# Appendices for Online Publication

# A Appendix Tables

Table A1: Characteristics of Unions and Employer Associations

| | Unions | | | | Employer associations | | | |
|---|---|---|---|---|---|---|---|---|
| | Count (Share) (1) | Share signing CBAs (2) | Count if signing (Share) (3) | Median CBAs signed per year (4) | Count (Share) (5) | Share signing CBAs (6) | Count if signing (Share) (7) | Median CBAs signed per year (8) |
| Local union | 6,249 | 0.80 | 4,976 | 3.1 | 3,463 | 0.81 | 2,804 | 1.0 |
| | (0.94) | | (0.95) | | (0.95) | | (0.97) | |
| (Con)federation | 382 | 0.69 | 263 | 3.3 | 171 | 0.58 | 99 | 2.4 |
| | (0.06) | | (0.05) | | (0.05) | | (0.03) | |
| Municipal-level | 4,582 | 0.83 | 3,782 | 3.4 | 1,884 | 0.83 | 1,555 | 1.0 |
| | (0.69) | | (0.72) | | (0.52) | | (0.54) | |
| State-level | 2,049 | 0.71 | 1,457 | 2.8 | 1,750 | 0.77 | 1,348 | 1.1 |
| | (0.31) | | (0.28) | | (0.48) | | (0.46) | |
| Industry-based | 5,164 | 0.86 | 4,460 | 3.5 | 3,326 | 0.84 | 2,803 | 1.0 |
| | (0.78) | | (0.85) | | (0.92) | | (0.97) | |
| Occupation-based | 1,467 | 0.53 | 779 | 1.6 | 308 | 0.32 | 100 | 1.7 |
| | (0.22) | | (0.15) | | (0.08) | | (0.03) | |
| North | 522 | 0.75 | 389 | 2.4 | 346 | 0.74 | 257 | 1.0 |
| | (0.08) | | (0.07) | | (0.10) | | (0.09) | |
| Northeast | 1,329 | 0.71 | 937 | 2.1 | 735 | 0.72 | 528 | 1.0 |
| | (0.20) | | (0.18) | | (0.20) | | (0.18) | |
| Southeast | 2,576 | 0.82 | 2,117 | 3.9 | 1,269 | 0.80 | 1,018 | 1.3 |
| | (0.39) | | (0.40) | | (0.35) | | (0.35) | |
| South | 1,619 | 0.86 | 1,387 | 3.6 | 984 | 0.87 | 854 | 1.4 |
| | (0.24) | | (0.26) | | (0.27) | | (0.29) | |
| Midwest | 640 | 0.75 | 481 | 3.0 | 388 | 0.82 | 318 | 1.0 |
| | (0.10) | | (0.09) | | (0.11) | | (0.11) | |
| Total | 6,631 | 0.79 | 5,239 | 3.1 | 3,634 | 0.80 | 2,903 | 1.0 |

*Notes:* Table shows descriptive statistics for labor unions (columns (1)-(4)) and employer associations (columns (5)-(6)) registered in the *Cadastro Nacional de Entidades Sindicais* (CNES). Each section of the table reports statistics for a mutually exclusive and collectively exhaustive characterization of union types. Hence, the shares reported in parentheses add to one within each section. Columns (1)-(2) and (5)-(6) consider all unions and employer associations, while columns (3)-(4) and (7)-(8) restrict to the entities that signed at least one CBA in 2009-2016 based on *Sistema Mediador* data. Unions can be involved in the negotiation of either sectoral or firm-level CBAs, while employer association only participate in sectoral negotiations. The sample excludes rural unions and public servants' unions.

## Table A2: Baseline Panel of Negotiations

| | Count (unique observations) | | | | Average over unique CBAs | | | |
|---|---|---|---|---|---|---|---|---|
| | CBAs | Workers covered | Establishments covered | Main unions | Clause count | Clause type count | Duration in months | Negotiation in months |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| **Panel A: Firm-level CBAs** | | | | | | | | |
| 2009 | 26,807 | 4,509,503 | 45,500 | 2,792 | 21.7 | 15.1 | 13.9 | 4.3 |
| 2010 | 28,987 | 5,144,924 | 53,025 | 2,896 | 22.5 | 15.8 | 13.9 | 4.1 |
| 2011 | 31,549 | 5,546,271 | 58,770 | 3,068 | 22.1 | 15.5 | 13.9 | 3.8 |
| 2012 | 33,977 | 6,011,904 | 69,348 | 3,168 | 22.5 | 15.4 | 13.7 | 3.9 |
| 2013 | 34,267 | 6,124,469 | 68,327 | 3,196 | 22.3 | 15.5 | 13.8 | 4.0 |
| 2014 | 34,396 | 6,031,869 | 70,509 | 3,187 | 22.6 | 15.7 | 13.6 | 4.1 |
| 2015 | 33,064 | 5,475,811 | 55,250 | 3,093 | 22.8 | 15.9 | 13.6 | 4.2 |
| 2016 | 33,364 | 5,374,480 | 65,228 | 3,047 | 24.4 | 17.0 | 13.8 | 4.2 |
| **Panel B: Sectoral CBAs** | | | | | | | | |
| 2009 | 4,976 | 9,691,834 | 429,231 | 2,287 | 42.5 | 28.4 | 13.4 | 3.8 |
| 2010 | 5,554 | 11,231,217 | 479,922 | 2,465 | 42.3 | 28.2 | 13.4 | 3.7 |
| 2011 | 5,996 | 12,247,152 | 524,207 | 2,587 | 43.3 | 28.6 | 13.6 | 3.4 |
| 2012 | 6,257 | 12,761,485 | 540,760 | 2,685 | 42.7 | 28.3 | 13.4 | 3.3 |
| 2013 | 6,095 | 12,995,819 | 542,883 | 2,594 | 43.8 | 29.0 | 13.3 | 3.8 |
| 2014 | 5,819 | 12,105,854 | 497,810 | 2,570 | 44.2 | 29.2 | 13.3 | 4.0 |
| 2015 | 5,653 | 11,287,985 | 494,380 | 2,529 | 44.3 | 29.4 | 13.5 | 4.1 |
| 2016 | 5,831 | 10,810,918 | 482,538 | 2,561 | 43.8 | 29.0 | 13.1 | 4.0 |

*Notes:* Table provides year-specific descriptive statistics for establishment-year panels that track negotiations from firm-level CBAs (Panel A) and sectoral CBAs (panel B). Columns (1)-(4) provide counts of the unique CBAs, workers, establishments, and unions in the panel. Coverage is based on end-of-year coverage status. Columns (5)-(8) are averages over unique CBA observations. "Clause count" is the number of clauses, while "clause type count" is the number of unique clause types (137 in total). "Duration" refers to the difference between the start and end date of a CBA, while "Negotiation" refers to the difference between the start and the filing date of a CBA. Creating this panel requires determining who in RAIS is covered by which CBA from *Sistema Mediador*. In the case of firm-level CBAs, I consider all establishments in the municipalities of coverage that belong to the signing firms. The assigned CBA is the one negotiated by the "main union", i.e., the modal union across establishment-CBA-union observations. In the case of sectoral CBAs, I consider all establishments in the municipalities of coverage that contribute to the signing employer association. The assigned CBA is the one negotiated by the "main union", i.e., the modal union across establishment-CBA-union observations among those negotiating CBAs with binding wage floors. Refer to Appendix C for more details.

Table A3: Clause Subgroups and Clause Types (continued in next page)

| Wage adjustments | Wage payment | Other wages | Other wage rules |
|---|---|---|---|
| Wage deductions | Wage payment | Weekly rest wage | Other wage rules |
| Wage isonomy | | Intern wage | |
| Wage floors | | Production wage | |
| Wage increases | | | |

| Separations | Contract types | Hiring | Other contracting |
|---|---|---|---|
| Advance notice | Part-time | Female workforce | Other contracting |
| Separations | Internships | Youth workforce | |
| Suspensions | | Outsourcing | |
| | | Advanced-age workforce | |
| | | Hiring rules | |
| | | Other workforces | |
| | | Special needs workforce | |

| Staffing | Working conditions | Employment protections | Workday |
|---|---|---|---|
| Work functions adaptation | Moral harassment | Abortion protections | Sundays and holidays |
| Work function assignment | Sexual harassment | Injury protections (work-related) | Workday compensation |
| Performance evaluation | Tools and equipment | Adoption protections | Workday controls |
| Other staffing rules | Equal opportunity | Retirement protections | Weekly rest |
| Worker participation in management | Disciplinary norms | Apprenticeship protections | Duration and schedule |
| Task-and-wage schedule | Other rules on working conditions | Employment protections | Absences |
| Vocational training | | Maternity protections | Break intervals |
| Transfers | | Paternity protections | Special shifts |
| | | Injury protections (nonwork-related) | Other workday provisions |
| | | Military service protections | Workday extension/reduction |
| | | Other employment protections | On-call shifts |
| | | Policy for dependents | Uninterrupted shifts |
| | | Policies for maintaining employment | |
| | | Employment protection program | |

*Notes:* Table shows the clause types that belong to each of the 24 clause subgroups provided by *Sistema Mediador*. The clause subgroups in blue refer to *wage-related* amenities, in green to *employment-related* amenities, and in gray to *other* amenities. There are a total of 137 clause types, unevenly distributed across the subgroups.

## Clause Subgroups and Clause Types (continued from previous page)

| Bonuses | Pays | Assistances | Other incentives |
|---|---|---|---|
| 13th month bonus | Night pay | Food assistance | Retirement |
| Work function bonus | Overtime pay | Childcare assistance | Fees |
| Other bonuses | Hazard pay (health) | Illness assistance | Loans |
| | Shift pay | Education assistance | Profit sharing |
| | Hazard pay (danger) | Housing assistance | Awards |
| | On-call pay | Maternity assistance | Family salary |
| | Seniority pay | Funeral assistance | Life insurance |
| | Other pays | Health assistance | |
| | | Transportation assistance | |
| | | Subsistence allowance | |
| | | Other assistances | |

| Injuries | Prevention | Union organization | Union-firm relations |
|---|---|---|---|
| Medical certificates | Accident prevention committee | Factory commission | Union access to information |
| Work-related injuries | Health education campaigns | Union dues | Union access to workplace |
| Nonwork-related injuries | Working environment conditions | Opposition to union dues | Guarantees to union officers |
| Other injury provisions | Equipments for individual safety | Other union provisions | Union activities leave |
| Rehabilitation of the injured | Safety equipment | Union representatives | Other union-firm provisions |
| | Medical exams | Unionization campaigns | Strikes |
| | Insalubrity | | |
| | Machine and equipment maintenance | | |
| | Other injury prevention standards | | |
| | Hazards | | |
| | First aid | | |
| | Health and safety professionals | | |
| | Training for injury prevention | | |
| | Uniforms | | |

| Vacations | Leaves | Other time off provisons | Bargaining |
|---|---|---|---|
| Vacation days and duration | Abortion leave | Other time off provisions | CBA enforcement |
| Collective vacations | Adoption leave | | CBA non-compliance |
| Vacation remuneration | Maternity leave | | Conflict resolution |
| | Paid leave | | Other bargaining provisions |
| | Unpaid leave | | CBA negotiation rules |
| | | | CBA renewal/termination |

Table A4: Differences in Firm-level CBAs by Industry

| | Mean CBA clause count | Mean share of clauses: broad groups | | | Top 5 overrepresented subgroups (by rank gain: industry vs. overall) |
|---|---|---|---|---|---|
| | | Wage | Employment | Other | |
| Manufacturing | 24.3 | 13.9% | 18.6% | 67.5% | Hiring<br>Other contracting<br>Prevention<br>Other incentives<br>Injuries |
| Trade/commerce | 20.1 | 13.3% | 17.1% | 69.6% | Other contracting<br>Pay<br>Separations<br>Wage payment<br>Other incentives |
| Banking/finance | 48.8 | 3.4% | 6.5% | 90.2% | Leaves<br>Vacations<br>Injuries<br>Other time off provisions<br>Bonuses |
| Transportation | 31.8 | 13.5% | 20.5% | 66.0% | Wage payment<br>Other contracting<br>Prevention<br>Workday<br>Employment protections |

*Notes:* Table shows differences in CBA content by four broad industries using establishments-level observations among those in the baseline panel that negotiated a firm-level CBA in 2009. These industries are the most represented and account for more than two-thirds of the analysis sample. For each industry, the table reports the mean number of CBA clauses, the share of clauses by broad groups (i.e., wage-related, employment-related, and other amenities), and the five most overrepresented clause subgroups in the industry. Overrepresented subgroups are defined by creating industry-specific and overall rankings based on mean clause count, calculating the "rank gain" by comparing the industry-specific to the overall ranking, and choosing the five subgroups experiencing the largest "rank gain."

Table A5: Changes in Clause Count by Subgroup in Firm-level CBAs

| | | Change in clause count (2009-2010) | | | |
|---|---|---|---|---|---|
| Rank | Clause subgroup | Mean | Std dev | 10th perc | 90th perc |
| 1 | Bargaining | 4.83 | 17.54 | -1 | 10 |
| 2 | Assistances | -1.46 | 4.76 | -12 | 2 |
| 3 | Workday | -0.51 | 3.42 | -7 | 2 |
| 4 | Wage adjustments | 0.08 | 2.52 | -1 | 1 |
| 5 | Prevention | -0.39 | 2.23 | -4 | 1 |
| 6 | Union organization | -0.84 | 2.19 | -6 | 1 |
| 7 | Other incentives | 0.28 | 2.05 | -1 | 2 |
| 8 | Pay | -0.51 | 1.83 | -4 | 1 |
| 9 | Working conditions | -0.36 | 1.77 | -2 | 1 |
| 10 | Union-firm relations | 0.05 | 1.76 | -1 | 1 |

*Notes:* Table reports statistics for the change in clause count by subgroup using establishments in the baseline panel that negotiated a firm-level CBA in 2009 and 2010. For each clause subgroup, the sample is restricted to the establishments where the corresponding CBA has a nonzero clause count in either year. The reported statistics are the mean, standard deviation, 10th percentile, and 90th percentile. The subgroups are ranked by standard deviation (descending order) and the table reports results for the top ten subgroups.

Table A6: Relation between Changes in Subgroup Clauses and Workplace Outcomes

| | Hiring rate (1) | Turnover rate (2) | Share with fixed contract (3) | Retention rate (4) | Share with safety events (5) | Share taking leave (6) |
|---|---|---|---|---|---|---|
| **Panel A: Contracting clauses** | | | | | | |
| Separations | -4.37% | -1.99% | 0.81% | 0.35% | 1.04% | 1.50% |
| | [0.000] | [0.018] | [0.435] | [0.044] | [0.638] | [0.057] |
| Contract types | 16.28% | 12.88% | -13.74% | 0.32% | 6.08% | 0.13% |
| | [0.007] | [0.073] | [0.017] | [0.721] | [0.508] | [0.965] |
| Hiring | -4.30% | -1.84% | -0.09% | 0.61% | 0.89% | -0.07% |
| | [0.002] | [0.075] | [0.951] | [0.003] | [0.671] | [0.931] |
| Other contracting | -1.57% | 2.47% | 0.36% | 0.05% | 4.51% | 0.50% |
| | [0.431] | [0.331] | [0.863] | [0.792] | [0.137] | [0.273] |
| **Panel B: Incentives clauses** | | | | | | |
| Bonuses | -8.30% | -3.84% | 2.24% | 0.79% | 2.00% | 1.07% |
| | [0.000] | [0.000] | [0.061] | [0.000] | [0.267] | [0.003] |
| Pays | -6.36% | -1.36% | 0.47% | 0.79% | 0.69% | 0.79% |
| | [0.000] | [0.137] | [0.513] | [0.000] | [0.650] | [0.018] |
| Assistances | -1.79% | -0.95% | 0.86% | 0.28% | 0.74% | 0.48% |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.153] | [0.001] |
| Other incentives | 0.92% | -0.47% | -1.45% | 0.18% | -2.25% | -0.58% |
| | [0.414] | [0.495] | [0.014] | [0.080] | [0.213] | [0.057] |
| **Panel C: Workplace safety clauses** | | | | | | |
| Injuries | -4.30% | -3.13% | 1.57% | 0.99% | -3.86% | -0.97% |
| | [0.000] | [0.003] | [0.539] | [0.000] | [0.276] | [0.169] |
| Prevention | -3.53% | -2.01% | -0.15% | 0.53% | 2.68% | 0.38% |
| | [0.000] | [0.002] | [0.843] | [0.000] | [0.124] | [0.217] |
| **Panel D: Time off clauses** | | | | | | |
| Vacations | -14.77% | -6.23% | 5.83% | 1.30% | 1.64% | 1.16% |
| | [0.000] | [0.000] | [0.000] | [0.000] | [0.623] | [0.058] |
| Leaves | -12.12% | -5.38% | -5.14% | 1.45% | 10.80% | 2.20% |
| | [0.000] | [0.004] | [0.249] | [0.000] | [0.020] | [0.007] |
| Other time off provisions | -11.25% | -5.98% | 1.31% | 1.41% | -2.70% | 0.43% |
| | [0.000] | [0.000] | [0.513] | [0.000] | [0.427] | [0.430] |

*Notes:* Table reports the association between changes in labor market statistics and changes in CBA clauses between 2009 and 2010. The sample is restricted to establishments in the baseline panel that negotiated firm-level CBAs in both 2009 and 2010. The estimated associations are reported in percentage terms relative to the 2009 mean (corresponding p-values are reported in brackets). Each estimate comes from a separate regression restricted to establishments that negotiated at least one of the corresponding clauses in either year. Highlighted cells are those selected from an adaptive lasso of the change in the labor market statistic on changes in mean log hourly earnings and changes in CBA clauses (for all 24 subgroups) regardless of whether any specific clause was negotiated in either year. Cells highlighted in green are those where the lasso coefficient is positive, whereas gray correspond to negative lasso coefficients. Each panel focuses on similar clause subgroups that are matched to outcomes of relevance (demarcated by partially enclosed columns). Specifically, Panel A contains subgroups linked to outcomes in columns (1)-(3); Panel B to column (4); Panel C to column (5); and Panel D to column (6).

