Oligopsony and Collective Bargaining *

Tirza J. Angerhofer † Allan Collard-Wexler ‡
Matthew C. Weinberg §

March 20, 2025

MOST RECENT VERSION

Abstract

Employers facing limited labor market competition may suppress wages below socially optimal levels. Unions can counteract this wage suppression, though they may also push wages above the socially optimal level. To assess these forces, we estimate a structural model of labor supply, labor demand, and Nash-in-Nash bargaining over wages between teacher unions and school districts in Pennsylvania's K-12 public school system from 2013 to 2020. Using the estimated parameters, we compare negotiated equilibrium wages and employment to the pure oligopsony scenario and the social planner scenario. On average, pure oligopsony reduces wages 7 percent below the social optimum, while collective bargaining raises them very close to the optimum. This average masks substantial district-level heterogeneity driven by variation in bargaining power. Thirty percent of schools have negotiated salaries below the social optimum due to cross-district externalities, where high salaries in one district reduce hiring and thereby increase labor supply in competing districts.

*We thank participants at Chicago IO+, UT-Austin, KU Leuven, UNC, and the Federal Trade Commission Microeconomics Conference for comments. This paper would not be possible without Hugh McCartney's contribution in sourcing the data and understanding the institutional background of teachers in Pennsylvania. We are grateful for his help. We have no financial interests to disclose.

†Duke University

‡Duke University and NBER

§Ohio State University

1 Introduction

A large and growing body of work measures employer monopsony power and considers its welfare implications on the labor market. An immediate implication of monopsony power is that, when firms face upward-sloping labor supply curves, wages and employment are depressed below the socially optimal level. In this paper, we consider the most traditional policy intervention for a monopsonized labor market, unionization, where workers form collective bargaining units to negotiate over wages with employers. As discussed by Galbraith (1954), and earlier by Robinson (1933), collective bargaining allows workers to have "countervailing power" when negotiating with large employers; unions act as labor supply monopolists to counter labor demand monopsony power on the employer side.[1] Ideally, unions would cause wages to rise to the socially optimal level, but unions also have the potential to raise wages beyond that.

In this paper, we use a structural model of bargaining, hiring, and labor supply to quantify the effects of collective bargaining relative to the monopsony setting without unions (posted wages) and evaluate how the settings compare to the socially efficient outcome. We use the model to assess the welfare effects of unionization, the spillover effects of unions on competing non-unionized employers, and a counterfactual salary policy in which the state sets a single wage for all employers.

We focus on the market for K-12 teachers in Pennsylvania, which has several characteristics that make it particularly well suited to understanding collective bargaining and monopsony power. First, the government is the predominant employer of schoolteachers, and at least in a fixed geographic area is a true monopsonist. Second, wages for schoolteachers are set using a uniform schedule, which greatly simplifies the monopsony distortion. In contrast, if wages are negotiated for each worker, it is possible that monopsony power causes no distortion in hiring, since the employer need not raise the wages of other workers in order to make a hire.[2] Third, in most school districts in the state, with the notable exception of most charter schools, wages are negotiated through a collective agreement between each school district

[1]See Chapter 25 of Robinson (1933).

[2]Indeed, in order to generate predictions of monopsony quantity distortions, the literature needs to make strong assumptions on the information sets of employers, such as assuming the econometrician and the employer have the same information on the reservation wages of individual workers.

and school district teacher union.

Given that many school districts compete for the same pool of workers, we use an oligopsony model with hiring. In practice, true employer monopsonists are rare; oligopsony, therefore, is a more relevant market structure. The model is flexible and can be applied in many settings.

At the heart of our model of collective bargaining with oligopsony are estimates of each school district's demand for and supply of teacher labor. Labor demand is determined by the school district's input allocation decision, in which it decides how to allocate a fixed budget to hiring teachers versus purchasing other inputs such as buildings, computers, other staff, etc. We estimate a constant elasticity of substitution (CES) production function that determines how the school district makes this tradeoff. We estimate an elasticity of labor demand of $\epsilon_D = 0.72$. Based on their demand and the negotiated wage, school districts send out offers to a subset of teachers. When negotiated wages are set above the level that clears the labor market, not all workers will receive a job offer.

For labor supply, we estimate a mixed-logit labor supply model that depends on wages and commuting distance to each school district as well as other school characteristics, such as the share of students who qualify for free-lunch and charter status. Since our model incorporates constrained choice sets that we do not observe, a traditional estimation strategy using all possible choices would bias our estimates (Crawford et al., 2021). Instead, we leverage the independence of irrelevant alternatives assumption found in logit models and estimate the model on a subset of choices that are always available for teachers, quitting or staying at the current job. We find that the estimated marginal rate of substitution between commuting time and wages implies that a teacher values each hour of commuting at \$ 49.22. The average elasticity of labor supply is around 5.74, but varies between 4.4 and 6, which matches values seen in the literature.[3] However, since teachers may only receive offers from a subset of schools in equilibrium, the elasticity of labor supply conditional on these offers is lower at about 3.76. We also estimate cross-elasticities of labor supply and find that teachers are more likely to substitute to nearby school districts than to the outside option. Thus, lowering wages in one school district leads teachers to switch school districts

[3]Using pre-determined salary schedules as an instrument, Ransom and Sims (2010) estimate a labor supply elasticity of 3.7 for teachers in Missouri. In a meta-study not specific to teacher labor markets, Sokolova and Sorensen (2021) report elasticities between 4-6.

rather than leaving the teaching profession.[4]

For collective bargaining, we use the Nash-in-Nash model commonly used in the bilateral oligopoly literature in empirical industrial organization. We estimate a bargaining power parameter for each school district-collective bargaining unit pair, and find substantial dispersion in this parameter, echoing previous work such as Grennan (2013) that performs this type of exercise.

After estimating the model, we generate equilibrium outcomes of wages and teachers for the social planner, posted wages, and collective bargaining environments. The model predicts average wages of \$56,042 for the social planner, while predicting average wages of only \$52,329 for the posted wage setting. Each school district always pays lower wages to teachers in the posted wage setting compared to the social planner setting. Moreover, overall employment only drops from 113,237 to 112,201 teachers when moving from the social planner to the posted wage environment. Because demand and supply for labor are relatively inelastic, the monopsony distortion has larger effects on wages than on employment. In the collective bargaining setting, average wages are \$55,904 which is 7% higher than the oligopsonist solution and quite close to the social planner solution. However, the number of teachers falls from 113,237 to 106,976 as job rationing occurs under collective bargaining.

Collective bargaining, however, does not lead to uniformly higher wages for teachers. For instance, thirty percent of school districts have lower wages under collective bargaining compared to the social planner setting. To make sense of this result, notice that monopsony is the limit of the collective bargaining solution as bargaining power to workers tends to 0. Thus, absent bargaining externalities between school districts, collective bargaining must raise wages. However, this is also because when wages at a school district increase, the school will reduce hiring, which expands the pool of teachers that other school districts can hire. Essentially, the reduction in hiring increases the labor supply curves at nearby school districts, which allows them to lower wages. Thus, the overall effect of collective bargaining depends on how negotiated wages interact with school district hiring decisions, which creates externalities on the labor supply of nearby schools. This possibility has been discussed since at least the pioneering work of Lewis (1963). The complexity of the effects of bargaining power on outcomes in markets with

[4]If a school is far away from other schools, teachers may prefer the outside option when wages fall at that school.

externalities also recalls the work of Ho and Lee (2017) who find effects of market power in the market for hospital services that can raise or lower prices to consumers due to complex externalities.

This paper relates to several strands of literature at the intersection of labor, monopsony, and industrial organization.

We contribute to the long-standing literature on the impact of unionization on the wage distribution. Many papers attempt to estimate the impact of unionization on the wage distribution. Notable studies include Lewis (1963), Freeman and Medoff (1984), Card (1996), Lemieux (1998), DiNardo et al. (1996), DiNardo and Lee (2004), and Farber et al. (2021). These studies largely focus on estimating the extent to which the union wage gap represents private sector workers across occupations and tend to find that unions raise wages. Recent estimates based on regression comparisons by Blanchflower and Bryson (2024) indicate that unionized workers earn around 17.5 percent higher hourly wages than observationally equivalent non-unionized workers.

Despite the fact U.S. public sector workers are about five times more likely to be unionized than private sector workers (Card et al., 2020), fewer studies focus on the impact of unions on public sector wages. Notable exceptions include Blanchflower and Bryson (2010), Card et al. (2020), and Baker et al. (2024). More narrowly, teaching occupations are the most heavily unionized occupations in the United States, yet we know of only two studies that consider union effects on the labor market for teachers with mixed conclusions. Hoxby (1996) uses panel data on district budgets per student and finds that spending per student increases after teachers unionize. She argues that the spending increase is due to increased salaries, as class sizes do not fall after unionization. Lovenheim (2009) uses a difference in differences design with data from teacher union election results in Iowa, Indiana, and Minnesota to estimate the impact of teacher unions on school budget outcomes. He finds no impact on teacher pay, but that unions lead to a small increase in employment.

These papers do not specify a model for labor supply, demand, or how wages are set, so they cannot provide insight into how unions may influence the gap between wages and the marginal revenue product of labor, provide a framework for the analysis of counterfactuals, or explore implications for welfare. We fill that gap by positing and estimating a model of union negotiations that relates to an older strand of labor and industrial relations theory surveyed in Farber (1986). As that literature preceded the development of structural methods in applied microeconomics, there are scant attempts to use this work for quantitative prediction. Some important recent exceptions are the work of Green et al. (2022) and Dodini et al. (2021).

Turning to the side of industrial organization, we draw on a recent literature that looks at bilateral oligopoly, that is, situations where both the buyers and sellers of an input have market power, and where prices are determined by Nash-in-Nash Bargaining (see Horn and Wolinsky (1988) and Collard-Wexler et al. (2019)). This literature starts with Crawford and Yurukoglu (2012) who study negotiations in cable television, or Grennan (2013) for stents, but has reached its most fully developed form in the work on negotiations between hospitals and health insurance companies, see Gowrisankaran et al. (2015) and Ho and Lee (2017), among others. One of the important innovations of this literature for our purposes is that the Nash-in-Nash concept allows us to deal with both bargaining and the oligopsonic interactions in the labor market that introduce externalities between negotiations.

This work in industrial organization is complemented by production side approaches to monopsony, such as Delabastita and Rubens (2025) and Rubens (2023). Indeed, our input allocation problem is an adaptation of this type of approach to the context of the public sector where the objective of profit maximization is nonsensical.

Third, there are a large number of recent papers looking into monopsony power, coming from a broad cross-section of economics. Some prominent papers are Berger et al. (2022) and Gottfries and Jarosch (2023).

Finally, there is extensive work that tries to understand the labor market for teachers using equilibrium models. Some recent work is Bates et al. (2025) and Biasi et al. (2021). However, while we broadly relate to this work in education, our focus is on understanding negotiations between school districts and teachers over wages and hiring rather than how bargaining affects outcomes for students. Indeed, baked into the model of collective bargaining over wages is a fundamental opposition between the interests of teachers and school districts.

The paper proceeds as follows. Section 2 presents a stylized version of the model we will estimate to build intuition for the components that need to be estimated. Section 3 discusses the institutional features of the market for teachers in Pennsylvania and the data that we will use. Section 4 presents some preliminary empirical evidence on wage setting. Section 5 estimates the structural model and Section 6 discusses counterfactuals.

2 Stylized Model

We begin by providing a stylized model of collective bargaining and monopsony power. This model will be extended in Section 5 to account more precisely for the institutional features of the market and to allow for estimation. We start with the classical analysis in Robinson (1933) in chapter 25, and then adapt the model to the context of teachers in Pennsylvania.

Consider the case of a monopsonist employer that has a labor demand curve given by $L^D(w)$, where w is a uniform wage paid to all employees. Workers have a labor supply curve given by $L^S(w)$. Panel a) of Figure 1 presents this case. The intersection of the labor supply and labor demand curve at point A yields the social planner solution, with wages, w^*, and L^* workers hired.

The marginal factor cost for a monopsonist is given by $MFC = w + L\frac{\partial w}{\partial L^S} = w\left(1 + (\epsilon^S)^{-1}\right)$, where ϵ^S is the elasticity of the labor supply curve. In order to hire additional workers the monopsonist also needs to pay inframarginal workers more as well, and thus its marginal factor cost curve is always above the demand curve.

As such, the monopsonist will choose to hire at point B, the intersection of the labor demand curve and the marginal factor cost curve, yielding L^0 workers and a lower wage w^0. Note that this depression of wages and employement below the competitive level implies a deadweight loss from monopsony power.

Now suppose that workers can collectively bargain with the employer. In this case, depending on the exact form of bargaining, wages will be higher than B (as long as workers care more about wages than employers do). In addition, if workers put more weight on higher wages rather than greater employment, wages could be bargained to w^b above w^* as indicated in Panel b) of Figure 1. In this case, labor supply at point D is greater than labor demand at point E; there will be excess supply of workers. Thus, we need to make an assumption on whether labor will be chosen by labor supply or labor demand. We follow the classic Dunlop (1944) approach used in Farber (1978), and assume that employers choose employment on the labor demand curve.[5] This implies that labor will be chosen on the labor demand curve

[5] The collective bargaining agreements for school districts in Pennsylvania do not specify employment levels in the contract, only wage schedules. The US Port workers, where 50 percent of the union membership is paid, but not assigned to a particular job, is an example of where employment levels are agreed upon in the union contract (see WSJ Article accessed January 25, 2025). See section 3 of Farber (1986) for more discussion on

at point E, and thus $L^b = L^S(w^b)$ workers will be hired. Again there is a deadweight loss, but with different incidence on workers and employers. The relative magnitudes of the deadweight losses in the case of monopsony in panel a) of Figure 1 versus the deadweight loss in panelb) of Figure 1 cannot be signed purely from theory. As such, we will use empirical methods to quantify the relative sizes of these deadweight losses.

