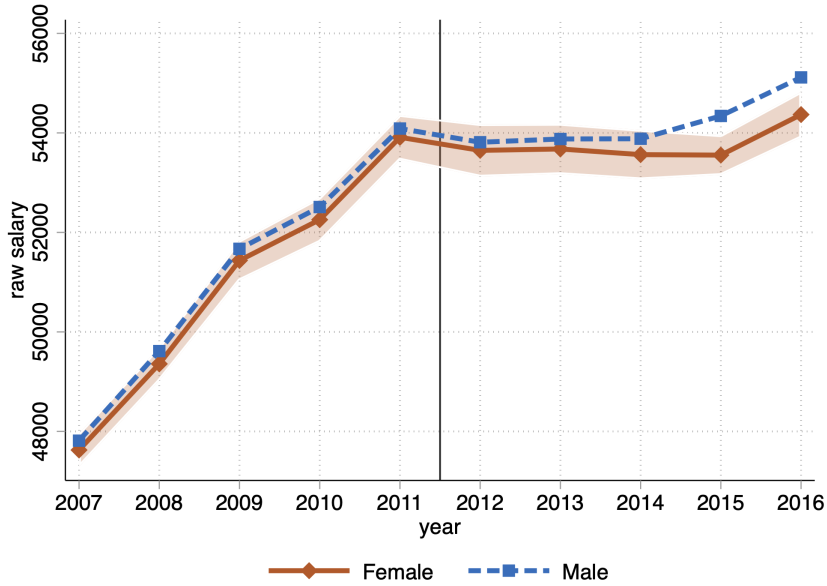
# Appendix

For online publication only

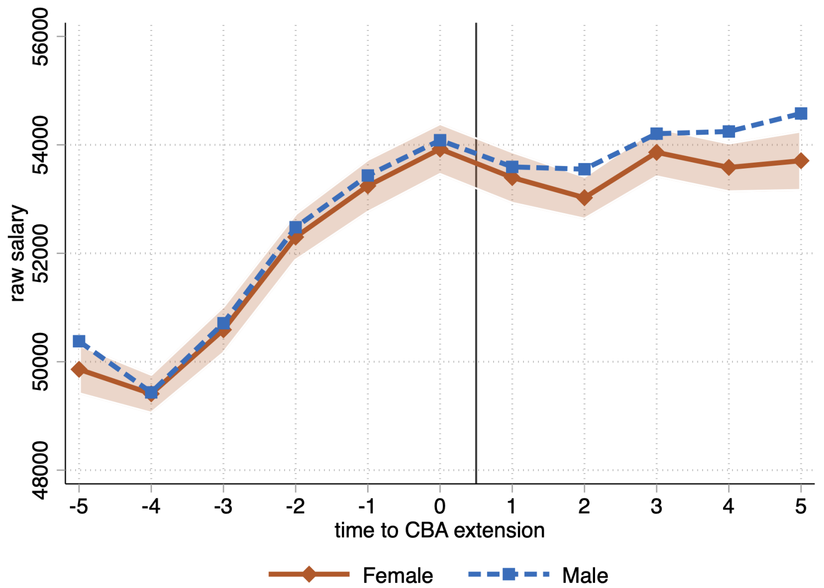
# Appendix A  Additional Tables and Figures

Figure AI: Raw Salaries of Men and Women

Panel A) Raw Salaries by Year

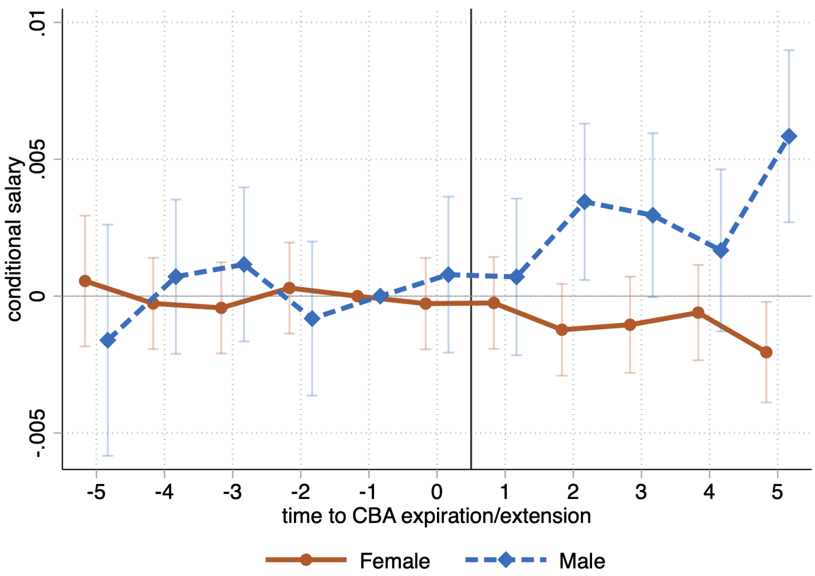



Panel B) Raw Salaries around Extension Date




*Note*: The figure shows the unconditional salaries of male and female teachers by calendar year (Panel A) and relative to the year a CBA or its extension expired ($t = 0$). (Panel B).

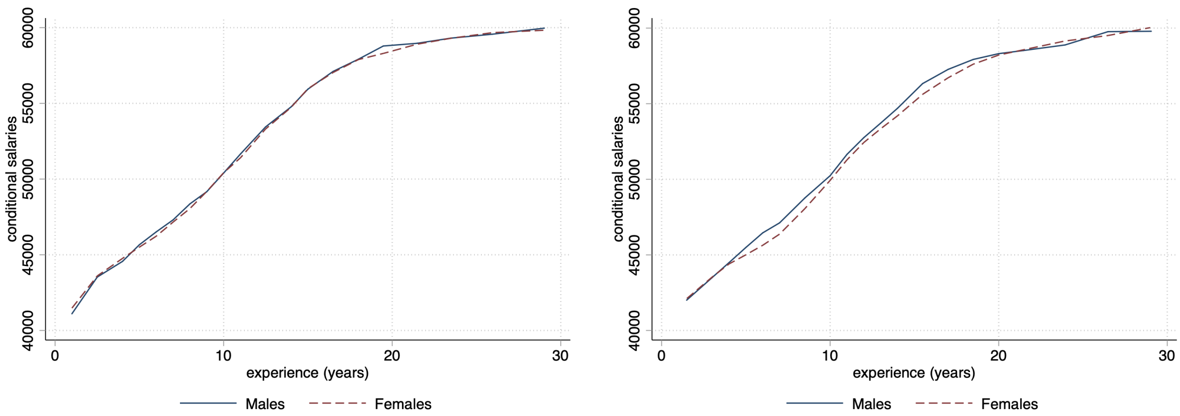
Figure AII: Salaries of Men and Women, by Time to Expiration of CBA



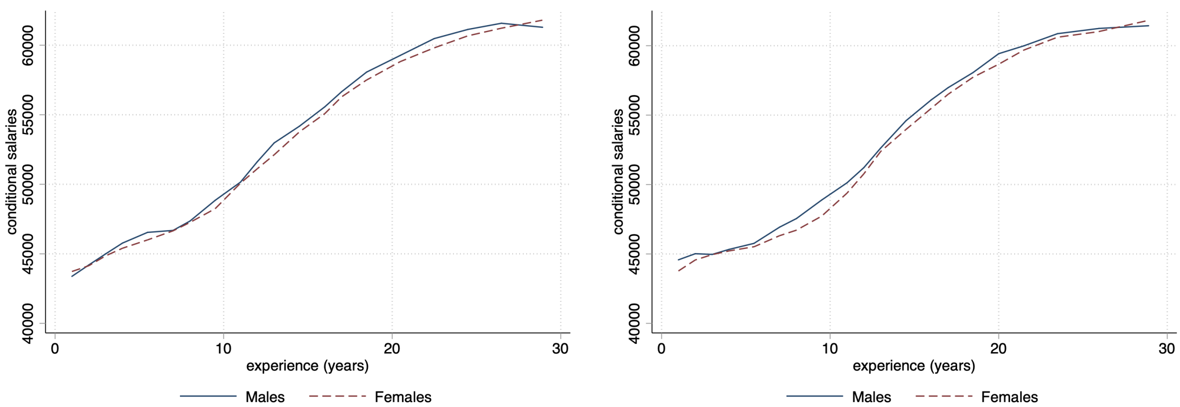

*Note*: The figure shows OLS point estimates and 90% confidence intervals of the coefficients $\delta_s$ in equation (2) in the paper, for $g$ =female (solid line) and $g$ =male (dashed line), and using CBA expiration dates (rather than extensions). All coefficients are plotted relative to the year a CBA expired ($t = 0$). Standard errors are clustered at the district level.

Figure AIII: Conditional Salaries of Men and Women, by Experience

Panel A) Years before a CBA expiration. Seniority pay (left) and flexible pay (right)

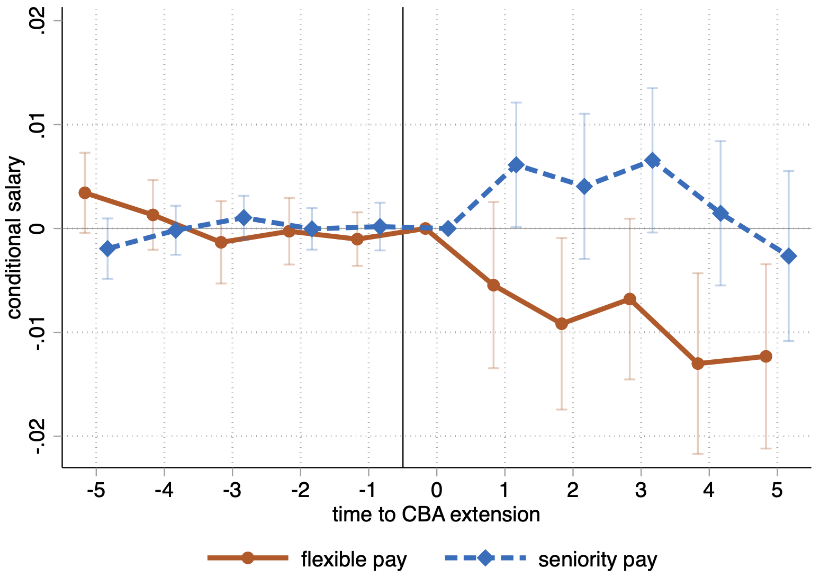



Panel B) Years after a CBA expiration. Seniority pay (left) and flexible pay (right)




*Note*: The figure shows conditional salaries per years of experience, separately for males and females; the top panel uses data prior to (and including) 2011, the bottom panel uses data after CBA extensions. Conditional salaries are obtained as residuals of a regression of salaries on education, district, and teaching assignment fixed effects, alone and interacted with an indicator for years following extensions, as well as year effects interacted by extension years.