Table A7: Falsification Tests for Confounding Trends

| | Establishment: change in employment | | Union: change in leave-out mean negotiation length | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Coverage gap at renewal | -0.871 | - | 0.125 | - |
| | (1.459) | - | (0.131) | - |
| Duration of original CBA | - | 0.157 | - | 0.030 |
| | - | (0.106) | - | (0.042) |
| Adjusted R2 | 0.015 | 0.015 | 0.572 | 0.574 |
| Observations | 37,030 | 37,030 | 34,634 | 34,634 |

*Notes:* Table provides falsification tests for whether characteristics that influence treatment assignment are correlated with trends in the pre-treatment period that could confound identification. The two characteristics considered are coverage gaps at renewal and the duration of the original CBA. Coverage gaps influence treatment assignment since not having CBA coverage in September 2012 determines treatment status. The duration of a CBA also influences treatment assignment since longer lasting CBAs increase the likelihood that one has CBA coverage when ultractivity is enacted. The two trends considered that could confound identification are changes in employment at the establishment level (columns (1)-(2)) and changes in the leave-out mean negotiation length of the union (columns (3)-(4)). The former captures establishment growth, while the latter proxies for trends in union effectiveness. Each column is a separate regression of the corresponding trend variable on the characteristic of interest, controlling for fixed effects by industry, microregion, and negotiation month. Standard errors are clustered at the establishment-level for columns (1)-(2) and at the union-level for columns (3)-(4). The sample consists of the establishments with a firm-level CBA negotiated in 2009 that is renewed.

Table A8: Impact of Union Bargaining Power (Union-level Clustering)

| **Panel A: Impact on wages** | | | | | | |
|---|---|---|---|---|---|---|
| | Log contracted wage (1a) | Log contracted wage (hourly) (2a) | Log average earnings (hourly) (3a) | Log December earnings (hourly) (4a) | 90-10 log wage ratio (5a) | 90-50 log wage ratio (6a) |
| $D_i \times \delta_{t\geq 2012}$ | 0.016 | 0.019 | 0.011 | 0.010 | -0.012 | -0.015 |
| | (0.007) | (0.008) | (0.005) | (0.005) | (0.012) | (0.010) |
| Baseline mean | 12.283 | 7.142 | 7.367 | 7.421 | 0.947 | 0.610 |
| Observations | 183,828 | 183,828 | 183,532 | 183,259 | 183,828 | 183,828 |
| **Panel B: Impact on employment** | | | | | | |
| | Log employment (1b) | Hiring rate (2b) | Retention rate (3b) | Turnover rate (4b) | Turnover rate: layoffs (5b) | Turnover rate: quits (6b) |
| $D_i \times \delta_{t\geq 2012}$ | 0.002 | -0.025 | 0.009 | -0.019 | -0.006 | -0.004 |
| | (0.017) | (0.019) | (0.005) | (0.014) | (0.006) | (0.003) |
| Baseline mean | 3.286 | 0.623 | 0.797 | 0.396 | 0.185 | 0.078 |
| Observations | 183,854 | 185,895 | 183,673 | 185,895 | 185,895 | 185,895 |
| **Panel C: Impact on CBAs** | | | | | | |
| | Clause count (1c) | Clause count: unique types (2c) | Wage clauses (3c) | Employment clauses (4c) | Other clauses (5c) | |
| $D_i \times \delta_{t\geq 2012}$ | 1.556 | 0.954 | 0.256 | 0.750 | 0.548 | |
| | (0.755) | (0.537) | (0.114) | (0.159) | (0.655) | |
| Baseline mean | 31.056 | 19.385 | 2.780 | 4.085 | 24.189 | |
| Observations | 136,912 | 136,912 | 136,912 | 136,912 | 136,912 | |

*Notes:* Table reports the pooled DID estimates based on Equation (2) for outcomes related to (A) wages; (B) employment; and (C) collective bargaining agreements. In other words, it reports the estimates for $\beta$ in the following specification $Y_{jt} = \beta(D_i \times \delta_{t\geq 2012}) + \alpha_j + \delta_t + \gamma X_{jt} + \varepsilon_{jt}$ along with standard errors clustered at the union-level in parentheses. Refer to Appendix C for definitions of the outcomes. The unit of observation is the establishment-year restricted to establishments in the analysis sample. The baseline mean refers to the average outcome in 2009.

Table A9: Impact of Union Bargaining Power on Workforce Composition

| **Panel A: Education and occupation** | | | | | | |
|---|---|---|---|---|---|---|
| | No high school degree | High school degree only | College degree | Managers | Professionals | Rank-and-file workers |
| | (1a) | (2a) | (3a) | (4a) | (5a) | (6a) |
| $D_i \times \delta_{t\geq 2012}$ | -0.005 | -0.005 | 0.010 | -0.008 | 0.008 | 0.000 |
| | (0.002) | (0.002) | (0.002) | (0.001) | (0.002) | (0.002) |
| Baseline mean | 0.272 | 0.462 | 0.266 | 0.075 | 0.170 | 0.755 |
| Observations | 183,854 | 183,854 | 183,854 | 183,854 | 183,854 | 183,854 |
| **Panel B: Gender, race, and age** | | | | | | |
| | Male workers | White workers | Age: below 30 | Age: 30-40 | Age: above 40 | |
| | (1b) | (2b) | (3b) | (4b) | (5b) | |
| $D_i \times \delta_{t\geq 2012}$ | 0.000 | 0.002 | -0.003 | 0.003 | 0.000 | |
| | (0.002) | (0.003) | (0.002) | (0.002) | (0.002) | |
| Baseline mean | 0.620 | 0.693 | 0.345 | 0.310 | 0.354 | |
| Observations | 183,854 | 183,854 | 183,854 | 183,854 | 183,854 | |

*Notes:* Table reports the pooled DID estimates based on Equation (2) for outcomes related to the composition of the workforce. In other words, it reports the estimates for $\beta$ in the following specification $Y_{jt} = \beta(D_i \times \delta_{t\geq 2012}) + \alpha_j + \delta_t + \gamma X_{jt} + \varepsilon_{jt}$ along with standard errors clustered at the establishment-level in parentheses. Refer to Appendix C for definitions of the outcomes. The unit of observation is the establishment-year restricted to establishments in the analysis sample. The baseline mean refers to the average outcome in 2009.

Table A10: Impact of Union Bargaining Power by Clause Count in Pre-period Renewal

| | Log employment | Log contracted wage | Clause count | Wage clauses | Employment clauses | Other clauses |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $D_i \times \delta_{t\geq 2012}$ | -0.011 | 0.018 | 1.216 | 0.123 | 0.503 | 0.591 |
| | (0.012) | (0.006) | (0.207) | (0.035) | (0.052) | (0.156) |
| $D_i \times \delta_{t\geq 2012} \times HighClause_i$ | 0.038 | -0.003 | 2.289 | 0.566 | 0.951 | 0.767 |
| | (0.020) | (0.009) | (0.487) | (0.082) | (0.124) | (0.342) |
| Baseline mean | 3.286 | 12.283 | 31.056 | 2.780 | 4.085 | 24.189 |
| Observations | 183,854 | 183,828 | 136,912 | 136,912 | 136,912 | 136,912 |

*Notes:* Table summarizes the heterogeneous effects of an increase in union bargaining power on employment (column (1)); wages (column (2)); and CBA clauses (columns (3)-(6)). Specifically, the table reports the estimates for $\beta_1$ and $\beta_3$ in the following fully interacted specification $Y_{jt} = \beta_1(D_i \times \delta_{t\geq 2012}) + \beta_2(H_i \times \delta_{t\geq 2012}) + \beta_3(D_i \times \delta_{t\geq 2012} \times H_i) + \alpha_j + \delta_t + \gamma X_{jt} + \varepsilon_{jt}$, where $H_i = HighClause_i$ is an indicator for having above median clauses in the CBA renewal prior to the shock. Refer to Appendix C for definitions of the outcomes. The unit of observation is the establishment-year restricted to establishments in the analysis sample. The baseline mean refers to the average outcome in 2009.

Table A11: Impact of Union Bargaining Power (IV Strategy)

| | Treat | Log employment | Log contracted wage | Clause count |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| **Panel A: First stage** | | | | |
| Fast pre-period renewal | 0.133 | - | - | - |
| | (0.010) | - | - | - |
| **Panel B: Second stage (pre-period)** | | | | |
| Treat | - | -0.075 | 0.027 | 2.161 |
| | - | (0.062) | (0.031) | (1.689) |
| **Panel C: Second stage (post-period)** | | | | |
| Treat | - | -0.037 | 0.048 | 3.580 |
| | - | (0.055) | (0.025) | (1.559) |
| F-statistic | 179.7 | - | - | - |
| Cragg-Donald F-stat | - | 201.0 | 200.8 | 196.6 |
| Kleibergen-Paap F-stat | - | 184.5 | 184.3 | 179.8 |
| Observations | 23,528 | 23,445 | 23,442 | 23,527 |

*Notes:* Table presents results for the complementary instrumental variable strategy to identify the effects of increased union bargaining power on employment, wages, and CBA clauses. The instrument for treatment is a dummy for "fast pre-period renewal", i.e., whether the negotiation of the 2009 CBA's renewal in the pre-period lasted less than 1 month. Panel A reveals a strong first stage. Panel B reports the 2SLS estimates from regressing the first differenced outcome (using pre-period observations) on treatment. Panel C reports the 2SLS estimates from regressing the first differenced outcome (using post-period observations) on treatment. The F-statistics for weak identification tests correspond to the estimates in Panel C. For columns (2) and (3), the pre-period first differences are from 2010 to 2011 and the post-period first differences are from 2011 to 2012. For column (4), the pre-period first difference is from the 2009 CBA to its renewal and the post-period first difference is from the 2009 CBA renewal to the CBA negotiated in 2013. All regressions include time-varying fixed effects for detailed industry, microregion, and negotiation month. Standard errors are clustered at the establishment level.

Table A12: Impact of Union Bargaining Power on Authority Measures

| **Panel A: Overall** | | | | | |
|---|---|---|---|---|---|
| | Provisions | Labor: entitlements and permissions | Labor: obligations and constraints | Capital: entitlements and permissions | Capital: obligations and constraints |
| | (1a) | (2a) | (3a) | (4a) | (5a) |
| $D_i \times \delta_{t\geq 2012}$ | 1.649 | 0.436 | 0.036 | 0.140 | 1.037 |
| | (0.184) | (0.043) | (0.041) | (0.023) | (0.115) |
| | | [26%] | [2%] | [8%] | [63%] |
| Baseline mean | 17.690 | 5.030 | 4.071 | 0.910 | 7.680 |
| Observations | 136,912 | 136,912 | 136,912 | 136,912 | 136,912 |
| **Panel B: Labor (workers and unions)** | | | | | |
| | Provisions | Entitlements | Permissions | Obligations | Constraints |
| | (1b) | (2b) | (3b) | (4b) | (5b) |
| $D_i \times \delta_{t\geq 2012}$ | 0.472 | 0.204 | 0.232 | 0.046 | -0.010 |
| | (0.075) | (0.032) | (0.019) | (0.038) | (0.009) |
| | | [43%] | [49%] | [10%] | -[2%] |
| Baseline mean | 9.100 | 3.880 | 1.149 | 3.472 | 0.599 |
| Observations | 136,912 | 136,912 | 136,912 | 136,912 | 136,912 |
| **Panel C: Capital (firms and managers)** | | | | | |
| | Provisions | Entitlements | Permissions | Obligations | Constraints |
| | (1c) | (2c) | (3c) | (4c) | (5c) |
| $D_i \times \delta_{t\geq 2012}$ | 1.177 | 0.045 | 0.094 | 1.005 | 0.032 |
| | (0.129) | (0.007) | (0.019) | (0.111) | (0.007) |
| | | [4%] | [8%] | [85%] | [3%] |
| Baseline mean | 8.590 | 0.215 | 0.695 | 7.473 | 0.207 |
| Observations | 136,912 | 136,912 | 136,912 | 136,912 | 136,912 |

*Notes:* Table reports the pooled DID estimates based on Equation (2) for the authority measures extracted from the text of CBAs. In other words, it reports the estimates for $\beta$ in the following specification $Y_{jt} = \beta(D_i \times \delta_{t\geq 2012}) + \alpha_j + \delta_t + \gamma X_{jt} + \varepsilon_{jt}$ along with standard errors clustered at the establishment-level in parentheses. Authority measures capture the authority of agents in a legal document where 1) *obligations* and *constraints* reduce authority; and 2) *entitlements* and *permissions* expand authority. These measures rely on an unsupervised extraction of these provisions among agents of capital (firms and managers; Panel C) and labor (workers and unions; Panel B) in the CBA. Refer to Appendix E for more details. The percentage of the overall effects in column (1) explained by the headings in columns (2)-(5) are reported in brackets. The unit of observation is the establishment-year restricted to establishments in the analysis sample. The baseline mean refers to the average outcome in 2009.

Table A13: Relation of PageRank Values to Wage Premiums and CBA Clause Count

| | PageRank: pre-period | | | PageRank: first differences | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Wage premium | 1.3296 | | 1.3291 | 0.2164 | | 0.2153 |
| | (0.0218) | | (0.0218) | (0.0243) | | (0.0243) |
| CBA clause count | | 0.0006 | 0.0005 | | 0.0009 | 0.0008 |
| | | (0.0002) | (0.0002) | | (0.0003) | (0.0003) |
| Adjusted R2 | 0.266 | 0.206 | 0.267 | 0.046 | 0.045 | 0.046 |
| Observations | 55,898 | 55,898 | 55,898 | 55,898 | 55,898 | 55,898 |

*Notes:* Table shows estimates of coefficients from regressing PageRank values on AKM wage premiums and CBA clause counts. The unit of observation is the establishment-period with $t \in [2007\text{-}2011, 2012\text{-}2016]$. Columns (1)-(3) focus on the pre-period and columns (4)-(5) use first-differences, i.e., changes between the two periods. Regressions include detailed industry, microregion, and negotiation month fixed effects. Standard errors are clustered at the establishment-level. The sample is restricted to the analysis sample constructed with coverage from sectoral CBAs, where the establishments also have PageRank values and AKM wage premiums in each $t$, i.e., they are in the strongly connected set of establishments linked by worker flows.