Notice as well that this model exhibits excess labor supply, since labor supply at point D given by $L^S(w^b)$, exceeds hired labor at L^b. Thus, we will need to specify a rationing rule for labor. In other words, not all workers will get job offers.

3 Institutional Background and Data

We study public school teachers for kindergarten to twelfth grade, which are employed in regular public schools or charter schools in the state of Pennsylvania from 2009 to 2019.[6]

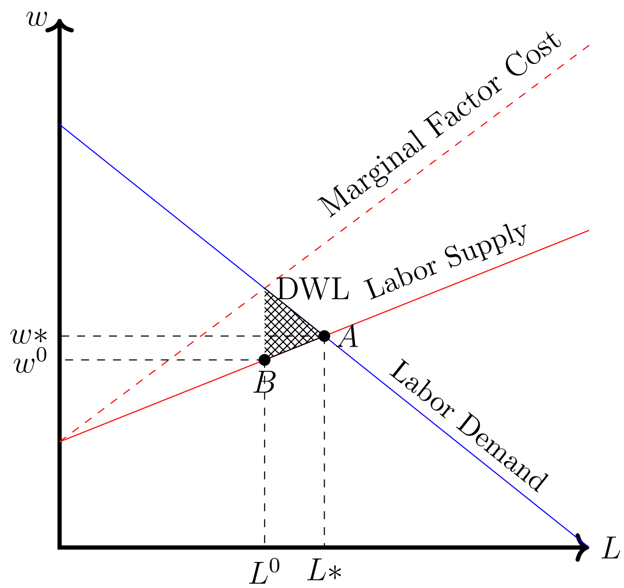
3.1 Teachers

Teachers are licensed professionals that are hired by schools to teach children. After a 3-year period on the job, teachers become tenured, which extends across jobs. As a practical matter, once teachers obtain tenure, they are very rarely fired, and indeed, are infrequently fired before obtaining tenure as well.

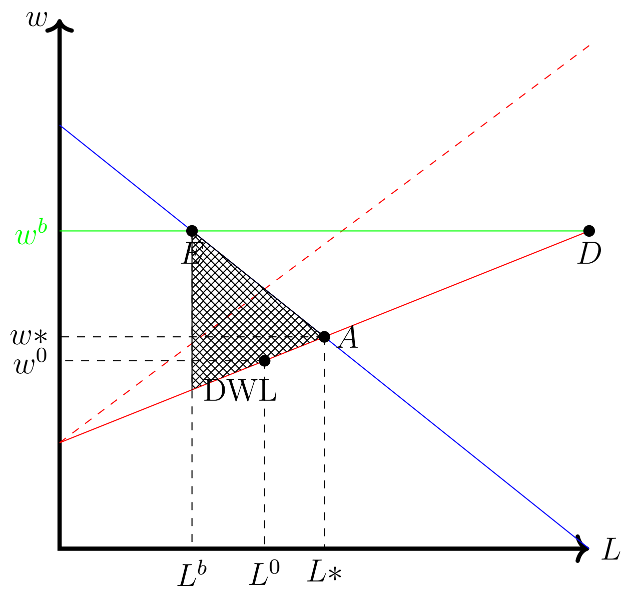
Teachers in Pennsylvania are paid through a uniform wage schedule that depends only on a) years of experience, typically until 10 or 15 year in yearly steps, and b) education level (whether they have a masters or Ph.D. degree). An example of such as schedule in shown in Figure 2 for Abington in the Appendix. This is important for two reasons. First, a uniform wage schedule is needed for monopsony power to distort hiring choices. If a school district can hire a marginal teacher by offering a higher wage only to this hire, the monopsony distortion on hiring would not exist. Second, uniform wage

this point.

[6]We have longitudinal panel ids for teachers from 2013-2023. Most of the analysis is done with panel ids. We avoid using data from 2020-2023 to avoid the Covid shock to schools.



(a) Posted Wage Monopsonist



(b) Collective Bargaining

Figure 1: Welfare Analysis of Collective Bargaining and Monopsony Power

schedules reveal the wage a teacher would earn in an alternative job, which is crucial for modeling labor supply.

Teachers are also compensated with benefits, such as retirement or health insurance plans. These benefits represent about 30 percent of the cost of hiring a teacher in 2023. These benefits are similar across school districts. For example, pension benefits are administered by the Pennsylvania Public School Employees' Retirement System; also, most schools offer health care from a Blue Cross Blue Shield plan with a fairly comprehensive network, often a PPO. Thus, job lock due to limited portability of benefits or employer differentiation due to benefits should not be significant forces. In addition, while wages depend on experience, this experience transfers across school districts.

3.1.1 Teacher Data

First, we received annual payroll data from Pennsylvania.[7] It contains the name, salary, experience, tenure, degree status, job description, demographic characteristics, and school district employer for each teacher and non-instructional staff member from the academic years 2009-2010 through 2022-2023 inclusive. Beginning in the 2013-2014 school year, the data includes stable employee identifiers, allowing us to track individuals over time.

Second, we use data from InfoUSA with names and home addresses for all Pennsylvania residents from 2010 to 2022.[8] In order to identify the commuting distance between teachers and school districts, we match the InfoUSA data to the payroll data on exact name and year. Since many teachers have common names — think Elizabeth Smith — we restrict attention to teachers who are uniquely matched in InfoUSA to compute commuting distance. We are able to obtain a unique match for about 35.8% of teachers in the payroll data. We assume that teachers with unique names do not have significantly different residential patterns than other teachers. See Appendix A.2 for more details.

We use the home address and work addresses to compute commuting times using the HERE georouting application,[9] both for the school a teacher

[7] Penn Staff Data, accessed January 25, 2025.

[8] Note that InfoUSA rebranded as Data Axle in 2020. Data Axle's address data is generated using public sources, voter registration data, utility company data, and real estate data.

[9] We used the stata georoute package which matches to the HERE API

	Mean	Standard Deviation	Observations
Charter Schools	0.07	0.26	113,224
Percent Female	0.74	0.44	113,224
Percent w/ Masters	0.56	0.50	113,224
Years in Education	13.91	8.44	113,224
Commute Time	23.32	19.18	28,492
Wage	68,923.94	17,572.32	112,945
Exit Rate	0.05	0.22	113,224
Switching Rate	0.02	0.14	113,224

Notes: The summary statistics were computed on the 2017 population of public schoolteachers in Pennsylvania who work for either a public school district or a charter school. We exclude teachers working at juvenile penitentiaries or technical high schools. The data was retrieved from the personnel files from the State of Pennsylvania. See Footnote 7 for details. Average commute time has missing data, because we were not able to match all teachers to InfoUSA addresses. In addition, the top 2.5% of commute time values were not included in the average computation.

Table 1: Summary Statistics of Teacher Data in 2017

currently works at, and the school districts that a teacher could potentially work at.

Table 1 presents summary statistics of the public schoolteachers in the data in 2017. We restrict our analysis to teachers that work at either school districts or charter schools, and avoid teachers who work in juvenile penitentiaries or technical high schools, which make up about 5% of the teachers. Public school districts hire a much greater percentage of teachers than charter schools. In addition, the profession is almost three-fourths female, which can exacerbate monopsony power if we believe that women find it more difficult to move locations. Most teachers are career teachers and stay for a long tenure. Thus, teachers have about 14 years of experience teaching, on average. They do not switch schools very often (only 2% of the time), and indeed tend to exit teaching rather than switching schools. About half of exits are due to retirement. We find that teachers tend to exit either very early in their career or at the end when they retire. Workers spend about 23 minutes commuting one-way to their jobs. The average wage of $68,923 hides a lot of heterogeneity across school districts and experience levels.

We also add potential teachers, i.e., qualified teachers who are not teaching, to the model. These teachers work outside the Pennsylvania public

https://www.here.com/platform/geocoding.

school system not necessarily as teachers and are important for the equilibrium outcomes of our counterfactuals. They allow us to capture the effect of wages on the extensive margin. We estimate the number of potential teachers by using the number of licensed teachers and the number of teachers who quit in the previous five years. We find that there are an additional 30.9% of teachers who are qualified but not teaching. See Appendix A.1 for details on the construction of potential teachers and robustness checks.

3.2 Public Schools in Pennsylvania

There are two types of public schools in Pennsylvania, regular school districts and independently run charter schools.[10] In 2017, the state had 499 regular school districts and 159 charter schools in Pennsylvania, with over 93% of public school teachers working for (and students attending) a regular school district. Each school district belongs to one of 29 intermediate unit which manages curriculum, education support, and other administrative tasks. Intermediate units do not affect wages, and thus we focus our model at the school district level. Figure 11 shows the geography of school districts and intermediate units in Pennsylvania.

School districts get their funding from federal, state, and local sources. There is wide dispersion in local funding since it is usually sourced from local property taxes. In addition there is wide dispersion in the size of schools, with the Philadelphia School District being an order of magnitude larger than almost all other school districts. We use data from the Annual Survey of School System Finances from the US Census Bureau.[11] We complement this data with information from the National Center for Education Statistics (henceforth NCES) from the common core of data.[12] The common core of data includes school characteristics like number of students, number of students who qualify for free lunch, and number of teachers.

Table 2 presents summary statistics of the school data. The table also includes 543 private schools, which are relatively small.[13] Due to lack of data,

[10]There are also penitentiary schools and technical career schools, but we do not consider them in our analysis.

[11]https://web.archive.org/web/20241203193341/https://www.census.gov/programs-surveys/school-finances.html accessed January 25, 2025.

[12]Accessed at https://nces.ed.gov/ccd/files.asp.

[13]Private school data comes from the National Center for Education Statistics accessible at https://nces.ed.gov/surveys/pss/pssdata.asp.

they are excluded from our model; however, they only make up about 12% of teachers in Pennsylvania. In our model, teachers who work for a private school have chosen the outside option.

Charter schools are generally much smaller than regular school districts with three times fewer students–842 students vs 3,128 students on average. About 65% of these charter schools are in in urban areas, such as Philadelphia, while only 16% of regular school districts are in urban areas. School districts hire about 215 schoolteachers, while charter schools only hire about 51 teachers. In addition, charter schools have a higher student-teacher ratio of 15.74 students for each teacher compared to the school district's 14.01. Charter schools also tend to have poorer students which we proxy by the percentage of students that are on free lunch. The summary statistics for private schools show that private schools are small and should not have a large impact on our results.

We also have data on school expenditures. Since school districts are much larger, they spend much more–$60.01 million compared to $11.16 million for charters. They also spend more per student–$18,230 vs $14,290 per student. School districts tend to have more diverse funding sources, while charter schools tend to get most of their funding from local sources. Overall, schools tend to spend about the same amount on teachers as a share of their budget, i.e., about 30% of expenditures go towards teachers, while the remainder goes toward other inputs. Teachers cost about $113,300 on average, which includes salary and benefits. About 2/3 of the cost of a teacher is wages and salary.

Due to some missing expenditure and teacher data, the sample that we use to estimate our model includes only 425 regular school districts and 106 charter schools.

3.3 Collective Bargaining

Teachers at every regular school district are represented by a union, the Pennsylvania State Education Association.[14] The union and schools negotiate a collective agreement every 3-5 years that incorporates teacher salary, retirement, health care, working hours, vacation, termination rules, among other topics. Charter schools rarely negotiate with teacher unions, and we will assume that charter school teachers are not unionized. In our model, we

[14]For more information, see https://www.psea.org/.

	Public			Private
	Full	School Dist	Charter	
Basic Facts				
Charter Schools	0.24	0.00	1.00	0.00
	(0.43)	(0.00)	(0.00)	(0.00)
Urban	0.28	0.16	0.65	0.42
	(0.45)	(0.37)	(0.48)	(0.49)
Num of Schools	4.42	5.44	1.00	-
	(9.21)	(10.29)	(0.00)	
Num of Students	2,599.77	3,128.31	841.51	258.99
	(5,535.61)	(6,183.38)	(1,195.34)	(221.33)
Num of Teachers	176.90	214.83	50.70	25.93
	(326.52)	(362.70)	(56.19)	(23.00)
Num of Other Staff	30.57	35.70	13.51	-
	(88.12)	(99.57)	(15.73)	
Student-Teacher Ratio	14.41	14.01	15.74	10.37
	(2.58)	(1.62)	(4.22)	(4.39)
Percent Students on Free Lunch	0.46	0.43	0.58	-
	(0.29)	(0.22)	(0.42)	
Expenditures				
Total Expenditures (mil)	48.20	60.01	11.16	-
	(151.01)	(171.51)	(17.02)	
Expenditure per Student (thou)	17.32	18.23	14.29	-
	(4.75)	(4.35)	(4.78)	
Percent Funding from Local Sources	0.61	0.52	0.90	-
	(0.23)	(0.18)	(0.06)	
Expenditure on Teachers (mil)	21.35	26.74	4.42	-
	(39.37)	(43.77)	(5.20)	
Teacher Share of Exp	0.31	0.30	0.34	-
	(0.10)	(0.04)	(0.18)	
Avg Teacher Cost	113,306.60	119,496.80	92,713.88	-
	(23,142.61)	(17,923.08)	(26,463.87)	
Avg. Salary Share of Teacher Exp	0.63	0.62	0.66	-
	(0.04)	(0.02)	(0.07)	
Obs	658	499	159	543

Notes: The table shows summary statistics for school districts, charter schools, and private schools in Pennsylvania in 2017 with 10 or more teachers employed. The data was retrieved from NCES. The public school data comes from the School District Finance Survey (F33) and the School District Universe Survey Data. Private school data comes from the Private School Universe Survey Data that is collected every two years by the NCES.

