The Labor Market for Teachers Under Different Pay Schemes

Barbara Biasi

Online Appendix

Appendix A Additional Tables and Figures

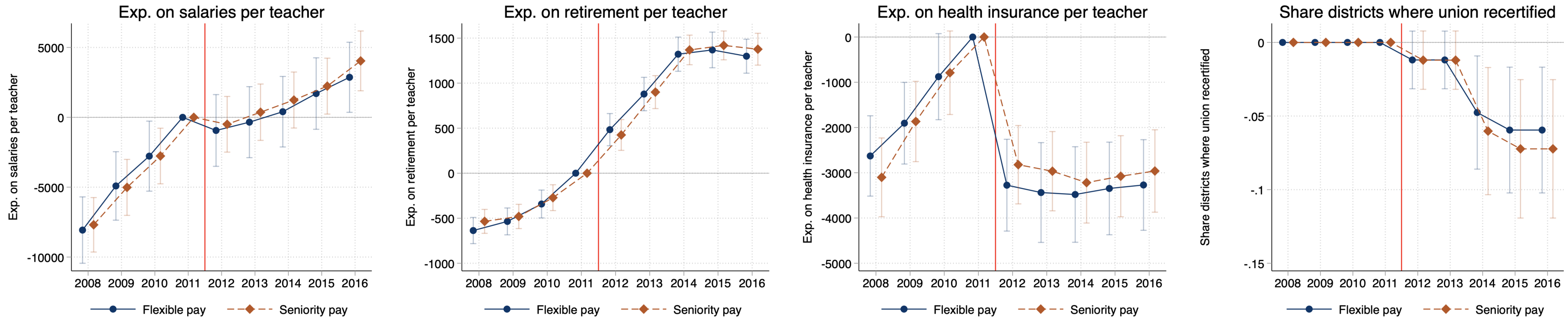
Figure A1: Salary Schedule - Racine School District, 2015

Step	BA	BA+12	BA+24	MA
1	40,593	42,784	44,976	47,169
2	41,526	43,717	45,909	48,516
3	42,459	44,651	46,842	49,864
4	43,392	45,584	47,775	51,211
5	44,325	46,517	48,709	52,560

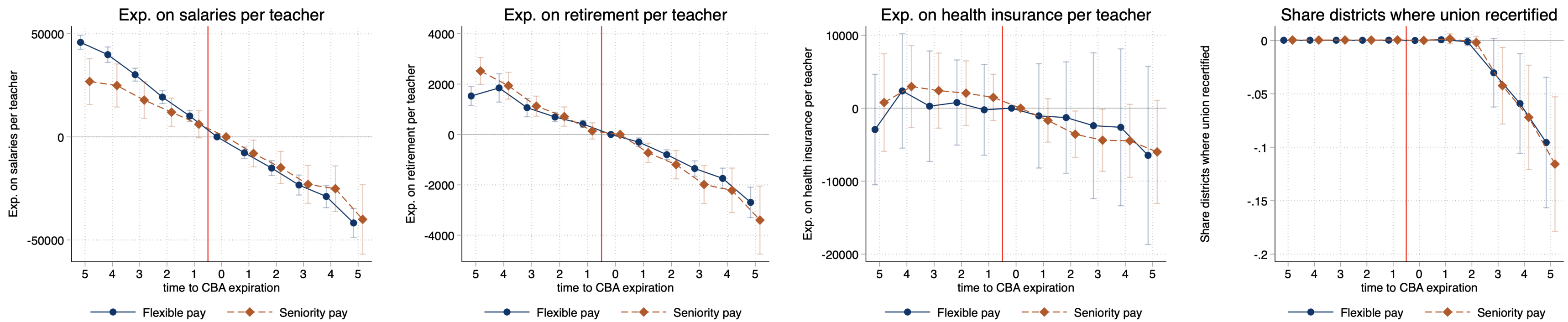
Notes: Portion of the salary schedule used in the school district of Racine in 2015. Source: http://www.rusd.org.

Figure A2: Districts' Characteristics: By Year (top panel) and Time-to-CBA (bottom panel)

Panel a) By year - FP and SP districts



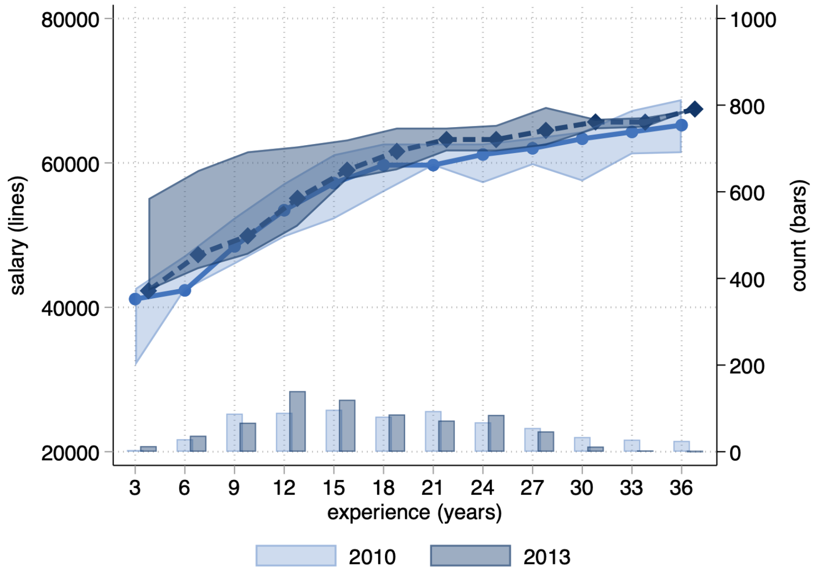
Panel b) By time-to-CBA, controlling for year fixed effects - FP and SP districts



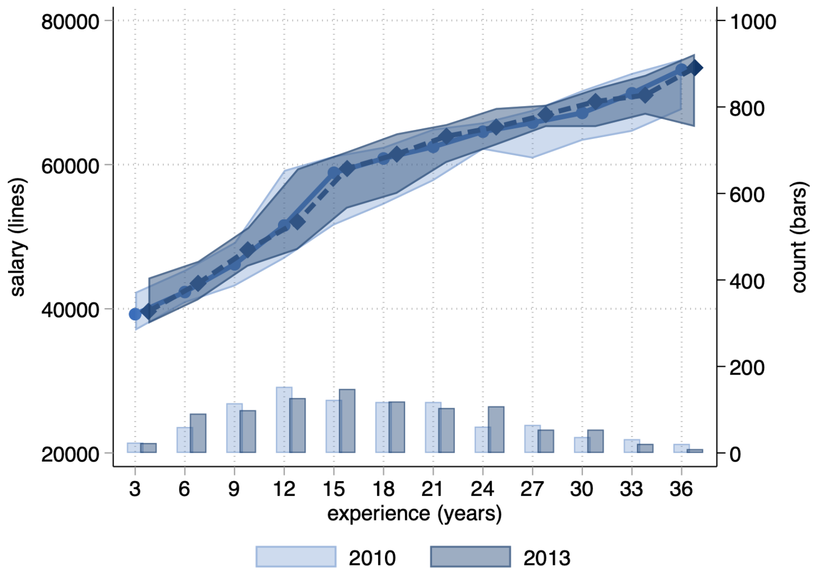
Notes: Panel a) shows trends in district characteristics between 2008 and 2016, relative to 2011, in FP and SP districts. Panel b) shows estimates of event studies of the same characteristics in a ten-years window around a CBA expiration, controlling for year fixed effects. Characteristics include, from left to right: Expenditure on salaries per teacher, expenditure on retirement benefits per teacher, expenditure on health insurance benefits per teacher, and an indicator for whether the union managed to recertify. Standard errors are clustered at the district level.