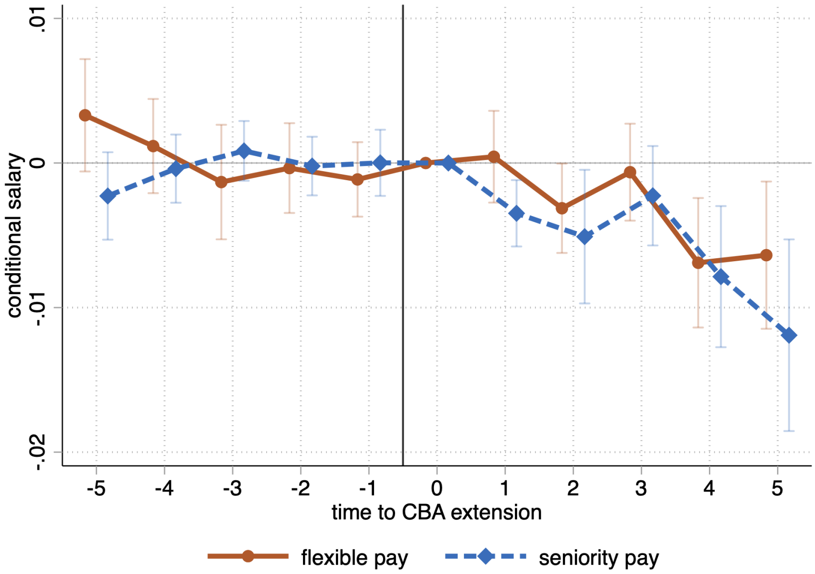
Figure AIV: Gender Gap in Salaries, by Time to Expiration/Extension of CBA and District type. Teachers with 19-20 years of experience and a master's degree



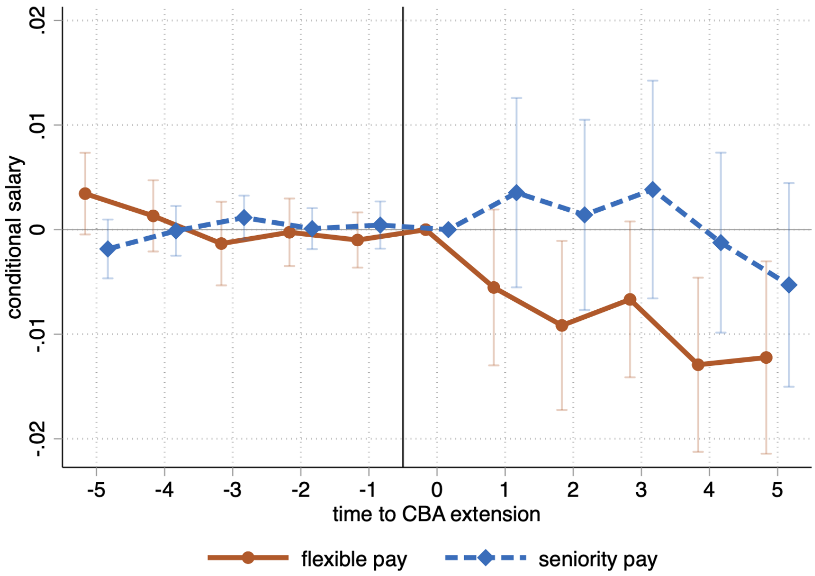

*Note*: The figure shows OLS point estimates and 90% confidence intervals of the coefficients $\delta_s$ in the equation $\ln(w_{ijt}) = \sum_{s=-4}^{3} \delta_s Female_i \mathbb{1}(t - Y_j = s) + \beta X_{it} + \varepsilon_{ijt}$, where $\ln(w_{ijt})$ is the natural logarithm of salaries for teacher $i$, working in district $j$ in year $t$; $Female_i$ equals 1 for women; $Y_j$ is either the year of expiration of district $j$'s CBA or the year in which the extension to the CBA ended; the vector $X_{it}$ contains district, seniority, and education fixed effects (alone and interacted for an indicator for years after a CBA expiration), and year fixed effects interacted with extension year fixed effects. We also control for seniority and education fixed effects interacted with $Female_i$ and with an indicator for years following $Y_j$; the plotted coefficients refer to teachers with 19 or 20 years of experience and a master's degree. The coefficients are estimated and shown separately for flexible-pay (FP) and seniority-pay (SP) districts. All coefficients are plotted relative to the year a CBA or its extension expired ($t = 0$). Standard errors are clustered at the district level.

Figure AV: Gender Gap in Salaries, by Time to Expiration/Extension of CBA and District type

Panel a) Baseline

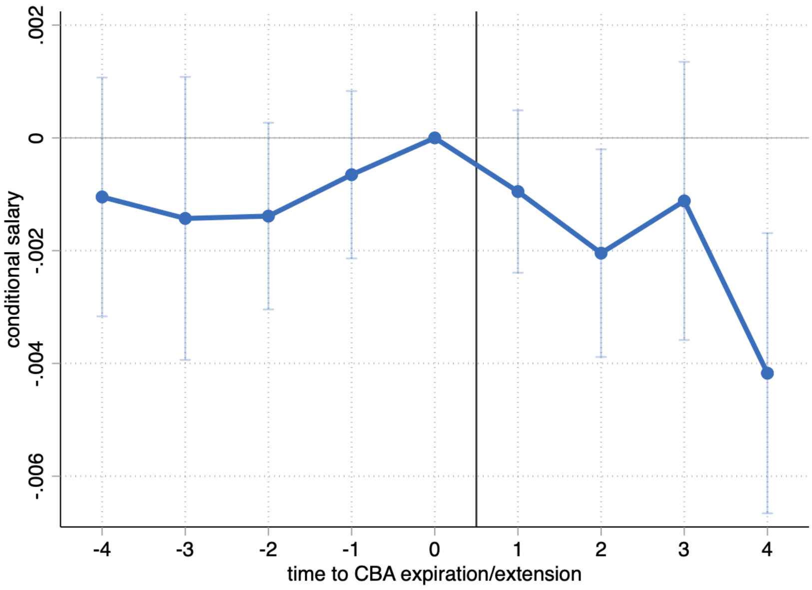



Panel b) With gender-specific experience returns, for teachers with 3-4 years of experience and a master's degree



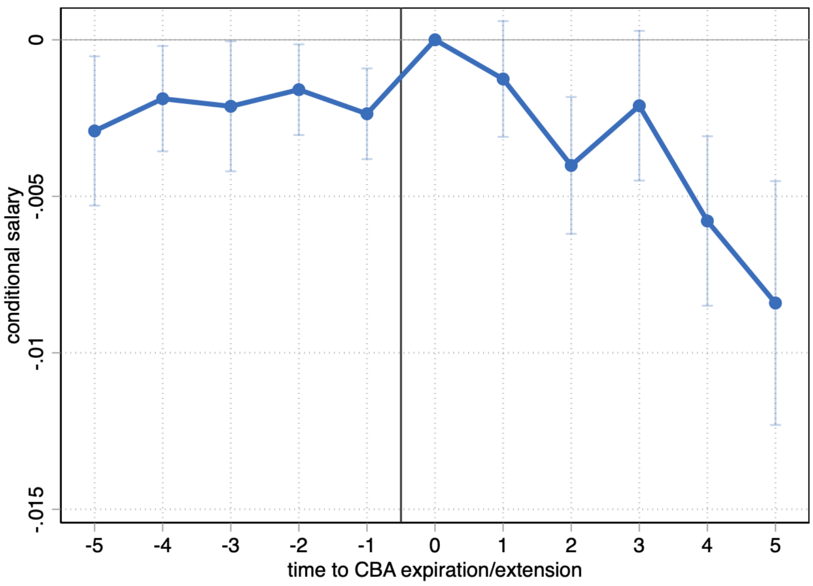

*Note*: The figure shows OLS point estimates and 90% confidence intervals of the coefficients $\delta_s$ in the equation $\ln(w_{ijt}) = \sum_{s=-4}^{3} \delta_s Female_i \mathbb{1}(t - Y_j = s) + \beta X_{it} + \varepsilon_{ijt}$, where $\ln(w_{ijt})$ is the natural logarithm of salaries for teacher $i$, working in district $j$ in year $t$; $Female_i$ equals 1 for women; $Y_j$ is either the year of expiration of district $j$'s CBA or the year in which the extension to the CBA ended; the vector $X_{it}$ contains district, seniority, and education fixed effects (alone and interacted for an indicator for years after a CBA expiration), and year fixed effects interacted with extension year fixed effects. The coefficients are estimated and shown separately for flexible-pay (FP) and seniority-pay (SP) districts. In the bottom panel, we further control for seniority and education fixed effects interacted with $Female_i$ and with an indicator for years following $Y_j$; the plotted coefficients refer to teachers with 3 or 4 years of experience and a master's degree. All coefficients are plotted relative to the year a CBA or its extension expired ($t = 0$). Standard errors are clustered at the district level.

Figure AVI: Gender Gap in Salaries, by Time to Expiration/Extension of CBA. Balanced Panel



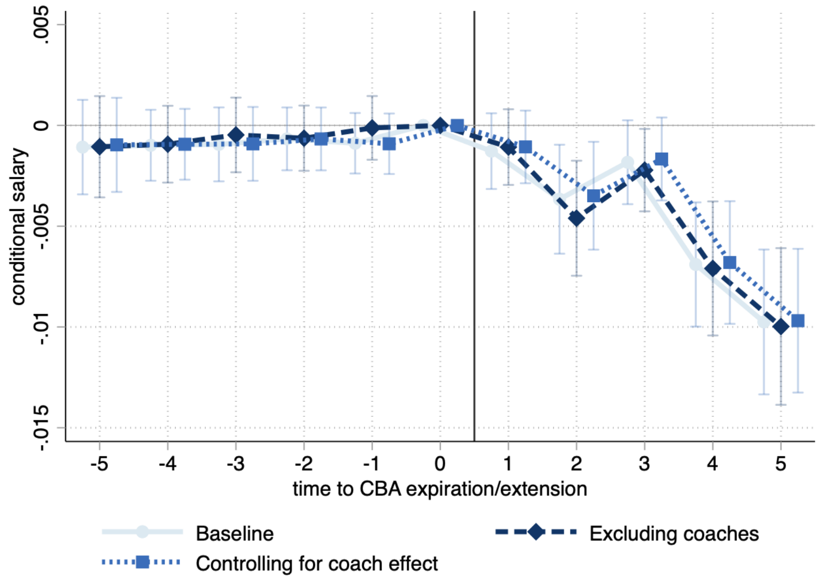

*Note*: This figure estimates equation (3) using a balanced panel. Teachers in this sample are working in the Wisconsin public school district three years before and three years after their district's extension date. All coefficients are plotted relative to the year a CBA or its extension expired ($t = 0$). Standard errors are clustered at the district level.