Table 2: 2017 Summary Statistics of School Data

APPENDIX A SALARY SCHEDULE

SCHOOL YEAR 2018-2019 (FY19)- 2.5%

Step	B	B+15	M	M+15	M+30	M+45	DOC
1	$41,998	$43,373	$45,322	$46,704	$48,286	$49,918	$51,549
2	$46,590	$48,123	$50,280	$51,740	$53,403	$55,034	$56,666
3	$49,059	$50,603	$53,328	$54,749	$56,705	$58,337	$59,969
4	$51,826	$53,532	$56,414	$57,874	$59,579	$61,211	$62,843
5	$54,422	$56,615	$59,579	$61,045	$62,711	$64,342	$65,974
6	$57,308	$59,870	$62,913	$64,332	$66,040	$67,672	$69,303
7	$60,468	$63,154	$66,156	$67,590	$69,287	$70,919	$72,551
8	$63,603	$66,484	$69,736	$71,235	$72,977	$74,609	$76,241
9	$66,728	$69,736	$73,024	$74,450	$76,109	$77,741	$79,373
10	$70,188	$73,228	$76,473	$77,899	$79,569	$81,201	$82,831
11	$74,467	$77,836	$81,655	$83,078	$84,780	$86,412	$88,043
12	$75,764	$79,167	$83,021	$84,460	$86,180	$87,812	$89,444
13	$76,655	$80,091	$83,986	$85,437	$87,172	$88,804	$90,436

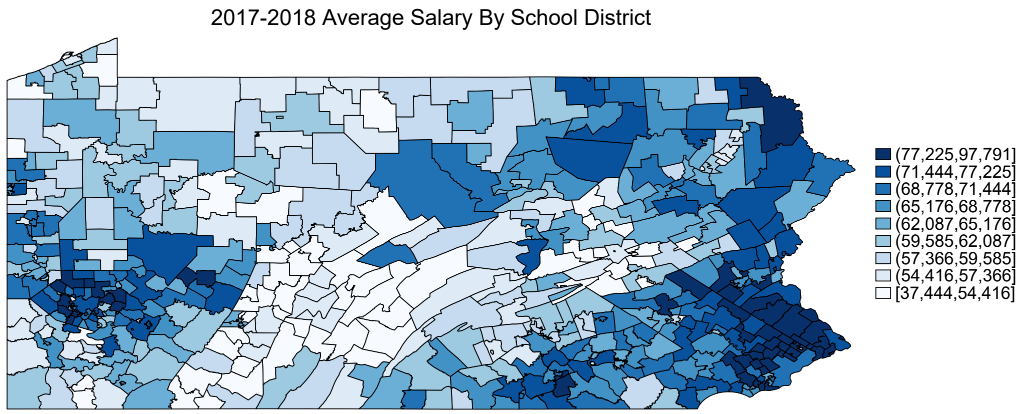
Notes: B, M and DOC refer to bachelor, masters, and doctoral degree, plus additional credit hours. Steps are years of experience. Source: Page 47 of collective bargaining agreement: "Agreement Between Abington School Committee and Abington Education Association Effective September 1, 2018 to August 31, 2021" Accessed Jan 28, 2025.

Figure 2: Salary Schedule for Abington School District

assume that the union negotiates individually with each school district and does not consider the effect negotiations at one school will have on any other school district.

Of importance is that unions are negotiating wages on a uniform wage schedule. Figure 2 shows the salary schedule for Abington School district in 2017-2018, where salary is based on steps, or years of experience in teaching, as well as having a bachelor, masters, or doctoral degree. Based on cross-referencing collective agreements with our data, we find that schools pay the wages as stipulated in the contracts. In the model, we will assume that each school pays a single wage to all teachers that is determined as the median wage of all teachers with 6-7 years of experience that are working at that school.[15]

[15] We did not find significant differences in the slope coefficient of the regression of experience on salary, which indicates that a single wage is sufficient to represent the distribution of wages across schools.



Notes: The map displays the average real salary of teachers during SY 2017-2018 for each school district in Pennsylvania.

Figure 3: Average Salaries Vary Across Districts

4 Stylized Facts and Suggestive Evidence

In this section we discuss some stylized facts that motivate our model of oligopsony power and collective bargaining.

4.1 Wage Variation

First, we want to understand the correlates of wages at different school districts. Figure 3 presents a map of average salaries at different school districts in 2017. Salaries vary greatly across the state, from 2017 average salaries of $\$37,444$ in Turkeyfoot Valley to over $\$97,000$ for Lower Merion. Some of this salary dispersion in salaries can be explained by compositional differences in the workforce, such as having teachers with more experience and more education. However, this shrinks the overall variance in salary dispersion by only one-third, indicating large differences in salaries across employers unrelated to worker attributes.

We also note that much of the salary dispersion in Pennsylvania is not explained exclusively by local labor markets, as locations such as Lower Merion and Philadelphia border each other and yet have average salaries that differ by over $\$20,000$ per year. We believe that differences in wages can be explained at least in part by differences in bargaining power.

4.2 Concentration

Next, we consider school district concentration in the market for teachers. We consider two approaches to measure concentration: number of schools within a geographic region and the share of teachers within 10 miles of the school district that are hired by that school district.

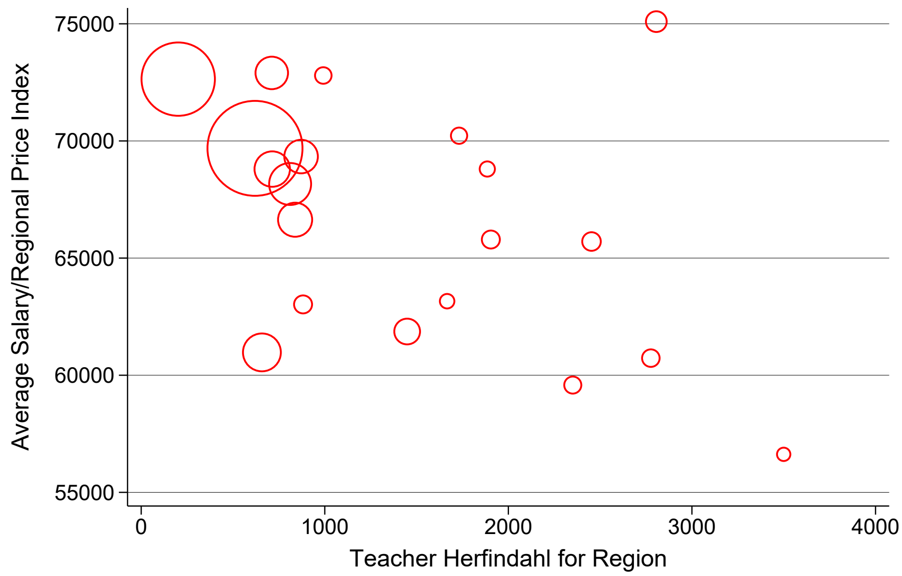
First, we assume that teachers are partitioned into local labor markets; e.g., a teacher in the State College region cannot look for work in the Altoona region, and vice-versa. These local labor markets are defined by Metropolitan Statistical Areas from the Bureau of Labor Statistics.[16] Note that we drop rural school districts, since they are not part of a MSA. Figure 4 shows a scatterplot of average wages at the regional level against regional concentration measured by the HHI, where the size of each circle is proportional to the number of teachers in each district. We deflate wages by a regional price index so that differences in the cost of living between regions do not drive the result. Figure 4 shows a negative relationship between wages and concentration. The OLS regression coefficient shows that an increase in HHI of 1,000 is associated with a drop in wages of $ 3,081 (with a t-stat of -2.5, but with only 20 regions). HHI varies from a low of 202 in Pittsburg, to 2,776 for Altoona.

Second, we construct a measure of concentration that varies at the school district level. We compute the share of teachers that live within 10 miles of the school district and that are employed by that school district. The higher the share of teachers a school hires, the more market power it should have. Figure 5 presents a map of Pennsylvania with these concentration measures indicated by school district.

We want to emphasize that these regressions show correlation, not causation. These regressions have a perilous history in empirical industrial organization as summarized by Miller et al. (2022), but have been treated with caution in the industrial organization literature since at least Demsetz (1973)'s work. One can get positive relationships between concentration and pricing without market power. Interpretation of these regressions should be done with care.

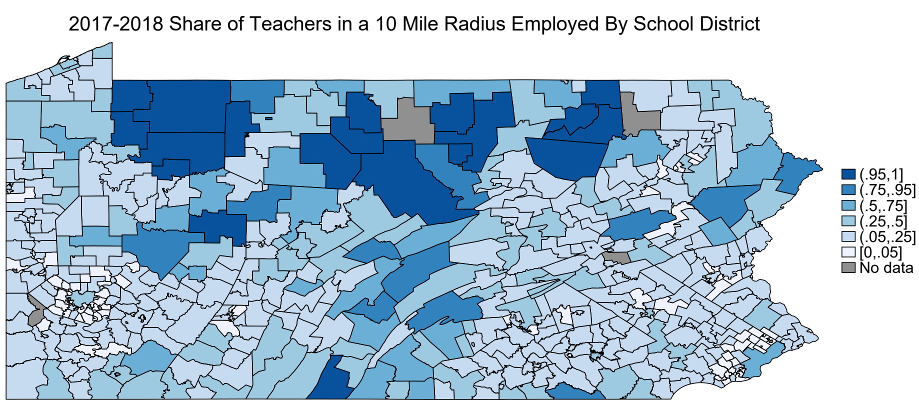
Table 3 runs the regression between log salaries at a school district and either a) the log of the number of school districts within 10 miles, or b) the number of districts within 10 miles being less than 5, 5 to 15, or more than 15.

[16]See BLS Definition of Metropolitan Statistical Area, accessed Jan 28, 2025.



Notes: Each circle in the figure represents one of the 21 primary statistical areas in Pennsylvania. The Herfindahl is the sum of squared share of each employers teachers in the overall market, scaled by 10,000.

Figure 4: Regional HHI and Wages



Notes: The map displays the share of teachers that live within 10 miles of each school district that are employed by that district for the state of Pennsylvania.

Figure 5: Concentration Varies Across Districts

	Salary Over Cost of Living Index		
	(1)	(2)	(3)
log(Districts Within 10 Miles)	0.06	0.05	
	(0.01)	(0.01)	
≤ 5 Districts within 10 Miles			-0.10
			(0.02)
> 5 and < 15 Districts within 10 Miles			-0.06
			(0.02)
Region FE		X	
Observations	1077398	1077398	1077398
R^2	0.52	0.54	0.59

Notes: All regressions include controls for teacher experience, education, and whether the school is a charter school. Standard errors are clustered at the school district level.

Table 3: Correlation between Concentration and Salaries

A one percent increase in the number of school districts is associated with an increase in salary of 0.06 percent. This effect is not driven by the previous regional regression in Figure 4, but instead operates within a region. However, if we include other controls for density, such as the number of students living nearby, we cannot rule out zero or positive effects of concentration on wages. Separating employer concentration and the economics of density is not clear cut.

4.3 Wages and Unionization: Charter Schools

Charter school teachers are rarely unionized . We find that they pay wages that are much lower than regular school districts. In particular, they pay $48,722 on average to teachers compared to an average of $69,125 for regular school districts. Table 4 shows that the charger school salary penalty is not explained by teacher's experience, degree status, and regional fixed effects. We believe the charter salary penalty arises due to the lack of a union rather than compensating differentials. The latter is unlikely, since teachers quit at much higher rates from charter schools, conditional on salary. Even so,

Table 4: Union Wage Gap

	Dependent Var: Log Real Wage		
Variable	(i)	(ii)	(iii)
1(Charter)	-0.340	-0.165	-0.326
	(0.016)	(0.018)	(0.004)
Teacher Covariates	No	Yes	Yes
Intermediate Unit F.E.'s	No	No	Yes
N	556394	556394	556394

Notes: The table summarizes point estimates and standard errors of a regression of the log of real salary on a charter school indicator. Column 2 adds a post-bachelor degree indicator and indicators for years of experience. Standard errors are clustered by Intermediate Unit.

it is still difficult to interpret the difference in salary as the causal effect of unions on salaries. This is because charter schools may compete with nearby standard school districts for teachers and set salaries that depend on local competitive conditions. However, the fact motivates a modeling assumption where we allow charter schools to post wages.