Table A14: Wage-equivalent Value of CBA Clauses (Broad Groups)

| | Estimate (1) | Std. error (2) | Value (3) |
|---|---|---|---|
| Wage clauses | 0.002 | (0.001) | 0.010 |
| Employment clauses | 0.002 | (0.001) | 0.010 |
| Other clauses | 0.001 | (0.000) | 0.003 |
| Wage premium | 0.214 | (0.024) | 1.000 |
| Adjusted R2 | 0.047 | | |
| Observations | 55,898 | | |

*Notes:* Table displays the estimated wage-equivalent value for the 3 broad groups of clauses. Column (1) reports the point estimates of $\alpha$ and $\beta^z$ from the specification in first-differences given by Equation (4). Column (2) report the standard error of the estimates. Column (3) reports the wage-equivalent value, which is defined as the impact of the clause type divided by that of the wage premium, i.e., $\widehat{\beta}^z/\widehat{\alpha}$. The sample includes establishments from an analysis sample constructed as in Section 3.1 that uses sectoral CBA coverage, rather than firm-level CBA coverage. Moreover, establishments must have PageRank values and AKM wage premiums in both periods, i.e.,they are in the strongly connected set of establishments linked by worker flows for each $t \in [2007\text{-}2011, 2012\text{-}2016]$.

Table A15: Wage-equivalent Value of CBA Clauses (Clause Types)

| | Estimate (1) | Std. error (2) | Value (3) |
|---|---|---|---|
| **Wage clauses** | | | |
| Wage floors | 0.005 | (0.002) | 0.026 |
| Intern wage | -0.039 | (0.019) | -0.184 |
| **Employment clauses** | | | |
| Internships | 0.029 | (0.016) | 0.138 |
| Hiring rules | 0.009 | (0.004) | 0.041 |
| Worker participation in management | 0.178 | (0.074) | 0.835 |
| Equal opportunity | -0.028 | (0.015) | -0.132 |
| Retirement protections | 0.030 | (0.012) | 0.139 |
| Apprenticeship protections | 0.166 | (0.088) | 0.778 |
| Paternity protections | 0.055 | (0.033) | 0.258 |
| **Other clauses** | | | |
| Overtime pay | -0.014 | (0.008) | -0.066 |
| Night pay | 0.028 | (0.014) | 0.130 |
| Illness assistance | -0.029 | (0.014) | -0.136 |
| Funeral assistance | -0.024 | (0.014) | -0.111 |
| Retirement | -0.021 | (0.012) | -0.099 |
| Workday compensation | 0.008 | (0.005) | 0.039 |
| Absences | 0.011 | (0.005) | 0.050 |
| Special shifts | 0.015 | (0.007) | 0.071 |
| Work-related injuries | -0.032 | (0.016) | -0.149 |
| Rehabilitation of the injured | -0.077 | (0.029) | -0.362 |
| Working environment conditions | 0.008 | (0.004) | 0.035 |
| Safety equipment | 0.034 | (0.013) | 0.158 |
| Medical exams | -0.020 | (0.010) | -0.096 |
| Other injury prevention standards | 0.023 | (0.008) | 0.107 |
| Uniforms | -0.029 | (0.013) | -0.136 |
| Collective vacations | -0.045 | (0.019) | -0.210 |
| Unpaid leave | 0.070 | (0.018) | 0.330 |
| Paid leave | 0.024 | (0.008) | 0.114 |
| Guarantees to union officers | -0.021 | (0.012) | -0.098 |
| Union dues | 0.013 | (0.003) | 0.059 |
| Opposition to union dues | 0.032 | (0.009) | 0.152 |
| Union representatives | -0.018 | (0.009) | -0.083 |
| Unionization campaigns | 0.031 | (0.012) | 0.147 |
| CBA enforcement | -0.011 | (0.004) | -0.053 |
| CBA renewal/termination | -0.008 | (0.004) | -0.036 |
| Wage premium | 0.213 | (0.024) | |
| Adjusted R2 | 0.050 | | |
| Observations | 55,898 | | |

*Notes:* Table displays the estimated wage-equivalent value for the 34 (out of 137) clause types that have an impact on workers' value of employment at an establishment. Column (1) reports the point estimates of $\alpha$ and $\beta^z$ from the specification in first-differences given by Equation (4). Column (2) report the standard error of the estimates. Column (3) reports the wage-equivalent value, which is defined as the impact of the clause type divided by that of the wage premium, i.e., $\widehat{\beta}^z/\widehat{\alpha}$. Sample restrictions are the same as in Table 6.

Table A16: Relation of Amenity Value with Authority Expansion

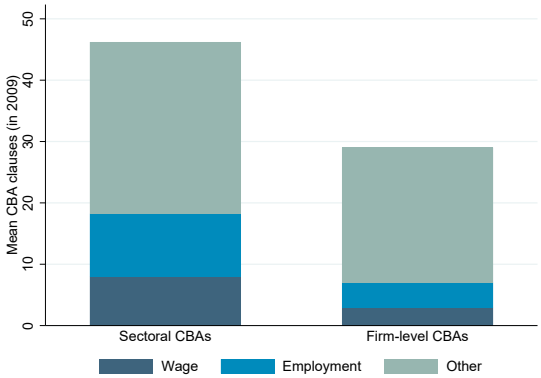
| | Amenity value: baseline panel | | | Amenity value: analysis sample | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Labor−Capital | 0.029 | | | 0.033 | | |
| | (0.001) | | | (0.001) | | |
| Labor | | 0.004 | | | 0.007 | |
| | | (0.001) | | | (0.002) | |
| Workers | | | 0.005 | | | 0.007 |
| | | | (0.001) | | | (0.002) |
| Union | | | -0.004 | | | 0.014 |
| | | | (0.003) | | | (0.006) |
| Capital | | -0.059 | | | -0.062 | |
| | | (0.001) | | | (0.002) | |
| Firms | | | -0.058 | | | -0.060 |
| | | | (0.001) | | | (0.002) |
| Managers | | | -0.164 | | | -0.216 |
| | | | (0.012) | | | (0.020) |
| Observations (K) | 130.158 | 130.158 | 130.158 | 46.697 | 46.697 | 46.697 |

*Notes:* Table shows the association between the amenity value in firm-level CBAs and measures of authority extracted from the text of these CBAs. The amenity value of a CBA is given by $\sum_{clause} 1\{clause = z\} \times (\hat{\beta}^z/\hat{\alpha})$. Authority measures capture the authority of agents in a legal document where 1) *obligations* and *constraints* reduce authority; and 2) *entitlements* and *permissions* expand authority. The independent variables in these regressions refer to the share of provisions in the CBA that expand authority minus those that reduce authority for agents of capital (firms and managers) and of labor (workers and unions). Refer to Appendix E for more details on extracting the provisions of authority from the text of CBAs. Columns (1)-(3) focus on the baseline panel, while columns (4)-(5) focus on the analysis sample. Observations are unique CBAs from the sample period: 2009-2016.

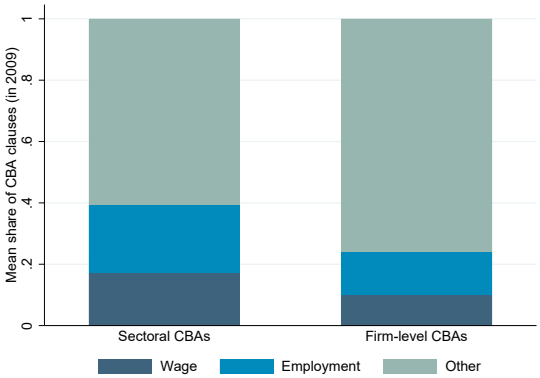
# B Appendix Figures

Figure B1: Content of CBAs by Broad Groups

(a) Clause counts



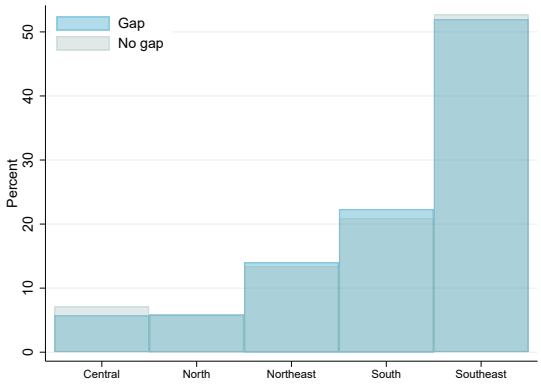

(b) Shares of clauses



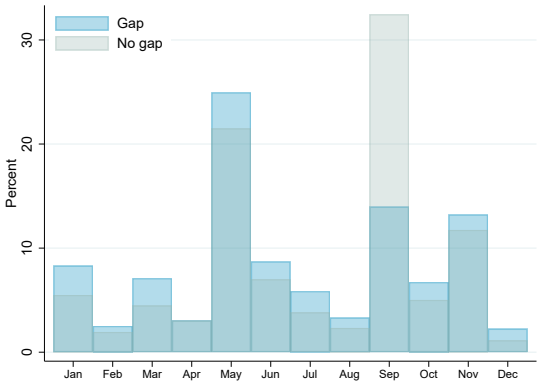

*Notes:* Figures display the content of sectoral and firm-level CBAs negotiated in 2009 in terms of clauses by broad groups, i.e., wage-related amenities, employment-related amenities, and other amenities. Figure B1a shows the mean clauses by broad groups across establishments covered by these CBAs. The mean total clauses for sectoral and firm-level CBAs is 46 and 29, respectively. Figure B1b shows the mean clauses by broad groups as a share of total clauses across establishments covered by these CBAs. Refer to Tables 2 and A3 for more details on the clause subgroups and clause types included in each broad group.

Figure B2: Establishment Characteristics by Presence of Coverage Gaps

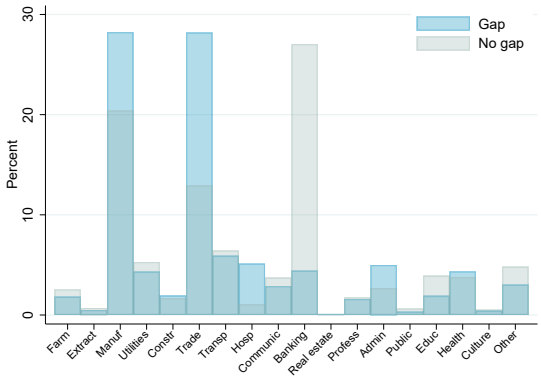
(a) Regions



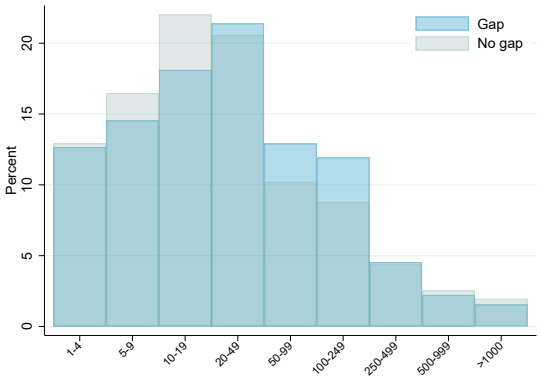

(b) Negotiation month




(c) Industry



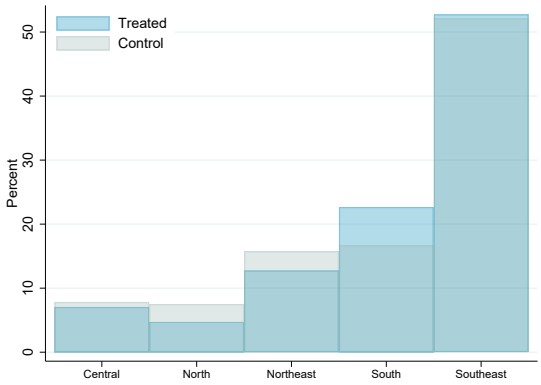

(d) Establishment size



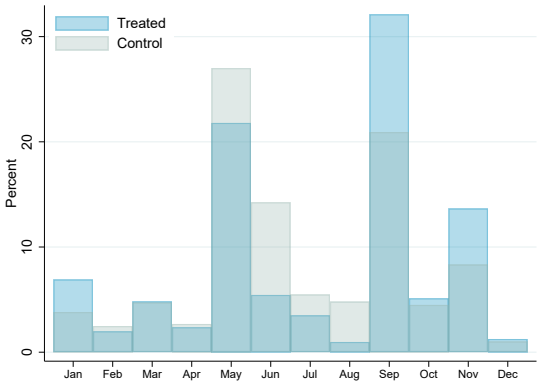

*Notes:* Figures show the distribution of various establishment characteristics among establishments covered by a CBA from 2009 that was renewed by whether the new CBA implied a coverage gap or not. Figure B2a focuses on the distribution across the 5 broad geographical regions in Brazil. Figure B2b concerns the month in which each establishment is due to negotiate with its union. Figure B2c depicts the distribution by broad industries. Figure B2d pertains to establishment size—the difference in average size between establishments with and without coverage gaps is small and statistically insignificant, i.e., 102 versus 109, respectively. The imbalances shown here map closely to those between treated and control establishments in the analysis sample (Figure B3), which motivate the inclusion of time-varying controls for micro-region, negotiation month, and detailed industry in Equation (2).