Figure A3: Empirical Salary Schedule - Median and 90-10 Percentile Range of Salaries, 2010 and 2013, School Districts of Green Bay and Madison

Panel A) Green Bay Area School District (FP)

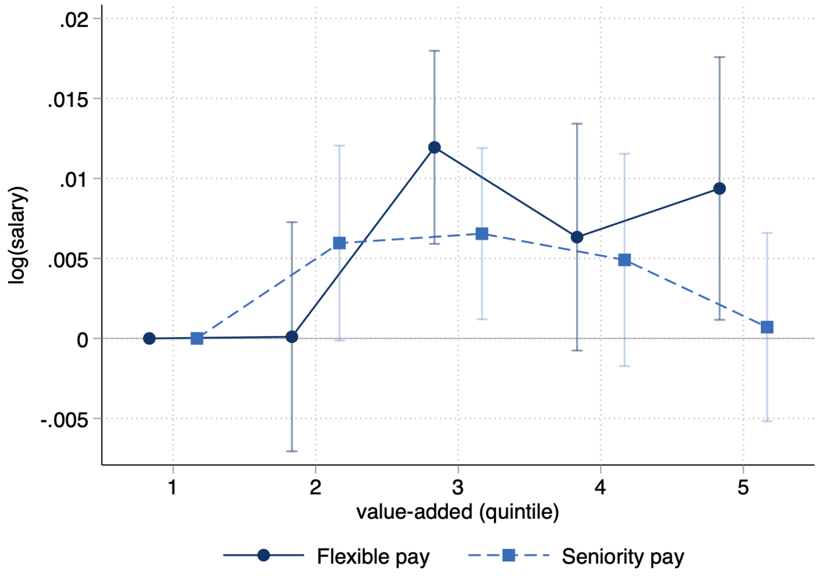


Panel B) Madison School District (SP)



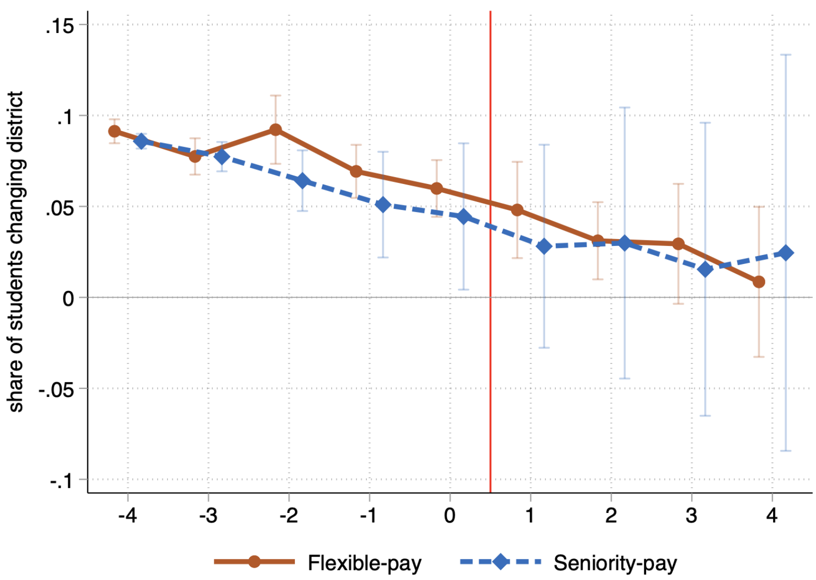
Notes: The figure shows the median and the 90-10 percentile range of salaries, by three-year experience classes, for teachers in the school districts of Green Bay (a flexible-pay district, panel A) and Madison (a seniority-pay district, panel B), for the years 2010 (lighter line and area) and 2013 (darker line and area). The bars correspond to counts of teachers in each experience bin. The sample is restricted to full-time teachers with a master's degree.

Figure A4: Salaries, by Quintile of Value-Added: FP and SP districts



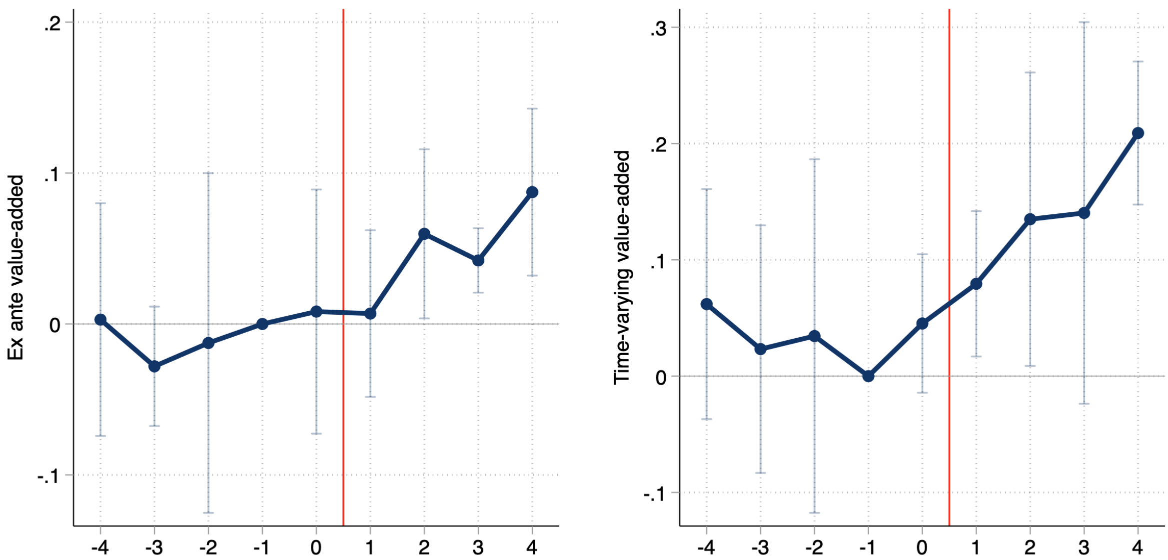
Notes: OLS estimates and 90% confidence intervals of the coefficients δ_s in the regression $log(w_{it}) = \sum_{s=1}^{5} \delta_s^0 \mathbb{1}(D(VA_{it}) = s) + \sum_{s=1}^{5} \delta_s \mathbb{1}(D(VA_{it}) = s) * \mathbb{1}(t > Exp_{j(it)}) + \beta X_{it}^w + \theta_{j(jt)t} + \varepsilon_{it}$. The variable $log(w_{it})$ is the natural logarithm of salary for teacher i in year t. The variable VA_{it} is teacher VA. The function $D(VA_{it})$ denotes the quintile in the distribution of value added, the variable Exp_j is the year of expiration of district j's CBA, and $\mathbb{1}(.)$ is an indicator function. The vector X_{it}^w includes a non-parametric function of years of experience, interacted with indicators for the highest education degree and with a dummy for years after 2011. The vector θ_{jt} contains district-by-year fixed effects. The coefficients δ_s are estimated separately for FP and SP districts. Bootstrapped standard errors are clustered at the district level.