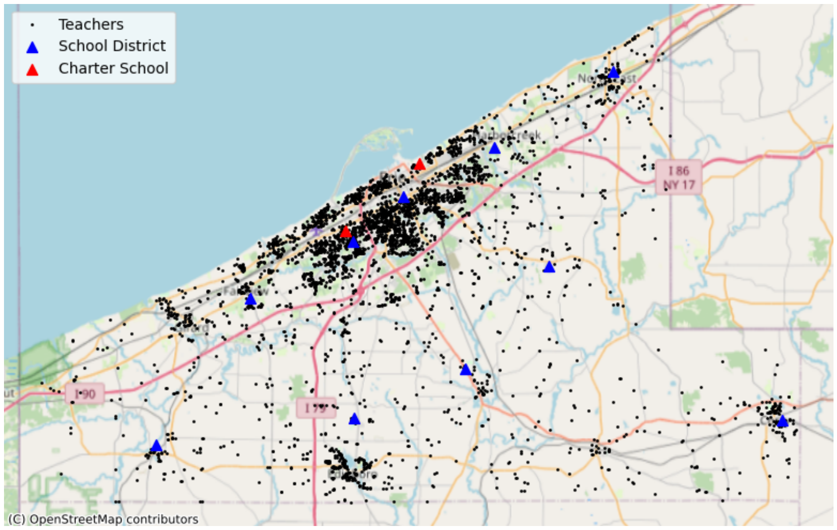
5 Structural Model

5.1 Overview of Model Components

We present a model of bargaining between school districts and teacher unions that incorporates a labor supply decision and an offer game. Figure 6 presents a visual representation of the data that we are using to estimate the model. The figure represents the location of teachers and school districts in Erie County in northwest Pennsylvania. Small black dots indicate the location of teachers,[17] while the blue circles represent regular school districts and the red circles are the two charter schools.

Our model proceeds in three stages. First, wages are negotiated through a Nash bargaining process between school districts and teacher unions. Since negotiated wages at one school affect the labor supply of neighboring schools, we find the wage solution that leads to a Nash equilibrium, i.e., we model

[17] The location also includes potential teachers, i.e., teachers who are qualified, but not teaching in 2017. See Appendix A.1 for details.



Notes: The map shows the location of school districts and charter schools in Erie County, Pennsylvania in 2017. Black dots represent teacher addresses, which were determined by merging names from the teacher data with InfoUSA data. Teacher addresses depicted here are from the years 2010-2020.

Figure 6: Schools and Teacher Location For Erie County, Pennsylvania

a Nash-in-Nash bargaining game. Note that the Nash bargaining model nests the posted price solution, since when the bargaining weight equals one, teacher unions have zero bargaining power. Second, given wages, school districts choose which teachers to send offers to based on their labor demand. Third, workers, with offers in hand, choose which school to work at.

Notice that this model is entirely static in two important ways. First, the model assumes that hiring and job choice happen at one point in time. This is at odds with the strong tenure protections that teachers enjoy, and the fact that teachers typically spend their entire career at a single school district. However, a model that would account for the entire life-cycle of a teacher's career would be challenging to incorporate into the labor supply. Second, negotiations happen simultaneously. The model implications would not change if we incorporated the fact that collective agreements are renegotiated every 4-5 years and are not negotiated simultaneously. But this would introduce dynamics into the bargaining problem which are very complex. Lee and Fong (2013) discussing the issues associated with it.

We will proceed through the three stages of the model in reverse order, that is labor supply, labor demand, and wage determination.

5.2 Labor Supply

We model the teacher labor supply decision using a mixed logit framework with constrained choice sets. Teachers $i = 1, \cdots, N$ either receive utility from working in a school district indexed by $j = 1, ..., J$:

$$u_{ij} = \psi\omega_j + x_j'\beta - \tau d_{ij} + \sigma\nu_i + \epsilon_{ij}, \tag{1}$$

or from leaving public education, denoted by $j = 0$:

$$u_{i0} = \psi r + \epsilon_{i0}.$$

The teacher's utility from working at school j depends on the salary offered ω_j and a vector of non-salary attributes x_j which includes the fraction of students on free lunch and an indicator for whether the employer is a charter school. The mean value of the outside option is given by r.

The choice model allows for both observable and unobservable heterogeneity in preferences across teachers. Observable heterogeneity is captured through d_{ij}, the commuting distance between teacher i's home address and the centroid of school district j.Unobservable heterogeneity depends on two terms: 1) ϵ_{ij}, which follows a Type 1 extreme value distribution that is independent across both i and j, and 2) ν_i, which follows a $N(0, \sigma^2)$ random variable that is independently distributed across teachers. The ν_i term is common across all public teaching jobs. This allows for the possibility that individuals in a public education job will be more likely to switch to another public education job given a salary reduction instead of switching to the outside option.[18]

Furthermore, this choice model allows teachers to have different choice sets. Each teacher i has a list of job offers summarized by the vector $\mathbf{O}_i$, with element $o_{ij} = 1$ if teacher i receives an offer from district j and zero otherwise.

Each teacher chooses the employment option that yields the highest utility, i.e., they can choose among their offers or take the outside option, not working. Since our ϵ errors are distributed Type 1 Extreme Value, given ν_i, we can represent the probability that teacher i chooses to work at district j as:

$$s_{ij}(\nu_i) = \frac{\exp(\delta_{ij}(\nu_i))o_{ij}}{\exp(r) + \sum_k \exp(\delta_{ik}(\nu_i))o_{ik}}. \tag{2}$$

[18] As is common in the IO literature on discrete choice in oligopoly, it is important to understand whether teachers are more likely to switch between schools or between the teaching profession and the outside option.

We integrate over ν_i to get:

$$s_{ij} = \int_{\nu} s_{ij}(\nu)\phi(\nu)d\nu. \tag{3}$$

Given a vector of wages and the estimated parameters, we can aggregate the individual choice probabilities to create a firm-level teacher supply curve for each school district:

$$T_j = \sum_{i \in \mathbf{O}_j} s_{ij}, \tag{4}$$

where $\mathbf{O}_j$ is the set of workers i that receive an offer from school j, i.e., $\mathbf{O}_j = \{i \text{ s.t. } o_{ij} = 1\}$.

5.3 Estimation with Unobserved Offers

The choice probability equation given by equation (3) is a very common mixed logit which is easy to estimate by maximum likelihood. However, we include an offer set, $\mathbf{O}$, which is unobserved. In most cases, except in the case of matching mechanisms, the entire set of offers would not be observable, since job search is often a sequential search process which terminates when one has a good enough offer. Thus papers such as Bates et al. (2025) show principals sending out very few offers. Instead, we rely on the Independence of Irrelevant Alternatives (IIA) assumption, to estimate the model on a subset of two choices that are known. We assume that teachers always have the option of either keeping their current job (stay) or leaving the public education sector (quit). Conditional on $\nu_i = \nu$, the realization of the idiosyncratic preference shock common to all public education jobs, the likelihood of quitting or staying in one's current job is given by:

$$\begin{aligned}
\Pr(\text{quit}|\text{quit or stay}) &= \frac{\Pr(\text{quit})}{\Pr(\text{quit or stay})} \\
&= \frac{\exp(r)}{\exp(r) + \sum_k \exp(\delta_{ik}(\nu_i))o_{ik}} \times \frac{\exp(r) + \sum_k \exp(\delta_{ik}(\nu_i))o_{ik}}{\exp(r) + \exp(\delta_{ij}(\nu_i))} \\
&= \frac{\exp(r)}{\exp(r) + \exp(\delta_{ij}(\nu_i))}.
\end{aligned} \tag{5}$$

Note that this is an immediate consequence of logit driven IIA as discussed by McFadden (1978) for estimation with a subset of choices. We will call $s_{iq}(\nu, \phi, \beta, \tau, \sigma) \equiv \Pr(\text{quit}|\text{quit or stay})$.

Thus, we can form a likelihood on quit decisions given by:

$$\begin{aligned}\mathcal{L}(\nu, \phi, \beta, \tau, \sigma) = \int_{\nu} \sum_{i} \log[1(q_i = 1) s_{iq}(\nu, \phi, \beta, \tau, \sigma) \\ + (1(q_i = 0)(1 - s_{iq}(\nu, \phi, \beta, \tau, \sigma))]\phi(\nu) d\nu,\end{aligned} \tag{6}$$

where $\phi(\cdot)$ is the density of the normal distribution.[19] Since we never use choices between school districts, it is difficult to separately identify σ, at least without relying on the functional form of the normal distribution over ν. As such, we will estimate r and σ using an indirect inference procedure that we discuss in section 5.7 later in the paper.

5.4 Labor Supply Quitting Regressions

Table 5 presents estimates of the quitting versus staying decision using the likelihood given by equation (6) conditional on different values of σ. The first three columns present maximum likelihood estimates on the sample of teachers with fewer than 15 years of experience. Column 1 corresponds to the case when $\sigma = 0$, where the model collapses to a binary logit. Each coefficient estimate is significant at the .05 level and has an intuitive sign. A \$10,000 increase in salary reduces the likelihood of quitting by half a percentage point, which is one-sixth of the mean quit rate. A one-hour increase in commute time increases the likelihood of quitting by 1.2 (60*.002*100) percentage points. A one standard deviation (.29) increase in the share of students on free lunch increases the likelihood of quitting by 17.6 percentage points, and quit rates are 2.2 percentage points higher at charter schools.

[19]The distribution of ν is $\mathcal{N}(0, 1)$. However, the distribution of ν conditional on either quitting or staying in your job might not be $\mathcal{N}(0, 1)$. Concretely, if ν is very large, then it is unlikely that this worker chooses to quit. We think this a real issue in terms of methods, but has small quantitative impact in our setting. In our data, the move rate is quite low, between 2 to 3 percent. As such, the truncation of the distribution of $F_{\nu}(\cdot|\text{quit or stay})$ is likely to be small. Furthermore, the marginal rates of substitution between attributes are quite similar for the logit model, which does not suffer from the conditional distribution of ν issue, as the mixed logit estimates for different values of ν. Finally, the indirect inference estimates of σ and ν do not have the same issue with conditional distributions of ν.

Table 5: MLE Estimates of the Quit Decision

Variable	(i)	(ii)	(iii)	(iv)
Real Salary	-0.224	-0.534	-1.127	-0.517
(10k)	(0.030)	(0.067)	(0.234)	(0.237)
	[-0.005]	[-0.005]	[-0.005]	[-0.003]
Commute Time	0.008	0.017	0.015	0.0152
(Minutes)	(0.002)	(0.005)	(0.007)	(0.007)
	[0.0002]	[0.0001]	[0.0001]	[0.0001]
Free Lunch Share	0.607	1.425	1.954	2.097
	(0.226)	(0.532)	(0.695)	(0.920)
	[0.014]	[0.0139]	[0.009]	[0.009]
Charter	0.948	2.026	2.311	2.214
	(0.113)	(0.259)	(0.348)	(0.274)
	[0.022]	[0.020]	[0.011]	[0.012]
σ	0	6.75	13.5	13.5
$Pr(quit)$	0.026	0.026	0.026	0.041
Experience	< 15 *years*	< 15 *years*	< 15 *years*	< 25 *years*
Observations	72,802	72,802	72,802	115,321

Notes: The table summarizes point estimates and standard errors of Maximum Likelihood Estimation of the Random Coefficient Model of quitting for different values of the random coefficient parameter σ. Marginal effects are included in brackets. The sample includes teachers with less than 15 years of experience and a commute time of less than 70 minutes. Year indicators are supressed. Standard errors clustered by Intermediate Unit.

Column 3 corresponds to the value of σ from our indirect inference procedure. The marginal effects (numbers in brackets under standard deviations) are essentially the same as those from the binary logit model. Assuming 180 school days per year and two trips to school per day, the coefficient on commute time and real salary implies a marginal rate of substitution between salary and commute time of $49.22 per hour. The second column corresponds to an intermediate value of σ and shows how the coefficients respond to other values of σ.

The last column expands the sample to all teachers with at least 25 years of experience. This increases the sample size by over 50 percent. The mean quit rate increases to 4.1 percent, since older teachers retire more frequently. Despite this, the marginal effects are roughly the same as in column 3.

5.4.1 Labor Supply Elasticities

Given the options available to each teacher, the estimated labor supply model can be used to construct the labor supply elasticity matrix to each school. The elasticity of supply to school district j with respect to salary at school k is given by[20]:

$$\begin{aligned} \epsilon_{jk} &= \frac{\partial log(T_j)}{\partial log(\omega_k)} = \frac{\omega_k}{T_j} \sum_{i \in N} \int_{\nu} \frac{\partial s_{ij}(\nu)}{\partial \omega_k} \phi(\nu) d\nu \\ &= \begin{cases} \frac{\omega_k}{T_j} \sum_{i \in N} \int_{\nu} \psi s_{ij}(\nu)(1 - s_{ik}(\nu))\phi(\nu) d\nu & \text{if } k = j, \\ \frac{\omega_k}{T_j} \sum_{i \in N} \int_{\nu} -\psi s_{ij}(\nu) s_{ik}(\nu) \phi(\nu) d\nu & \text{if } k \neq j \end{cases} \end{aligned} \tag{7}$$

Using 2017 wages and teachers from our data, we find that the labor supply elasticities given that teachers receive offers from all schools range between 4.4-6 which is consistent with other values in the literature that range from 3.7-6.[21] Table 6 illustrates the elasticity matrix for some of the schools in Erie, including charter schools. The entry in column j and row k is equal to ϵ_{jk}. Notice that many of the cross-wage elasticities between school districts are similar, so for instance, the elasticity with respect to a change in the wage in Erie City School District ($S - ER$) is -0.053 for Millcreek ($S - MI$), and -0.049 for the North East ($S - NE$) school district which is a bit further away. This is what we would expect from a model close to logit. What is not so similar is the elasticity for the outside good which is an order of magnitude smaller at -0.006. Thus, the mixed logit structure is very important: when wages rise, most of the teachers that take up offers are coming from existing teachers at nearby schools, not from people working in other occupations.

Note that these results depend crucially on the set of job offers each teacher has. When jobs are rationed in some districts because salaries are negotiated above that which equates labor supply and labor demand, these relationships can break down as highlighted in Section 5.8.

