-(1+\eta)} \left(\frac{w_{jt}}{b_{jt}} \right) \quad \forall j \in \Omega_{t,t-1} \tag{20}$$

. Following Redding and Weinstein, 2016, we take logs of both sides, difference over time, and sum over all $j \in \Omega_{t,t-1}$ to get

$$\ln \left(\frac{\tilde{W}_t^*}{\tilde{W}_{t-1}^*} \right) = \ln \left(\frac{\bar{w}_t^*}{\bar{w}_{t-1}^*} \right) - \frac{1}{1+\eta} \ln \left(\frac{\bar{S}_t^*}{\bar{S}_{t-1}^*} \right) - \ln \left(\frac{\bar{b}_t^*}{\bar{b}_{t-1}^*} \right) \tag{21}$$

where the bars indicate a geometric average and the last term is zero because we assume quality remains the same for these common firms. Thus, a change in welfare depends only on three terms that are observed in the data and η:

$$\ln \phi_{t-1,t} = -\frac{1}{1+\eta} \ln \left(\frac{\lambda_t}{\lambda_{t-1}} \right) - \frac{1}{1+\eta} \ln \left(\frac{\bar{S}_t^*}{\bar{S}_{t-1}^*} \right) + \ln \left(\frac{\bar{w}_t^*}{\bar{w}_{t-1}^*} \right)$$

Microfoundation of CES demand using discrete choices Following the CES demand in (Berger, Herkenhoff, and Mongey, 2022), workers' utility for working at a firm has a component that is common across workers, encompassing wages and a common taste for the firm amenities, and an idiosyncratic shock that follows a logit distribution. Firms post utility offers—we don't model the source of firm heterogeneity and assume that they exogenously differ. There is a unit measure of workers indexed by $i \in [0,1]$. Each worker has a disutility for working at firm j:

$$\nu_{ij} = \exp^{-\xi_{ij}} h_{ij} b_j$$

with ξ_{ij} iid across workers and drawn from a multivariate Gumbel distribution with parameter η. Each worker must earn $y \sim F(y)$, where earnings $y_i = w_j h_{ij}$. The worker chooses firm j to minimize disutility:

$$\min_j \{\log h_{ij} + \log b_j - \xi_{ij}\} = \max_j \{\log w_j - \log y_i - \log b_j + \xi_{ij}\}$$

Following McFadden (1973) on logit, the probability that worker i chooses to work at firm j is:

$$p_i(\tilde{w}) = \frac{\tilde{w}_j^{1+\eta}}{\sum_k \tilde{w}_k^{1+\eta}}$$

where $\tilde{w}_j := \frac{w_j}{b_j}$. The aggregate labor supply to firm j is then found by integrating the probability that a worker works at that firm times the hours supplied by that worker, over the mass of all workers:

$$n_j = \int p_i(\tilde{w}) \cdot h_{ij} \cdot dF(y) \quad \text{where} \quad h_{ij} = y_i/w_j$$

$$n_j = \frac{\tilde{w}_j^{1+\eta}}{\sum_k \tilde{w}_k^{1+\eta}} \frac{1}{w_j} \int y_i dF(y)$$

$$= \left(\frac{w_j}{\tilde{W}}\right)^{\eta} \frac{1}{b_j^{1+\eta}} N$$

This is exactly the aggregate labor supply to firm j as in the representative worker's problem with CES demand. The last line follows from the fact that in equilibrium:

$$Y = \int y_i dF(y) = \sum_{j \in \mathcal{J}} w_j n_j^* = \tilde{W} N$$

F.C   Estimation

We separately estimate Equation (17) for men and women. Our estimates employ the establishment sample from Section V.C. Years 2012-2014 comprise the pre-reform period $(t-1)$ and 2015-2017 the post-reform period (t). We calibrate an estimate of the cross-firm elasticity of substitution (η) from Felix (2022), but assess robustness to other reasonable values.

We estimate the log change in $\bar{w}^*$ and in $\bar{S}^*$ using average changes across non-CUT establishments between $t-1$ and t, estimated via the following regression:

$$y_{jt} = \alpha + \beta Post_t + \mu_j + \epsilon_{jt} \tag{22}$$

where y_{jt} is either the average log earning at establishment j ($\log w_{jt}$) or the log of the

share of labor income among non-CUT establishments at that establishment ($\log s_{jt}$).[22] The specification includes establishment fixed effects μ_j. The coefficient of interest, β, captures the average within-establishment change in the dependent variable between between $t-1$ and t.[23]

To estimate the change in λ we take a first order approximation of λ_t around λ_{t-1}. This allows us to map the market-level change in the share of labor income at CUT establishments (the desired object) to changes in quantities that are estimable through establishment-level regressions as in Equation (22).