Figure A5: Share of Students Changing District, by Time-to-CBA Expiration: FP and SP Districts



Notes: Fraction of students who change district in each year, by type of district and time-to-CBA expiration, controlling for year fixed effects.

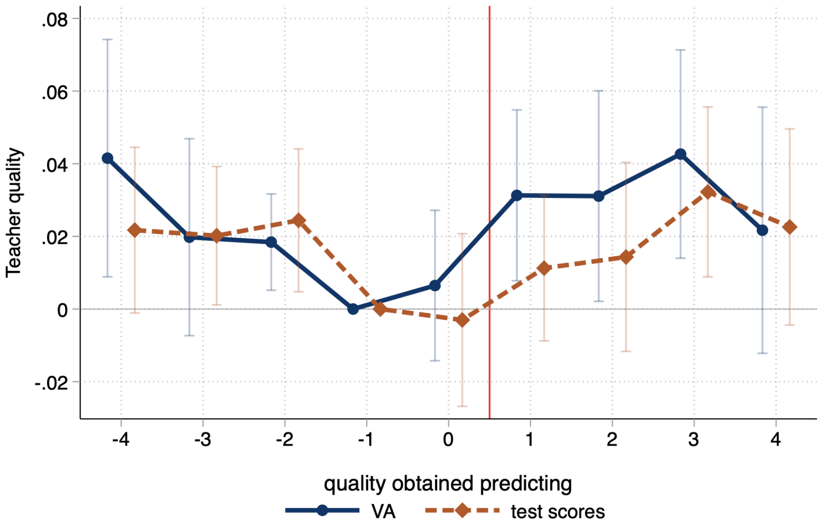
Figure A6: Changes in Workforce Composition and Effort Around a CBA Expiration. Sample Restricted to Districts With CBA Expirations in 2012 and 2013



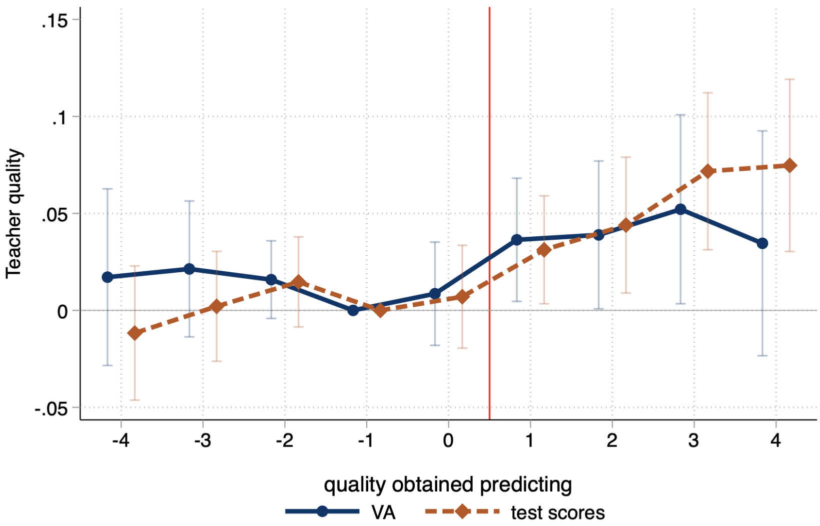
Notes: Point estimates and confidence intervals of the parameters β_k in the equation $VA_{it} = \sum_{-4}^{4} \beta_k FP_{j(it)} * \mathbb{1}(t - Exp_{j(it)} = 4) + \theta_{j(it)} + \tau_t + \varepsilon_{ijt}$, where VA_{it} is either (a) ex ante teacher VA (left panel) or (b) time-varying VA (right panel); Exp_j is the year of expiration of district j's CBA; and the vectors θ_j and τ_t contain district and year fixed effects, respectively. Ex ante VA is calculated using test score data for the years 2007–2011. Time-varying VA is calculated separately for each teacher using using test score data for the years 2007–2011 and 2012–2016. Standard errors are clustered at the district level.

Figure A7: Changes in Teacher Quality, With Quality Predicted Using Teacher Observables

Panel A) Baseline



Panel A) Controlling for district-specific pre-trends



Notes: Point estimates and confidence intervals of the parameters β_k in the equation $Y_i = \sum_{-4}^{4} \beta_k FP_{j(it)} * \mathbb{1}(t - Exp_{j(it)} = 4) + \theta_{j(it)} + \tau_t + \varepsilon_{ijt}$, where Y_i is a measure of teacher quality obtained predicting either VA (solid line) or conditional test scores (dashed line) using teacher observables in the NYC data and applying the corresponding estimates in the Wisconsin data. Exp_j is the year of expiration of district j's CBA; and the vectors θ_j and τ_t contain district and year fixed effects, respectively. In panel B, I remove district-specific pre-trends (estimated using data prior to each expiration) from each dependent variable. Standard errors are clustered at the district level.

Table A1: Expiration Dates of CBAs and Districts' Observable Characteristics. OLS, Dependent Variables are Indicators for Years of Expiration

	CBA expires in		
	(1)	(2)	(3)
	2011	2011	2011
enrollment	-0.0000	0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)
share econ. disadv. students	-0.0674	0.0062	0.0612
	(0.0942)	(0.0699)	(0.0641)
share Black students	-0.8868	0.8092	0.0777
	(0.7153)	(0.6750)	(0.2232)
share Hispanic students	0.0330	-0.0678	0.0349
	(0.2097)	(0.1648)	(0.1142)
teacher experience	-0.0018	0.0005	0.0012
	(0.0052)	(0.0040)	(0.0033)
share teachers w/ Master's	0.0044	-0.0471	0.0427
	(0.0593)	(0.0420)	(0.0449)
teacher value-added	0.0739	-0.1094	0.0355
	(0.0783)	(0.0714)	(0.0370)
urban district	-0.0055	-0.1142***	0.1197
	(0.0903)	(0.0340)	(0.0933)
suburban district	0.0010	-0.0227	0.0217
	(0.0410)	(0.0279)	(0.0309)
N	208	208	208
R^2	0.21	0.22	0.09

Notes: The dependent variables are indicators for districts' CBA agreements expiring in 2011, 2012, and 2013 (columns 1, 2, and 3 respectively). The explanatory variables are measured in 2010 and averaged at the district level (one observation is a district). Robust standard errors in parentheses.

Table A2: Summary Statistics, Districts with and without Handbook Information

	without Handbook	with Handbook	Difference
enrollment	288.8	1089.6	-800.8***
			(147.9)
share Black students	0.014	0.033	-0.018***
			(0.0054)
share Hispanic students	0.031	0.049	-0.018***
			(0.0050)
share econ. disadv. students	0.38	0.33	0.051***
			(0.016)
Math test scores	-0.0041	0.060	-0.064**
			(0.026)
teacher salary ($)	46633.0	50417.7	-3784.7***
			(510.7)
teacher experience (yrs)	15.8	15.3	0.54**
			(0.23)
teachers w/ BA	0.59	0.48	0.11***
			(0.016)
teachers w/ MA	0.41	0.51	-0.11***
			(0.016)
teachers w/ PhD	0.00047	0.0012	-0.00071**
			(0.00033)
urban district	0	0.071	-0.071***
			(0.018)
suburban district	0.050	0.24	-0.19***
			(0.034)
value-added	-0.015	-0.011	-0.0044
			(0.015)
expenditure p.p. ($)	45055.8	42315.3	2740.5
			(1786.1)
state aid/expenditure (share)	0.32	0.34	-0.016
			(0.014)

Notes: Means, difference in means, and standard deviations of the difference in means (in parentheses) of district-level characteristics for 102 FP and 122 SP districts with non-missing handbook information, and 203 districts with missing handbook information, for the years 2009–2011.