[20]Note that a change in wage may affect the offers a school decides to distribute. However, since the offer game is discrete, an infinitesimal increase in the wage should not affect the offer matrix. In addition, if we allow schools to change their offers, we would be estimating demand elasticities, rather than labor supply elasticities.

[21]Using a quasi-random experiment, Ransom and Sims (2010) find a labor supply elasticity of 3.7 for teachers in Missouri. Other monopsony papers (e.g., Sokolova and Sorensen (2021)) find most elasticities between 4-6 across labor markets.

	S-ER	S-MI	S-NE	S-NW	C-MO	C-RO	Out Opt
All Offers							
S-ER	5.55	−0.053	−0.049	−0.047	−0.053	−0.054	−0.006
S-MI	−0.529	6.001	−0.461	−0.485	−0.553	−0.506	−0.059
S-NE	−0.324	−0.305	5.749	−0.318	−0.307	−0.347	−0.039
S-NW	−0.175	−0.179	−0.178	5.397	−0.179	−0.173	−0.022
C-MO	−0.003	−0.003	−0.003	−0.003	4.369	−0.003	−0.0
C-RO	−0.026	−0.024	−0.025	−0.023	−0.025	5.153	−0.003
Out Opt	−0.319	−0.322	−0.319	−0.329	−0.322	−0.318	1.485

Notes: The table shows own- and cross-elasticities of a subset of schools in Erie County when teachers get offers from every school district. S refers to a school district, while C refers to a charter school. See the map in Figure 12 for the location of the schools. Each element shows the percentage increase in teachers at the schools in the columns willing to work after a 1% increase in the wage paid by the schools in the rows.

Table 6: Elasticities Erie

5.5 Labor Demand

Each school simultaneously makes job offers O_j in an effort to maximize an objective function $W(T_j(O_j, \mathcal{O}_{-j}), X_j)$ that depends upon two inputs: the number of teachers employed, T_j, and all other educational inputs, denoted by X_j. Other inputs X_j could include things like non-teacher staff, equipment, building, and land. The number of teachers hired $T_j(O_j, \mathcal{O}_{-j})$ depends not only on the offers school j makes but also on the offers made by competing schools, denoted by the matrix $\mathcal{O}_{-j}$. We allow the school's welfare function to have a Constant Elasticity of Substitution (CES) form, which allows the elasticity of labor demand to be somewhat flexibility. Imposing the budget constraint $B_j = X_j + w_j T_j$ yields the problem:

$$\max_{O_j} W(T_j(O_j, \mathcal{O}_{-j}), X_j) = \Big(\gamma T_j(O_j, \mathcal{O}_{-j})^\rho + (1-\gamma)(B_j - w_j T_j(O_j, \mathcal{O}_{-j})))^\rho\Big)^{\frac{1}{\rho}}$$

$$\begin{aligned} s.t.\ T_j(O_j, \mathcal{O}_{-j}) &= \sum_{i=1}^{O_j} s_{ij}(\mathcal{O}_{-j}), \\ O_j &\leq N, \end{aligned} \tag{8}$$

where B_j is district j's budget, w_j is the cost of hiring a teacher, the price of other goods X_j is normalized to one, and N is the number of teachers in the labor market.[22]

If Equation 8 is not binding, the district is able to hire enough teachers to meet labor demand at the prevailing wage. In this case, the optimal number of teachers hired equates the marginal rate of technical substitution between teachers and other inputs to the relative price of hiring a teacher:

$$\frac{\frac{\partial W}{\partial T}}{\frac{\partial W}{\partial x}} = \frac{\gamma}{1-\gamma} \cdot \left(\frac{T_j}{X_j}\right)^{\rho-1} = w_j. \tag{9}$$

Imposing the budget constraint and rearranging yields school j's labor demand curve:

$$T_j(w_j) = \frac{w_j^{\frac{1}{1-\rho}} B_j}{\left(\frac{\gamma}{1-\gamma}\right)^{\frac{1}{\rho-1}} + w_j^{\frac{\rho}{\rho-1}}} \tag{10}$$

Note that when Equation 8 does not bind, labor supply to the district exceeds labor demand. Then, schools send out limited job offers. We assume that schools send out offers sequentially to the closest teacher. This is an intuitive rationing rule that is also efficient given the labor supply model. In particular, we avoid introducing misallocation due to travel costs.

The other possibility is that Equation 8 binds, and the district cannot hire as many teachers as it would like to at the prevailing wage. In this case the benefit of hiring another teacher exceeds the cost:

$$\frac{\frac{\partial W}{\partial T}}{\frac{\partial W}{\partial x}} = \frac{\gamma}{1-\gamma} \cdot \left(\frac{T_j}{X_j}\right)^{\rho-1} > w_j \tag{11}$$

and the district makes offers to every teacher in the market so $T_j = \sum_{i=1}^{N} s_{ij}$.

An equilibrium in the offer game is defined by a set of offers for each district $\mathcal{O}^*$ such that each district's offers $\mathcal{O}_j^*$ maximizes its payoff function given the offers made by other districts $\mathcal{O}_{-j}^*$. We compute the equilibrium using best-response function iteration. The details of the algorithm 2 are given in the appendix.[23]

[22] In practice, only teachers within 60 minutes of a school district are part of the school's labor market.

[23] Although we cannot prove that there is a unique offer game equilibrium, starting with different offer guesses had no significant impact on the final offer matrix. In addition, we randomized the order of school districts and found no differences in the final offer matrix.

We obtain a reduced form estimating equation for the labor demand parameters by transforming Equation 22:

$$\begin{aligned} \log(w_jT_j) - \log(X_j) &= \frac{1}{\rho - 1}\left[-log(\frac{\gamma}{1-\gamma}) + \rho \log(w_j)\right] \\ &= \gamma_1 + \rho_1 \log(w_j) \end{aligned} \tag{12}$$

Table 7 presents estimates constructed by fitting Equation 22 to the data using OLS. The independent variable, labeled teacher marginal factor cost, is total spending on teachers over the number of teachers hired. We exclude observations from charter schools, because they are not unionized and thus never offer salaries high enough for labor demand to determine hiring. Column 1 presents the results from a simple bivariate regression. Column 2 adds year fixed effects and Column 3 includes both year and region fixed effects. Our preferred specification is Column 4 which includes year and school district fixed effects. In each case, the coefficient on the teacher marginal factor cost is significantly different from zero. This rejects Cobb-Douglas preferences. The implied average elasticity of demand ranges from −0.68 to −0.80, and is −0.72 in our preferred specification.

The first four columns of Table 7 present results estimated on all traditional school districts. This is despite the fact that the first-order condition used to derive the estimating condition holds only for districts where the wage is high enough that labor demand determines employment. Column 5 presents estimates of labor demand conditional on the set of schools we classify as being on the labor demand curve given the OLS estimate of ρ in Column 4. First, note that of the 422 school districts only 16 were not on the labor demand curve. Second, the parameter and elasticity estimates are essentially unchanged.[24]

[24]We could also estimate ρ within the indirect inference procedure. Conditional on labor supply, given a guess of ρ and γ we could classify whether labor demand or supply bind for each district. We could then construct a moment inequality estimator that minimizes the criterion function:

$$\begin{aligned} Q(\gamma_1, \rho_1) = &\sum_{j \in \text{non-binding labor supply}} \left[\log(\frac{w_jT_j}{X_j}) - \gamma_1 - \rho_1 \log(w_j)\right]^2 \\ &+ \sum_{j \in \text{binding labor supply}} 1\left(log(\frac{w_jT_j}{X_j}) < \gamma_1 + \rho_1 \log(w_j)\right)\left[log(\frac{w_jT_j}{X_j}) - \gamma_1 - \rho_1 \log(w_j)\right]^2, \end{aligned} \tag{13}$$

Dependent Variable : Log of Teacher to non-Teacher Expenditure Ratio	(1)	(2)	(3)	(4)	(5)
Regression					
Teacher Marginal Factor Cost	0.3230	0.4455	0.5294	0.4127	0.4585
	(0.0529)	(0.1458)	(0.1065)	(0.0857)	(0.1050)
Average Demand Elasticity					
ϵ_D	-0.7991	-0.7287	-0.6832	-0.7472	-0.7215
Obs	4218	4218	4218	4218	4059
Year FE		✓	✓	✓	✓
Region FE			✓		
District FE				✓	✓

Notes: We regress the log of the teacher marginal factor cost on the log of the ratio of teacher ($mc_T \cdot T$) to non-teacher expenditure (X) for all regular school districts in our sample. Teacher Marginal Factor Cost is determined by taking total spending on teachers and dividing it by the number of teachers a school has hired. We do not use charter schools in the regression, since due to our assumptions, charter schools hire teachers on the labor supply rather than on their demand curve. Regression (1) is a simple ols regression. Regression (2) includes year fixed effects and regression (3) includes year and region fixed effects, where regions are defined as intermediate units. Regression (4) includes year and district fixed effects. Finally, regression (5) only includes schools that are in the labor demand binds case. There were 406 schools in the labor demand binds case (i.e., only 16 regular school districts were not hiring labor on their demand curve). All standard errors are clustered at the school district level.

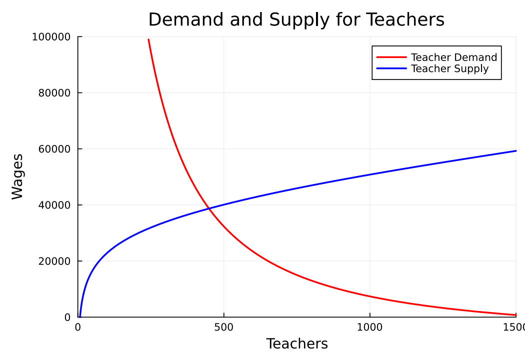
Table 7: CES School Production Function Estimates

Figure 7 presents our estimated labor supply and demand curves for the Millcreek Township school district, which is a mid-sized school district in Erie County. We keep wages fixed at other schools based on the data.

5.6 Wage Negotiation

The wages in our model are determined via simultaneous negotiations between each school district and the union representing teachers in that district, with the outcome of each negotiation solving a bilateral Nash bargaining problem. The bilateral bargaining outcomes must be consistent and constitute a Nash equilibrium so that no district or union wants to renegotiate wages given the other bargaining outcomes. Specifically, salaries are set to maximize the generalized Nash Product conditional on other negotiated

We would search over ρ and γ to minimize this criterion function. While this procedure is consistent, we expect it to give very similar results as nearly all schools are on the labor demand curve.



Notes: The graph shows the labor supply and demand curves for Millcreek Township School District (S-MI) in Erie County, Pennsylvania. The demand curve comes directly from Equation 10. The labor supply curve is determined by keeping wages and offers fixed at all other schools, while allowing S-MI to give offers out to everyone. Then we sum up how many teachers are willing to work at S-MI at each wage. Wages at other schools were kept fixed at the wages seen in the data.

Figure 7: Labor Supply and Labor Demand

salaries w_{-j}:

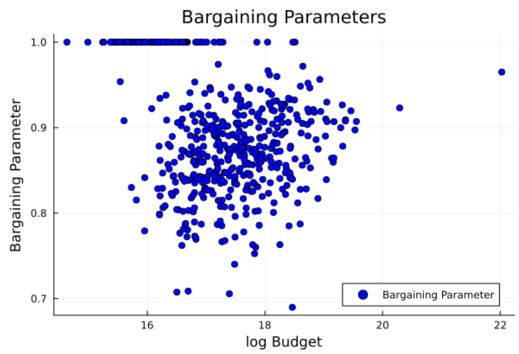
$$\begin{aligned}\mathcal{N}(w_j|w_{-j}) &= [W(T_j, B_j - w_j T_j))]^{\alpha_b(j)} \, [T_j(w_j - r_o)]^{1-\alpha_b(j)} \\ &= [W(T^*(w_j, w_{-j}), B_j - w_j T^*(w_j, w_{-j}))]^{\alpha_b(j)} \times \\ &\quad [T^*(w_j, w_{-j})(w_j - r_o)]^{1-\alpha_b(j)}\,,\end{aligned} \tag{14}$$

where the parameters $\alpha_b(j)$, $0 \leq \alpha_b(j) \leq 1$ determine the bargaining power of district j. The payoff to the district is given by the school's objective function $W(T, X)$. In the event of disagreement, the district hires no teachers and gets a payoff of 0. The payoff to the union is the earnings of teachers in the district $T_j w_j$.[25][26] In the event of disagreement, the union payoff is $T_j r_0$, where we set $r_0 = 30,000$.[27]

[25] In each negotiation, the union only puts weight on the teachers hired at the school district with which it is negotiating. Collard-Wexler et al. (2019) show that, given this assumption, a union negotiating simultaneously or sequentially leads to the same equilibrium outcomes.

[26] We have also experimented with other union objective functions, such as w_j. We found that sensible union objective functions, in which unions prioritize high wages, lead to similar results. The key element in the union's objective function is to ensure that the union cares more about wages, and the school cares more about hiring additional teachers.

[27] Note that this is not the same r that is the outside option wage in the labor supply.



Notes: The figure shows the relationship between log school budget and the estimated bargaining parameter for all schools in Pennsylvania. Charter schools have bargaining parameters equal to one.