Figure AVII: Gender Gap in Salaries, by Time to Expiration/Extension of CBA. Intent-to-Treat Estimates



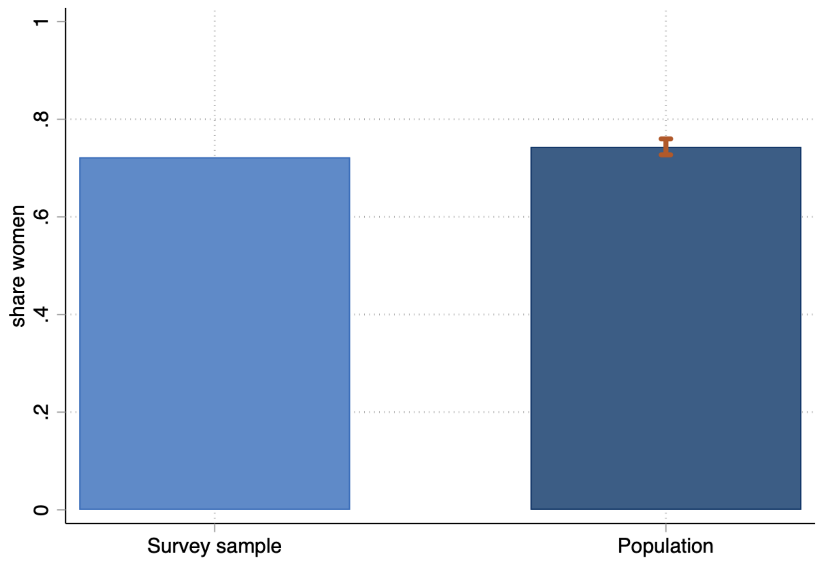

*Note*: This figure shows the ITT estimates from equation (3). We assign teachers to the district they taught in the year before Act 10 and hold this constant. All coefficients are plotted relative to the year a CBA or its extension expired ($t = 0$). Standard errors are clustered at the district level.

Figure AVIII: Gender Gap in Salaries, by Time to Expiration/Extension of CBA. Controlling for Extra Duties (Coaching a Sports Team)



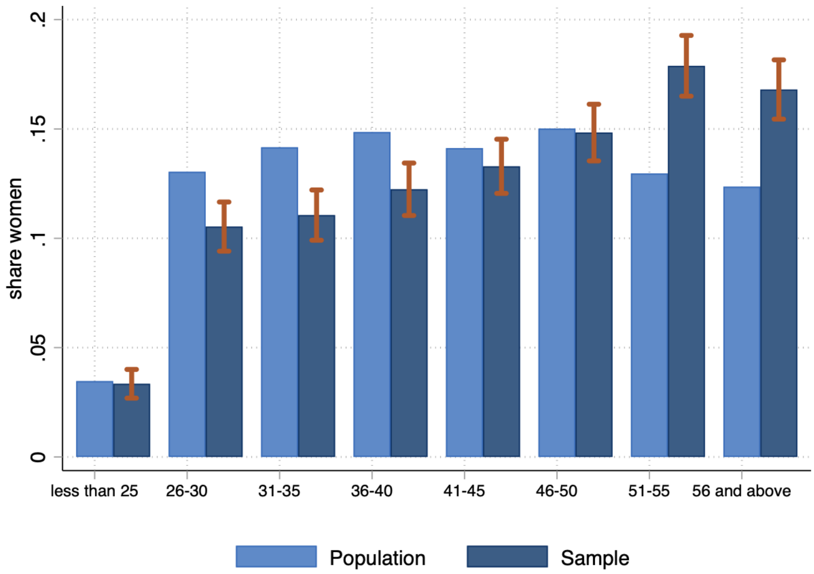

*Note*: The figure shows OLS point estimates and 90% confidence intervals of the coefficients $\delta_s$ in equation (3). The dashed series is obtained excluding teachers who serve as sports coaches. The dotted series is obtained further controlling for an indicator for *coach* interacted with an indicator for years following a CBA expiration. All coefficients are plotted relative to the year a CBA or its extension expired ($t = 0$). Standard errors are clustered at the district level.

Figure AIX: Share of Women: Survey Sample vs. Population



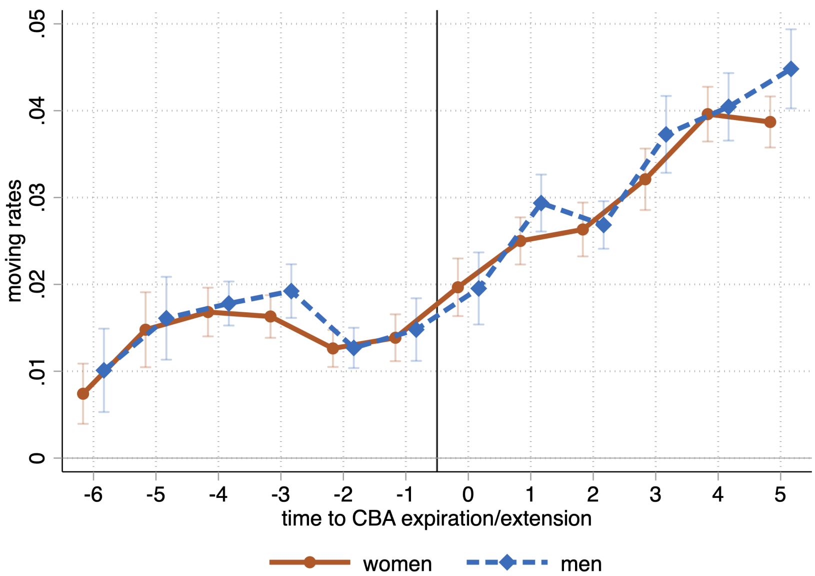

*Note*: Share of female teachers in the survey sample and in the 2016 population. Spikes represent confidence intervals for the difference in mean shares across the two groups. Standard errors are clustered at the district level.

Figure AX: Age Distribution: Survey Sample vs. Population




*Note*: Share of teachers in each age group, in the survey sample and in the 2016 population. Spikes represent confidence intervals for the difference in mean shares across the two groups.

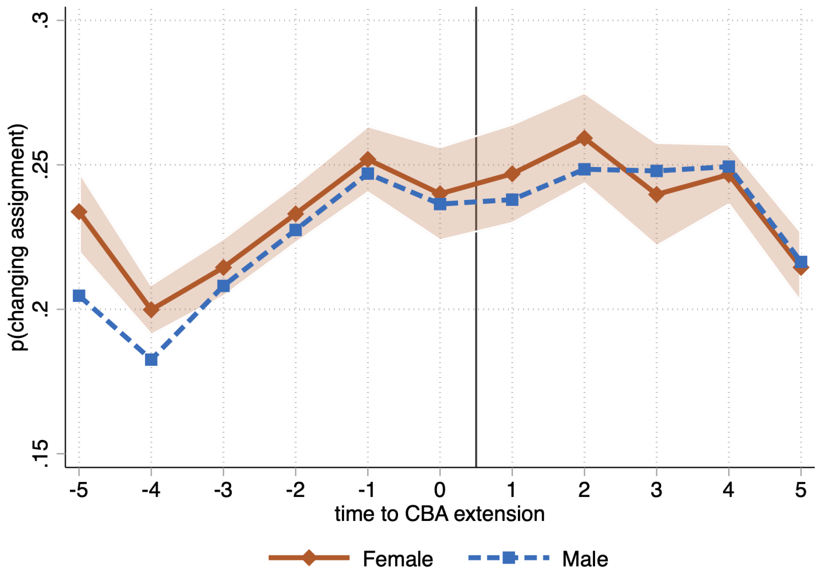
Figure AXI: Mobility Rates, Men and Women



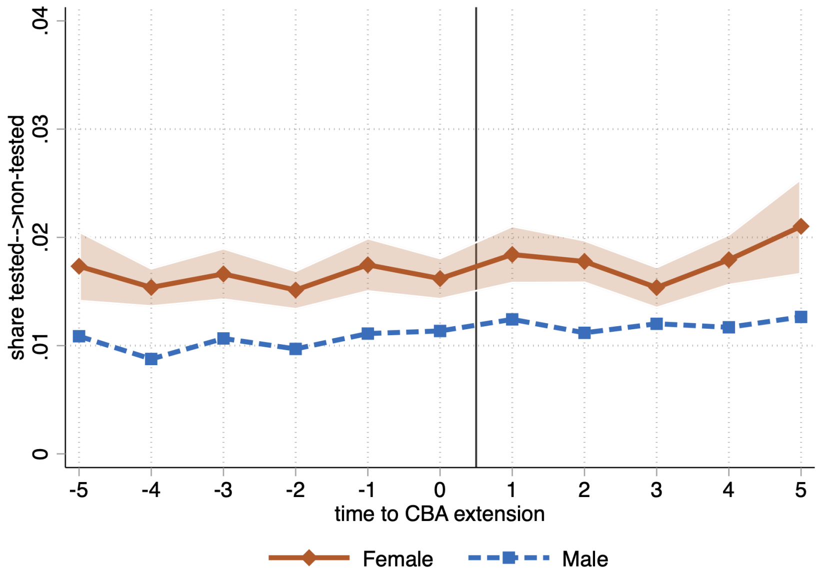

*Note*: Share of teachers who change district (with district-clustered confidence intervals) by time-to-expiration of a district's CBA or its extension. Rates are shown separately for men and women.