Table A3: Summary Statistics, Wisconsin Teachers

	(1) Full sample		(2) FP/SP		(3) FP/SP with CBA exp. date	
	2007-11	2012-15	2007-11	2012-15	2007-11	2012-15
female	0.733	0.741	0.736	0.744	0.740	0.748
	(0.443)	(0.438)	(0.441)	(0.436)	(0.438)	(0.434)
experience (years)	14.46	13.94	14.31	13.78	14.20	13.74
	(9.769)	(9.184)	(9.647)	(9.048)	(9.575)	(8.999)
highest ed = BA	0.497	0.467	0.482	0.452	0.477	0.449
	(0.500)	(0.499)	(0.500)	(0.498)	(0.499)	(0.497)
highest ed = Master	0.496	0.525	0.510	0.539	0.514	0.542
	(0.500)	(0.499)	(0.500)	(0.498)	(0.500)	(0.498)
highest ed = PhD	0.00187	0.00205	0.00214	0.00239	0.00224	0.00251
	(0.0432)	(0.0452)	(0.0462)	(0.0488)	(0.0473)	(0.0500)
salary ($)	50341.1	53878.4	51179.3	54488.4	51522.5	54838.8
	(11545.0)	(12351.5)	(11692.2)	(12402.7)	(11728.8)	(12424.9)
mover	0.0151	0.0300	0.0141	0.0299	0.0135	0.0291
	(0.122)	(0.171)	(0.118)	(0.170)	(0.115)	(0.168)
value-added	-0.000392	-0.000311	-0.000475	-0.000372	-0.000192	-0.000592
	(0.0744)	(0.0487)	(0.0771)	(0.0507)	(0.0790)	(0.0515)

Notes: Means and standard deviations (in parentheses) of teachers' observable characteristics for the years 2007–2011 and 2012–2015, for all Wisconsin districts (columns 1), for FP and SP districts (columns 2), and for FP and SP districts with non-missing CBA expiration information (columns 3). The sample only includes teachers for whom VA estimates are available.

Table A4: Summary Statistics, Wisconsin Teachers, With and Without *ex ante* Value-Added

	Full sample			FP/SP		
	w/ *ex ante* VA	w/out *ex ante* VA	Difference	w/ *ex ante* VA	w/out *ex ante* VA	Difference
female	0.79	0.72	0.074***	0.80	0.72	0.078***
			(0.0013)			(0.0015)
experience (years)	14.8	14.1	0.71***	14.5	13.9	0.60***
			(0.029)			(0.031)
highest ed = BA	0.42	0.50	-0.076***	0.41	0.48	-0.071***
			(0.0015)			(0.0017)
highest ed = Master	0.57	0.49	0.076***	0.58	0.51	0.071***
			(0.0015)			(0.0017)
highest ed = PhD	0.0012	0.0022	-0.00095***	0.0014	0.0025	-0.0011***
			(0.00013)			(0.00016)
salary ($)	53090.5	51570.3	1520.2***	53771.9	52320.5	1451.4***
			(36.2)			(40.8)
mover	0.017	0.024	-0.0076***	0.016	0.023	-0.0070***
			(0.00048)			(0.00053)
value-added	-0.00040	-0.00012	-0.00029	-0.00050	-0.000084	-0.00042
			(0.00045)			(0.00052)

Notes: Means and standard deviations (in parentheses) of teachers' observable characteristics for the years 2007–2011, for all Wisconsin districts (columns 1), for FP and SP districts (columns 2), and for FP and SP districts with non-missing CBA expiration information (columns 3), and separately for teachers with and without *ex ante* VA, calculated as the average over the years 2007–2011. The sample includes teachers for whom VA estimates are available.

Table A5: Teacher Salaries and Value-Added. OLS, Dependent Variable is log(Salary). Sample of Tenured Teachers

	All	FP	SP	Difference
	(1)	(2)	(3)	(4)
VA	0.001	-0.000	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
VA × post-CBA expiration	0.001	0.004**	-0.001	-0.001
	(0.001)	(0.002)	(0.002)	(0.002)
VA × FP				-0.001
				(0.001)
VA × FP × post-CBA expiration				0.005**
				(0.002)
District × year FE	Yes	Yes	Yes	Yes
Edu, exp × post-2011	Yes	Yes	Yes	Yes
N	91205	36336	47680	84018
# districts	210	74	89	163

Notes: The dependent variable is the natural logarithm of salaries. The variable *VA* is teacher VA, normalized to have mean 0 and standard deviation 1. The variable *post-CBA expiration* equals 1 for years after the expiration of each district's CBA. All the specifications include district-by-year fixed effects, as well interactions between indicators for years of experience, indicators for the highest education degree, and an indicator for years after 2011. VA is calculated separately for the years 2007–2011 and 2012–2016. The sample is restricted to tenured teachers (with at least three years of experience). Bootstrapped standard errors in parentheses are clustered at the district level. *** $p < .01$, ** $p < .05$, * $p < .1$.

Table A6: Teacher Salaries and Value-Added. OLS, Dependent Variable is log(Salary). Excluding Milwaukee and Madison

	All	FP	SP	Difference
	(1)	(2)	(3)	(4)
VA	0.001	-0.000	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
VA × post-CBA expiration	0.001	0.004**	-0.001	-0.001
	(0.001)	(0.002)	(0.002)	(0.002)
VA × FP				-0.001
				(0.002)
VA × FP × post-CBA expiration				0.005**
				(0.003)
District × year FE	Yes	Yes	Yes	Yes
Edu, exp × post-2011	Yes	Yes	Yes	Yes
N	85799	40142	37804	77946
# districts	209	74	88	162

Notes: The dependent variable is the natural logarithm of salaries. The variable *VA* is teacher VA, normalized to have mean 0 and standard deviation 1. The variable *post-CBA expiration* equals 1 for years after the expiration of each district's CBA. All the specifications include district-by-year fixed effects, as well interactions between indicators for years of experience, indicators for the highest education degree, and an indicator for years after 2011. VA is calculated separately for the years 2007–2011 and 2012–2016. The sample excludes teachers in the school districts of Milwaukee and Madison. Bootstrapped standard errors in parentheses are clustered at the district level. *** $p < .01$, ** $p < .05$, * $p < .1$.