Figure 8: Estimated Bargaining Power Parameters and School Budget

The first-order condition of Equation 14 with respect to w_j is:

$$\frac{\partial \mathcal{N}_j}{\partial w_j} = \frac{\alpha_b(j)\frac{\partial W_j(X_j,T_j)}{\partial w_j}}{W_j(X_j,T_j)} + \frac{(1-\alpha_b(j))}{T_j}\frac{\partial T_j}{\partial w_j} + \frac{(1-\alpha_b(j))}{w_j - r_0} = 0. \tag{15}$$

The derivative, $\frac{\partial T_j}{\partial w_j}$, depends on whether labor demand or labor supply binds. Rearranging this equation allows the bargaining parameter to be expressed as a function of data and our estimated labor demand and labor supply parameters:

$$\alpha_b(j) = \frac{W_j(X_j,T_j)(T_j + (w_j - r)\frac{\partial T_j}{\partial w_j})}{(W_j(X_j,T_j))(T_j + (w_j - r)\frac{\partial T_j}{\partial w_j}) - (T_j\frac{\partial W_j}{\partial X_j} + \frac{\partial T_j}{\partial \omega_j}\left[\frac{\partial W_j}{\partial T_j} - \omega_j\frac{\partial W_j}{\partial X_j}\right])T_j(\omega_j - r_o)}. \tag{16}$$

Thus, we can recover each bargaining power parameter, $\alpha_b(j)$. Since charter schools are not unionized, we assume they post wages and set their bargaining parameter to 1.

Figure 8 presents our estimates of the bargaining power parameter for each school district on the vertical axis, and school budget on the horizontal axis. The unweighted correlation between budget and bargaining parameters is 0.01 and the weighted correlation is 0.41.[28] The lack of strong correlation indicates that wealthy schools are not significantly stronger bargainers than

[28] The top five schools make up 13% of the weight and account for the majority of the correlation.

other schools. Thus, our results are not being driven by budget or size of schools.

5.7 Indirect Inference Estimation

We estimate all but two parameters using regressions and first order conditions directly from the data. The remaining two parameters, σ and r, which appear in the labor supply model, are estimated via indirect inference. In particular, the parameter σ governs the importance of the random coefficient ν_i that affects the choice between teaching and not teaching. This should be estimated separately from the ϵ terms that determine how teachers choose between schools. The parameter, r, meanwhile, is the value of the outside option and rationalizes the proportion of teachers that choose the outside option. Since we do not observe the salaries of teachers working outside teaching, we estimate r through indirect inference.

We begin by estimating the model *conditional* on (σ, r). We then form moments of the model given $\mathbb{M}(\sigma, r)$ that we use to form an indirect inference criterion given by the usual GMM formula, that is

$$Q(\sigma, r) = \left(\mathbb{M}(\sigma, r) - \hat{\mathbb{M}}\right)' \mathbb{W} \left(\mathbb{M}(\sigma, r) - \hat{\mathbb{M}}\right),$$

where $\mathbb{W}$ is a weighting matrix, in particular the identity matrix. We choose the σ and r that minimize $Q(\sigma, r)$.

Our moments to match are the number of teachers hired and wages predicted by the model for each school district. Since we have fixed the bargaining parameter for charter schools to one, only labor supply and labor demand parameters predict the wages set by charter schools (conditional on the wages of other school districts). We compute standard errors by bootstrapping the moment matrix $\hat{\mathbb{M}}$ at the school district level.[29] For computational convenience, we restrict our procedure to Erie County in Pennsylvania.[30]

Table 8 presents the estimates of σ and r by indirect inference. In our model, where the majority of teachers are employed, σ primarily plays a role in determining how difficult it becomes to hire teachers as schools go up the supply curve (i.e., the slope of the supply curve). Higher sigmas make it

[29] Standard errors will be in the next draft.

[30] We have tested various (σ, r) pairs in other regions of Pennsylvania and found that the same range of parameters does well everywhere. For example, in the Harrisburg region of Pennsylvania, we found that $(12, 105k)$ was optimal.

Parameter	Estimate
σ	13.5
r	102,000

Table 8: Indirect Inference Estimates

more difficult to attract teachers from the outside option. Meanwhile, the reservation value, r which is common across all teachers, affects the level of the supply curve. The higher the r, the more desirable is the outside option. Although r may seem extraordinarily large for an outside option, this is because it includes the intrinsic value of being a teacher and incorporates switching costs that could prevent teachers from moving.

5.8 Labor Supply Elasticities with Constrained Offers

We revisit the labor supply elasticities while conditioning on the set of offers made to teachers in equilibrium. The panel labeled *Constrained Offers* in Table 9 presents labor supply elasticities given equilibrium job offer sets. The top panel contains the same results as presented in Table 6 for ease of comparison. Three observations are worth noting. First, the elasticity of supply to each school is lower when jobs are rationed, as expected because teachers have fewer options to switch to given a salary reduction. On average, the own-salary elasticites are 73 percent of elasticities when schools make offers to all teachers. Relatedly, the elasticity of the outside option with respect to each school's salary increases. Second, a wage change can have no impact on supply to other schools when the teachers impacted by the wage change do not get offers from potential substitute schools. Third, charter schools become closer substitutes to regular school districts under restricted offers. In equilibrium, this allows charter schools to pay lower salaries.

6 Counterfactuals

6.1 Model Validation

Before showing counterfactual exercises, we first assess different measures of how well our model fits the data. Table 10 presents average salary and

	S-ER	S-MI	S-NE	S-NW	C-MO	C-RO	Out Opt
All Offers							
S-ER	5.55	−0.053	−0.049	−0.047	−0.053	−0.054	−0.006
S-MI	−0.529	6.001	−0.461	−0.485	−0.553	−0.506	−0.059
S-NO	−0.324	−0.305	5.749	−0.318	−0.307	−0.347	−0.039
S-NO	−0.175	−0.179	−0.178	5.397	−0.179	−0.173	−0.022
C-MO	−0.003	−0.003	−0.003	−0.003	4.369	−0.003	−0.0
C-RO	−0.026	−0.024	−0.025	−0.023	−0.025	5.153	−0.003
Out Opt	−0.319	−0.322	−0.319	−0.329	−0.322	−0.318	1.485
Constrained Offers							
	S-ER	S-MI	S-NO	S-NO	C-MO	C-RO	Out Opt
S-ER	3.522	−0.528	−0.661	−0.988	−2.071	−2.115	−0.204
S-MI	−0.492	1.863	0.0	0.0	−0.449	−0.513	−0.246
S-NO	−0.147	0.0	2.907	0.0	−0.159	−0.139	−0.055
S-NO	−0.19	0.0	0.0	3.506	−0.167	−0.196	−0.048
C-MO	−0.127	−0.03	−0.044	−0.053	4.244	−0.129	−0.012
C-RO	−1.083	−0.282	−0.319	−0.523	−1.076	4.076	−0.104
Out Opt	−1.204	−1.561	−1.471	−1.483	−1.198	−1.204	2.644

Notes: The table shows own- and cross-elasticities of a subset of schools in Erie County when teachers get offers from every school district. S refers to a school district, while C refers to a charter school. See the map in Figure 12 for the location of the schools. Each element shows the percentage increase in teachers at the schools in the columns willing to work after a 1% increase in the wage paid by the schools in the rows.

Table 9: Elasticities Erie

total number of teachers hired in the data and under Collective Bargaining for the entire state of Pennsylvania. Our model predicts average wages of \$ 55,904 versus average wages of \$ 54,319 in the data. Likewise, our model predicts that 106,948 teachers will be hired, which is close to the 104,748 that were actually hired. Moving beyond averages, Figure 9a plots Collective Bargaining wages predicted by the model against the wages in the data. The correlation between predicted and real wages is 0.58. Figure 9b shows that the fit is substantially higher for districts on the labor demand curve, with a correlation coefficient of 0.86, versus schools on the labor supply curve where the correlation is 0.23. When labor demand binds, the Nash bargaining weights absorb much of the wage variation. In the labor supply binding case, the majority of schools are charter schools, which post wages. In addition, the model has difficulty explaining why districts with large free lunch shares do not pay better.

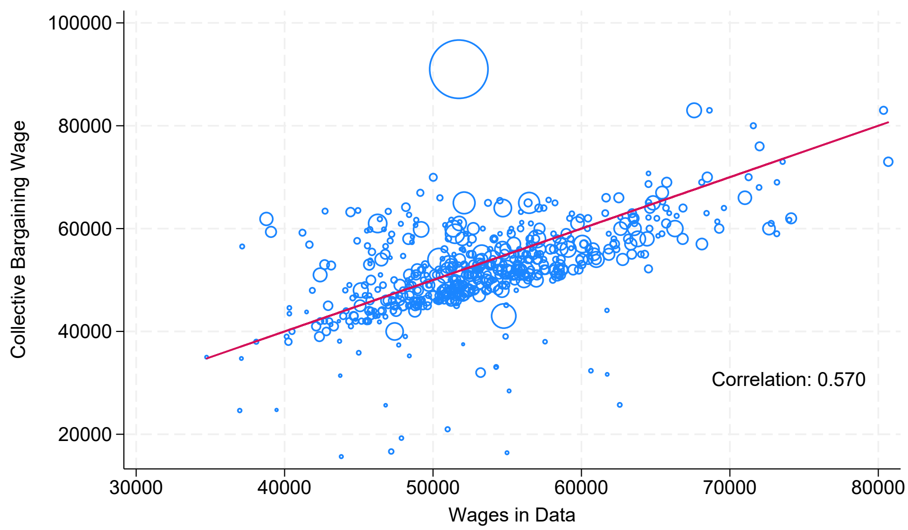
As the number of teachers hired by a district is strongly correlated with budgets, the model does a better job of explaining the number of teachers hired. The correlation between the model's prediction and actual teachers hired is essentially 1. More interesting is that our model predicts a student to teacher ratio of 15.8 versus 14.9 in the data.

6.2 Welfare of Collective Bargaining and Oligopsony

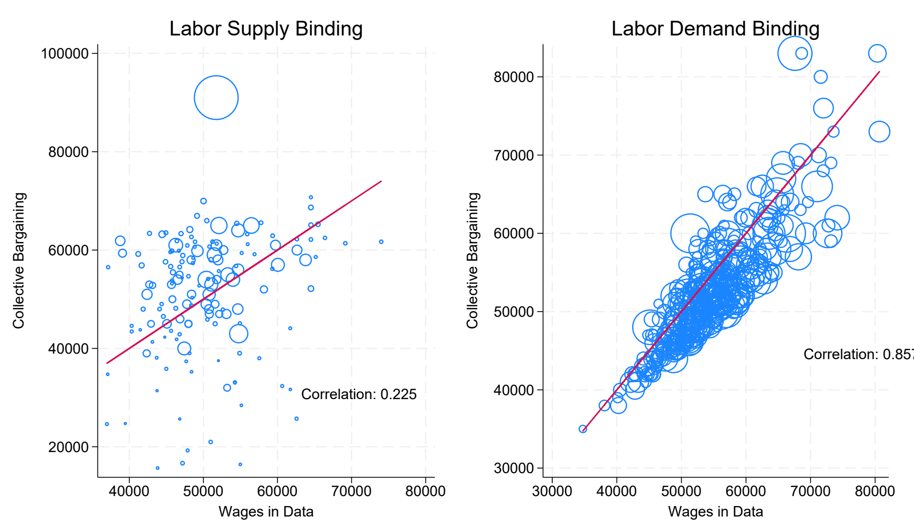
Our first counterfactuals address the challenge posed in section 2; i.e., we compute the social planner equilibrium and compare it to the collective bargaining setting. We also compare the planner and collective bargaining setting to the posted wage setting, which determines the equilibrium under pure oligposony. Thus, we can separately understand how oligopsony and collective bargaining distort the market.

In the oligopsony and social planner solutions, all teachers receive offers from every school; otherwise, schools and the social planner leave welfare on the table.[31] Note that the oligopsony solution is nested inside the Nash-in-

[31] Consider the oligopsony scenario. Consider a school j that does not give an offer to teacher, $\tilde{i}$. If this teacher would receive an offer and accept with probability $s_{\tilde{i}j} > \epsilon$, then the school district could hire T_j teachers by making this additional offer to $\tilde{i}$, and lowering wages to $\tilde{w}_j = w_j - \delta$ for some small δ, and still hire T_j teachers. Thus, not making an offer to $\tilde{i}$ cannot maximize school welfare W_j, which is a contradiction. The proof for the planner problem proceeds in a similar way with the observation that there are values of the ϵ shock such that this teacher $\tilde{i}$ has the highest value for a job at school j, and thus,



(a) Predicted and Actual Wages



(b) Predicted and Actual Wages

Nash solution for $\alpha_j^b = 1$ for all j.

The second through fourth rows of Table 10 present the predictions of the model both in terms of average wages and the total number of employed teachers, for Collective Bargaining, the Social Planner, and Oligopsony with posted wages.

We find that the planner would set wages of \$ 56,042, which is above the oligopsony model prediction of wages of \$ 52,329, or a seven percent decrease in wages due to oligopsony power. Associated with this decrease in wages is a decrease in the number of teachers from 113,237 to 112,201, or a 1 percent fall. As teacher supply is fairly inelastic in terms of substitution between teaching and the outside job rather than between schools, a seven percent drop in wages only results in a one percent change in employment. Thus, the oligopsony distortion mostly leads to wage transfers from teachers to schools, rather than large quantity distortions and accompanying deadweight loss. Second, the average elasticity of teacher supply is $\epsilon_S = -5$, which would imply a markdown in the monopsonist model in section 2 of 20 percent. However, wages are not depressed by this amount because labor demand is also fairly inelastic, with a labor demand elasticity of $\epsilon_D = -0.72$. Thus, to understand the effect of monopsony power on wages, at least relative to the efficient level, it is important to understand wage sensitivity on both supply and demand. In addition, the wage distortion due to oligopsony power is fairly uniform across school districts in the data.