Figure AXII: Switches Across Teaching Posts, By Gender

Panel A) Share of teachers who switch teaching post, by gender

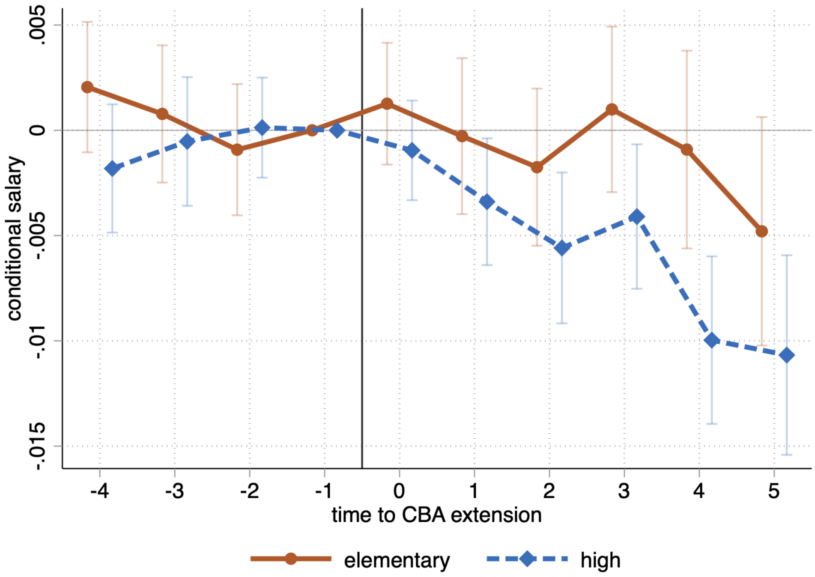



Panel B) Share of teachers who switch from a tested to a non-tested post, by gender




*Note*: The top panel shows the share of teachers who switch teaching position (i.e., grade or subject), by time-to-CBA expiration and gender. The bottom panel shows the share of teachers who switch from a tested to a non-tested post, by time-to-CBA expiration and gender. Shaded areas represent confidence intervals for the female-male difference in the shares.

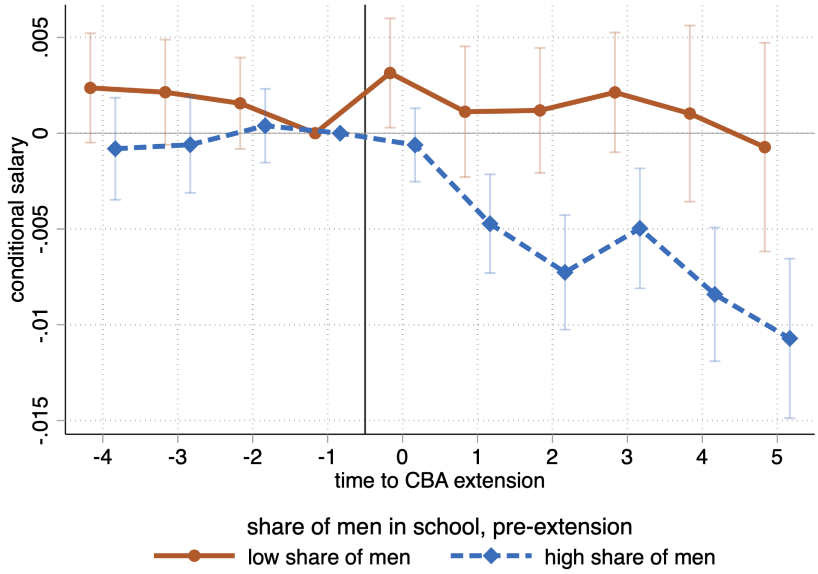
Figure AXIII: Gender Gap in Salaries, for Elementary vs High School Teachers



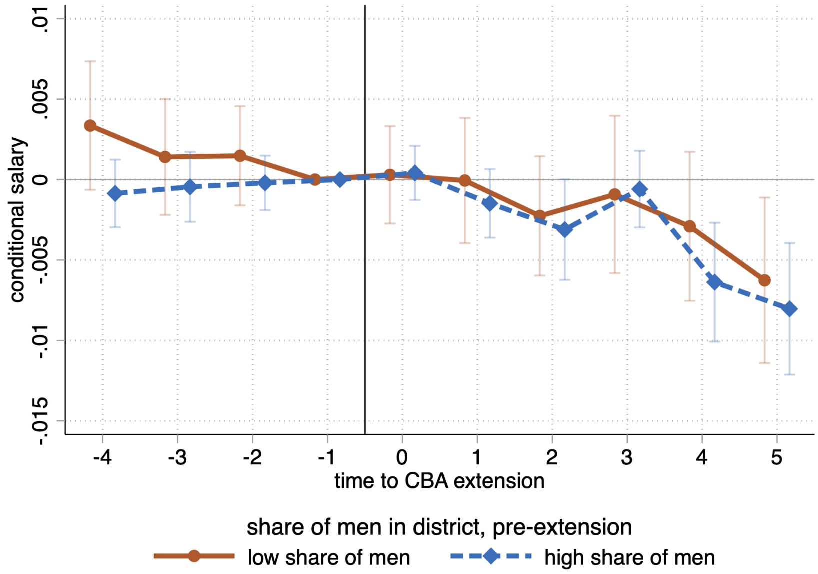

*Note*: OLS point estimates and 90% confidence intervals of the coefficients $\delta_s$ in equation (3), estimated separately for teachers in elementary school (solid line) and in high school (dashed line). All coefficients are plotted relative to the year a CBA or its extension expired ($t = 0$). Standard errors are clustered at the district level.

Figure AXIV: Gender Gap in Salaries, By Share of Men in School and District

Panel A: By Share of Male Colleagues in School

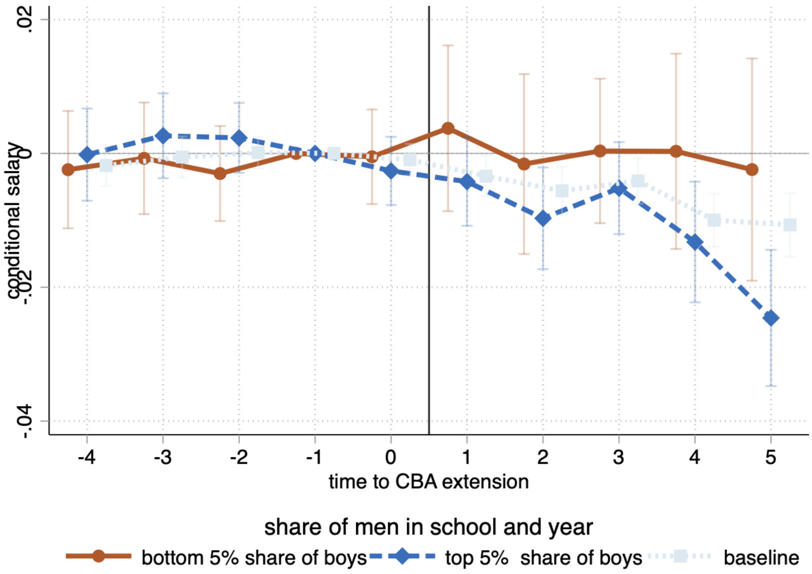



Panel B: By Share of Male Colleagues in District



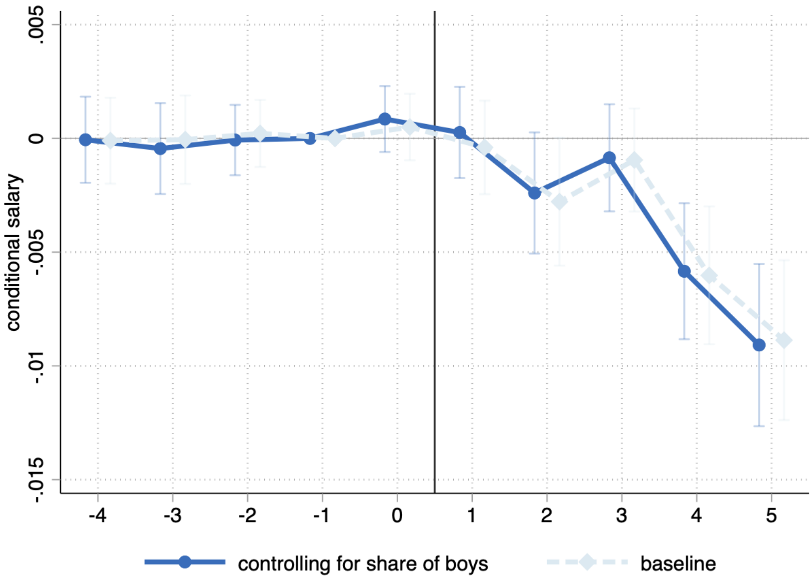

*Note*: Panel A shows OLS point estimates and 90% confidence intervals of the coefficients $\delta_s$ in equation (3), estimated separately for teachers in schools in the top quartile of the share of men (i.e., with more than 30 percent of men, solid line), and teachers in all other schools (dashed line). Panel B shows OLS point estimates and 90% confidence intervals of the coefficients $\delta_s$ in the equation (3), estimated separately for teachers in districts in the top quartile of the share of men (i.e., with more than 30 percent of men, solid line), and teachers in all other districts (dashed line). All coefficients are plotted relative to the year a CBA or its extension expired ($t = 0$). Standard errors are clustered at the district level.

Figure AXV: Gender Gap in Salaries, By Share of Boys in School




*Note*: OLS point estimates and 90% confidence intervals of the coefficients $\delta_s$ in equation (3), estimated separately for teachers in schools in the top and bottom 5 percent of the share of boys. "Baseline" refers to the gap estimated on all schools. All coefficients are plotted relative to the year a CBA or its extension expired ($t = 0$). Standard errors are clustered at the district level.

Figure AXVI: Gender Gap in Salaries, Controlling for The Share of Boys in The School




*Note*: OLS point estimates and 90% confidence intervals of the coefficients $\delta_s$ in equation (3), controlling for the share of boys in each school (alone and interacted with an indicator for years after a CBA expiration). "Baseline" refers to the gap estimated on all schools. All coefficients are plotted relative to the year a CBA or its extension expired ($t = 0$). Standard errors are clustered at the district level.

Table AI: Gender Gap in Salaries, Prior to CBA Expirations/Extensions

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Female | -0.0087** | -0.0055*** | -0.0046*** | -0.0011 | -0.0011 |
| | (0.0037) | (0.0015) | (0.0013) | (0.0010) | (0.0010) |
| Distr and year FE | Yes | Yes | Yes | Yes | Yes |
| Experience FE | No | Yes | Yes | Yes | Yes |
| Education FE | No | No | Yes | Yes | Yes |
| Teaching assignm | No | No | No | Yes | Yes |
| Subject | No | No | No | Yes | Yes |
| N | 307525 | 307522 | 307355 | 307355 | 307355 |
| # districts | 428 | 428 | 428 | 428 | 428 |

*Note*: This table shows how the pre-Act 10 gender salary gap changes as we control for observable characteristics that go into district salary schedules. Estimates are obtained using data on years prior to each district's CBA expiration. The dependent variable is the natural logarithm of salary per year, in full-time equivalency units. The variable *Female* equals one for female workers. All specifications include district and year fixed effects; columns 2-5 include years of experience fixed effects, columns 3-5 include fixed effects for the highest education degree, columns 4-5 include fixed effects for the school level (elementary, middle, high school), and column 5 includes fixed effects for subjects taught. Standard errors in parentheses are clustered at the district level. $^{*} \leq 0.1$, $^{**} \leq 0.05$, $^{***} \leq 0.01$.