Table A7: Teacher Salaries and Value-Added. OLS, Dependent Variable is log(Salary)

	Schools w/ max 3 teachers/grade			Teachers whose VA is identified		
	(1)	(2)	(3)	(4)	(5)	(6)
VA	-0.0006	0.0007	0.0007	-0.0005	0.0009	0.0010
	(0.0013)	(0.0012)	(0.0015)	(0.0017)	(0.0009)	(0.0009)
VA × post-CBA expiration	0.0063**	0.0006	0.0005	0.0025	0.0002	0.0001
	(0.0026)	(0.0021)	(0.0024)	(0.0024)	(0.0013)	(0.0013)
VA × FP			-0.0013			-0.0014
			(0.0023)			(0.0018)
VA × FP × post-CBA expiration			0.0061			0.0026
			(0.0045)			(0.0025)
District × year FE	Yes	Yes	Yes	Yes	Yes	Yes
Edu, exp × post-2011	Yes	Yes	Yes	Yes	Yes	Yes
N	18131	24078	42217	18360	25656	44029
# districts	73	89	162	73	89	162

Notes: The dependent variable is the natural logarithm of salaries. The variable *VA* is teacher VA, normalized to have mean 0 and standard deviation 1. The variable *post-CBA expiration* equals 1 for years after the expiration of each district's CBA. All the specifications include district-by-year fixed effects, as well interactions between indicators for years of experience, indicators for the highest education degree, and an indicator for years after 2011. VA is calculated separately for the years 2007–2011 and 2012–2016. The sample is restricted to teachers in schools and grades with at most 3 teachers per subject (columns 1-3) and to teachers with identified VA (columns 4-6). VA is calculated separately for the years 2007–2011 and 2012–2016. Bootstrapped standard errors in parentheses are clustered at the district level. *** $p < .01$, ** $p < .05$, * $p < .1$.

Table A8: Teacher Salaries and Value-Added. OLS, Dependent Variable is log(Salary). Controlling for Teaching Assignment

	All districts	FP	SP	Difference
	(1)	(2)	(3)	(4)
VA	0.001	-0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
VA × post-CBA expiration	0.001	0.004**	-0.001	-0.000
	(0.001)	(0.002)	(0.001)	(0.001)
VA × FP				-0.001
				(0.001)
VA × FP × post-CBA expiration				0.004*
				(0.002)
District × year FE	Yes	Yes	Yes	Yes
Edu, exp × post-2011	Yes	Yes	Yes	Yes
Grade, subject × post-2011	Yes	Yes	Yes	Yes
N	100750	40130	52745	92887
# districts	211	74	90	164

Notes: The dependent variable is the natural logarithm of salaries. The variable *VA* is teacher VA, normalized to have mean 0 and standard deviation 1. The variable *post-CBA expiration* equals 1 for years after the expiration of each district's CBA. All the specifications include district-by-year fixed effects, as well interactions between indicators for years of experience, indicators for the highest education degree, and an indicator for years after 2011, and controls for subject and grade fixed effects interacted with an indicator for years after 2011. VA is calculated separately for the years 2007–2011 and 2012–2016. Bootstrapped standard errors in parentheses are clustered at the district level. *** $p < .01$, ** $p < .05$, * $p < .1$.

Table A9: Changes in the Composition of the Teaching Workforce and in Effort, by Year of CBA Expiration. OLS, Dependent Variable is Teacher Value-Added

	Ex ante VA				Time-varying VA			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	2011	2011	2012-13	2012-13	2011	2011	2012-13	2012-13
FP	-0.023		-0.009		-0.019		-0.007	
	(0.025)		(0.068)		(0.020)		(0.056)	
			[-0.134]				[-0.134]	
FP × post-CBA expiration	0.027**	0.024*	0.020	0.050*	0.057	0.053	0.054	0.080
	(0.013)	(0.013)	(0.011)	(0.021)	(0.043)	(0.043)	(0.068)	(0.064)
			[1.775]	[2.379]			[0.793]	[1.252]
post-CBA expiration			0.011	0.031			0.185*	0.113
			(0.029)	(0.026)			(0.078)	(0.061)
			[0.370]	[1.200]			[2.366]	[1.855]
FP × post-2011			-0.017	-0.033			0.010	-0.017
			(0.042)	(0.029)			(0.101)	(0.104)
			[-0.414]	[-1.135]			[0.096]	[-0.159]
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
District FE	No	Yes	No	Yes	No	Yes	No	Yes
N	71179	71179	18521	18521	74699	74699	19340	19340
# districts	157	157	7	7	157	157	7	7

Notes: The dependent variable is *ex ante* teacher VA (columns 1-4) or time-varying VA, allowed to vary before and after 2011 (columns 5-8). The variable *FP* equals 1 for FP districts. The variable *post-CBA expiration* equals one for years following each district's CBA expiration, and the variable *post-2011* equals one for years following 2011. All the specifications include year fixed effects; columns 2, 4, 6, and 8 also include district fixed effects. Ex ante VA is calculated using test score data for the years 2007–2011. In columns 1, 2, 5, and 6 the sample is restricted to teachers in districts with agreements expiring in 2011. In columns 3, 4, 7, and 8, the sample is restricted to teachers in eight school districts with agreements expiring in 2012 and 2013; t-statistics obtained using a wild bootstrap and clustering at the district level are shown in brackets. Standard errors in parentheses are clustered at the district level. *** $p < .01$, ** $p < .05$, * $p < .1$.

Table A10: Changes in the Composition of the Teaching Workforce. OLS, Dependent Variable is Ex Ante Teacher Value-Added. Excluding Milwaukee and Madison

	(1)	(2)	(3)	(4)	(5)
FP	-0.024				
	(0.023)				
FP × post-CBA expiration	0.024**	0.021*	0.093***	0.034**	0.077**
	(0.012)	(0.013)	(0.021)	(0.014)	(0.034)
post-CBA expiration	-0.055**	-0.026	-0.038**	-0.016	-0.027
	(0.021)	(0.019)	(0.018)	(0.028)	(0.034)
FP × post-2011			-0.072***		-0.044
			(0.020)		(0.033)
Year FE	Yes	Yes	Yes	Yes	Yes
District FE	No	Yes	Yes	Yes	Yes
Edu, exp FE	Yes	Yes	Yes	Yes	Yes
Budget, CB controls	No	No	No	Yes	Yes
N	74229	74229	74229	61573	61573
# districts	162	162	162	159	159

Notes: The dependent variable is *ex ante* teacher VA. The variable *FP* equals 1 for FP districts. The variable *post-CBA expiration* equals one for years following each district's CBA expiration, and the variable *post-2011* equals one for years following 2011. All the specifications include year, year of experience, and higher education degree fixed effects. Columns 2-5 also include district fixed effects. *CB controls* include an indicator for whether the district had a union recertification election in year *t* and whether the election was successful. *Budget controls* are district-year-level controls for the level of state aid as a share of total revenues, as well as per-teacher expenditure on salaries, retirement, health, life, and other insurance, and other employee benefits. Ex ante VA is calculated using test score data for the years 2007–2011. The sample excludes teachers in the school districts of Milwaukee and Madison. Standard errors in parentheses are clustered at the district level. *** $p < .01$, ** $p < .05$, * $p < .1$.

Table A11: Changes in the Composition of the Teaching Workforce. OLS, Dependent Variable is Ex Ante Teacher Value-Added (Columns 1-2) or Time-Varying Value-Added (Columns 3-4)

	Ex ante VA		Time-varying VA	
	(1)	(2)	(3)	(4)
	3 teachers max	T's w/identified VA	3 teachers max	T's w/identified VA
FP × post-CBA expiration	0.067**	0.196***	0.170***	0.198***
	(0.029)	(0.051)	(0.035)	(0.048)
post-CBA expiration	-0.045*	-0.012	-0.036	0.003
	(0.027)	(0.029)	(0.033)	(0.040)
FP × post-2011	-0.070**	-0.180***	-0.129***	-0.130***
	(0.031)	(0.033)	(0.030)	(0.048)
Year FE	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes
Edu, exp FE	Yes	Yes	Yes	Yes
N	39695	36193	42918	44640
# districts	162	162	162	162

Notes: The dependent variable is *ex ante* teacher VA (columns 1 and 2) and time-varying VA (columns 3 and 4). The variable *FP* equals 1 for FP districts. The variable *post-CBA expiration* equals one for years following each district's CBA expiration, and the variable *post-2011* equals one for years following 2011. All the specifications include district and year fixed effects and indicators for years of experience and for whether a teacher has a Master's or a PhD. Ex ante VA is calculated using test score data for the years 2007–2011. Standard errors in parentheses are clustered at the district level. The sample is restricted to teachers in schools and grades with at most 3 teachers per subject (columns 1 and 3) and to teachers with identified VA (columns 2 and 4). *** $p < .01$, ** $p < .05$, * $p < .1$.