The Collective Bargaining model (or Nash-in-Nash Bargaining) predicts wages of \$ 55,904 which is close to the social planner wage of \$ 56,042, and far above the situation where there are no unions, where the posted wage would be \$ 52,329 on average. This means that collective bargaining is nearly sufficient to eliminate the oligopsony distortion on wages, at least on average. However, the Collective Bargaining model leads to substantially fewer teachers being employed, with 106,948 hired teachers, versus 112,201 in the Posted solution, or a 5 percent decrease in employment given a 6 percent increase in wages. This is because under Collective Bargaining rationing occurs in many districts.

Considering the entire distribution of salaries in the two scenarios reveals a more complete picture. Figure 10 presents the collective bargaining wages on the vertical axis, and the oligopsony (or no-union) wages on the

not offering them a job would not maximize total surplus. This is because schools do not care about which teacher accepts a job, so higher utility of accepting a job is enough.

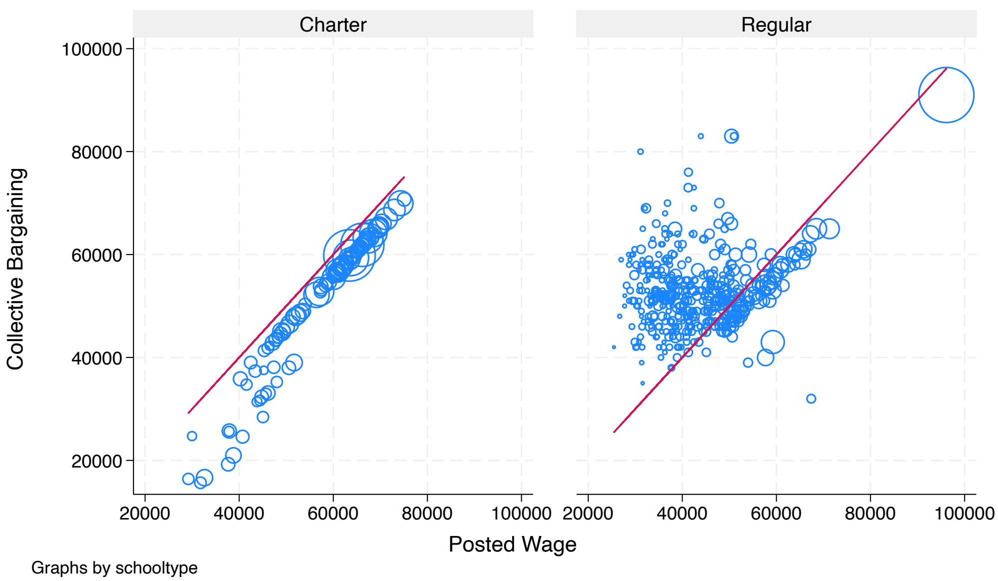
horizontal axis for all the 525 school districts in our data. The 45 degree line shows the case where these two wages would be the same. A striking fact is that Collective Bargaining leads to increased wages at 333 schools, but *decreased* wages at 194 schools. The direct effect of collective bargaining increasing wages is simple to understand: increased bargaining power for teachers unions increases wages. Thus, eliminating unions should decrease wages, since it decreases bargaining weight for teacher unions given by $1-\alpha_j$. However, eliminating unions also means that all teachers will now receive offers, and this will make the labor market for teachers more competitive, as was shown in Table 9 comparing labor supply elasticities with limited offers versus unconstrained offers. Consider the set of charter schools who are assumed to be wage posting before and after unions are eliminated. Figure 10 shows that banning unions *raises* the wages for all charter schools. In other words, by banning unions, previously unionized school districts lower wages and expand hiring making the labor market for teachers more competitive. This pushes wages up. It is because of the opposing forces coming from bargaining and equilibrium that banning unions has such a modest effect on wages.

Model	Mean Wage	Teachers Hired
Data	54,103	104,748
Collective Bargaining	55,904	106,976
Posted	52,329	112,201
Planner	56,042	113,237

Notes: Average Wage over all teachers shown. Posted corresponds to the situation where employers set wages with no unions, planner to the social planner solution, where labor demand meets labor supply, and Collective Bargaining to the situation where school districts and teacher unions bargain collectively over wages.

Table 10: Welfare, Collective Bargaining and Oligopsony Distortion

Collective bargaining substantially changes the distribution of wages. In both the posted wage equilibrium and planner problem, which have highly correlated predicted wages, with a correlation coefficient of 0.99, wages would be higher at both charter schools and schools that have a substantial fraction of poor students as measured by the free lunch share. Wages compensate teachers for lower amenity value at these schools. In contrast, in both the data and the collective bargaining model, we do not observe a positive correlation between the factors that induce teachers to quit in Table 5 and wages.



Notes: The figure shows the model predicted wages from collective bargaining on the y-axis and the predicted wages from the posted wage scenario on the x-axis. Circle size is proportional to the number of teachers at the school. The 45 degree line is plotted in red.

Figure 10: Collective Bargaining and Posted Wage Predictions

Instead, the collective bargaining model rationalizes this pattern by assigning high bargaining power to schools that pay relatively low wages.

7 Conclusion

Oligopsony power is likely present in many labor markets, but is exacerbated in the public sector labor market for the simple reason that the government is the sole employer. In response, the public sector has the highest percentage of workers represented by unions in the American economy. This pattern also holds in other countries such as Canada where 77% of public sector workers are unionized.[32]

We find potential for substantial monopsony power in the market for public schoolteachers in Pennsylvania, with wages under a pure wage posting arrangement 7% percent below the social planner solution on average. However, collective bargaining allows public teachers to have countervailing power that pushes their wages 6.8% higher on average than in the oligopsony scenario, and very close to average salaries under the social planner. Thus, we find that collective bargaining can correct the distortions from oligopsony. The effect of collective bargaining, however, depends crucially on bargaining power. In some cases, where bargaining power is too strong, collective bargaining could add an additional distortion that pushes wages too high. In particular, there is a vector of bargaining parameters between 0 and 1 that would lead to the socially optimal equilibrium. In the schoolteacher environment, the union's bargaining power seems almost optimal since the collective bargaining wages match the planner wages.

Much of the wage differences that we find are not about greater or lower monopsony power, but instead about the ability to negotiate higher or lower wages. Moreover, raising wages, say due to a higher bargaining power parameter, has ambiguous effects on workers, since either labor supply or labor demand might be binding. In particular, newly unemployed workers might put downward pressure on wages at rival school districts. More vexing is the implication that extrapolation from existing empirical studies on the impact of unions on wages might be quite difficult.

Further counterfactuals are being developed, such as an evaluation of

[32] See Statistics Canada. Table14-10-0132-01https://doi.org/10.25318/1410013201-eng Union status by industry.

statewide uniform wage settings and a comparison of minimum salary regulations versus collective bargaining.

Overall, this paper highlights the challenges of predicting the effect of policies in the labor market when oligopsony power and unions are present, especially when complex externalities between employers are taken into account.

References

Baker, Michael, Yosh Halberstam, Kory Kroft, Alexandre Mas, and Derek Messacar. 2024. "The Impact of Unions on Wages in the Public Sector: Evidence from Higher Education." Working Paper 32277, National Bureau of Economic Research.

Bates, Michael, Michael Dinerstein, Andrew C Johnston, and Isaac Sorkin. 2025. "Teacher labor market policy and the theory of the second best." *Quarterly Journal of Economics*.

Berger, David, Kyle Herkenhoff, and Simon Mongey. 2022. "Labor market power." *American Economic Review* 112 (4): 1147–1193.

Biasi, Barbara, Chao Fu, and John Stromme. 2021. "Equilibrium in the market for public school teachers: District wage strategies and teacher comparative advantage."Technical report, National Bureau of Economic Research.

Blanchflower, David G., and Alex Bryson. 2010. "The Wage Impact of Trade Unions in the UK Public and Private Sectors." *Economica* 77 (305): 92–109.

Blanchflower, David G, and Alex Bryson. 2024. "Unions, wages and hours." Working Paper 32471, National Bureau of Economic Research.

Card, David. 1996. "The Effect of Unions on the Structure of Wages: A Longitudinal Analysis." *Econometrica* 64 (4): 957–979, http://www.jstor.org/stable/2171852.

Card, David, Thomas Lemieux, and W. Craig Riddell. 2020. "Unions and wage inequality: The roles of gender, skill and public sector employment." *Canadian Journal of Economics/Revue canadienne d'économique* 53 (1): 140–173.

Collard-Wexler, Allan, Gautam Gowrisankaran, and Robin S Lee. 2019. ""Nash-in-Nash" bargaining: a microfoundation for applied work." *Journal of Political Economy* 127 (1): 163–195.

Crawford, Gregory S., Rachel Griffith, and Alessandro Iaria. 2021. "A survey of preference estimation with unobserved choice set heterogeneity." *Journal of Econometrics* 222 (1): 4–43, https://EconPapers.repec.org/RePEc:eee:econom:v:222:y:2021:i:1:p:4-43.

Crawford, Gregory S, and Ali Yurukoglu. 2012. "The welfare effects of bundling in multichannel television markets." *American Economic Review* 102 (2): 643–685.

Delabastita, Vincent, and Michael Rubens. 2025. "Colluding against workers." *Journal of Political Economy.*

Demsetz, Harold. 1973. "Industry structure, market rivalry, and public policy." *The Journal of Law and Economics* 16 (1): 1–9.

DiNardo, John, Nicole Fortin, and Thomas Lemieux. 1996. "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica* 64 (5): 1001–44, https://EconPapers.repec.org/RePEc:ecm:emetrp:v:64:y:1996:i:5:p:1001-44.

DiNardo, John, and David S. Lee. 2004. "Economic Impacts of New Unionization on Private Sector Employers: 1984–2001." *The Quarterly Journal of Economics* 119 (4): 1383–1441, https://EconPapers.repec.org/RePEc:oup:qjecon:v:119:y:2004:i:4:p:1383-1441.

Dodini, Samuel, Kjell Salvanes, and Alexander Willen. 2021. "The Dynamics of Power in Labor Markets: Monopolistic Unions versus Monopsonistic Employers." *CESifo Working Paper No. 9495.*

Dunlop, John T. 1944. *Wage Determination Under Trade Unions.*. New York, NY: Macmillan.

Farber, Henry S. 1978. "Individual preferences and union wage determination: the case of the united mine workers." *Journal of Political Economy* 86 (5): 923–942.

Farber, Henry S. 1986. "The Analysis of Union Behavior." In *Handbook of Labor Economics*, edited by Ashenfelter, Orley, and Layard Volume 2. Chap. 18 1039–1089, North Holland Publishing.

Farber, Henry S, Daniel Herbst, Ilyana Kuziemko, and Suresh Naidu. 2021. "Unions and Inequality over the Twentieth Century: New Evidence from Survey Data*." *The Quarterly Journal of Economics* 136 (3): 1325–1385.

Freeman, Richard, and James Medoff. 1984. *What Do Unions Do.* N.Y.: Basic Books, , http://www.amazon.com/What-Do-Unions-Richard-Freeman/dp/0465091334/ref=sr_1_1?s=books&ie=UTF8&qid=1305753795&sr=1-1.

Galbraith, John Kenneth. 1954. "Countervailing Power." *The American Economic Review* 44 (2): 1–6, http://www.jstor.org/stable/1818317.

Gottfries, Axel, and Gregor Jarosch. 2023. "Dynamic Monopsony with Large Firms and Noncompetes."Technical report, National Bureau of Economic Research.

Gowrisankaran, Gautam, Aviv Nevo, and Robert Town. 2015. "Mergers when prices are negotiated: Evidence from the hospital industry." *American Economic Review* 105 (1): 172–203.

Green, David A, Benjamin Sand, and Iain G Snoddy. 2022. "The impact of unions on nonunion wage setting: Threats and bargaining."Technical report, Working Paper UBC.

Grennan, Matthew. 2013. "Price discrimination and bargaining: Empirical evidence from medical devices." *American Economic Review* 103 (1): 145–177.

Ho, Kate, and Robin S Lee. 2017. "Insurer competition in health care markets." *Econometrica* 85 (2): 379–417.

Horn, Henrick, and Asher Wolinsky. 1988. "Bilateral monopolies and incentives for merger." *The RAND Journal of Economics* 408–419.

Hoxby, Caroline. 1996. "How Teachers' Unions Affect Education Production." *The Quarterly Journal of Economics* 111 (3): 671–718, https://EconPapers.repec.org/RePEc:oup:qjecon:v:111:y:1996:i:3:p:671-718.

Lee, Robin S, and Kyna Fong. 2013. "Markov perfect network formation: An applied framework for bilateral oligopoly and bargaining in buyer-seller networks."

Lemieux, Thomas. 1998. "Estimating the Effects of Unions on Wage Inequality in a Panel Data Model with Comparative Advantage and Nonrandom Selection." *Journal of Labor Economics* 16 (2): 261–91, https://EconPapers.repec.org/RePEc:ucp:jlabec:v:16:y:1998:i:2:p:261-91.

Lewis, H. Gregg. 1963. *Unionism and Relative Wages in the United States.*. Chicago, IL: University of Chicago Press.

Lovenheim, Michael F. 2009. "The Effect of Teachers' Unions on Education Production: Evidence from Union Election Certifications in Three Midwestern States." *Journal of Labor Economics* 27 (4): 525–587.

McFadden, D. 1978. "Modeling the choice of residential location." *Transportation Research Record* 672 72–77.