Table AII: Gender salary gap after a CBA expiration: Robustness checks. OLS, dependent variable is log(salaries)

| | Balanced | Teacher FE | ITT | District sched. |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Female | -0.0007 | 0.0016 | -0.0011 | -0.0007 |
| | (0.0010) | (0.0050) | (0.0010) | (0.0010) |
| Female × Post Extension | -0.0043*** | -0.0047*** | -0.0060*** | -0.0067*** |
| | (0.0012) | (0.0012) | (0.0012) | (0.0012) |
| Distr × Post exp | Yes | Yes | Yes | Yes |
| Educ, Exper, Teaching Assign × Post exp | Yes | Yes | Yes | Yes |
| Yr × Exp yr | Yes | Yes | Yes | Yes |
| N | 327687 | 569111 | 490644 | 576135 |
| # districts | 428 | 428 | 428 | 428 |

*Note*: The dependent variable is the natural logarithm of salary per year, in full-time equivalency units. The variable *Female* equals one for female workers, the variable *Post Extension* equals one for years following the expiration of a CBA or its extension. All specifications include fixed effects for the district, number of years of seniority, highest education degree, grade level (elementary, middle, high), and subject (math, reading, and others), alone and interacted with an indicator for years after the extension of a CBA. Column 1 is estimated on a balanced sample of teachers in the 3 years before and after each expiration; column 2 includes teacher fixed effects; column 3 assigns teachers to the districts where they were teaching in 2011; and column 4 controls for indicators for years of experience and highest education degree, interacted with district fixed effects and for an indicator for years after the extension of a CBA. All specifications also include year fixed effects interacted with extension year effects. Standard errors in parentheses are clustered at the district level. $^{*} \leq 0.1$, $^{**} \leq 0.05$, $^{***} \leq 0.01$.

Table AIII: Differences in school district characteristics by gender of the leadership

| | Principal (school level) | | | Superintendent (district level) | | |
|---|---|---|---|---|---|---|
| | Female | Male | Diff. | Female | Male | Diff. |
| Female teachers | 0.8 | 0.7 | 0.09*** | 0.7 | 0.7 | 0.02** |
| | | | (0.008) | | | (0.009) |
| Black teachers | 0.02 | 0.01 | 0.01*** | 0.003 | 0.002 | 0.001 |
| | | | (0.003) | | | (0.001) |
| Salary ($) | 53230.3 | 52485.3 | 745.0*** | 52083.7 | 50624.5 | 1459.1** |
| | | | (281.9) | | | (681.8) |
| Value-added | -0.002 | -0.002 | -0.00002 | -0.001 | -0.002 | 0.001 |
| | | | (0.002) | | | (0.002) |
| Cross-district mover | 0.01 | 0.02 | -0.003 | 0.02 | 0.02 | -0.002 |
| | | | (0.002) | | | (0.007) |
| Leaves sample | 0.09 | 0.09 | -0.002 | 0.09 | 0.10 | -0.004 |
| | | | (0.004) | | | (0.006) |

*Note*: This table shows average characteristics, measured in 2011, of schools (left panel, "Principals") and districts (right panel "Superintendent") by school and district leadership. Columns 1-3 show average school characteristics by principal gender and a t-test of differences by principal gender (one observation is a school). Columns 4-6 show average district characteristics by superintendent gender and a t-test of differences by superintendent gender (one observation is a district).

Table AIV: Survey Answers: Likelihood of Negotiating, OLS Estimates. No controls

Panel A) Ever negotiated with:

| | Previous employer | Current empl., at start | Current empl, after start |
|---|---|---|---|
| Female | -0.090*** | -0.085*** | -0.043** |
| | (0.020) | (0.021) | (0.019) |
| N | 2836 | 2836 | 2836 |
| Y mean, males | 0.379 | 0.306 | 0.245 |

Panel B) Negotiated successfully conditional on negotiating, with:

| | Previous employer | Current empl., at start | Current empl., after start |
|---|---|---|---|
| Female | -0.080*** | -0.080*** | -0.106* |
| | (0.028) | (0.028) | (0.056) |
| N | 902 | 902 | 614 |
| Y mean, males | 0.904 | 0.814 | 0.572 |

Panel C) Reasons for not negotiating (current employer, at start)

| | Not possible | Not comfortable | Useless | Fear backlash | Satisfied w/pay |
|---|---|---|---|---|---|
| Female | -0.028 | 0.083*** | 0.038 | 0.010 | -0.051** |
| | (0.027) | (0.029) | (0.025) | (0.018) | (0.021) |
| N | 2222 | 2222 | 2222 | 2222 | 2222 |
| Y mean, males | 0.565 | 0.210 | 0.215 | 0.131 | 0.189 |

Panel D) Likelihood of negotiating in the future, over:

| | Salary | Classroom assignment | Non-teaching duties |
|---|---|---|---|
| Female | -0.563*** | 0.271* | -0.160 |
| | (0.165) | (0.148) | (0.131) |
| N | 2836 | 2836 | 2836 |
| Y mean, males | 3.889 | 4.539 | 4.579 |

*Note*: All regressions include controls for age class, self-reported job performance (above/below average), a measure of people skills, an indicator for whether the respondent knows someone who negotiated his/her salary, an indicator for whether the respondent knows his/her colleagues' salaries, and district fixed effects. Standard errors in parentheses are clustered at the district level. * $\leq 0.1$, ** $\leq 0.05$, *** $\leq 0.01$.

Table AV: Survey Answers: Likelihood of Negotiating, OLS Estimates. No controls

| | Neg. beginning (1) | Neg. after (2) | Neg. future (3) | Successful neg (4) | Not confident (5) |
|---|---|---|---|---|---|
| Female | -0.077*** | -0.026 | -0.383** | -0.105** | 0.112*** |
| | (0.025) | (0.020) | (0.162) | (0.042) | (0.021) |
| Knows colleague pay | 0.013 | 0.083*** | 0.230 | 0.066 | -0.094*** |
| | (0.035) | (0.030) | (0.238) | (0.052) | (0.023) |
| Female * knows colleague pay | -0.001 | -0.006 | -0.335 | 0.005 | -0.033 |
| | (0.041) | (0.036) | (0.262) | (0.070) | (0.029) |
| Female | -0.151** | -0.063 | -0.914* | -0.190 | 0.200** |
| | (0.076) | (0.076) | (0.494) | (0.184) | (0.079) |
| Female * People skills | 0.074 | 0.025 | 0.415 | 0.084 | -0.088 |
| | (0.081) | (0.077) | (0.497) | (0.190) | (0.079) |
| People skills | 0.028 | 0.025 | 0.084 | -0.045 | -0.074 |
| | (0.067) | (0.062) | (0.426) | (0.098) | (0.057) |
| Female | 0.005 | -0.046 | -0.672** | -0.225*** | 0.130*** |
| | (0.037) | (0.043) | (0.320) | (0.086) | (0.045) |
| Confident talking | 0.179*** | 0.045 | 0.149 | -0.114 | -0.104** |
| | (0.039) | (0.045) | (0.324) | (0.076) | (0.043) |
| Female * Confident talking | -0.083* | 0.017 | 0.231 | 0.122 | -0.037 |
| | (0.043) | (0.048) | (0.333) | (0.090) | (0.046) |
| Female | -0.088 | 0.028 | -0.307 | -0.111 | 0.230*** |
| | (0.062) | (0.057) | (0.459) | (0.129) | (0.056) |
| Understand feelings | 0.048 | 0.107** | 0.088 | -0.016 | 0.049 |
| | (0.054) | (0.049) | (0.400) | (0.092) | (0.038) |
| Female * Understand feelings | 0.010 | -0.073 | -0.219 | -0.001 | -0.127** |
| | (0.063) | (0.060) | (0.488) | (0.131) | (0.058) |
| Female | -0.093*** | -0.019 | -0.577*** | -0.110** | 0.151*** |
| | (0.033) | (0.023) | (0.184) | (0.055) | (0.027) |
| Perf > avg | 0.016 | 0.130*** | -0.121 | -0.028 | -0.024 |
| | (0.036) | (0.033) | (0.221) | (0.049) | (0.026) |
| Female * Perf > avg | 0.023 | -0.038 | 0.112 | 0.004 | -0.059* |
| | (0.040) | (0.033) | (0.249) | (0.071) | (0.033) |
| N | 2810 | 2809 | 2801 | 701 | 2810 |
| Y mean, males | 0.306 | 0.245 | 3.889 | 0.814 | 0.128 |

*Note*: The dependent variable is an indicator for whether a teacher negotiated with the current employer at the beginning or after the start of the work relationship (columns 1, 2, respectively); whether the teacher plans to negotiate pay in the future (column 3); whether past negotiations were successful (column 4); and whether a teacher did not negotiate in the past because she did not feel comfortable doing so (column 5). Each column and panel is a separate regression. Standard errors in parentheses are clustered at the district level. $^* \leq 0.1$, $^{**} \leq 0.05$, $^{***} \leq 0.01$.