Table A12: Combined Changes in Teacher Composition and Effort. OLS, Dependent Variable is Teacher Value-Added. Excluding Milwaukee and Madison

	(1)	(2)	(3)	(4)	(5)
FP	-0.020				
	(0.019)				
FP × post-CBA expiration	0.062	0.058	0.211***	0.070	0.082
	(0.041)	(0.040)	(0.041)	(0.047)	(0.054)
post-CBA expiration	-0.009	-0.038	-0.061***	-0.027	-0.049
	(0.034)	(0.035)	(0.021)	(0.059)	(0.036)
FP × post-2011			-0.154***		
			(0.035)		
Year FE	Yes	Yes	Yes	Yes	Yes
District FE	No	Yes	Yes	Yes	Yes
Edu, exp FE	Yes	Yes	No	Yes	Yes
Budget, CB controls	No	No	No	Yes	Yes
Teacher FE	No	No	No	No	Yes
N	78050	78050	78050	64934	76412
# districts	162	162	162	159	162

Notes: The dependent variable is teacher VA, allowed to vary before and after 2011. The variable *FP* equals 1 for FP districts. The variable *post-CBA expiration* equals one for years following each district's CBA expiration, and the variable *post-2011* equals one for years following 2011. All the specifications include year, year of experience, and higher education degree fixed effects. Columns 2-6 include district fixed effects and column 6 controls for teacher fixed effects. *CB controls* include an indicator for whether the district had a union recertification election in year t and whether the election was successful. *Budget controls* are district-year-level controls for the level of state aid as a share of total revenues, as well as per-teacher expenditure on salaries, retirement, health, life, and other insurance, and other employee benefits. VA is calculated separately for the years 2007–2011 and 2012–2016. The sample excludes teachers in the school districts of Milwaukee and Madison. Standard errors in parentheses are clustered at the district level. *** $p < .01$, ** $p < .05$, * $p < .1$.

Table A13: Changes in Student Characteristics, Movers to FP and SP

	Econ disadv	Female	Black	Hispanic	Disabled
	(1)	(2)	(3)	(4)	(5)
FP	-0.059	-0.005**	-0.033	-0.003	-0.001
	(0.040)	(0.002)	(0.026)	(0.019)	(0.006)
FP × post-CBA expiration	0.016	0.005	0.001	0.017*	-0.008
	(0.022)	(0.003)	(0.015)	(0.009)	(0.006)
post-CBA expiration	-0.169***	-0.000	-0.101**	-0.107***	-0.009
	(0.034)	(0.006)	(0.045)	(0.027)	(0.008)
Year FE	Yes	Yes	Yes	Yes	Yes
Mean of Dep. Var.	0.36	0.49	0.07	0.09	0.13
N	4920	4920	4920	4920	4920
# districts	162	162	162	162	162

Notes: The dependent variables are average characteristics of students of the teachers who move to FP and SP districts. The variable *FP* equals 1 for FP districts, the variable *post-CBA expiration* equals 1 for years after a CBA expiration. All specifications contain year fixed effects. Standard errors in parentheses are clustered at the district level.

Table A14: Predicting VA and Test Scores Using Teacher Observables - NYC Data

	VA	Test scores
	(1)	(2)
Elementary	0.029∗ ∗ ∗	0.041∗ ∗ ∗
	(0.003)	(0.004)
Math	0.038∗ ∗ ∗	0.031∗ ∗ ∗
	(0.012)	(0.006)
Reading	0.037∗ ∗ ∗	0.033∗ ∗ ∗
	(0.005)	(0.005)
Experience:		
2 years	0.015∗ ∗ ∗	0.018
	(0.003)	(0.037)
3 years	0.016∗ ∗ ∗	-0.001
	(0.003)	(0.031)
4-5 years	0.019∗ ∗ ∗	0.027∗ ∗ ∗
	(0.003)	(0.004)
6-10 years	0.017∗ ∗ ∗	0.021∗ ∗ ∗
	(0.003)	(0.003)
11-15 years	0.017∗ ∗ ∗	0.016∗ ∗ ∗
	(0.004)	(0.005)
16-20 years	0.008	0.006
	(0.005)	(0.006)
>20 years	-0.014∗	-0.005
	(0.008)	(0.006)
Female	0.014∗ ∗ ∗	0.044∗ ∗ ∗
	(0.003)	(0.004)
Asian	0.009	0.014∗∗
	(0.007)	(0.007)
Black	-0.004	-0.023∗ ∗ ∗
	(0.003)	(0.004)
Hispanic	0.008∗∗	-0.000
	(0.003)	(0.005)
N	33114	2443049

Notes: The dependent variables are teacher VA (column 1) and student test scores in Math and Reading (column 2). Column 2 includes all controls included in the estimation of teacher VA. Estimates obtained using NYC data. Standard errors in parentheses are clustered at the teacher level.

Table A15: Changes in Teacher Quality, With Quality Predicted Using Teacher Observables

	VA		Test scores	
	(1)	(2)	(3)	(4)
FP	-0.048		0.001	
	(0.040)		(0.020)	
FP × post-CBA expiration	0.171***	0.085***	0.196***	0.059*
	(0.043)	(0.019)	(0.061)	(0.035)
post-CBA expiration		-0.006	-0.097***	-0.026
		(0.016)	(0.031)	(0.033)
FP × post	-0.138***	-0.060***	-0.178***	-0.049
	(0.044)	(0.016)	(0.061)	(0.038)
Year FE	Yes	Yes	Yes	Yes
District FE	No	Yes	No	Yes
N	399830	399830	399830	399830
# districts	164	164	164	164

Notes: The dependent variables are measures of teacher quality obtained predicting either VA (columns 1-2) or conditional test scores (columns 2-4) using teacher observables in the NYC data and using the corresponding estimates in the Wisconsin data. The variable *FP* equals 1 for FP districts. The variable *post-CBA expiration* equals one for years following each district's CBA expiration, and the variable *post-2011* equals one for years following 2011. All the specifications include year fixed effects; columns 2 and 4 also include district fixed effects. Standard errors in parentheses are clustered at the district level.

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):

$$A_{kt} = \beta X_{kt} + \nu_{kt} \tag{1}$$
$$\text{where } \nu_{kt} = \mu_{i(kt)} + \theta_{c(kt)} + \varepsilon_{kt}$$

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.[1] The residual ν_{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 ε_{kt} is an idiosyncratic student-specific component which varies over time. VA is an estimate of the teacher effect μ_i.