Figure B3: Establishment Characteristics by Assignment of Treatment

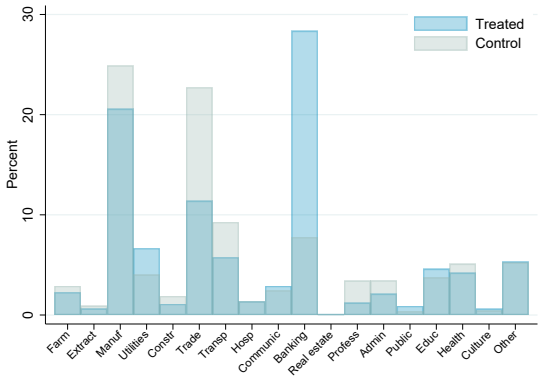
(a) Regions



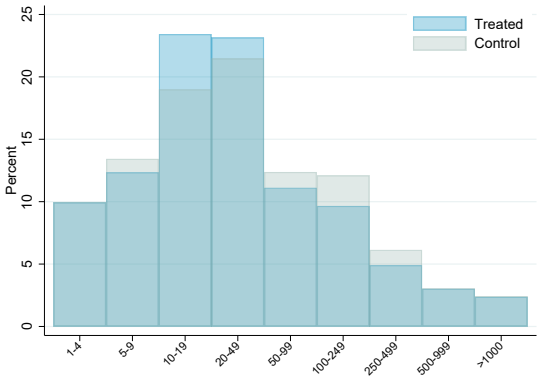

(b) Negotiation month




(c) Industry

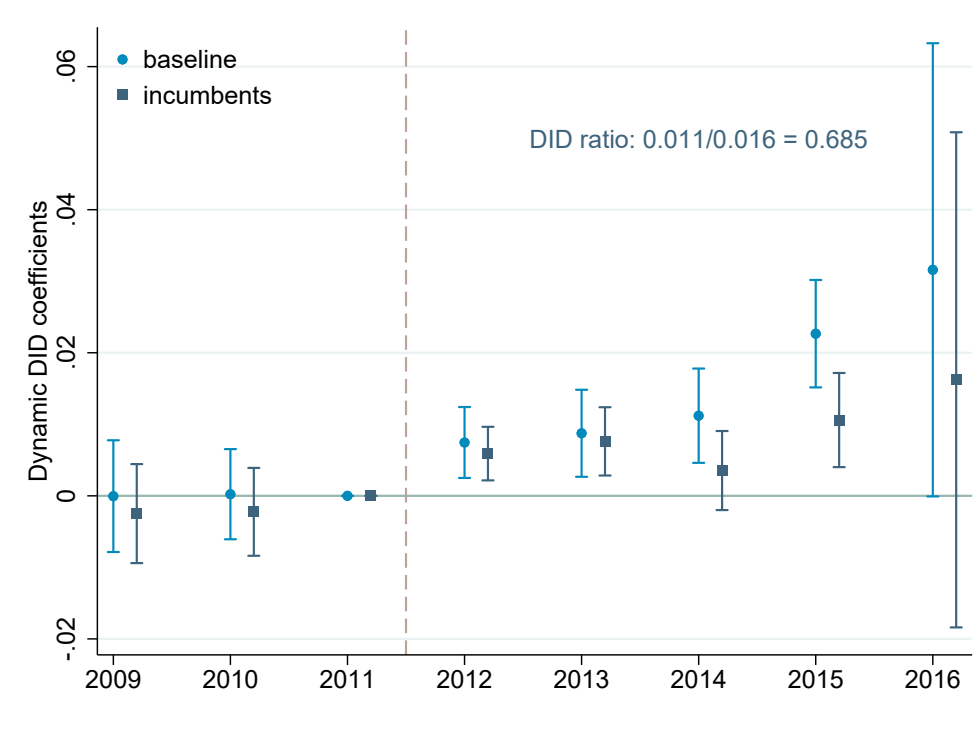



(d) Establishment size



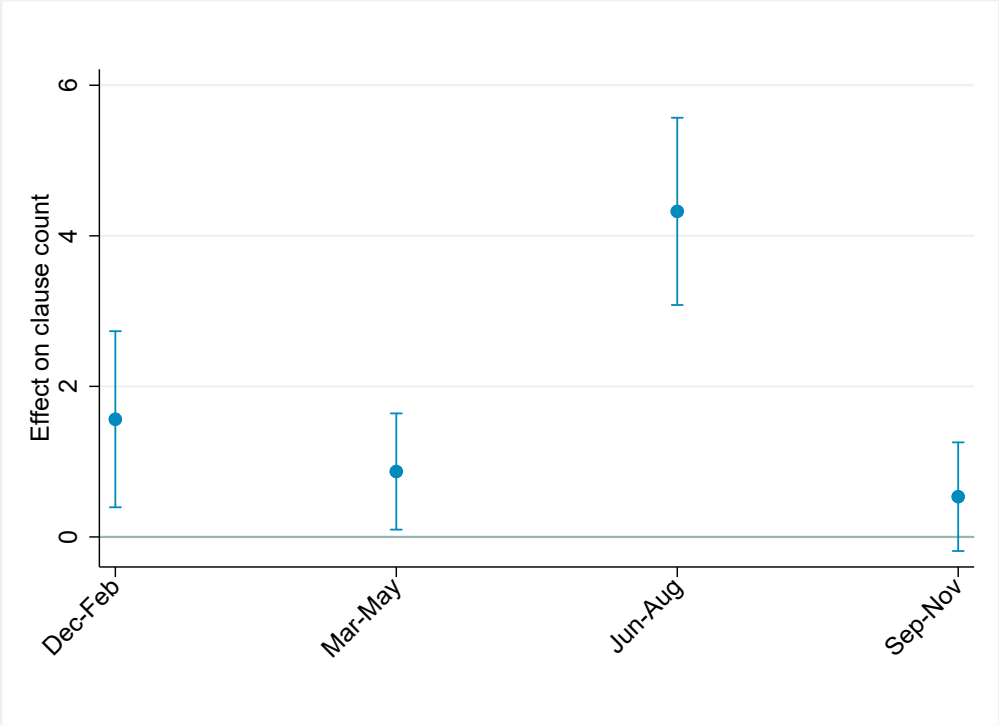

*Notes:* Figures show the distribution of the treated and control establishments in the analysis sample by different characteristics. Figure B3a focuses on the distribution across the 5 broad geographical regions in Brazil. Figure B3b concerns the month in which each establishment is due to negotiate with its union. Figure B3c depicts the distribution by broad industries. Figure B3d pertains to establishment size—the difference in average size between treated and control establishments is small, i.e., 129 versus 133, respectively (Table C1). The imbalances between treated and control motivate the inclusion of time-varying controls for micro-region, negotiation month, and detailed industry in Equation (2).

Figure B4: Event Study on Wages (Baseline versus Incumbents)




*Notes:* The figure overlays the event studies for the impact of union bargaining power on wages with our baseline DID specification (Equation (2)) and one that captures the effects on incumbent workers. For the latter, each year's coefficient is estimated using the first difference in wages relative to 2011 among workers employed at the same establishment in both years. In addition to depicting the estimates of the $\beta^t$ coefficients, the figure reports the incumbent-to-baseline DID ratio, where both the numerator and denominator are the corresponding difference between the average of the post-treatment point estimates and the average of the pre-treatment estimates. The omitted year is 2011, i.e., the year prior to the shock to union bargaining power. Refer to Section 3.1 for more details. The unit of observation is the establishment-year restricted to establishments in the analysis sample.

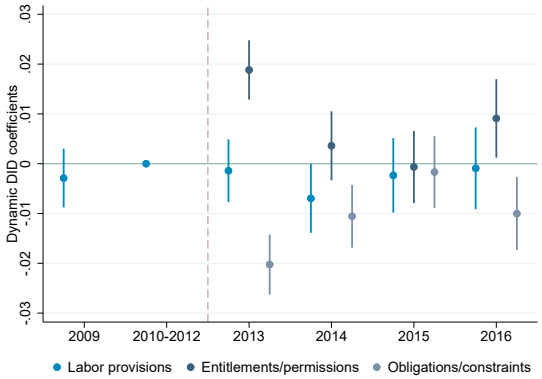
Figure B5: Impact on Clauses by Negotiation Month



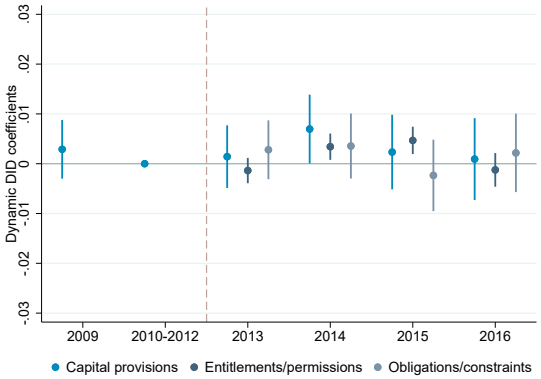

*Notes:* Figure shows the impact of union bargaining power on CBA clauses by negotiation month, i.e., the calendar month when each establishment is due to negotiate with its union. Each marker represents the pooled DID estimate based on Equation (2)—that is, the estimate for the $\beta$ coefficient in the following specification $Y_{jt} = \beta(D_i \times \delta_{t \geq 2012}) + \alpha_j + \delta_t + \gamma X_{jt} + \varepsilon_{jt}$, with standard errors clustered at the establishment-level. Each estimate is calculated on the subset of establishments in the analysis sample with negotiation months falling in December-February, March-May, June-August, or September-November. Since ultractivity was introduced in September 2012, those with negotiation months in June-August were allowed at most 3 months of negotiation delays to be considered treated.

Figure B6: Event Studies for Share of Provisions with Authority

(a) Labor: workers and unions



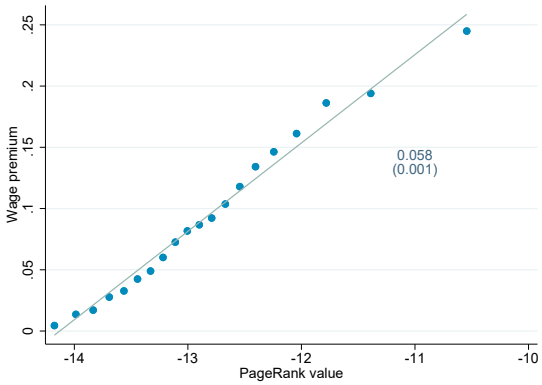

(b) Capital: firms and managers



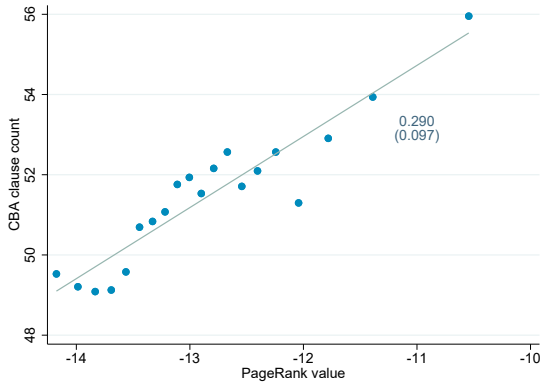

*Notes:* Figures show event studies for the impact of union bargaining power on authority measures using the DID specification in Equation (2). That is, each marker depicts the estimates of the $\beta^t$ coefficients with standard errors clustered at the establishment-level. Authority measures capture the authority of agents in a legal document where 1) *obligations* and *constraints* reduce authority; and 2) *entitlements* and *permissions* expand authority. These measures rely on an unsupervised extraction of these provisions among agents of capital (firms and managers; Figure B6b) and labor (workers and unions; Figure B6a) in the CBA. Refer to Appendix E for more details. Each figure reports event studies for 3 outcomes measured as a share of provisions in the CBA: 1) provisions corresponding to either capital or labor; 2) their respective entitlements and permissions; and 3) their respective obligations and constraints. The unit of observation is the establishment-year restricted to establishments in the analysis sample.

Figure B7: Relation of Wage Premiums and CBA Clauses with PageRank Values

(a) Wage premium and $V_j$



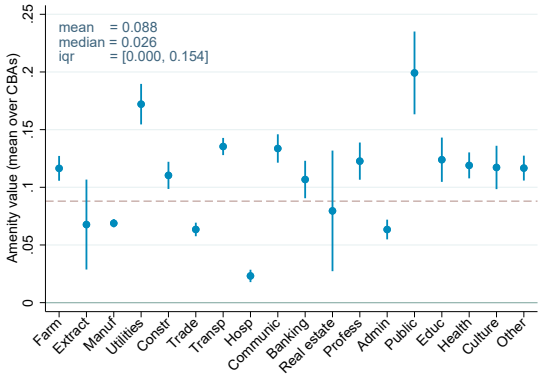

(b) Clause count and $V_j$



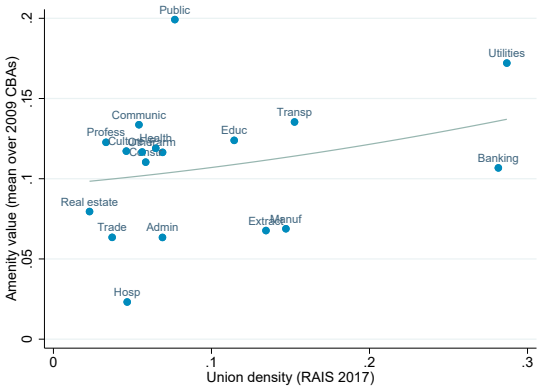

*Notes:* Figures provide motivation for the linear relation in Equation (3). Figures B7a and B7b are binned scatter plots of wage premiums and number of CBA clauses on PageRank values (respectively) using information from 2007-2011. Each binned scatter plot includes fixed effects by detailed industry, microregion, and negotiation month, and reports the point estimate (and standard error) for the slope of the fitted line. The unit of observation is the establishment. The sample is restricted to the analysis sample constructed with coverage from sectoral CBAs, where the establishments also have PageRank values and AKM wage premiums in $t \in [2007\text{-}2011, 2012\text{-}2016]$, i.e., they are in the strongly connected set of establishments linked by worker flows for each $t$.

Figure B8: Mean Amenity Value of Firm-level CBAs by Industry

(a) Amenity value



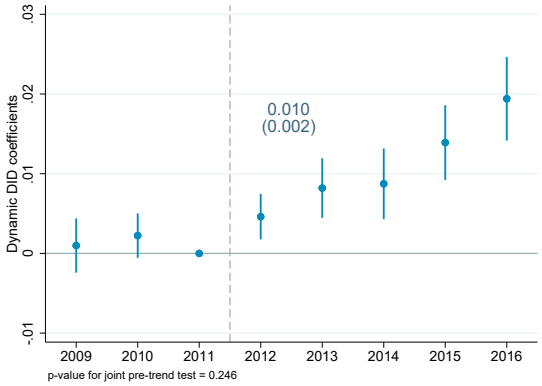

(b) Relation to union density



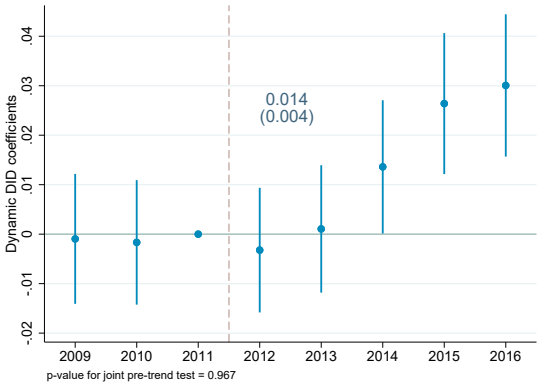

*Notes:* Figures show descriptive statistics of the amenity value of firm-level CBAs negotiated in 2009. The amenity value of a CBA is given by $Amenity_{cba} = \sum_{clause} 1\{clause = z\} \times (\hat{\beta}^z/\hat{\alpha})$. Refer to section 4.2 for more details. Figure B8a displays the estimate of the mean amenity value by broad industry groups. The dashed horizontal line corresponds to the overall mean, which is reported on the figure along with the median and the inter-quartile range. Figure B8b shows the relationship between the mean amenity value by industry and the industry-specific union density. Union density is calculated as the share of workers who are union members and is only reported in RAIS starting 2017. Since coverage is universal in Brazil, union membership consists of workers willing to pay dues in exchange for additional benefits that are not in CBAs, e.g., recreational facilities and private health insurance plans.

Figure B9: Event Studies on Workforce Composition

(a) Share workers: college




(b) Retention rate: college




*Notes:* Figures show the event studies for the impact of union bargaining power on outcomes related to workforce composition using the DID specification in Equation (2). In addition to depicting the estimates of the $\beta^t$ coefficients, each figure reports the pooled DID estimate—that is, the estimates for $\beta$ in the following specification $Y_{jt} = \beta(D_i \times \delta_{t \geq 2012}) + \alpha_j + \delta_t + \gamma X_{jt} + \varepsilon_{jt}$ along with standard errors clustered at the establishment-level in parentheses. Refer to Section 3.1 for more details. The unit of observation is the establishment-year restricted to establishments in the analysis sample.

# C Data Appendix

## C.1 Sample construction

To study the relation between the wage and amenity components of compensation under collective bargaining, I construct three main samples: 1) initial sample; 2) baseline panel; and 3) analysis sample. The first is a sample at the establishment-CBA level intended to provide general statistics on collective bargaining coverage. The second is a panel at the establishment-year level that allows me to track CBA negotiations over time in order to provide descriptives on changes to the content of CBAs, the negotiation process, and the role of ultractivity. The third is a subset of the baseline panel focusing on bargaining units with strong attachment to collective bargaining where ultractivity introduced a wedge in union bargaining power.

**Initial sample** This sample focuses on providing general statistics on CBA coverage and the starting point for subsequent samples used in the paper. To do so, I needs to assign each CBA in *Sistema Mediador* to an establishment in RAIS. I first extract the CNPJ identifiers of the employer counterparts as well as the municipalities of coverage in each CBA. Assignment is different for firm-level CBAs and sectoral CBAs. For firm-level CBAs, I use the first 8 digits of the CNPJ identifiers to find all establishments in RAIS that are part of the signing firms in the CBA, have nonzero employment by end-of-year in the start year of the CBA, and are located in one of the municipalities of coverage of the CBA.[1] For sectoral CBAs, I implement the same procedure but I first match the employer associations signing the CBA to the firms being represented in the agreement. The crosswalk that allows me to create these links was secured through a public information request, from which I obtained the CNPJ identifiers of firms contributing to individual employer associations.[2] Through this procedure, each establishment is assigned to its CBAs such that in a given year there can be multiple, unique, or no CBAs negotiated. The statistics in Table 1 report the share of establishments and workers that negotiated at least one CBA in the private sector.

**Baseline panel** This sample focuses on tracking firm-level CBA negotiations over time for a given establishment. Although workers in an establishment are almost always represented

[1] The CNPJ is a 16-digit unique identifier for establishments where the first 8 digits correspond to the firm.

[2] These data corroborate an important aspect of the union context: that workers in an establishment are generally represented by a single union and that establishments are generally represented by a single employer association. Specifically, 92% of firms contributing to an employer association are 1-to-1 matches, and 95% of firms sending union contributions on behalf of workers are 1-to-1 matches. One-to-many matches are possible when industry-based and occupation-based categories overlap.

by a single union, exceptions occur when industry-based and occupation-based categories overlap. Hence, each establishment is assigned a "main union" which corresponds to the modal union across establishment-CBA-union observations in the initial sample. As such, the goal is for each establishment-year observation to only include information on the CBAs negotiated by the main union in a given year.