Table AVI: Gender Differences in Mobility, by Type of District and Value-Added

| | Move to FP | | | Move to SP | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | All teachers | High VA | Low VA | All teachers | High VA | Low VA |
| Female | -0.0005 | -0.0012 | -0.0009 | 0.0003 | 0.0011 | 0.0017 |
| | (0.0005) | (0.0011) | (0.0012) | (0.0003) | (0.0013) | (0.0012) |
| Post Extension | -0.0025** | -0.0032 | -0.0029 | -0.0013 | -0.0025 | -0.0027 |
| | (0.0011) | (0.0025) | (0.0034) | (0.0013) | (0.0038) | (0.0033) |
| Female × Post Extension | -0.0015** | -0.0035 | -0.0016 | -0.0012 | 0.0023 | -0.0001 |
| | (0.0007) | (0.0027) | (0.0028) | (0.0008) | (0.0023) | (0.0027) |
| District FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Experience, education FE | Yes | Yes | Yes | Yes | Yes | Yes |
| N | 430916 | 48789 | 51358 | 430916 | 48789 | 51358 |
| # districts | 224 | 222 | 221 | 224 | 222 | 221 |
| Mean of dep. var. | | | | | | |

| | Move from FP | | | Move from SP | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | All teachers | High VA | Low VA | All teachers | High VA | Low VA |
| Female | -0.0002 | 0.0002 | 0.0004 | -0.0003 | -0.0002 | -0.0006 |
| | (0.0003) | (0.0011) | (0.0009) | (0.0004) | (0.0011) | (0.0012) |
| Post Extension | -0.0027*** | -0.0040 | -0.0109*** | -0.0008 | 0.0019 | 0.0029 |
| | (0.0009) | (0.0029) | (0.0026) | (0.0011) | (0.0033) | (0.0038) |
| Female × Post Extension | -0.0000 | -0.0011 | 0.0051*** | -0.0011 | 0.0006 | -0.0014 |
| | (0.0007) | (0.0023) | (0.0020) | (0.0008) | (0.0024) | (0.0028) |
| District FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Experience, education FE | Yes | Yes | Yes | Yes | Yes | Yes |
| N | 430310 | 48629 | 51179 | 430310 | 48629 | 51179 |
| # districts | 411 | 257 | 263 | 411 | 257 | 263 |
| Mean of dep. var. | | | | | | |

*Note*: The dependent variable is an indicator for a teacher moving to a flexible-pay district (panel a, columns 1-3), to a seniority-pay district (panel a, columns 4-6), out of a flexible-pay district (panel b, columns 1-3), and out of a seniority-pay district (panel b, columns 4-6), and separately for all teachers (columns 1 ad 4), teachers with value-added above the median ("High VA", columnns 2 and 5), and teachers with value-added below the median ("Low VA", columns 3 and 6). The variable *Female* equals one for female teachers and the variable *Post Extension* equals one for years following the expiration of a CBA or its extension. All columns 2-5 include district and year fixed effects, as well as fixed effects for years of experience and for the highest education degree. Standard errors in parentheses are clustered at the district level. $^{*} \leq 0.1$, $^{**} \leq 0.05$, $^{***} \leq 0.01$.

Table AVII: Outside Options and the Gender Gap in Salaries

| | Log Salary | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Female | -0.004*** | -0.004** | -0.001 |
| | (0.001) | (0.002) | (0.002) |
| Post Extension | -0.012*** | -0.012*** | -0.013*** |
| | (0.004) | (0.004) | (0.004) |
| Female × Post Ext | -0.002 | -0.005** | -0.001 |
| | (0.001) | (0.002) | (0.002) |
| Female × Post Ext × Num Schools | -0.001 | | |
| | (0.001) | | |
| Female × Num Schools | | 0.001 | |
| | | (0.001) | |
| Female × Post Ext × Num High Schools | | -0.001 | |
| | | (0.001) | |
| Female × Num Elem Schools | | | 0.001 |
| | | | (0.001) |
| Female × Post × Num Elem Schools | | | -0.001 |
| | | | (0.001) |
| Distr × Post Exp FE | Yes | Yes | Yes |
| Educ, Exper, Teaching Assign × Post Exp FE | Yes | Yes | Yes |
| Observations | 579,331 | 184,060 | 247,500 |
| R-squared | 0.801 | 0.791 | 0.810 |

*Note*: The variable $NumSchools$ is the number of schools in a teacher's commuting zone. In column 2, $NumHighSchools$ is the number of high schools in a teacher's commuting zone and the sample is restricted to high school teachers. In column 3, $NumElemSchools$ is the number of elementary schools in a teacher's commuting zone and the sample is restricted to elementary school teachers. Standard errors in parentheses are clustered at the district level. $^{*} \leq 0.1$, $^{**} \leq 0.05$, $^{***} \leq 0.01$.

# Appendix B Estimating Teacher Value-Added With Grade-School Links

Teacher value-added (VA) is defined as the contribution of each teacher to achievement (or achievement growth), once all other determinants of student learning have been taken into account. The starting model is the following (Kane and Staiger, 2008):

$$
\begin{aligned}
A_{kt} &= \beta X_{kt} + \nu_{kt} \\
\text{where } \nu_{kt} &= \mu_{i(kt)} + \theta_{c(kt)} + \varepsilon_{kt}
\end{aligned}
\tag{8}
$$

and where $A_{kt}$ is a standardized measure of test scores (or test score gains) for student $k$ in year $t$, and $X_{kt}$ is a vector of student, grade, and school observables which could affect achievement, including: school and grade-by-year fixed effects; cubic polynomials of past scores interacted with grade fixed effects; cubic polynomials of average past scores for the students in the same grade and school, interacted with grade fixed effects; student $k$'s demographic characteristics, including gender, race and ethnicity, disability, English-language earner status, and socioeconomic status; the same demographic characteristics, averaged for all students in the same grade and school as student $k$ in year $t$; and the student's socioeconomic status interacted with the share of low-socioeconomic status in her grade and school in $t$.[45] The residual $\nu_{kt}$ can be decomposed into three parts: The error term component $\mu_{i(kt)}$ is the individual effect of teacher $i$, teaching student $k$ in year $t$; the component $\theta_{c(kt)}$ is an exogenous classroom shock; and $\varepsilon_{kt}$ is an idiosyncratic student-specific component which varies over time. VA is an estimate of the teacher effect $\mu_i$.

A range of techniques have been proposed to estimate $\mu_i$, including fixed effects (**?**) and two-steps procedures based on the decomposition of test score residuals (Kane and Staiger, 2008; Chetty et al., 2014). Here, we consider the two-steps estimator of Kane and Staiger (2008), a special case of the more general framework of Chetty et al. (2014) which allows for the correction of noise in the estimates using a Bayes "shrinkage" approach. The estimation procedure can be summarized as follows:

1. Estimate $\beta$ in equation (8) via OLS;
2. Construct residuals $\hat{\nu}_{kt} = A^*_{kt} - \hat{\beta} X_{kt}$, where $\hat{\beta}$ is the OLS estimate of $\beta$;
3. Estimate VA as $\bar{\nu}_i \left( \frac{\sigma_\mu}{Var(\bar{\nu}_i)} \right)$, where
   (a) $\bar{\nu}_i = \sum_t w_{it} \bar{\nu}_{it}$ is a weighted average of average test score residuals $\bar{\nu}_{it}$ for teacher $i$ in year $t$;
   (b) $w_{it} = \frac{h_{it}}{\sum_t h_{it}}$, with $h_{it} = \frac{n_{it}}{n_{it}\sigma^2_\theta + \sigma^2_\varepsilon}$, are the weights, function of class size $n_{it}$, the variance of the classroom component $\sigma^2_\theta$ and of the residual component $\sigma^2_\varepsilon$;
   (c) the variance of the teacher effect is $\sigma^2_\mu = Cov(\bar{\nu}_{it}, \bar{\nu}_{it-1})$; the variance of the residual component is $\sigma^2_\varepsilon = Var(\nu_{kt} - \bar{\nu}_{it})$; the variance of the classroom component is $\sigma^2_\theta = Var(\nu_{kt}) - \sigma^2_\varepsilon - \sigma^2_\mu$.

Constructing an estimate of teacher VA thus requires correctly estimating $\bar{\nu}_{it}$, which in turn requires linking each teacher with the students she taught in each year. The WDPI started to record classroom identifiers, which allow to link students to teachers, only in 2017; data from previous years only contain identifiers for schools and grades. This means that, in a given year, a student can be linked to all the teachers in her school and grade, but not to the specific teacher who taught her (and conversely, a teacher can be linked to all students attending her grade in her school, but not to her own pupils). The lack of information on classroom identifiers is common

[45] This specification largely follows Chetty et al. (2014).

to teacher-student datasets from several other states and/or districts (Rivkin et al., 2005, for example, face a similar issue with data from Texas).

How to identify teacher effects in the absence of classroom links? A simple approximation is given by grade-level average test score residuals. Rivkin et al. (2005), however, show that in the presence of teacher turnover across grades or schools one can obtain a more accurate measure of teacher effects than grade residuals. The intuition behind the identification of these effects is as follows. In the absence of teacher turnover, teachers in grade $g$ and school $s$ would have the same $\bar{\nu}_{it}$ for every $t$, and separately identifying their individual effects would be impossible. With data on test scores for multiple years and in the presence of turnover, teachers switches across schools or within schools and grades allow to isolate the effect of the individual teacher through the comparison of test score residuals before and after her arrival in a given grade and school. Importantly, teacher turnover allows a more precise identification of the effects not only of the teacher who switches school or grade, but also of the teachers teaching in her same grade and school at any point in time.

To incorporate this feature of the data, we proceed as follows.

a. We calculate the grade-school-year average residuals $\bar{\nu}_{gst}$ for each $g$, $s$, and $t$;

b. We construct the "teams" of teachers in each grade and school in each year;

c. Given these teams, we identify teachers or groups of teachers whose value added can be separately identified, either because they move or because other teachers in their team move. For these teachers we can identify a $\bar{\nu}_{it}$; in the Wisconsin data, these teachers represent 70 percent of the whole sample. For 10 percent of the sample, $\bar{\nu}_{it}$ will not be separately identifiable from that of another teacher, and for 20 percent of the sample $\bar{\nu}_{it}$ will not be separately identifiable from that of two or more teachers.

d. Given these $\bar{\nu}_{it}$, we can calculate VA from step 3 above. This strategy does not allow to separately identify $\theta_c$; we therefore assume $\theta_c$ and $\sigma_\theta$ to be zero.

Two features of this identification strategy are worth highlighting:

1. While my VA estimates are more precise than grade-school residuals, they contain more noise relative to estimates obtained with teacher-student links: Even in the presence of turnover, teachers always teaching the same grade-school would have the same $\bar{\nu}_{it}$ for every $t$, and hence the same estimate.

2. The aggregation of teacher effects at the grade level overcomes a problematic form of selection, which occurs within schools and grades and across classrooms when some parents manage to have their children assigned to specific teachers. The (forced) use of grade-school estimates circumvents this form of selection, and is in practice equivalent to an instrumental variable estimator based on grade rather than on classroom assignment (Rivkin et al., 2005).