A range of techniques have been proposed to estimate μ_i, including fixed effects (Aaronson, Barrow and Sander, 2007) and two-steps procedures based on the decomposition of test score residuals (Kane and Staiger, 2008; Chetty, Friedman and Rockoff, 2014). Here, I use an estimator largely based on the two-steps estimator of Kane and Staiger (2008), a special case of the more general framework of Chetty, Friedman and Rockoff (2014) which allows for the correction of noise in the estimates using a Bayes "shrinkage" approach. The estimation procedure of Kane and Staiger (2008) can be summarized as follows:

1. Estimate β in equation (1) via OLS;

[1] This specification largely follows Chetty, Friedman and Rockoff (2014).

2. Construct residuals $\hat{\nu}_{kt} = A^*_{kt} - \hat{\beta} X_{kt}$, where $\hat{\beta}$ is the OLS estimate of β;

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 $h_{it} = \frac{n_{it}}{n_{it}\sigma_\theta^2 + \sigma_\varepsilon^2}$, are the weights, function of class size n_{it}, the variance of the classroom component σ_θ^2 and of the residual component σ_ε^2;

   (c) the variance of the teacher effect is $\sigma_\mu^2 = Cov(\bar{\nu}_{it}, \bar{\nu}_{it-1})$; the variance of the residual component is $\sigma_\varepsilon^2 = Var(\nu_{kt} - \bar{\nu}_{it})$; the variance of the classroom component is $\sigma_\theta^2 = Var(\nu_{kt}) - \sigma_\varepsilon^2 - \sigma_\mu^2$.

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 to teacher-student datasets from several other states and/or districts (Rivkin, Hanushek and Kain, 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, Hanushek and Kain (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 not only of the effects of the teacher who switches school or grade, but also of those teachers teaching in her same grade and school at any point in time.

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

a. I impose θ_c to be zero for all c;

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

c. I construct the "teams" of teachers in each grade and school in each year;

d. Given these teams, I 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 I 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.

e. Given these $\bar{\nu}_{it}$, I can calculate VA from step 3 above. This strategy does not allow to separately identify θ_c; I therefore assume θ_c and σ_θ 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, Hanushek and Kain, 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. I define A_{kt} as the average test score residual for students of teacher k in period t, and $\bar{A}^g_t$ the average test score residuals of students in grade g in period t. Following Chetty, Friedman and Rockoff (2014) I can write the VA estimate for each teacher as follows (I suppress the hats on the VA estimates for ease of notation and I consider 3 lags):

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

$$\mu_{B3} = \begin{bmatrix} A^2_{B1} & A_{B1}A_{B2} \\ A_{B1}A_{B2} & A^2_{B2} \end{bmatrix}^{-1} \begin{bmatrix} A_{B1}A_{B3} \\ A_{B2}A_{B3} \end{bmatrix} \tag{3}$$

$$\mu_{C3} = \begin{bmatrix} A^2_{C1} & A_{C1}A_{C2} \\ A_{C1}A_{C2} & A^2_{C2} \end{bmatrix}^{-1} \begin{bmatrix} A_{C1}A_{C3} \\ A_{C2}A_{C3} \end{bmatrix} \tag{4}$$

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

$$\bar{A}^4_1 = \frac{1}{2}(A_{A1} + A_{B1}) \tag{5}$$

$$\bar{A}^5_1 = A_{C2} \tag{6}$$

$$\bar{A}^4_2 = \frac{1}{2}(A_{B2} + A_{C2}) \tag{7}$$

$$\bar{A}^5_2 = A_{A2} \tag{8}$$

$$\bar{A}^4_3 = \frac{1}{2}(A_{A3} + A_{C3}) \tag{9}$$

$$\bar{A}^5_3 = A_{B3} \tag{10}$$

My VA estimator implies:

$$\mu_{A3} = \begin{bmatrix} (\bar{A}^4_1)^2 & \bar{A}^4_1\bar{A}^5_2 \\ \bar{A}^4_1\bar{A}^5_2 & (\bar{A}^5_2)^2 \end{bmatrix}^{-1} \begin{bmatrix} \bar{A}^4_1\bar{A}^4_3 \\ \bar{A}^5_2\bar{A}^4_3 \end{bmatrix} \tag{11}$$

$$\mu_{B3} = \begin{bmatrix} (\bar{A}^4_1)^2 & \bar{A}^4_1\bar{A}^4_2 \\ \bar{A}^4_1\bar{A}^4_2 & (\bar{A}^4_2)^2 \end{bmatrix}^{-1} \begin{bmatrix} \bar{A}^4_1\bar{A}^5_3 \\ \bar{A}^4_2\bar{A}^5_3 \end{bmatrix} \tag{12}$$

$$\mu_{C3} = \begin{bmatrix} (\bar{A}^5_1)^2 & \bar{A}^5_1\bar{A}^4_2 \\ \bar{A}^5_1\bar{A}^4_2 & (\bar{A}^4_2)^2 \end{bmatrix}^{-1} \begin{bmatrix} \bar{A}^5_1\bar{A}^4_3 \\ A_{C2}\bar{A}^4_3 \end{bmatrix} \tag{13}$$

Equations (2)-(13) represent a system of 12 equations in 12 unknowns: μ_{A3}, μ_{B3}, μ_{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 I call GL) against the standard Kane and Staiger estimator with classroom links (CL), I use teacher and student data from the New York City Department of Education (NYCDOE) from the years 2006-07 to 2009-10.[2] This dataset contains classroom, grade, and school identifiers, which allow me to estimate both CL and GL measures. I 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 B1 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 B2 shows the density of GL for Wisconsin teachers. Its standard deviation is equal to 0.10 for Math teachers.

Forecast Bias of GL as a Proxy for CL Next, I text whether GL is a forecast-unbiased estimate for CL. Figure B3 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 proxy for $\hat{\mu}_i^{CL}$ can be defined based on the best linear predictor of $\hat{\mu}_i^{CL}$ given $\hat{\mu}_i^{GL}$:

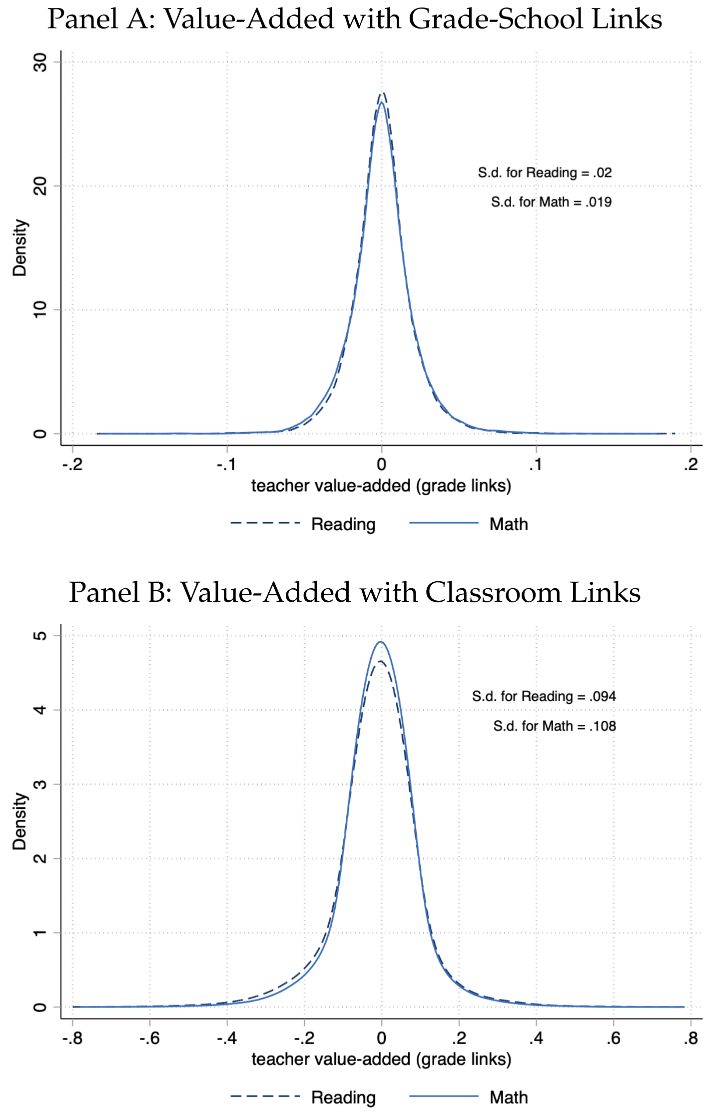
$$\hat{\mu}_i^{CL} = \alpha + \gamma\hat{\mu}_i^{GL} + \chi_i \tag{14}$$