Formally:

$$\Delta\lambda_t = \lambda_t - \lambda_{t-1} = \sum_{j \in \mathcal{J}} \frac{\partial\lambda}{\partial w_j} \cdot dw_j + \sum_{j \in \mathcal{J}} \frac{\partial\lambda}{\partial n_i} \cdot dn_j \Bigg|_{w_{t-1}, n_{t-1}}$$

$$= \frac{\sum_{j \in (\mathcal{J}\setminus\Omega)} w_{jt-1} n_{jt-1}}{(\sum_{j \in \mathcal{J}} w_{jt-1} n_{jt-1})^2} \left(\sum_{j \in \Omega} n_{jt-1} \cdot dw_j + \sum_{j \in \Omega} w_{jt-1} \cdot dn_j \right)$$

$$- \frac{\sum_{j \in \Omega} w_{jt-1} n_{jt-1}}{(\sum_{j \in \mathcal{J}} w_{jt-1} n_{jt-1})^2} \cdot \left(\sum_{i \in \mathcal{J}\setminus\Omega} n_{jt-1} \cdot dw_j + \sum_{j \in \mathcal{J}\setminus\Omega} w_{jt-1} \cdot dn_j \right)$$

where to simplify notation we use Ω in place of $\Omega_{t,t-1}$ to denote the set of non-CUT firms (of measure N_Ω) and $\mathcal{J}\setminus\Omega$ to denote the set of CUT-affiliated firms (of measure $N_{\mathcal{J}\setminus\Omega}$).

We define $\tilde{s}_{jt} = \frac{w_{jt-1} n_{jt}}{\sum_{k \in \mathcal{J}} w_{kt-1} n_{kt-1}}$ and $\hat{s}_{jt} = \frac{w_{jt} n_{jt-1}}{\sum_{k \in \mathcal{J}} w_{kt-1} n_{kt-1}}$ and re-write the expression above as

$$\Delta\lambda_t = N_\Omega (1 - \lambda_{t-1}) \left(\mathbb{E}[\Delta\tilde{s}_{jt} | j \in \Omega] + \mathbb{E}[\Delta\hat{s}_{jt} | j \in \Omega] \right)$$

$$- N_{\mathcal{J}\setminus\Omega} \lambda_{t-1} \left(\mathbb{E}[\Delta\tilde{s}_{jt} | j \in \mathcal{J}\setminus\Omega] + \mathbb{E}[\Delta\hat{s}_{jt} | j \in \mathcal{J}\setminus\Omega] \right)$$

where $\mathbb{E}[.]$ denotes an average across firms. Finally, because $\log\left(\frac{\lambda_t}{\lambda_{t-1}}\right) = \log\left(\frac{\Delta\lambda_t}{\lambda_{t-1}} + 1\right) \approx$

---

[22]Note that the difference in the log of the geometric mean of a variable x is equivalent to the average change in $\log(x)$ between t and $t-1$ across units.

[23]As any pre-post strategy, we recognize that this approach might also pick up the effect of any other shock in addition to the CUT reform that might affect wages or employment within establishments over time. Unfortunately, however, we lack the ideal experiment to estimate the causal effect of the CUT reform on these quantities, which would be to randomly shock some labor markets with the reform while leaving other markets unaffected.

$\frac{\Delta \lambda_t}{\lambda_{t-1}}$, we can write:

$$\log\left(\frac{\lambda_t}{\lambda_{t-1}}\right) \approx$$
$$N_\Omega \frac{(1-\lambda_{t-1})}{\lambda_{t-1}}\left(\mathbb{E}[\Delta\tilde{s}_{jt}|j\in\Omega]+\mathbb{E}[\Delta\hat{s}_{jt}|j\in\Omega]\right)-N_{\mathcal{J}\setminus\Omega}\left(\mathbb{E}[\Delta\tilde{s}_{jt}|j\in\mathcal{J}\setminus\Omega]+\mathbb{E}[\Delta\hat{s}_{jt}|j\in\mathcal{J}\setminus\Omega]\right)$$

.

We estimate the average change in $\hat{s}_{jt}$ and $\tilde{s}_{jt}$ with a within-establishment pre-post comparison similar to Equation (22).[24]

Finally, we combine these estimates with λ_{t-1}, N_Ω, and $N_{\Gamma\setminus\Omega}$ which are directly computed from the data. To obtain standard errors, we bootstrap the entire estimation exercise 1000 times, each time drawing with replacement a new establishment sample.