### Identification of Teacher Value-Added With Turnover

To understand the identification argument, consider a simple example of 3 teachers $(A, B, C)$ observed in 3 periods ($t = 1, 2, 3$) and in 2 possible grades ($g = 4, 5$). The teaching assignments are as follows.

| period | grade | |
|---|---|---|
| 1 | A,B | C |
| 2 | B,C | A |
| 3 | A,C | B |

The objective is to calculate VA of the three teachers in period 3. We define $A_{kt}$ as the average test score residual for students of teacher $k$ in period $t$, and $\bar{A}_t^g$ the average test score residuals of students in grade $g$ in period $t$. Following Chetty et al. (2014) we can write the VA estimate for each teacher as follows (we suppress the hats on the VA estimates for ease of notation and we consider 3 lags):

$$\mu_{A3} = \begin{bmatrix} A_{A1}^2 & A_{A1}A_{A2} \\ A_{A1}A_{A2} & A_{A2}^2 \end{bmatrix}^{-1} \begin{bmatrix} A_{A1}A_{A3} \\ A_{A2}A_{A3} \end{bmatrix} \tag{9}$$

$$\mu_{B3} = \begin{bmatrix} A_{B1}^2 & A_{B1}A_{B2} \\ A_{B1}A_{B2} & A_{B2}^2 \end{bmatrix}^{-1} \begin{bmatrix} A_{B1}A_{B3} \\ A_{B2}A_{B3} \end{bmatrix} \tag{10}$$

$$\mu_{C3} = \begin{bmatrix} A_{C1}^2 & A_{C1}A_{C2} \\ A_{C1}A_{C2} & A_{C2}^2 \end{bmatrix}^{-1} \begin{bmatrix} A_{C1}A_{C3} \\ A_{C2}A_{C3} \end{bmatrix} \tag{11}$$

Assuming a constant number of students in each classroom, one can write:

$$\bar{A}_1^4 = \frac{1}{2}(A_{A1} + A_{B1}) \tag{12}$$

$$\bar{A}_1^5 = A_{C2} \tag{13}$$

$$\bar{A}_2^4 = \frac{1}{2}(A_{B2} + A_{C2}) \tag{14}$$

$$\bar{A}_2^5 = A_{A2} \tag{15}$$

$$\bar{A}_3^4 = \frac{1}{2}(A_{A3} + A_{C3}) \tag{16}$$

$$\bar{A}_3^5 = A_{B3} \tag{17}$$

My VA estimator implies:

$$\mu_{A3} = \begin{bmatrix} (\bar{A}_1^4)^2 & \bar{A}_1^4\bar{A}_2^5 \\ \bar{A}_1^4\bar{A}_2^5 & (\bar{A}_2^5)^2 \end{bmatrix}^{-1} \begin{bmatrix} \bar{A}_1^4\bar{A}_3^4 \\ \bar{A}_2^5\bar{A}_3^4 \end{bmatrix} \tag{18}$$

$$\mu_{B3} = \begin{bmatrix} (\bar{A}_1^4)^2 & \bar{A}_1^4\bar{A}_2^4 \\ \bar{A}_1^4\bar{A}_2^4 & (\bar{A}_2^4)^2 \end{bmatrix}^{-1} \begin{bmatrix} \bar{A}_1^4\bar{A}_3^5 \\ \bar{A}_2^4\bar{A}_3^5 \end{bmatrix} \tag{19}$$

$$\mu_{C3} = \begin{bmatrix} (\bar{A}_1^5)^2 & \bar{A}_1^5\bar{A}_2^4 \\ \bar{A}_1^5\bar{A}_2^4 & (\bar{A}_2^4)^2 \end{bmatrix}^{-1} \begin{bmatrix} \bar{A}_1^5\bar{A}_3^4 \\ A_{C2}\bar{A}_3^4 \end{bmatrix} \tag{20}$$

Equations (9)-(20) represent a system of 12 equations in 12 unknowns: $\mu_{A3}$, $\mu_{B3}$, $\mu_{B3}$, $A_{A1}$, $A_{A2}$, $A_{A3}$, $A_{B1}$, $A_{B2}$, $A_{B3}$, $A_{C1}$, $A_{C2}$, $A_{C3}$. In this case, VA can be perfectly identified for all teachers because at least one teacher switches grade each year.

### Validation Exercise: Value-Added with Classroom Links and with Grade-School Links in the NYC data

To validate the VA estimator with grade-school links described above (which we call GL) against the standard Kane and Staiger estimator with classroom links (CL), we use teacher and student data from the New York City Department of Education (NYCDOE) from the years 2006-07 to 2009-10. This dataset contains classroom, grade, and school identifiers, which allow me to estimate both CL and GL measures. We estimate teacher VA for 15,469 teachers of Math and English-Language-Arts (ELA) using the procedure of Kane and Staiger (2008).

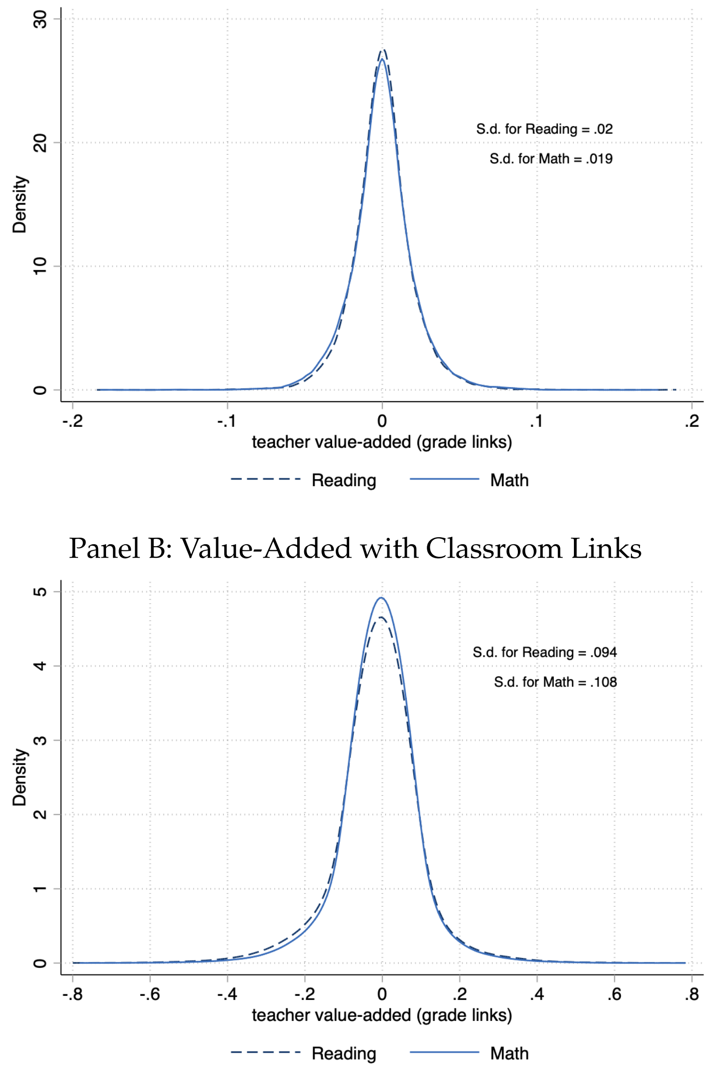
**Measurement Error** The main limitation of GL relative to CL is measurement error. Since students are linked to teachers at the grade-school level, the VA of a teacher will also be a function of test scores of students she never taught.

Classic measurement error will push VA estimates towards zero. To quantify the extent of this problem, Figure BI shows the kernel density of the distribution of GL (top panel) and CL (bottom panel). As expected, the distribution of GL is more concentrated around zero compared to CL. In spite of this, GL is able to explain a significant amount of variance in test scores. Its standard deviation (measured in test scores standard deviation units) is equal to 0.02 for Math teachers; by comparison, the standard deviation of CL is equal to 0.11. Figure BII shows the density of GL for Wisconsin teachers. Its standard deviation is equal to 0.10 for Math teachers.

Figure BI: Empirical Distribution of Value-Added Estimates: New York City, 2007-2010

Panel A: Value-Added with Grade-School Links

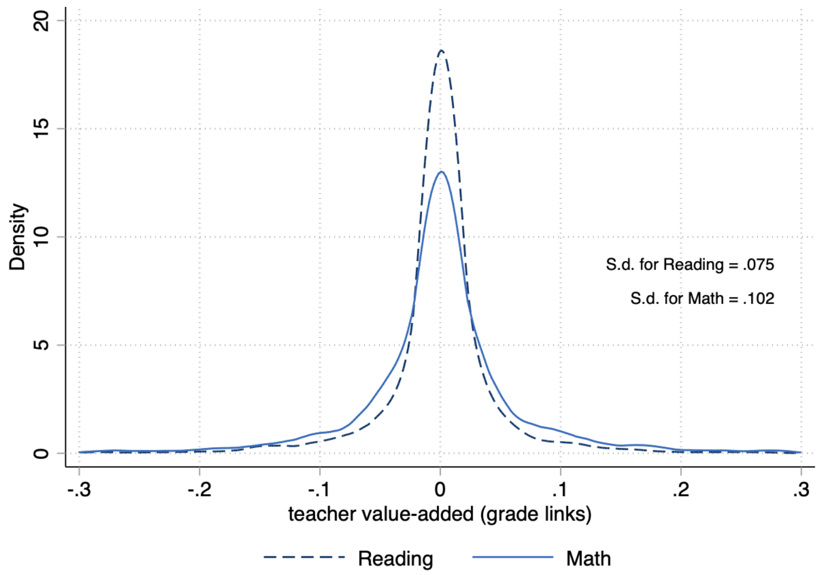
Panel B: Value-Added with Classroom Links




*Notes*: Kernel densities of the empirical distribution of VA estimates for NYC math and ELA teachers, for each subject. Estimates are averaged across years for each teacher. Each density is weighted by the number of student test scores observations used to estimate each teacher's VA, and estimated using a bandwidth of 0.05. The figure also reports the standard deviations of these empirical distributions.

**Forecast Bias of GL as a Proxy for CL** Next, we test whether GL is a forecast-unbiased estimate for CL. Figure BIII shows a binned scatterplot of the two estimates in the NYC data, averaged across the four years for each teacher. Their correlation is 0.62. The forecast bias of $\hat{\mu}_i^{GL}$ as a

Figure BII: Empirical Distribution of Value-Added Estimates: Wisconsin, 2007-2015




*Notes*: Kernel densities of the empirical distribution of VA estimates for Wisconsin math and reading teachers, for each subject. Estimates are averaged across years for each teacher, separately for years before and after Act 10. Each density is weighted by the number of student test scores observations used to estimate each teacher's VA, and estimated using a bandwidth of 0.05. The figure also reports the standard deviations of these empirical distributions.

proxy for $\hat{\mu}_i^{CL}$ can be defined based on the best linear predictor of $\hat{\mu}_i^{CL}$ given $\hat{\mu}_i^{GL}$:

$$\hat{\mu}_i^{CL} = \alpha + \gamma \hat{\mu}_i^{GL} + \chi_i \tag{21}$$

Assuming $\chi_i$ to be uncorrelated with $\hat{\mu}_i^{GL}$, the forecast bias $f$ is zero if $\gamma = 1$: $f = 1 - \gamma$. We can estimate the slope coefficient $\gamma$ via OLS on equation (21). The 95% confidence interval for $\gamma$, whose point estimate is equal to 0.99, includes 1, which implies that the forecast bias $f$ is equal to 0.01 and it is indistinguishable from zero (Figure BIII).