An establishment is treated if any CBA was filed before and expired after ultractivity. When multiple CBAs by the main union exist for a given establishment-year, the following rules are applied. I use the minimum start date and the maximum end date, keeping the maximum CBA duration. To avoid double counting from multiple entries of the same CBA or amendments, I also take the maximum of the clause counts and provisions of authority. For the months of negotiation (i.e., the difference between the start and filing dates), I use the mean across the CBAs in a given establishment-year.

The resulting panel only contains establishment-year observations for years when a CBA was negotiated. Using this panel, I calculate CBA renewals for each establishment and gaps in coverage (see Figure 3). I then create a balanced panel for the years 2009 to 2016. Coverage by year for each establishment is given if at end-of-year there was an active CBA and the establishment had nonzero workers.

I repeat this same procedure using sectoral CBAs.[3] Hence, there are two baseline panels: 1) the one tracking firm-level CBA negotiations over time, which is the starting point for the analysis sample; and 2) the one tracking sectoral CBA negotiations over time, which is used to construct the sample for estimating the wage-equivalent value of clauses.

**Analysis sample** This sample focuses on restricting the baseline panel for firm-level CBAs to bargaining units with strong attachment to collective bargaining where ultractivity introduced a wedge in union bargaining power. The idea is to restrict the nontreated establishments to those that provide a credible counterfactual (Section 3.1).

The analysis sample imposes two sets of restrictions. The first set of restrictions is that establishments must have negotiated a CBA in 2009 (the first year of the sample period) and at least another one prior to 2012 (the year ultractivity is enacted). The intention of these restrictions is to focus on bargaining units engaged in frequent negotiations. More practically, having a minimum of two negotiations prior to treatment implies that analyzing pre-trends in CBA content will always be feasible.

The second set of restriction nets out the fact that ultractivity can lead to differential exit from the market and/or negotiations. Hence, I rely on the data to determine the post-

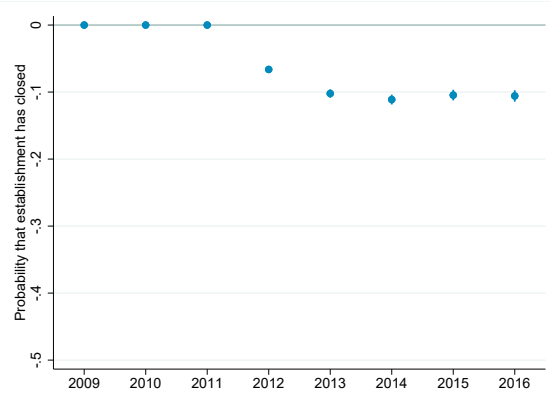
[3]The only difference relative to the firm-level CBA procedure is that the "main union" is the modal union across establishment-CBA-union observations where at least one worker in the establishment is paid the wage floor negotiated in the CBA.

treatment years by when this differential exit disappears. In terms of market exit, I restrict the sample to establishments that are open 2009-2014 (Figure C1a). This provides balance in the 3 pre-treatment years and guarantees labor market outcomes are observable in at least 3 post-treatment years. In terms of negotiation exit, I restrict to establishments covered by at least one CBA negotiated in or after 2012. (Figure C1b). This essentially guarantees that there is some CBA content to analyze under ultractivity. Note that both treated and control establishments can have a CBA negotiated in 2012 because treatment requires the filing date (rather than the start date) to be before September.[4]

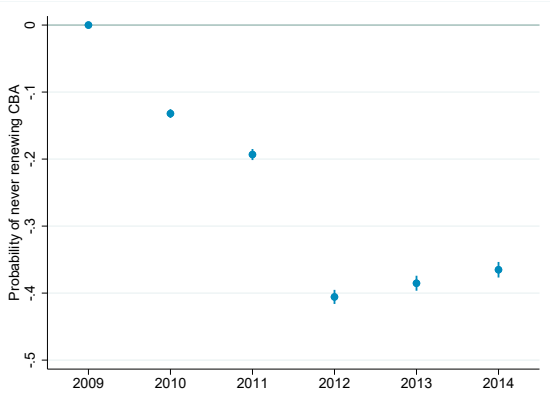
As shown in Table C1, there are roughly 23.6K establishments that satisfy these restrictions and they tend to be larger (130 workers on average) even compared to those signing a firm-level CBA in 2009 (104 workers).

Figure C1: Motivation for Analysis Sample Restrictions

(a) Establishment closure




(b) Never renewing CBA




*Notes:* Figures provide evidence motivating the restrictions for the analysis sample. In these analyses, the sample includes establishments open throughout the pre-period (2009-2011) that negotiate a firm-level CBA in 2009 which is renewed prior to 2012. Each figure plots the estimates of the dynamic DID coefficients from specifications of the following form $Y_{jt} = \sum_{k=2009}^{T} \beta^{t=k}(D_j \times \delta_{t=k}) + \alpha_j + \delta_t + \varepsilon_{jt}$ with standard errors clustered at the establishment-level. The unit of observation is the establishment-year and $D_j$ is an indicator for being treated. In Figure C1a the outcomes is an indicator for whether the establishment has closed, i.e., it no longer appears in RAIS. The omitted year is 2011 given that all establishments are open from 2009 to 2011. In Figure C1b the outcomes is an indicator for whether the establishment never renews its CBA. The omitted year is 2009 given that all establishments renew their CBA from 2009.

[4]The filing date captures when the agreement is reached, which due to negotiation delays usually occurs after the start date.

Table C1: Establishments, Workers, and Unions Across Samples

| | Has 2009 CBA | Analysis sample | Treat | Control |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Establishments | 43,209 | 23,565 | 17,726 | 5,839 |
| (share treated) | (0.49) | (0.75) | (1.00) | (0.00) |
| Workers | 4,500,553 | 3,055,484 | 2,281,443 | 774,041 |
| (mean estab size) | (104) | (130) | (129) | (133) |
| Unions | 2,734 | 2,100 | 1,724 | 1,156 |

*Notes:* Table reports the number of unique establishments, workers, and unions across samples. The sample in column (1) includes establishments in the baseline panel that negotiated a firm-level CBA in 2009. Column (2) restricts to the analysis sample, while columns (3) and (4) focus on the treated and control units, respectively. The share of treated establishments and the mean establishment size are reported in parentheses.

For estimating the wage-equivalent value of clauses (Section 4.2), an analogous version of the analysis sample is constructed but using the baseline panel for sectoral CBAs. This sample is further restricted to establishments with PageRank values and AKM wage premium estimates for both $t \in [2007 - 20011, 2012 - 2016]$. For several of the out-of-sample sense checks (Section 4.3), the analysis sample is restricted to establishments with PageRank values and AKM wage premium estimates.

## C.2 Construction of variables

The following is a comprehensive list of the variables used in my analyses along with a description of how they are constructed. I categorize these variables as 1) wage outcomes; 2) employment outcomes; 3) CBA outcomes; and 4) other variables.

### Wage outcomes

- Log contracted wage: For every spell-year recorded in RAIS, employers report the contracted wage of the worker. Based on the type of wage contract (e.g., monthly, biweekly, daily, etc.) I ensure that the contracted wage reflects a monthly contracted wage, which is the most common form of wage contract.[5] I then convert the contracted wage to real terms using Brazil's CPI with base year 2015 and take logs. Using workers employed

[5] An additional adjustment is needed for 2016 since the decimal on the contracted wage for a non-negligible portion of employers was incorrect, such that the reported value is the correct one divided by 100. Hence, I multiply by 100 any contracted wage in this year that is below the minimum wage.

at end-of-year and with tenure greater than one month (i.e., employed throughout the month of December), I take the mean of log contracted wage by establishment-year. When this outcome is reported as "hourly," I divide the real contracted wage by monthly contracted hours before taking logs and calculating the mean.

- 90-10 and 50-10 wage ratio: Using workers employed at end-of-year and with tenure greater than one month (i.e., employed throughout the month of December), I calculate the 90th, 50th, and 10th percentile of the log contracted wage by establishment-year. The 90-10 ratio is simply the difference between the 90th and the 10th percentile, while the 50-10 ratio is the difference between the 50th and the 10th percentile.

- Log average earnings: For every spell-year recorded in RAIS, employers report the mean monthly earnings across months in which the worker was employed. I convert these earnings to real terms using Brazil's CPI with base year 2015 and take logs. Using workers employed at end-of-year and with tenure greater than one month (i.e., employed throughout the month of December), I take the mean of log average earnings by establishment-year. When this outcome is reported as "hourly," I divide the average earnings by monthly contracted hours before taking logs and calculating the mean.

- Log December earnings: For every spell-year recorded in RAIS, employers report the total earnings in the month of December. I convert these earnings to real terms using Brazil's CPI with base year 2015 and take logs. Using workers employed at end-of-year and with tenure greater than one month (i.e., employed throughout the month of December), I take the mean of log December earnings by establishment-year. When this outcome is reported as "hourly", I divide the average earnings by monthly contracted hours before taking logs and calculating the mean.

**Employment outcomes**

- Log employment: For each establishment-year, I calculate total employment as the total workers employed at end-of-year with tenure greater than 1 month, i.e., worker employed at the establishment for the entirety of December. I then simply take logs of total employment.

- Hiring rate: For each establishment-year, I calculate the total number of workers who were hired by the establishment throughout the year. I then divide the hires by the average between the total employment at end-of-year for the current and previous year.

- Retention rate: For each establishment-year, I take the share of workers employed throughout the month of December among those who were employed throughout the month of January. When this outcome is reported for workers with a college education, I restrict to workers that have that level of education.

- Turnover rate: For each establishment-year, I calculate the total number of workers who separated from the establishment throughout the year. I then divide the separated workers by the average between the total employment at end-of-year for the current and previous year. When this statistic refers to layoffs (or quits), I restrict to separations that are classified as layoffs (or quits).

- Fixed contract: These are workers whose employment contracts have a fixed duration. I calculate the share of workers with such contracts in each establishment-year.

- Safety events: These are occurrences that reflect a lack of safety in the work environment. They include work-related deaths, separations due to injuries and disabilities, as well as leaves motivated by injuries. I can then calculate the share of workers with such events in each establishment-year.

- Taking leave: These are instances where workers take leave regardless of the motivation. I calculate the share of workers with such events in each establishment-year.

- Education groups: I generate three groups based on the worker-level education variable provided in RAIS, i.e., no high school degree, high school degree only, and college degree. I calculate the share of workers belonging to each group in each establishment-year.

- Occupation groups: I generate three groups based on the worker-level occupation variable provided in RAIS (*Classificação Brasileira de Ocupações*), i.e., managers, professionals, and rank-and-file. I calculate the share of workers belonging to each group in each establishment-year.

**CBA outcomes**

- Clause count: For each establishment-year, I take the assigned CBA of coverage and calculate the number of clauses in the CBA. When multiple CBAs exist for a given establishment year, I use the maximum for each clause type and then sum across clause types to get the total number of clauses. When this outcome refers to "unique types", I simply create an indicator for each clause type negotiated in the CBA and sum across these binary variables.

- Wage/employment/other clauses: For each establishment-year, I take the assigned CBA of coverage and calculate the number of clauses in the CBA that are classified as wage-related, employment-related or other. When multiple CBAs exist for a given establishment year, I use the maximum for each clause type and then sum across clause types that are classified as either wage-related, employment-related, or other. Table A3 details the clause types and subgroups that belong to each of these three classifications provided by *Sistema Mediador*.

- Provisions: For each establishment-year, I take the assigned CBA of coverage and calculate the number of subject-verb prefixes in the text of the CBA that use deontic modal verbs and concern agents of capital (firms and managers) or labor (workers and unions). I then classify these provisions as obligations, constraints, permissions, or entitlements (see Appendix E for more details). When multiple CBAs exist for a given establishment year, I use the maximum for each provision type and then sum across these types to get the total number of provisions.

- Authority measures: For each establishment-year, I calculate the share of provision types (i.e., obligations, constraints, permissions, or entitlements) by agent of interest (i.e., firms, managers, workers, or unions) among all provisions. The shares corresponding to obligations and constraints reduce authority, while the shares corresponding to permissions and entitlements expand authority. Hence, a summary measure of authority for a given agent is the sum of the respective share of provisions expanding authority minus the share of provisions reducing authority. Finally, the relative authority of labor to capital is calculated by taking the authority measure for agents of labor and subtracting the authority measure for agents of capital.

- Amenity value: For each establishment-year, I take the assigned CBA of coverage and calculate the weighted sum of clauses in the CBA, where the weights are the wage-equivalent amenity values assigned to clause subgroups in Table 6. That is, the amenity value of a CBA is given by $\sum_{clause} 1\{clause = z\} \times (\hat{\beta}^z/\hat{\alpha})$. When multiple CBAs exist for a given establishment year, I use the maximum for each clause type and then take the weighted sum to get the amenity value.

**Other variables**

- Wage premium: 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 in two periods (2007-2011 and 2012-2016). For each period, 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 ($\alpha_i$) and individual establishments ($\psi_j$), year dummies interacted with five education dummies, and quadratic and cubic terms in age interacted with the education dummies ($X_{it}$)—see Appendix F. 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 period are normalized relative to the restaurant industry.

- PageRank value: 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 in two periods (2007-2011 and 2012-2016). For each period, 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 F), we follow Morchio and Moser (2024) 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. Results are robust to (i) assuming the multiplicative factor is the same for both periods; (ii) choosing the establishment with the smallest wage premium gap as the normalizing establishment, and then assume it maintained a time invariant PageRank value of 1; and (iii) choosing the establishment with the smallest wage premium gap as the normalizing establishment, and then assume it maintained a time invariant PageRank value and that the multiplicative factor in 2011-2016 equals 1.

- Industry: Each establishment in RAIS has an industry classification code based on *Classificação Nacional de Atividades Econômicas* (CNAE). The broad industries—used in the heterogeneity analyses, for example—consist of 18 groupings created from the sectoral classification of CNAE (*seção*) provided by provided by the Brazilian statistical agency IBGE. The detailed industries—used as time-varying fixed effects in

Equation (2), for example—consist of 285 groupings that consist of the first three digits of the CNAE code (referred to as *grupo* by IBGE). While broad industries distinguish between manufacturing and transportation, the detailed industries distinguish between steel and non-steel metalwork (or between train and track transportation).

- Region: Each establishment in RAIS reports the municipality where it is located. In my descriptives, I assign each municipality to one of the five major regions provided by the Brazilian statistical agency IBGE. In my analyses, however, I use time-varying fixed effects by microregion. These microregions are neighboring municipalities grouped into 543 units by IBGE that capture local labor markets.

- Negotiation month: Each bargaining unit has an assigned negotiation month (or *data-base*). This corresponds to the calendar month in which CBAs for the bargaining unit tend to start and end, and it marks the time when negotiations for the unit should be taking place. Each CBA in *Sistema Mediador* reports the negotiation month, which I then assign to each establishment using the mode across establishment-year observations with CBA information.

- Union density: The 2017 files for RAIS include indicators for whether a workers is a union member. I calculate union density as the share of workers that are union members—restricted to workers employed at end-of-year with tenure greater than 1 month, i.e., worker employed at the establishment for the entirety of December.

# D Understanding *Sistema Mediador*

In 2007, Brazil's Ministry of Labor (MTE) launched a website for writing, filing, and registering collective bargaining agreements (CBAs) called *Sistema Mediador*. On August 2008, MTE announced that all CBAs from 2009 onward needed to be submitted electronically through this website in order to be registered by the ministry, and therefore become legally binding.[6] I've scraped all CBAs from this website as of September 2018. The objective of this appendix is to understand how these data are produced. There are three general steps needed to register a CBA: 1) origination, 2) submission, and 3) revision.

## D.1 Origination

The process starts when one of the counterparts in the negotiation of a CBA creates a new entry for the agreement, i.e., originate the CBA through *Sistema Mediador*. The applicant must provide the following information to complete this part of the process:

- Group represented by the applicant: workers or employers
- Applicant's unique identifier: CNPJ—this information is validated by the website which is then used to automatically fill-in other fields, e.g., name, address, etc.
- Level of bargaining: firm-level or sectoral
- Type of CBA: original or amendment to an original—in the latter case, the application number of the original CBA must be provided

After providing this basic information, the applicant receives an origination number that can be used to edit information about the CBA before submitting the agreement. The applicant can share this number with the negotiating counterpart, allowing either party to make online edits to the CBA.