F.D   Results

Table F1 reports results. Women's welfare increases by 0.059 log points (or 6.1%), consistent with our reduced form results that women are more likely to remain at, and comprise a larger share of new workers among, CUT establishments.[25] Worker moves following the reform account for over half of the increase in welfare. Women become more likely to work at CUT establishments, accounting for 15% of the welfare gain (a 1.8% rise in the share of CUT wage bill). In addition, the dispersion in the labor income across non-CUT firms rises (i.e., S^*

---

[24]Further simplifying the model to only two firms, one affiliated and one not affiliated to the CUT, allows us to directly use the differences-in-differences estimates of the effect of the reform on employment to estimate the change in λ, instead of using pre-post comparisons. Using w_1 and n_1 for wages and employment in the non-CUT firm and similarly w_2 and n_2 for the CUT-affiliated firm, we can write a first order approximation of the change in λ as:

$$\Delta\lambda_{t,t-1} = \underbrace{\frac{\partial\lambda}{\partial n_1}dn_1 + \frac{\partial\lambda}{\partial n_2}dn_2}_{-r(1-\lambda_{t-1})\lambda_{t-1}} + \underbrace{\frac{\partial\lambda}{\partial w_1}dw_1 + \frac{\partial\lambda}{\partial w_2}dw_2}_{-x(1-\lambda_{t-1})\lambda_{t-1}}$$

where $r = \frac{dn_2}{n_2} - \frac{dn_1}{n_1}$ is the difference in the percentage change in employment between the CUT and non-CUT firm and $x = \frac{dw_2}{w_2} - \frac{dw_1}{w_1}$ is the difference in the percentage change in wages. We can then plug in place of r and x our diff-in-diff estimates of the effect on log wages and log employment. Alternatively, we could use our diff-in-diff estimate on the effect of the reform on retention rates among incumbents as an estimate of r. Changes in retention could better capture change in employment due to voluntary transitions, which in this model entirely explain changes in employment. Plugging in $x \approx 0$ and $r \approx 1.8\%$ and the baseline value of $\lambda_{t-1} \approx 0.65$, we would find a percentage change in λ of approximately: $\frac{\Delta\lambda_{t,t-1}}{\lambda_{t-1}} = -r(1-\lambda_{t-1}) - x(1-\lambda_{t-1}) \approx -1.8 \cdot 0.35\% = -0.63\%$. This figure is consistent in sign with the one in Table F1, although smaller in magnitude. It relies, however, on a heavily stylized model that does not exploit fully the richness of the firm-level data.

[25]As predicted by the model, workers' elasticity of substitution across employers amplifies (or dampens) the welfare effect due to the shifts in employment across firms induced by the reform. For other reasonable values of η in the literature, ranging from 0.1 (Staiger, Spetz, and Phibbs, 2010) to 10.9 (Berger, Herkenhoff, and Mongey, 2022), women's welfare increases by between 2.8% and 9.5%.

falls), accounting for 48% of the increase in welfare.

The remaining 37% of the welfare gain is accounted for by higher wages among non-CUT employers. To the extent that these wage increases reflect pro-competitive responses to the CUT reform, any change in welfare from them can also be attributed to the reform. We recognize, however, that the increase in real wages at non-CUT employers following 2015 could be driven by a host of factors that are unrelated to the CUT reform. We therefore only view the change in welfare due to worker moves across firms—amounting to a 3.8% increase—as the credible estimate of the reform's welfare impact for women workers. For men, welfare is slightly higher (1.3%), but remains essentially unchanged if one only considers the component due to worker moves across firms (0.2%). Thus, the CUT reform improves women's welfare without reducing men's welfare.

TABLE F1
Welfare Estimation

		Women 20-35	All women	Men 20-35	All men
		(1)	(2)	(3)	(4)
$ln\phi_{t-1,t}$		0.044	0.059	-0.005	0.013
		(0.006)	(0.007)	(0.005)	(0.005)
Components breakdown:					
	$ln(\lambda_{t,t-1}) - ln(\lambda_{t-1,t})$	-0.012	-0.018	-0.005	-0.006
	$ln(\bar{S}_t^*) - ln(\bar{S}_{t-1}^*)$	-0.046	-0.058	0.013	0.001
	$ln(\bar{w}_t^*) - ln(\bar{w}_{t-1}^*)$	0.015	0.022	-0.001	0.011
η (calibrated)	1.015				
N establishments		58,417	60,651	59,438	60,651
N establishments in $\Omega_{t,t-1}$		45,331	47,195	46,182	47,195

Notes: Table reports the estimated welfare change for different groups of workers: women between 20 and 35 years old, all women, men between 20 and 35 years old, all men. It also reports estimates of the three components that make the welfare index, namely the Feenstra "new varieties" term $ln(\lambda_{t,t-1}) - ln(\lambda_{t-1,t})$, the change in the geometric average of the labor income shares of non-CUT firms $ln(\bar{S}_t^*) - ln(\bar{S}_{t-1}^*)$, and the change in the geometric average of the wages of non-CUT firms $ln(\bar{w}_t^*) - ln(\bar{w}_{t-1}^*)$. Standard errors in parenthesis come from the bootstrap procedure described in the text.