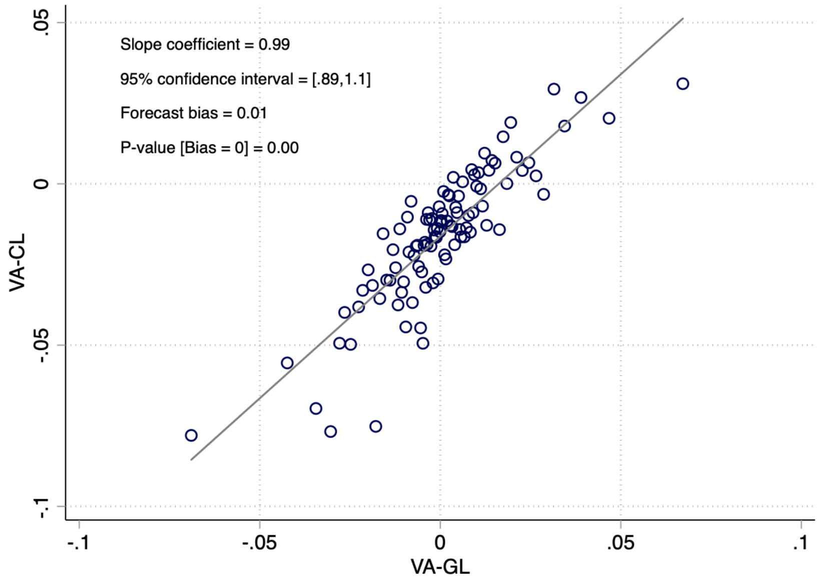
**Teacher Switches as a Quasi-Experiment** As an additional test for the unbiasedness of GL estimates we exploit teacher switches across grades as a quasi-experiment, as in Chetty et al. (2014). If VA is an unbiased measure of teacher quality, changes in average VA of teachers in a given school and grade (driven by teacher switches) should predict changes in average student test score residuals one-by-one. To understand the rationale behind this test suppose that, in a given school with three 4th-grade classrooms (and hence three 4th-grade math teachers), one of these teachers leaves and is replaced by a teacher with a 0.3 higher VA (measured in standard deviations of test scores). If VA is an unbiased measure of teacher effectiveness, test scores should raise by 0.3/3 = 0.1 standard deviations due to this switch (Chetty et al., 2014).

We estimate the degree of forecast bias for the Wisconsin GL measures by estimating the following first-differences equation (we restrict attention to the years 2007 to 2011 to parse out any changes in teacher effort, as done in the paper):

$$\Delta A^*_{gst} = a + b\Delta Q_{gst} + \Delta \chi_{gst} \tag{22}$$

where $A^*_{gst}$ are test score residuals of students in grade $g$, school $s$, and year $t$, $Q_{gst}$ is average

Figure BIII: Binned scatterplot: $\hat{\mu}_i^{CL}$ and $\hat{\mu}_i^{GL}$




*Notes*: The figure shows the relationship between $\hat{\mu}_i^{CL}$, estimate of teacher VA obtained using the procedure of Kane and Staiger (2008) and teacher-student links, and $\hat{\mu}_{GL}$, its analogous obtained discarding these links. Estimates are obtained using data from New York City students and teachers of math and ELA for the years 2007-2010.

teacher VA, and $\Delta W_{gst} = W_{gst} - W_{gst-1}$ for any variable $W_{gst}$. The forecast bias is defined as $\lambda = 1 - b$. Table BIII shows estimates of $b$ and $\lambda$, obtained using either mean residual test scores or mean actual test scores, and controlling for school-by-year fixed effects (as in Chetty et al., 2014).[46] Estimates of $b$ are all close to 1 both over the full sample period and in the years after Act 10. While slightly larger than Chetty et al. (2014), who estimate it to be between 0.003 and 0.026, estimates of $b$ are small and indistinguishable from zero, both over the full sample period and in the years after Act 10.

**Non-Classical Measurement Error** A possible concern with the GL version of VA is non-classical measurement error, which occurs when the precision of the estimates is related to characteristics of the teachers or the students. This issue could arise, for example, if teachers who switch across schools or grades (and, analogously, the grades and schools employing these teachers) are selected on the basis of observable and/or unobservable characteristics.

In Table BII we use the GL and CL estimates of VA from the NYC data to investigate the extent of measurement error. Specifically, we correlate the difference between GL and CL (a proxy for measurement error) with a range of student and teacher observable characteristics. These estimates reveal no discernible relationship between the error and these characteristics, with the exception of the share of special education students. Importantly, the measurement error does not appear to be systematically different between teachers who switch across grades (i.e., those with "switcher" equal to 1) and teachers who do not switch. While only suggestive of the lack of non-classical measurement error, this evidence reassuringly shows no systematic patterns of correlations between VA and student and teacher observables.

[46] The fact that using test scores as a regressor instead of test score residuals yields similar results further confirms that selection of students across teachers is unlikely to generate substantial bias in the estimates (Chetty et al., 2014).

Table BI: Forecast bias in teacher VA

| | Δ test scores | Δ test score residuals |
|---|---|---|
| | (1) | (2) |
| $\Delta VA_{gst}$ | 0.978 | 1.055 |
| | (0.290) | (0.377) |
| School-by-year FE | Yes | Yes |
| Observations | 13684 | 13684 |
| # districts | 414 | 414 |
| $\lambda$ | 0.022 | -0.055 |
| p-value $\lambda$=0 | 0.94 | 0.88 |

*Notes*: The dependent variable is the first difference in grade-school average test score residuals (from a regression of test scores on student characteristics, school, and grade fixed effects, column 1) or in average test scores at the grade, school, and year level (column 2). The variable $\Delta VA_{gst}$ is the first difference in average teacher VA in school $s$ and grade $g$. VA is calculated using data from Wisconsin for the years 2007-2011. All regressions include school-by-year fixed effects, and observations are weighted by the number of students. Standard errors in parentheses are clustered at the district level.

Table BII: Correlations Between the Difference [GL-CL ] and Student and Teacher Observables

| | (1) |
|---|---|
| experience | -0.0003 |
| | (0.0002) |
| switcher | 0.0013 |
| | (0.0024) |
| Black | -0.0014 |
| | (0.0026) |
| Hispanic | 0.0033 |
| | (0.0029) |
| % low SES students | -0.0042 |
| | (0.0028) |
| % Black students | 0.0052 |
| | (0.0035) |
| % Hispanic students | 0.0009 |
| | (0.0037) |
| % special Ed students | -0.0060 |
| | (0.0107) |
| % disabled students | -0.0414*** |
| | (0.0103) |
| Observations | 8077 |

*Notes*: OLS regression of the difference between GL and CL and a range of student and teacher characteristics, averaged at the teacher-year level. VA is calculated using data from NYC. Robust standard errors in parentheses.

# Appendix C Survey Details

## Survey Questionnaire

### General Questions

1. What is your age? (select one)
   - less than 25
   - 25-30
   - 31-35
   - 36-40
   - 41-45
   - 46-50
   - 51-55
   - over 55
2. What is your gender?
   - Male
   - Female
   - Other
3. Did you work in another industry before teaching in public schools?
   - Yes
   - No
4. Did you work in another industry before teaching in public schools?
   - Yes
   - No
5. Which industry did you work in before teaching in public schools?
   - Other job in public sector
   - Other job in private education
   - Other job in different sector

### Negotiation

6. Have you ever negotiated your pay with any of your past employers?
   - Yes, successfully
   - Yes, unsuccessfully
   - No, it was not a possibility
   - No, it was a possibility but I chose not to
   - No, it was a possibility but I did not feel I could negotiate without repercussions
7. When you started your current job, did you negotiate your pay?
   - Yes, successfully

- Yes, unsuccessfully
- No

8. Why didn't you negotiate your pay? [choose all that apply]
    - It was not a possibility
    - I would not have gotten anything out of it I was worried about backlash
    - I didn't feel comfortable negotiating
    - I was satisfied with my offered salary
    - I did not know that I could negotiate
9. Since starting your current job, have you ever asked for a pay increase?
    - Yes, successfully
    - Yes, unsuccessfully
    - No
10. Why haven't you asked for a pay increase? [choose all that apply]
    - I would not have gotten anything out of it It is not a possibility
    - I am worried about backlash
    - I don't feel comfortable asking
    - I am satisfied with my salary
11. How likely is it that you will negotiate any of the following in the future? [for each item, choose a number from 1 (not at all likely) and 10 (very likely))
    - Salary
    - Classroom assignment
    - Non-teaching duties
12. Do you know what your colleagues earn?
    - Yes
    - Only some of them
    - No
13. Do you know any public sector teachers who have negotiated their salary?
    - Yes, among my colleagues
    - Yes, outside of my colleagues
    - Yes, both among and outside of my colleagues
    - No
14. How would you rate your performance relative to your colleagues' performance?
    - Below average
    - Average
    - Above average
15. Are you confident about talking to people you don't know?

- Yes
- No

**Please state whether you agree or disagree with the following statements.**

16. I pick up the subtle signals of feelings from another person.
    - Agree
    - Disagree
17. I am astute at reading people's reactions and feelings.
    - Agree
    - Disagree
18. I have good people skills.
    - Agree
    - Disagree

Figure CI: Survey Email

**From:** Heather Sarsons

**To:** [TEACHER'S EMAIL]

**Subject**: A short survey for a Yale and UChicago study

Good evening,

We are a team of researchers at The University of Chicago and Yale University, and we are conducting a research study on public sector employees' perceptions about their jobs. As part of this study, we would like to ask you to fill in a **very short survey (length < 5 mins).** This survey is confidential, completely anonymous, and has been approved by the Institution Review Boards at Yale and the University of Chicago. Your participation is invaluable for our research.

If you would like to take the survey, please click here:

**Follow this link to the Survey:**
[LINK]

Or copy and paste the URL below into your internet browser:
[URL]

We sincerely appreciate your time and participation, and please feel free to contact us with any questions. Thank you!

Best regards,

Barbara Biasi
(email: barbara.biasi@yale.edu, website: www.barbarabiasi.com )

Heather Sarsons
(email: heather.sarsons@chicagobooth.edu, website: sites.google.com/view/sarsons/)

Follow the link to opt out of future emails:
(click here to unsubscribe}