## D.2 Submission

The next step in the process is to provide additional information, write the CBA, and submit it through the online portal. The additional information required includes:

- Identifier of the counterpart (CNPJ)

[6] Specifically, *Instrução Normativa SRT nº 11/2009* revoked *nº 6/2007* which had mandated submissions on paper. Although a 2012 decision removed the ban on paper submissions (*Processo nº RR - 3895000-45.2009.5.09.0003*), virtually all agreements are registered online through *Sistema Mediador*.

- Identifier of any additional entity signing the CBA (CNPJ)
- Name and title of the representatives for all parties involved
- Validity period of the CBA, i.e., a start and end date (cannot exceed two years)
- *Data-base* of the category, i.e., the reference date for negotiations
- Category of coverage, i.e., free text describing the workers covered
- Geographic coverage, i.e., the municipalities in which the CBA is binding

Once this information is provided, *Sistema Mediador* makes some validation checks, e.g., that the geographic coverage corresponds to the labor union's area of representation.

Once the additional information is validated, one can start writing the clauses of the CBA. Before writing each clause, one must classify it into one of 137 different clause types chosen from a predetermined list that is broken down by broad groups and subgroups (see Table A3). Once the clause is classified, the applicant can write the clause title and its content. Although the title and description of the clause is free text, one can import clauses from previous CBAs registered in *Sistema Mediador* by providing the corresponding origination number.

The final task in this step of the process is to submit the agreement. Additional validation checks are made after clicking submit, e.g., that the CBA has at least one clause. *Sistema Mediador* keeps track of the exact date when a CBA is submitted, which I refer to as the filing date. Once the agreement is submitted, it cannot be edited. This implies that negotiations should have finalized by the filing date, so that the difference between filing date and start date capture the duration of the negotiation process. If the submission is successful, a registration form becomes available. This form must be signed by all parties involved and sent to the regional offices of the MTE. A filing number is provided to track the registration status of the CBA.

## D.3 Revision

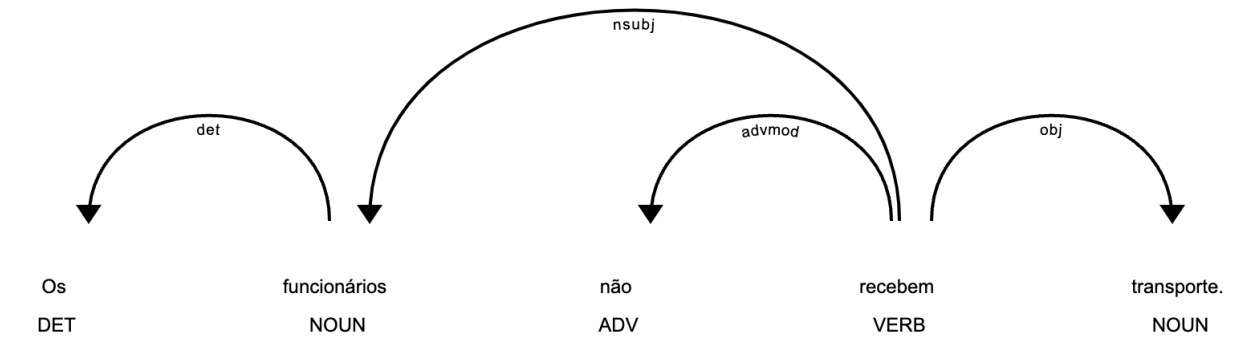
After submitting the registration form, it is possible that MTE notifies the parties that some corrections are required before formalizing the registration of the CBA. When that is the case, the applicant can access the CBA through *Sistema Mediador* and make the corrections. There are two modalities for corrections. One is for editing the CBA's category, which can be done completely online. The other is to provide additional requirements. In this last case, the applicant must print the requirement form, fulfill the request, and file the form at the regional MTE offices. Once all corrections (if any) are approved by the MTE, the CBA is registered. A registration number is provided as proof of the legally binding agreement.

# E Authority Measures

My analysis uses authority measures extracted from the text of the CBAs to characterize whether provisions in the agreement are expanding or reducing the authority of capital (firms and managers) and labor (workers and unions). I obtain these authority measures by following the documentation in Ash et al. (2020), introducing modification that are specific to Portuguese. This appendix details how I extract these authority measures.

## E.1 Subject-verb prefix identification

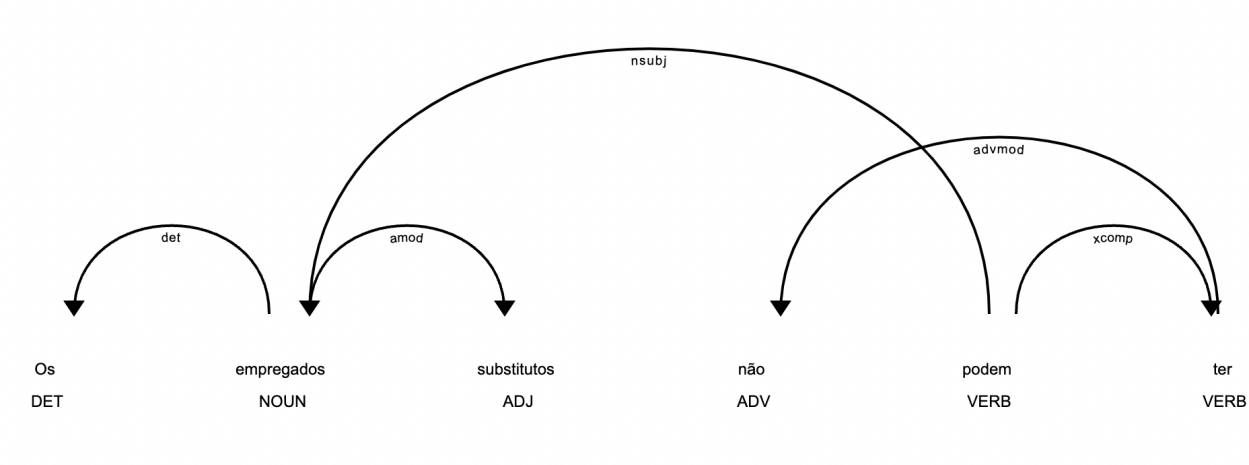
I build an NLP pipeline to parse sentences in each CBA into parts-of-speech using spaCy's Portuguese parser.[7] As a starting example, consider the following sentence: "Os funcionários não recebem transporte" (i.e., employees do not get transportation). Once processed by the NLP pipeline, spaCy provides attributes for each word describing its part-of-speech.




In the above parse tree, spaCy identifies "funcionários" as the subject of the sentence (nsubj), and the syntactic parent of "funcionários" is the verb "recebem." The arrows in the parse tree indicate the children of a given word. The syntactic relationships of the words can be accessed through the token.dep_ attribute, while the part-of-speech can be accessed through the token.pos_ or token.tag_ attributes. In order to access the parent of a word, tokens are assigned an attribute token.head.

When a modal verb is present, subject-verb prefix identification is slightly more complex. For instance, consider the following sentence: "Os empregados substitutos não podem ter direito à pensão" (i.e., substitute employees do not have the right to a pension).

[7] The spaCy model used in the pipeline is "pt_core_news_sm" (see documentation). Though spaCy has larger models potentially offering improved accuracy, they suffer from decreased efficiency. The largest spaCy model, "pt_core_news_lg," has a 0.97 accuracy for both lemmatization and part-of-speech tags, and the model used in the pipeline has a 0.97 accuracy for lemmatization and 0.96 for part-of-speech tags.




The parser correctly identifies "empregados" as the subject (nsubj) of the sentence. However, the syntactic parent of "empregados" is the modal verb "podem," rather than the verb of interest "ter." Hence, I iterate through the syntactic children of the verb "podem" in order to determine whether an open clausal complement (xcomp), in this case, "ter," is present.

**Differences in dependency parsing** There is one key difference between the spaCy Portuguese and English dependency parsers. The English parser has a more descriptive tag schema, providing "fine-grained" part-of-speech tags alongside "coarse-grained" part-of-speech tags from the Universal POS tag set, where the former can be obtained through the token.tag_ attribute and the latter from the token.pos_ attribute. The Portuguese dependency parser only provides "coarse-grained" part-of-speech tags, which are less descriptive.

In extracting subject-verb prefixes in Portuguese, modal verbs are not explicitly marked, and the syntactic parent of parsed subjects may not be the intended verb. This issue is resolved with an algorithmic determination of subject-verb prefixes, as explained above, which is more complicated than the original implementation in Ash et al. (2020). For instance, verbs are checked to see whether "ter que" is present in their children. In the example from the section above, the English spaCy dependency parser would identify the English equivalent "can" as a modal verb (MD) in the tag_ attribute.

The original implementation uses Hugging Faces NeuralCoref 4.0, a pipeline extension for spaCy that resolves coreference clusters using a neural network. No easily accessible and integrative alternative exists for Portuguese, so our implementation is unable to resolve coreferences. Nonetheless, the 2009 firm-level CBAs only contains 5.05% pronoun subjects, of which some reference already identified agents and other reference no agent at all.

*Portuguese Parsing of the Sentence "Eu devo ler este livro."*

| token.text | token.head | token.dep_ | token.tag_ | token.pos_ | token.children | token.ancestors |
|---|---|---|---|---|---|---|
| Eu | devo | nsubj | PRON | PRON | [] | [devo] |
| devo | devo | ROOT | VERB | VERB | [Eu, ler] | [] |
| ler | devo | xcomp | VERB | VERB | [livro, .] | [devo] |
| este | livro | det | DET | DET | [] | [livro, ler, devo] |
| livro | ler | obj | NOUN | NOUN | [este] | [ler, devo] |
| . | ler | punct | PUNCT | PUNCT | [] | [ler, devo] |

*English Parsing of the Sentence "I must read this book."*

| token.text | token.head | token.dep_ | token.tag_ | token.pos_ | token.children | token.ancestors |
|---|---|---|---|---|---|---|
| I | read | nsubj | PRP | PRON | [] | [read] |
| must | read | aux | MD | AUX | [] | [read] |
| read | read | ROOT | VB | VERB | [I, must, book, .] | [] |
| this | book | det | DT | DET | [] | [book, read] |
| book | read | dobj | NN | NOUN | [this] | [read] |
| . | read | punct | . | PUNCT | [] | [read] |

**Language differences** There is no one-to-one translation of modal verbs from English to Portuguese, though strict and permissive modals are mutually exclusive in both languages. The English "will" expresses the future tense and expectations, but Portuguese does not possess an equivalent modal verb. Rather, Portuguese utilizes future tense conjugations and the modal verb "dever." As a result, verbs must be systematically inspected for their tense by checking for endings "-rá" and "-rão."

Whereas the passive voice in English only requires the auxiliary verb "to be" followed by the past participle of the main verb, Portuguese employs two distinct structures to convey the passive voice. In Portuguese, the passive voice is formed by using the auxiliary verb "ser" (to be) followed by the past participle of the main verb. The passive voice can also be formed using the reflexive pronoun "se" in combination with a verb conjugated in the third person. Provided these alternatives, the dependency parser is not able to identify all passive subjects, so the code checks for the passive voice in addition to the parser's analysis. This

is done so by determining whether the pronoun "se" is present in the syntactic children of a verb and by determining whether a verb is preceded by "ser." As such, errors may arise with the parsing and labeling of reflexive verbs, as the program makes use of active verbs, i.e., non-passive and not special verbs.

Moreover, the usage of verb forms ending in "-se-á" and "-se-ão" are unique to Portuguese to express, for example, the future tense in a passive voice construction. In addition to checking for the future tense and the pronoun "se" individually, the code also identifies this structure, which is denoted as a strict modal and passive. As the parser is unable to properly lemmatize verbs with "-se," "-se-á," and "-se-ão" endings, the code removes such endings in all cases and then re-lemmatizes the verb in the case of "-se."

**Shared limitations** The pipeline is unable to distinguish between independent and dependent clauses. Take for example the sentence "Se os empregados, na citada semana, trabalharem no horário estabelecido na cláusula, deverá a empresa pagar-lhes com os adicionais de horas extras" (i.e., if employees, in the reference week, work the shifts established in the clause, the employer must pay them overtime for additional hours). In this case, the parser identifies two subject-verb prefixes: "empresa deverá pagar-lhes" and "empregados trabalharem". The annotation scheme also fails with exterior framing clauses or contingencies, as described in the original implementation.

Sentences written with indirect objects are not identified by the parser. Consider the sentence "Fica garantida ao aposentado a inscrição de novos beneficiários" (i.e., it is guaranteed to retirees the registration of new beneficiaries). The parser identifies "inscrição" as the subject of the sentence despite "aposentado" being the intended subject. Run through the pipeline, the above sentence would output "inscrição fica garantida" as the subject-verb prefix. This is a limitation also present in the original English implementation.

One of the most significant changes required for the pipeline to function similarly to the English implementation is the marking of verbs with "se" as passive, whether they are used in a reflexive or passive manner. The passive construction is more common; however, verbs such as "comprometer-se" (i.e., commit oneself to) are still prevalent in the agreements. Due to this difference, the provision verb dictionaries include some reflexive verbs in the passive category, and the provision assignment patterns are modified to accommodate reflexive verbs.

Another limitation in both the original implementation and this one is that the pipeline does not examine predicate adjectives. The sentence "Employees will be responsible. . . " is broken down into the subject-verb prefix "employees will be" since responsible is an adjective rather than a verb. In this example, the subject-verb prefix is marked as an obligation, which is correct. However, consider the sentence "Employees will be free to. . . ," which is

an entitlement. Since the pipeline does not use the predicate adjective "free," this clause is recorded as an obligation as well.

Both implementations also fail to account for multiple subjects within one provision. For instance, in the sentence "Employees, employers, and unions are subject to the following agreement," the parser can only identify one of the subjects instead of capturing all three subjects. Ideally, the parser should be able to identify and count all three subjects, increasing the count for each agent by one, but the current parser is limited to working with one subject at a time for each provision.

## E.2 Subjects

Agent dictionaries are used to identify whether the subject in a provision is an agent of capital (firms and managers) or an agent of labor (workers and unions). The agent dictionaries are derived from the list of 1,000 most common lemmatized subjects in the 2009 corpus of firm-level CBAs, which is approximately equivalent to all agents that appear more than 100 times in these agreements. Feminine and plural forms of nouns are included within dictionaries to accommodate possible lemmatization errors. Below are the dictionaries used for each agent (in Portuguese).

**Firms** companhia, companhias, concessionária, concessionárias, concessionário, concessionários, contratante, contratantes, corporação, corporações, corporativa, corporativas, corporativo, corporativos, empregador, empregadora, empregadoras, empregadores, empresa, empresar, empresária, empresárias, empresário, empresários, empresas, escola, escolas, estabelecimento, estabelecimentos, firma, firmas, hospital, hospitais, patrão, patroa, patroas, patrões, proprietária, proprietárias, proprietário, proprietários

**Managers** chefe, chefes, conselho, conselhos, diretor, diretora, diretoras, diretores, diretoria, diretorias, gerência, gerências, gerenciador, gerenciadora, gerenciadoras, gerenciadores, gerente, gerentes, manager, managers, superintendência, superintendente, superintendentes, supervisor, supervisora, supervisoras, supervisores

**Workers** acidentada, acidentadas, acidentado, acidentados, admitida, admitidas, admitido, admitidos, aposentada, aposentada, aposentadas, aposentadas, aposentado, aposentado, aposentados, aposentados, aprendiz, aprendizes, auxiliar, auxiliares, beneficiário, beneficiários, cobrador, cobradores, colaborador, colaboradora, colaboradoras, colaboradores, comissionista, comissionistas, contratada, contratadas, contratado, contratados, dependente,

dependentes, dimitida, dimitidas, dimitido, dimitidos, docente, docentes, empregada, empregadas, empregado, empregados, empregar, estagiária, estagiárias, estagiário, estagiários, estudante, estudantes, funcionária, funcionárias, funcionário, funcionários, gestante, gestantes, jornalista, jornalistas, mãe, mães, motorista, motoristas, operador, operadora, operadoras, operadores, operária, operárias, operário, operários, pai, pais, participante, participantes, pessoal, professor, professora, professoras, professores, servidor, servidora, servidoras, servidores, substituta, substitutas, substituto, substitutos, suplente, suplentes, trabalhador, trabalhadora, trabalhadoras, trabalhadores, vendedor, vendedora, vendedoras, vendedores

**Unions** assembleia, assembleias, confederação, confederações, cooperativa, cooperativas, delegado, delegados, dirigente, dirigentes, federação, federações, grêmio, líder, líderes, representante, representantes, sindicato, sindicatos, cipa, cipeiro, sindicalizado, sindicalizados, sindicalizada, sindicalizadas

## E.3 Verbs

Verb dictionaries are used to create a comprehensive list of verbs to be considered for classifying into provision types. The provision verb dictionaries are derived from the list of 1,000 most common lemmatized verbs in 2009 corpus of firm-level CBAs, which is approximately equivalent to all verbs that appear more than 100 times in these agreements. Verb conjugations are included within dictionaries to accommodate possible lemmatization errors. Adding all verb forms to dictionaries is excessive, so only commonly mis-lemmatized verbs within the 1,000 most common lemmatized verbs are included.