Assuming χ_i to be uncorrelated with $\hat{\mu}_i^{GL}$, the forecast bias f is zero if $\gamma = 1$: $f = 1 - \gamma$. I can estimate the slope coefficient γ via OLS on equation (14). The 95% confidence interval for γ, 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 B3).

[2] NYCDOE (2015).

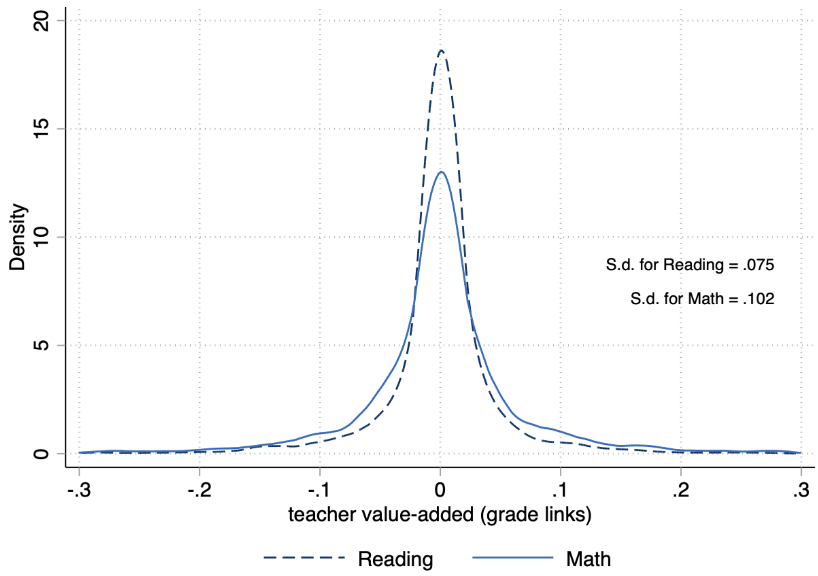
Teacher Switches as a Quasi-Experiment As an additional test for the unbiasedness of GL estimates I exploit teacher switches across grades as a quasi-experiment, as in Chetty, Friedman and Rockoff (2014). If VA is an unbiased measure of teacher quality, changes in average VA of teachers in a given school 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

Figure B1: Empirical Distribution of Value-Added Estimates: New York City, 2007-2010



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.

Figure B2: 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.

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, Friedman and Rockoff, 2014).

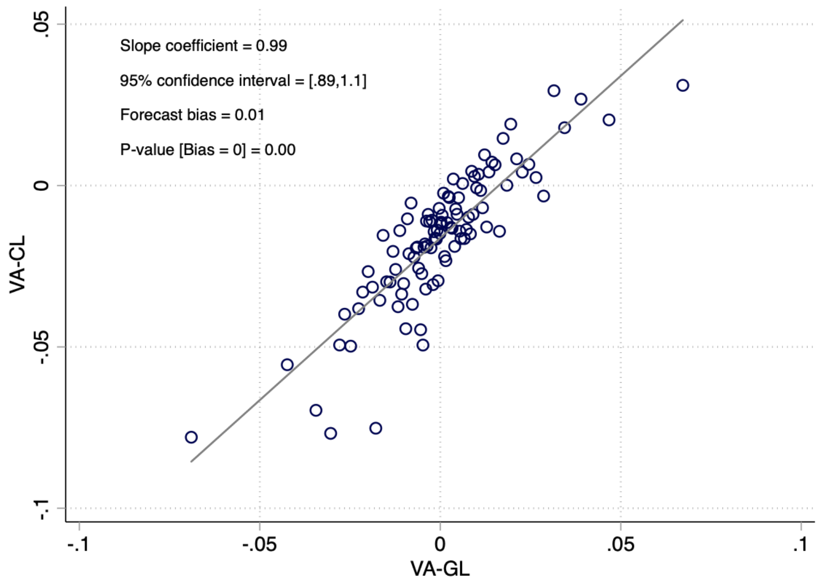
I estimate the degree of forecast bias for the Wisconsin GL measures by estimating the following first-differences equation (I restrict attention to the years 2007 to 2011 to parse out any changes in teacher effort, as done in the paper):

$$\Delta A^*_{gst} = a + b\Delta Q_{gst} + \Delta\chi_{gst} \tag{15}$$

where A^*_{gst} are test score residuals of students in grade g, school s, and year t, Q_{gst} is average 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 B3 shows estimates of b and λ, obtained using either mean residual test scores or mean actual test scores, and controlling for school-by-year fixed effects (as in Chetty, Friedman and Rockoff, 2014).[3] Estimates of b are all close to 1 both over the full sample period and in the

[3]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, Friedman and Rockoff, 2014).

Figure B3: 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.

years after Act 10. While slightly larger than Chetty, Friedman and Rockoff (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 B2 I use the GL and CL estimates of VA from the NYC data to investigate the extent of measurement error. Specifically, I 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. 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.

Table B1: Forecast bias in teacher VA

	Δ test scores	Δ test score residuals
	(1)	(2)
ΔVA_{gst}	0.978	1.055
	(0.290)	(0.377)
School-by-year FE	Yes	Yes
Observations	13684	13684
# districts	414	414
λ	0.022	-0.055
p-value λ=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 ΔVA_{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 B2: Correlations Between the Difference [GL-CL] and Student and Teacher Observables

	(1)
experience	-0.0002
	(0.0003)
switcher	0.0005
	(0.0022)
Black	-0.0019
	(0.0025)
Hispanic	0.0027
	(0.0028)
% low SES students	-0.0049
	(0.0030)
% Black students	0.0002
	(0.0046)
% Hispanic students	-0.0044
	(0.0051)
Observations	7032

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.

References

Aaronson, Daniel, Lisa Barrow, and William Sander. 2007. "Teachers and Student Achievement in the Chicago Public High Schools." *Journal of Labor Economics*, 25(1): 95–135.

Chetty, Raj, John N Friedman, and Jonah E Rockoff. 2014. "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates." *American Economic Review*, 104(9): 2593–2632.

Kane, Thomas J, and Douglas O Staiger. 2008. "Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation." NBER working paper n. 14607.

NYCDOE, New York City Department of Public Instruction. 2015. "*Student and Pedagogue Longitudinal Data*."

Rivkin, Steven G, Eric A Hanushek, and John F Kain. 2005. "Teachers, Schools, and Academic Achievement." *Econometrica*, 73(2): 417–458.