Relative to the original implementation, there is one improvement that involves an additional provision type, i.e., negative provisions. This provision type is included to solve an issue described by Ash et al. (2020) in their code for entitlement patterns. Some of the most common examples of non-negative, passive clauses with a strict modal include "employees will be fired." In addition, clauses such as "employees will not lose" and "employees will not work" are marked as constraints but should be labeled as entitlements. The negative provision type provides a simple way to search for these clauses to be properly labeled.[8] The original implementation attempts to solve this problem with an extension of entitlement verbs, but this method results in an extended list of verbs with some false positive for entitlements, e.g., "select" and "train".

[8]There are only a few "negative verbs" used in the CBAs, making it a small addition to the code.

**PROVISION DICTIONARIES**

*Provision Verb Dictionaries*

| Provision Type | Passive or Reflexive Verbs | Non-Passive Verbs |
|---|---|---|
| Obligation | exigir, esperar, coagir, compelir, obrigar, obrigado, forçar, requerer, comprometar, comprometer, responsabilizar | garantir, assegurar |
| Constraint | proibir, vedar, banir, impedir, impeder, restringir, proscrever, limitar, impossibilitar, negar, abster | |
| Permission | permitir, autorizar, aprovar, habilitar | |
| Entitlement | conceder, dar, outorgar, fornecer, garantir, garantido, proteger, cobrir, informar, notificar, assegurar, facultar, proporcionar, prestar, propiciar, providenciar, fornecir, avisar | ter, receber, ganhar, obter, gozar, beneficiar, repousar |
| Promise | | reconhecer, consentir, afirmar, segurar, estipular, assumir, concordar, prometer, aquiescer |
| Negative | despedir, despeder, dispensar, dispensado, dispensados | trabalhar, sofrer, perder |

*Verb Categories and Sample Verbs*

| Category | Verbs |
|---|---|
| Obligation | garantir, assegurar, ser exigido, ser obrigado, ser requerido, comprometer-se |
| Constraint | ser proibido, ser impedido, ser negado |
| Permission | poder, ser permitido, ser autorizado, ser aprovado, ser habilitado |
| Entitlement | ter, gozar, receber, ser concedido, ser dado, ser garantido, ser fornecido |
| Promise | concordar, prometer, consentir, aquiescer, reconhecer, segurar, assumir |
| Negative | trabalhar, sofrer, perder, ser dispensado |

The table above shows the verb dictionary by provision type, as well as some examples. The table below provides the deontic logic rules used to categorize a subject-verb prefixes

as obligations, constraints, permissions, or entitlements.

**PROVISION TYPE ASSIGNMENT**

*Pseudocode for Provision Type Assignment*

| Provision | Pattern |
|---|---|
| Obligation | not negative + strict modal + active verb<br>not negative + not permissive modal + obligation or promise verb |
| Constraint | negative + modal + not obligation, not negative, and not constraint verb<br>not negative + strict modal + constraint verb<br>negative + permission verb |
| Permission | not negative + permissive modal + not special verb<br>not negative + permission verb<br>negative + constraint verb |
| Entitlement | not negative + entitlement verb<br>not negative + strict modal + passive + not special and not negative verb<br>negative + obligation or negative verb |

*Provision Type Assignment Examples*

| Provision | Pattern |
|---|---|
| Obligation | empresa fornecerá, sindicato obriga-se |
| Constraint | empresa não poderá impedir, empresa se absterá |
| Permission | empregado poderá deixar, empregados serão autorizados |
| Entitlement | empregados receberão, supervisores serão custeados,<br>empresa não será responsabilizada |

# F PageRank Values and AKM Wage Premiums

The underlying models for PageRank values and AKM wage premiums have a direct connection to the model used to discuss the main results (Section 5.2). Specifically, they all rely on assumptions about $u_{ij}$, i.e., the indirect utility of worker $i$ from working at firm $j$. In this appendix, I walk through the PageRank and AKM models and then discuss assumptions in light of how I value CBA clauses (Section 4.1) and the posting model under collective bargaining (Appendix G).

## F.1 PageRank values

This is an abridged exposition of the model in Sorkin (2018). The starting point is $u_{ij} = V_j + \varepsilon_{ij}$, where $\epsilon_{i,j}$ refers to the idiosyncratic preferences of worker $i$ from working at establishment $j$. In a market with only two firms and independently distributed type I Extreme Value $\varepsilon_{ij}$ across workers, the probability that a worker prefers firm $j$ over $k$ is given by $\frac{\exp(V_j)}{\exp(V_j)+\exp(V_k)}$. 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 $M_{kj}/M_{jk} = \exp(V_k)/\exp(V_j)$.

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

$$M_{kj}\exp(V_j) = M_{jk}\exp(V_k), \forall j \in \mathcal{J}. \tag{F1}$$

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 (F1) across all employers and rearranging terms gives

$$\frac{\overbrace{\sum_{j\in\mathcal{J}} M_{kj}\exp(V_j)}^{\text{value-weighted entry}}}{\underbrace{\sum_{j\in\mathcal{J}} M_{jk}}_{\text{exits}}} = \underbrace{\exp(V_k)}_{\text{value}}, \tag{F2}$$

which implies a single linear restriction per establishment.

The intuition behind Equation (F2) is that a valuable firm tends to be chosen over other valuable firms and has fewer workers leave it. This recursive definition of $\exp(V_j)$ is closely

linked to Google's PageRank algorithm for ranking web-pages in a search. Along these lines, one can solve for $\exp(V_j)$ 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.

## F.2 AKM wage premiums

This is a version of the model in Card et al. (2018) extended to account for common amenity value. The starting point is $u_{ij} = \beta \log(w_j - b) + \eta \log(a_j - q) + \varepsilon_{i,j}$, which is the same as the posting model, i.e., Equation (G1). Hence, imposing the same set of assumptions as in Appendix G, one can derive the compensation bundles under monopsony provided in Equations (G4) and (G5).[9]

Rearranging equations (G4) and (G5), one can write the posted wage and amenity value as weighted averages of the marginal revenue product of labor and their respective reference values $b$ and $q$, i.e.,

$$w_j = \left(\frac{\beta}{1+\beta+e^L_{aj}}\right) T_j f'(L_j)\mu_j + \left(\frac{1+e^L_{aj}}{1+\beta+e^L_{aj}}\right) b \tag{F3}$$

$$a_j = \left(\frac{\eta}{\xi_j(1+\eta+e^L_{wj})}\right) T_j f'(L_j)\mu_j + \left(\frac{1+e^L_{wj}}{1+\eta+e^L_{wj}}\right) q. \tag{F4}$$

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

$$\log(w_j) = \log\left(\frac{\theta\bar{b}(1+e^L_{aj})}{1+\beta+e^L_{aj}}\right) + \log\left(1+\beta R^w_j\right) \tag{F5}$$

$$\log(a_j) = \log\left(\frac{\theta\bar{q}(1+e^L_{wj})}{1+\eta+e^L_{wj}}\right) + \log\left(1+\eta R^a_j\right), \tag{F6}$$

where $R^w_j = T_j P_j/[(1+e^L_{aj})\bar{b}]$ and $R^a_j = T_j P_j/[\xi_j(1+e^L_{wj})\bar{q}]$. With relatively small values of $\beta R^w_j$ and $\eta R^a_j$, compensation in the form of wages and amenity value are functions of a fixed worker component and a fixed establishment component as in Abowd et al. (1999).

[9] In this model, monopsony power stems from information asymmetries, i.e., employers cannot discriminate on the basis of workers' idiosyncratic preferences $\{\epsilon_{i,j}\}$. The model also abstracts from strategic complementarities by assuming a large number of firms $\mathcal{J}$.

Specifically,

$$\log(w_j) = \log\left(\frac{\bar{b}(1+e^L_{aj})}{1+\beta+e^L_{aj}}\theta\right) + \beta R^w_j \tag{F7}$$

$$\log(a_j) = \log\left(\frac{\bar{q}(1+e^L_{wj})}{1+\eta+e^L_{wj}}\theta\right) + \eta R^a_j. \tag{F8}$$

Therefore, Equation (F7) implies that workers' compensation in wages can be written as $\log(w_j) = \alpha^w + \psi^w_j$, where $\psi^w = \beta R^w_j$ is an establishment-specific wage premium. Similarly, Equation (F8) implies that workers' compensation in amenity value can be written as $\log(a_j) = \alpha^a + \psi^a_j$, where $\psi^a = \eta R^a_j$ 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.[10] The linear additivity of the wage and amenity value components implicitly assumed in Equation (G1) imply no changes to the typical estimation of AKM wage premiums is required. Interestingly, if one could observe variation in amenity value at the worker level across firms, estimating AKM-style amenity premiums would be feasible.

## F.3 Reconciling modeling assumptions

The previous sections specify the models and assumptions underlying PageRank values and AKM wage premiums. I now highlight differences in these assumptions relative to those made in the paper, specifically with regards to the posting model and valuing CBA clauses.

I start with the case of AKM since it is the most similar to the posting model. Both models set $u_{ij} = \beta\log(w_j - b) + \eta\log(a_j - q) + \varepsilon_{i,j}$ and make the same assumptions to get closed-form solutions for the compensation bundles in Equations (G4) and (G5). After this point, the AKM model relies on additional simplifying assumptions to write these expressions in a way that separates worker and firm components. Hence, the posting model nests an AKM model that accounts for amenity value.

In the case of the PageRank values, there are two modeling assumptions that seem at odds. First, setting $u_{ij} = V_j + \varepsilon_{ij}$ along with the same distributional assumptions about $\varepsilon_{ij}$ as the posting model, implies a direct mapping between $V_j$ and $\beta\log(w_j - b) + \eta\log(a_j - q)$. Second, by assuming $V_j = \alpha\psi_j + \sum_{z=1}^{Z}\beta^z A(z)_j$ for the estimation of the wage-equivalent value of clauses, $V_j$ is now linked to a completely different expression. If the models are internally consistent, both $\alpha\psi_j + \sum_{z=1}^{Z}\beta^z A(z)_j$ and $\beta\log(w_j - b) + \eta\log(a_j - q)$ should

[10] Unbiased estimates rely on an exogenous mobility condition. Gerard et al. (2021) provides falsification tests in the Brazilian setting.

equal each other. While I never observe either object to confirm this claim, I now turn my attention to explain how this is conceptually possible.

On one hand, the expression $\alpha\psi_j + \sum_{z=1}^{Z} \beta^z A(z)_j$ concerns the objects that generate value to all workers in a firm, rather than the values themselves. That is, it breaks down $V_j$ into the wage premium and multiple common amenities (e.g., CBA clauses), each of which carry some weight—i.e., $\alpha$ and $\beta^z$ for $z \in \{1, 2, ..., Z\}$—in determining how much workers value working at a specific firm. In a way, this expression is like an accounting assumption rather than an economic one. As such, workers' labor supply is responsive to changes in these objects (allowing for identification from worker flows), but meaningful expressions for elasticities are not possible without a common metric for value.

On the other hand, the expression $\beta \log(w_j - b) + \eta \log(a_j - q)$ concerns the values themselves as offered by firms, adding the assumption that workers' labor supply decisions are different depending on the form this compensation takes, i.e., wages versus amenities. By being explicit about how workers react to changes in compensation, this expression is about the underlying economic model rather than an accounting identity. Hence, CBA clauses have a wage-equivalent amenity value coming from their relative weight in determining $V_j$—that is, $\beta^z/\alpha$—but the economic model bundles all those clauses into a single amenity value metric, allowing workers to differentially respond to wages and amenity value. By having amenity value in the same metric as wages, it is possible to meaningfully speak about elasticities.

# G Posting Model Under Collective Bargaining

The increase in compensation caused by the boost in union bargaining power occurs without a reduction in employment. To explain this, I formalize a collective bargaining model that allows for imperfect competition in the labor market. The initial building block is a posting model where asymmetric information about workers' heterogeneous preferences over workplaces enables firms to post compensation bundles made up of wages and amenity values (Card et al., 2018). Firm posting is then constrained by collective bargaining in a right-to-manage framework. i.e., Nash bargaining between firms and unions pins down the compensation bundle, taking employment as given by the binding curve in the labor market (Nickell and Andrews, 1983).

## G.1 Posting wages and amenity values

Suppose there are $\mathcal{J}$ firms in a market, where each firm $j \in \{1, ..., \mathcal{J}\}$ posts a pair $(w_j, a_j)$ of wages and amenity values that workers observe at no cost.[11] The market has workers with heterogeneous preferences over the work environment offered by firms. For each worker $i$, the indirect utility of working at firm $j$ is

$$u_{ij} = \beta \log(w_j - b) + \eta \log(a_j - q) + \varepsilon_{i,j}, \tag{G1}$$

where $b$ and $q$ are the workers' reference wage and amenity value levels from their outside option, and $\varepsilon_{i,j}$ refers to the idiosyncratic preferences from working at firm $j$ arising from match factors such as commute time and coworker collegiality, among others.

Assuming that the $\{\varepsilon_{i,j}\}$ are independent draws from a Type I Extreme Value distribution and the number of firms $\mathcal{J}$ is very large (McFadden, 1973), workers' choice probabilities are closely approximated by exponential probabilities.[12] Hence, the firm-specific labor supply functions are approximated by

$$\log(L_j) = \log(\lambda) + \beta \log(w_j - b) + \eta \log(a_j - q). \tag{G2}$$

Hence, the key difference between wages and amenity values in this model (and the reason they are not perfect substitutes) is that labor supply responds differently to changes in these two forms of compensation.

Firms generate revenue in a perfectly competitive product market (whose price is used

[11] Since my objective is to characterize labor supply, I assume (for now) that labor demand is never binding.

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

as the numeraire) by producing output $Y_j = T_j f(L_j)$, where $T_j$ represents a firm-specific productivity shifter and $f(\cdot)$ is the production function—assumed to be a twice continuously differentiable and homogeneous of degree $\alpha \in (0, 1)$. The market is further simplified by ignoring capital and intermediate inputs.

The firm's problem is to post the wages and amenity values that minimize production costs given labor supply in Equation (G2). The posted bundles are common to all workers since firms cannot discriminate on the basis of idiosyncratic preferences $\{\varepsilon_{i,j}\}$. This asymmetry in information, rather than labor market concentration, is the source of monopsony power. The optimal choice is given by the following cost-minimization problem

$$\min_{w,a}(w_j + \xi_j a_j)L(w_j, a_j) \quad \text{s.t. } T_j f(L(w_j, a_j)) \geq \bar{Y}, \tag{G3}$$

where $\xi_j$ captures heterogeneity in the marginal cost of amenity provision across firms, which could be due to differences in HR capabilities for implementing new benefits, for example.

The first order conditions imply that the optimal compensation package of the "monopsonistic" firm $(w_j^M, a_j^M)$ is given by

$$w_j^M = T_j f'(L_j)\mu_j \left( \frac{e_{wj}^L}{1 + e_{wj}^L + e_{aj}^L} \right) \tag{G4}$$

$$a_j^M = T_j f'(L_j)\mu_j \left( \frac{e_{aj}^L}{\xi_j(1 + e_{wj}^L + e_{aj}^L)} \right), \tag{G5}$$

where $e_{wj}^L = \frac{\beta w_j}{w_j - b}$ and $e_{aj}^L = \frac{\eta a_j}{a_j - q}$ represent the firm-specific labor supply elasticities with respect to wages and amenities, respectively. The $\mu_j$ term captures the marginal cost of production, which the firm will equate to marginal revenue at its optimal choice for $Y_j$.

To guide the discussion of the posting model, it is helpful to compare the monopsonistic compensation bundle to that in a perfectly competitive market, i.e.,

$$w_j^* = T_j f'(L_j)\mu_j \left( \frac{e_{wj}^L}{e_{wj}^L + e_{aj}^L} \right) \tag{G6}$$

$$a_j^* = T_j f'(L_j)\mu_j \left( \frac{e_{aj}^L}{\xi_j(e_{wj}^L + e_{aj}^L)} \right). \tag{G7}$$

The solution $(w_j^*, a_j^*)$ is pinned down by two conditions: 1) marginal cost equals marginal revenue; and 2) marginal rate of substitution (MRS) equals marginal rate of transformation (MRT). I now discuss the monopsonistic solution in light of these two conditions.

In the monopsonistic case, the first condition does not hold since compensation is marked

down from the marginal revenue product of labor by $\frac{e^L_{wj}+e^L_{aj}}{1+e^L_{wj}+e^L_{aj}}$.[13] Hence, if $e^L_{aj}$ is non-negligible, markdown estimates that ignore amenity value may be biased.[14] Interestingly, wages are below marginal product not only because of employer market power, but also simply because amenity compensation exists. In general, the more inelastic workers are to a given type of compensation, the more it will be marked down from their marginal product.

The second condition, surprisingly, holds under monopsony since MRS equals MRT is implied by the ratio of wages-to-amenities under monopsony.[15] Therefore, although compensation has a markdown, the composition of the wage and amenity value bundle is optimal for the marginal worker. Lastly, the ratio of wages-to-amenities equals $\xi_j \left(e^L_{wj}/e^L_{aj}\right)$, implying that compensation tilts toward the type of compensation that labor supply is more responsive to, as well as the type which is less costly to provide. For example, if amenities are prohibitively expensive ($\xi_j \to \infty$) or do not influence workers decisions ($e^L_{aj} \to 0$) compensation will solely be in the form of wages.

## G.2 Bargaining over wages and amenity values

I now introduce collective bargaining into the model as a labor market institution that constrains firm posting. This follows a right-to-manage framework, i.e., firms and unions negotiate the compensation bundle through Nash bargaining considering employment as given by either the labor supply or the labor demand curve—whichever is binding at the resulting bundle. By seeking to improve on-the-job returns for workers, unions can push compensation above the monopsonistic bundle but can also exceed the efficient bundle under perfect competition.

The Nash bargaining problem

$$\max_{w,a} \left[\pi(w,a) - \pi(g_w, g_a)\right]^{1-\gamma} \left[\left(v(w) - v(g_w)\right) L(w,a)\right]^{\epsilon\gamma} \left[\left(v(a) - v(g_a)\right) L(w,a)\right]^{(1-\epsilon)\gamma} \quad \text{(G8)}$$

maximizes wages and amenity values over the weighted product of the firm's objective and the union's objective.[16] The firm is maximizing profits $\pi(w,a)$—relative to the inside option

[13] Total posted compensation is $w^M_j + \xi_j a^M_j = T_j f'(L_j)\mu_j \left(\frac{e^L_{wj}+e^L_{aj}}{1+e^L_{wj}+e^L_{aj}}\right)$, where the term in parentheses is the traditional Lerner markdown modified to incorporate amenities.

[14] Specifically, if $e^L_{aj} > 0$ and the variation used to estimate markdowns implies changes in amenities in the opposite (same) direction as wages, the markdown will be biased upward (downward).

[15] MRS equals MRT means that $\frac{e^L_{wj} a_j}{e^L_{aj} w_j} = \frac{1}{\xi_j}$ which implies $\frac{w_j}{a_j} = \frac{\xi_j e^L_{wj}}{e^L_{aj}}$. The MRS also implies homotheticity and hence linear Engel curves for both wages and amenities. Homotheticity and constant provision costs of amenities are key assumptions to this optimality result—also shown in Dube et al. (2022).

[16] The maximization problem's constraints are: $\pi(w,a) \geq 0$, $L(w,a) \geq 0$, $w \geq g_w$, and $a \geq g_a$. The subscript $j$ is dropped since the model is now focusing on negotiations between a given firm and its union.

$\pi(g_w, g_a)$—with bargaining power $1-\gamma$. The union is modeled as being split into two committees: one maximizing wage gains and the other maximizing amenity gains among the employed $L(w,a)$ using some increasing and concave preference function $v(\cdot)$. The union's bargaining power $\gamma \in [0,1]$ is allocated to each committee based on $\epsilon \in [0,1]$.

The contract curve tracing the Nash bargaining solution $(w^{NB}, a^{NB})$ for all values of $\gamma$ can be characterized using the problem's first order conditions. Assuming $v(\cdot) = \log(\cdot)$, $f(L) = L^{\alpha}$, and constant elasticities, the slope of the contract curve is given by

$$\frac{da}{dw} = \left(\frac{1-\epsilon}{\epsilon}\right)\left(\frac{e^{\pi}_{w}}{e^{\pi}_{a}}\right)\left(\frac{a}{w}\right), \tag{G9}$$

which is always positive, i.e., compensation increases with the union's bargaining power.[17] The steepness of this slope captures the amenity-wage tradeoff occurring in CBA negotiations. The union's choice affecting this tradeoff is $\epsilon$, where higher values tilt compensation toward wages by flattening the slope of the contract curve.

I define efficient bargaining over wages and amenities as a union whose contract curve produces compensation bundles that are always optimal for the marginal worker. If one requires that $(w^{NB}, a^{NB})$ imply MRS equals MRT for any $\gamma$, this means that $\epsilon = \frac{e^{L}_{w}}{e^{L}_{w}+e^{L}_{a}}$. Intuitively, under efficient bargaining, unions set their effort to bargain over wages relative to amenities according to workers' labor supply responsiveness to wages relative to amenity values. Hence, assuming efficient bargaining implies that the union's collective bargaining objectives are aligned with workers' preferences over the wage and amenity components.[18]

## G.3 Characterization

The Nash bargaining solution $(w^{NB}, a^{NB})$ can be either supply constrained or demand constrained. Figure G1 provides an example characterizing the set of solutions on the amenity-wage space. Any bundle below the curve where marginal cost equals marginal revenue is supply constrained and any bundle above this curve is demand constrained. Naturally, the outcome under perfect competition is at the boundary where MC equals MR, while the monopsony outcome is supply constrained, i.e., MC<MR.

At $\gamma = 0$, the solution is given by the inside option $(g_w, g_a)$. Assuming efficient bargaining, the contract curve starts at the inside option traveling toward $(w^*, a^*)$ maintaining a slope that guarantees MRS=MRT. (Equivalently, the contract curve falls on a line that crosses the

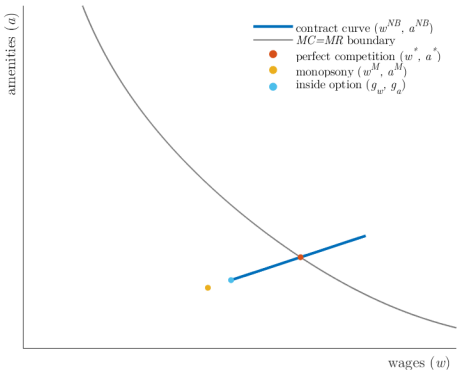
[17] $e^{\pi}_{w} < 0$ and $e^{\pi}_{a} < 0$ are the elasticities of profits with respect to wages and amenity value, respectively.

[18] To distinguish between optimality for the *marginal* worker as opposed to the *average* worker (or any subset of workers) would require knowing the different elasticities for these groups of workers. Since nothing in $e^{L}_{wj} = \frac{\beta w_j}{w_j - b}$ and $e^{L}_{aj} = \frac{\eta a_j}{a_j - q}$ is worker-specific, the model abstracts from this distinction.

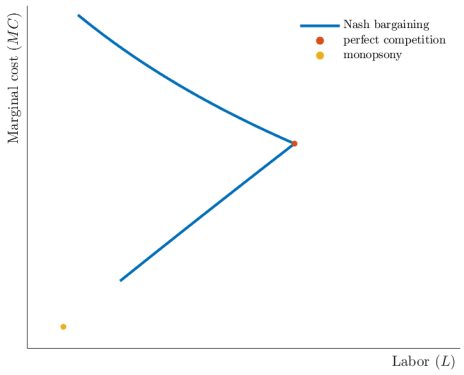
origin and maintains a ratio of amenities-to-wages of $\frac{e^L_{aj}}{\xi_j e^L_{wj}}$). Employment is increasing with compensation in this initial segment of the contract curve. However, once bargaining power is such that the Nash bargaining solution becomes demand constrained—that is, above the perfectly competitive outcome—employment will then decrease with compensation.

Figure G1: Illustration of Posting Model under Collective Bargaining

(a) Contract curve




(b) Labor supply and demand




*Notes:* Figure G1a shows the Nash bargaining solution as traced by the bargaining power parameter $\gamma$. The figure depicts how the two forms of compensation—i.e., wages and amenity value—evolve from the inside option (where $\gamma = 0$) and move along the contract curve as $\gamma \to 1$. The compensation bundles under full monopsony and a perfectly competitive labor market are provided as reference. The fact that the contract curve aligns with these bundles and the origin implies efficient bargaining. As the contract curve crosses the competitive boundary—i.e., the locus of $(w, a)$ pairs at which marginal cost equals marginal revenue—the employment effect switches from positive to negative. This is shown explicitly in Figure G1b. The figures are generated with $f(L) = L^\alpha$, $v(w) = \ln(w)$, $T = 40$, $\alpha = 0.5$, $\xi = 1.2$, $\beta = 0.750$, $\eta = 0.375$, and inside options at the first fourth of the segment from the monopsonistic to the perfectly competitive outcome.

In the model, firms would never pay higher wages or provide better amenities than those negotiated. Even if bargaining over wages and amenities is inefficient—i.e., $(w^{NB}, a^{NB})$ is such that MRS does not equal MRT—the firm would not adjust either type of compensation to correct this imbalance. This is because the negotiated wage and amenity levels act as minimums, i.e., the employer can only adjust upward on either dimension. Hence, this imposes a constraint on the usual profit maximization problem, such that the optimal solution lies on the boundary of that constraint. Any upward adjustment to restore MRS=MRT would move the firm further away from its profit maximizing bundle. While the firm is forced to accept any potential inefficiency introduced by the constraint, it is important to recall that, as a counterpart in the collective bargaining process, the employer is likely to push for a bundle where MRS equals MRT, regardless of the union's preferences.

Lastly, another important feature to highlight about the model is that collective bargain-

ing acts as a labor market institution that constrains posting. Unlike *individual* bargaining, which only takes place after workers match to firms, *collective* bargaining is not conditional on job matches. In fact, the changes in negotiated compensation bundles relative to those that would have resulted without collective bargaining affects the allocation of workers across firms. Specifically, unions shift $w_j$ and $a_j$ through collective bargaining, which affects each worker's indirect utility of working at some firm (Equation (G1)), which would then materialize in the firm-specific labor supply functions (Equation (G2)).

## G.4 Additional Derivations

This section shows derivations of the bargaining model excluded from the discussion above.

**First order conditions** The Nash bargaining problem is

$$\max_{w,a} \left[\pi(w,a) - \pi(g_w, g_a)\right]^{1-\gamma} \left[\left(v(w) - v(g_w)\right) L(w,a)\right]^{\epsilon\gamma} \left[\left(v(a) - v(g_a)\right) L(w,a)\right]^{(1-\epsilon)\gamma}.$$

Taking logs of the expression above, compute the derivative with respect to $w$ and $a$ to get

$$\begin{aligned} \frac{v(w) - v(g_w)}{wv'(w)} &= \frac{\epsilon\gamma}{(1-\gamma)(-e^{\pi}_w) + \gamma(-e^{L}_w)} \\ \frac{v(a) - v(g_a)}{av'(a)} &= \frac{(1-\epsilon)\gamma}{(1-\gamma)(-e^{\pi}_a) + \gamma(-e^{L}_a)}. \end{aligned}$$

The left hand-side of both equations will always be greater than zero since the union would never agree to anything below its inside option $(g_w, g_a)$. This condition places some bounds on the elasticities of labor to ensure that the right-hand side is also positive, i.e., $e^L_w < \left(-\frac{1-\gamma}{\gamma}\right) e^{\pi}_w$ and $e^L_a < \left(-\frac{1-\gamma}{\gamma}\right) e^{\pi}_a$.

The signs of the elasticities depend on whether the labor supply or the labor demand curve are binding. When the labor supply curve is binding, the labor elasticities are positive, i.e., $e^L_w = \frac{\beta w}{w-b}$ and $e^L_a = \frac{\eta a}{a-q}$. When the labor demand curve is binding, the labor elasticities are negative, i.e., $e^L_w = \sigma\left(-\frac{1}{1-\alpha}\right)$ and $e^L_a = (1-\sigma)\left(-\frac{1}{1-\alpha}\right)$ where $\sigma \equiv \frac{w}{w+\xi a}$ denotes the proportion of marginal costs of labor accounted for by the wage component. Plugging in these conditions, the profit elasticities expressions change but remain negative. In the demand constrained case, they are $e^{\pi}_w = -\frac{\sigma\alpha}{1-\alpha}$ and $e^{\pi}_a = -\frac{(1-\sigma)\alpha}{1-\alpha}$. In the supply constrained case, they are $e^{\pi}_w = \alpha\left(\frac{\beta - x(\beta+\sigma)}{1-\alpha x}\right)$ and $e^{\pi}_a = \alpha\left(\frac{\eta - x(\eta+(1-\sigma))}{1-\alpha x}\right)$ where $x$ represent the markdown on the marginal revenue product of labor. As $x \to 1$, the profit elasticities approach the solution from the demand constrained case.

**Contract curve** Take the first order condition with respect to wages—assuming that $v(\cdot) = \log(\cdot)$—to obtain

$$\gamma = \frac{e_w^\pi(\log(w) - \log(g_w))}{(e_e^\pi - e_w^L)(\log(w) - \log(g_w)) - \epsilon}.$$

Plug this into the first order condition with respect to amenities to get

$$\begin{aligned}(e_w^L e_a^\pi - e_w^\pi e_a^L)(\log(a) - \log(g_a))(\log(w) - \log(g_w)) + \\ \epsilon e_a^\pi(\log(a) - \log(g_a)) - (1-\epsilon)e_w^\pi(\log(w) - \log(g_w)) = 0.\end{aligned}$$

This is a function $F(w, a) = 0$ for which implicit differentiation implies $\frac{da}{dw} = -\frac{dF/dw}{dF/da}$.

Taking these derivatives while holding $\sigma$ and the elasticities constant gives

$$\frac{da}{dw} = \left(-\frac{a}{w}\right)\left(\frac{(e_w^L e_a^\pi - e_w^\pi e_a^L)(\log(a) - \log(g_a)) - (1-\epsilon)e_w^\pi}{(e_w^L e_a^\pi - e_w^\pi e_a^L)(\log(w) - \log(g_w)) + \epsilon e_a^\pi}\right).$$

One can show that $(e_w^L e_a^\pi - e_w^\pi e_a^L) = 0$ under both the demand constrained and the supply constrained cases. Hence, the slope of the contract curve can be simplified to

$$\frac{da}{dw} = \left(\frac{1-\epsilon}{\epsilon}\right)\left(\frac{e_w^\pi}{e_a^\pi}\right)\left(\frac{a}{w}\right),$$

which is always positive.

**Negotiation effort on wages** Assume bargaining only occurs over efficient combinations of $(w, a)$ so that $\frac{w}{a} = \frac{\xi e_w^{L_S}}{e_a^{L_S}}$. This implies that the chosen $\epsilon$ would be such that the slope of the contract curve satisfies this condition. In other words, $\frac{da}{dw} = \frac{e_a^{L_S}}{\xi e_w^{L_S}}$ which implies

$$\epsilon^{NB} = \frac{\xi e_w^{L_S} e_w^\pi a}{\xi e_w^{L_S} e_w^\pi a + e_a^{L_S} e_a^\pi w}.$$

Plugging in the demand constrained profit elasticities simplifies to $\epsilon^{NB} = \sigma$, while plugging in the supply constrained profit elasticities simplifies to $\epsilon^{NB} = e_w^{L_S}/(e_w^{L_S} + e_a^{L_S})$. Under the efficient bargaining assumption, it turns out that $\sigma$ is also equal to $e_w^{L_S}/(e_w^{L_S} + e_a^{L_S})$. Hence, the same negotiation effort on wages holds in both supply and demand constrained labor market conditions.