Miller, Nathan et al. 2022. "On the misuse of regressions of price on the HHI in merger review." *Journal of Antitrust Enforcement* 10 (2): 248–259.

Ransom, Michael, and David P. Sims. 2010. "Estimating the Firm's Labor Supply Curve in a "New Monopsony" Framework: Schoolteachers in Missouri." *Journal of Labor Economics* 28 (2): 331–355, https://EconPapers.repec.org/RePEc:ucp:jlabec:v:28:y:2010:i:2:p:331-355.

Robinson, Joan. 1933. *The economics of imperfect competition.*

Rubens, Michael. 2023. "Management, productivity, and technology choices: evidence from US mining schools." *The RAND Journal of Economics* 54 (1): 165–186.

Sokolova, Anna, and Todd Sorensen. 2021. "Monopsony in Labor Markets: A Meta-Analysis." *ILR Review* 74 (1): 27–55.

A Data Construction

A.1 Potential Teachers

Unfortunately, we do not have data on the number of teachers who are qualified and could teach, but are not teaching. We refer to this group as the potential teachers. Instead, we use a back of the envelope condition to get the number of potential teachers using data on certifications and the number of teachers who leave the profession. We find that the potential teachers are about 30.9% of teachers. We do sensitivity analysis on this number and show that it does not have a large effect on results. After calculating the percentage of potential teachers, we assign addresses to these teachers by sampling with replacement from our address dataset.

A.1.1 Calculating the Share of Potential Teachers

There are two types of non-teachers. There are teachers who get certified and never teach and there are teachers who exit teaching. We use both of these groups.

Certifications/New Teachers: The state of Pennsylvania publishes data on the number of people who have received teacher certificates from 2010 to the present.

Teachers need an Instructional I certificate to be licensed teachers.[33] Once received, it lasts for 6 years. The number of Instructional I certificates issued gives us a good sense of how many new teachers are qualified to work in Pennsylvania and have put in effort to pursue a teaching career. Even though certificates are valid for 6 years, we assume that after 4 years a teacher who does not teach will never teach and has exited the potential pool.

Next, we calculate the number of teachers who started teaching (i.e., have years of experience equal to one) in each year. By subtracting the number of starting teachers from the number of newly licensed teachers, we can get a value of the number of teachers in each year who are licensed, but not teaching. In 2017, this adds 17,179 potential teachers.

Exiting Teachers: The next part of the pool of non-teachers are teachers who exit teaching. Of course, there are many teachers who retire and who would no longer be part of the pool of non-teachers. To prevent

[33] There are some emergency certificates that are given out when there is a shortage of teachers. We ignore these.

this possibility, we assume that if a teacher has greater than 20 years of experience in education, they permanently exit the pool of non-teachers. We assume that after 5 years out of the labor market, the teacher will not return to teaching.

For each year with longitudinal IDs, we calculate the number of teachers who permanently exit and have less than 20 years of experience. For each additional year until five years, I add the exited teachers to the pool as long as their years of experience had they continued teaching would not be greater than 20. For 2017, this adds 17,214 teachers to the pool. We also add teachers who take gap years. This is a negligible amount for each year (between 1k-2k), and only 1,205 teachers in 2017.

Results: For 2017, after summing these values, we get that there are 35,598 non-teachers. Our data contain 119,141 total teachers in 2017, but only 106,941 are considered in the final sample.[34] Thus, we add the additional 12,200 teachers. We get that the share of non teachers is $\frac{47{,}798}{47{,}798+106{,}941} = 0.309$ of all teachers. We believe this number matches well to other statistics. The number of Pennsylvania private school teachers in 2017 was 23,630, which is about half of our non-teachers.[35] In addition, the Bureau of Labor Statistics says that the number of teachers in Pennsylvania is 152,980, which would imply that teachers working outside the school system would be a 22.1% share of the total teachers.[36]

A.1.2 Potential Teachers in the Data

For the structural model, we need to add potential teachers to our teacher data. In the model, teachers enter as addresses. We have collected teacher address information using InfoUSA. We are not able to get all of the addresses since some people were not merged or were double counted in the InfoUSA data. However, if we use unique addresses for teachers from 2010-2021, we have 151,890 unique addresses. We use all of these unique addresses. When using a duplicate share of 0.309, we need 154,739 teachers (since there are a

[34] We did not include teachers who worked at Technical Career High schools, penitentiary schools, or were hired by the intermediate unit rather than a school district. Finally, we did not include teachers who worked at schools which had missing data.

[35] https://nces.ed.gov/surveys/pss/tables1314.asp

[36] https://web.archive.org/web/20250129185820/https://www.bls.gov/oes/tables.htm. This number includes elementary, middle, and secondary school teachers, along with special education teachers and career technical education teachers. We do not include preschool teachers

total of 106,941 teachers hired at the schools in our sample). Thus, there are another 2,849 teachers that are unaccounted for.

In order to get additional teacher addresses, we clone teacher addresses to make up the remaining teachers. We sample all of the teacher addresses with replacement. Since this is a very small share of the total addresses that we have, we do not believe that cloning the teachers has a significant impact on our results.

A.2 InfoUSA Teacher Matching

We use InfoUSA data from 2009-2023 that includes names and addresses of people living in Pennsylvania. InfoUSA compiled this data from public sources, voter registration data, utility company data, and real estate data. We match our public school data to the Infousa data using exact name and year. About 22% of the sample had no match, 32.9% of the sample had a unique match, and 45.1% had multiple matches. We only include unique matches and make the assumption that teachers with unique names do not have residential patterns that are very different from their peers.

Table 11 shows summary statistics across the three match possibilities. Teachers that were matched once or multiple times have very similar statistics. However, there seems to be selection for teachers who were not matched. These teachers tend to have less experience, lower pay, and are more likely to work for charter schools.

B Proofs and Derivations

B.1 Nash Bargaining Derivations

B.2 Union utility function: w*T

The problem is:

$$\max_{w_j} \quad \Big(W_j(X_j, T_j)\Big)^{\alpha_j} \Big(T_j(w_j - r)\Big)^{1-\alpha_j}$$
$$s.\ t.\ B_j = w_j T_j + X_j.$$

	Unique	None	Multiple
Percent Female	0.79	0.84	0.66
	(0.40)	(0.37)	(0.47)
Percent White	0.95	0.91	0.95
	(0.22)	(0.28)	(0.21)
Percent w/ Masters	0.59	0.48	0.57
	(0.49)	(0.50)	(0.50)
Avg Years in Education	14.24	10.84	14.40
	(8.30)	(8.40)	(8.43)
Charter Schools	0.06	0.12	0.06
	(0.24)	(0.32)	(0.24)
Average Wage	69,849.90	62,936.39	69,871.95
	(17,481.44)	(17,196.92)	(17,329.97)
Obs	42,664	16,859	59,618

Notes: The table shows summary statistics for teachers who were matched once, zero, or more than once to the InfoUSA data in 2017.

Table 11: 2017 Statistics Of Teachers Matched to InfoUSA

Let's starting by transforming the problem into logs:

$$\alpha_j log\Big(W_j(X_j,T_j)\Big) + (1-\alpha_j)log\Big(T_j\Big) + (1-\alpha_j)log\Big(w_j - r\Big)$$

FOC:

$$\frac{\alpha_j \frac{\partial W_j(X_j,T_j)}{\partial w_j}}{W_j(X_j,T_j)} + \frac{(1-\alpha_j)}{T_j}\frac{\partial T_j}{\partial w_j} + \frac{(1-\alpha_j)}{w_j - r} = 0$$

We use the same trick as above and use:

$$m = (-k_1 T_j \frac{\partial W_j}{\partial X_j} + \frac{\partial T_j}{\partial \omega_j}\Big[\frac{\partial W_j}{\partial T_j} - (k_2 + k_1\omega_j)\frac{\partial W_j}{\partial X_j}\Big]).$$

Then we have:

$$\frac{\alpha_j m}{W_j(X_j,T_j)} + (1-\alpha_j)[\frac{1}{w_j - r} + \frac{1}{T_j}\frac{\partial T_j}{\partial w_j}] = 0$$

Solving step by step to get α's:

$$
\begin{aligned}
\frac{\alpha_j m}{W_j(X_j,T_j)} &= -(1-\alpha_j)[\frac{1}{w_j - r} + \frac{1}{T_j}\frac{\partial T_j}{\partial w_j}], \\
\frac{\alpha_j m}{W_j(X_j,T_j)} &= -(1-\alpha_j)\frac{T_j + (w_j - r)\frac{\partial T_j}{\partial w_j}}{T_j(w_j - r)}, \\
\alpha_j m T_j(w_j - r) &= -(1-\alpha_j)W_j(X_j,T_j)(T_j + (w_j - r)\frac{\partial T_j}{\partial w_j}), \\
\alpha_j m T_j(w_j - r) - \alpha_j W_j(X_j,T_j)(T_j + (w_j - r)\frac{\partial T_j}{\partial w_j}) &= -W_j(X_j,T_j)(T_j + (w_j - r)\frac{\partial T_j}{\partial w_j}), \\
\alpha_j(m T_j(w_j - r) - W_j(X_j,T_j)(T_j + (w_j - r)\frac{\partial T_j}{\partial w_j})) &= -W_j(X_j,T_j)(T_j + (w_j - r)\frac{\partial T_j}{\partial w_j}), \\
\alpha_j &= \frac{-W_j(X_j,T_j)(T_j + (w_j - r)\frac{\partial T_j}{\partial w_j})}{m T_j(w_j - r) - W_j(X_j,T_j)(T_j + (w_j - r)\frac{\partial T_j}{\partial w_j})}, \\
\alpha_j &= \frac{W_j(X_j,T_j)(T_j + (w_j - r)\frac{\partial T_j}{\partial w_j})}{(W_j(X_j,T_j))(T_j + (w_j - r)\frac{\partial T_j}{\partial w_j}) - m T_j(w_j - r)}, \\
\alpha_j &= \frac{W_j(X_j,T_j)(T_j + (w_j - r)\frac{\partial T_j}{\partial w_j})}{(W_j(X_j,T_j))(T_j + (w_j - r)\frac{\partial T_j}{\partial w_j}) - (-k_1 T_j \frac{\partial W_j}{\partial X_j} + \frac{\partial T_j}{\partial \omega_j}\Big[\frac{\partial W_j}{\partial T_j} - (k_2 + k_1\omega_j)\frac{\partial W_j}{\partial X_j}\Big])}
\end{aligned}
\tag{17}
$$

C Code Appendix

C.1 Labor Supply

The labor supply function $s_{ij}(\mathbf{w}, \mathcal{O})$, where $\mathbf{w} = \{w_j\}_{j=1}^{J}$, and $\mathcal{O} = \{o_{ij} \forall i, j\}$ is given by the following share:

$$
\begin{aligned}
s_{ij} &= \int_{\nu} \frac{o_{ij}\exp(\psi\omega_j + x_j'\beta - \tau d_{ij} + \sigma\nu)}{\exp(r) + \sum_{k=1}^{k} o_{ik}\exp(\psi\omega_j + x_j'\beta - \tau d_{ij} + \sigma\nu)}\phi(\nu)d\nu \\
&\approx \sum_{n=1}^{N} w_n \frac{o_{ij}\exp(\psi\omega_j + x_j'\beta - \tau d_{ij} + \sigma\nu_n)}{\exp(r) + \sum_{k=1}^{k} o_{ik}\exp(\psi\omega_j + x_j'\beta - \tau d_{ij} + \sigma\nu_n)}
\end{aligned}
\tag{18}
$$

where w_n are Gaussian Quadrature weights, and ν_n are Gaussian Quadrature nodes.

It will also be convenient to define the unconstrained offer supply function given by $s_{ij}(\mathbf{w})$, which is just the supply function if all workers get offers, that is $\mathcal{O}$ such that $o_{ij} = 1$ for all i and j.[37]

The firm makes offers to the closest workers first. Thus, there is an ordering of workers $\iota(j) = \{i_1, i_2, \cdots, i_I\}$ where $\iota(j)_5$, say, is the worker i that is the fifth closest to school j and receives the fifth offer.

Labor Supply is thus $L_j^S(\mathbf{w}, \mathcal{O}_{-j}) = \sum_{\iota(j)=1}^{I} s_{\iota(j)j}(\mathbf{w}, (o_{\iota(j)j} = 1, \mathcal{O}_{\iota(j),-j}))$. Labor supply if all firms post wages, i.e. everyone gets an offer, is just $L_j^S(\mathbf{w}) = \sum_{i=1}^{I} s_{ij}(\mathbf{w})$.

C.2 Labor Demand

Labor Demand $L_j^D(w_j)$ is given by:

$$L_j^D(w_j) = \frac{w_j^{\frac{1}{\rho-1}} B_j}{\left(\frac{\gamma}{1-\gamma}\right)^{\frac{1}{\rho-1}} + w_j^{\frac{\rho}{\rho-1}}}. \tag{19}$$

The elasticity of labor demand is given by:

$$\frac{\partial log(L_j^D)}{\partial log(w_j)} = \frac{1}{\rho - 1} - \frac{\frac{\rho}{\rho-1} w^{\frac{\rho}{\rho-1}}}{(\frac{\gamma}{1-\gamma})^{\frac{1}{\rho-1}} + w^{\frac{\rho}{\rho-1}}}. \tag{20}$$

C.3 Offer Game

To hire $L_j^D(w_j)$ workers, the firm needs to make offers to more than L^D workers, since some of these workers will choose to work at another school district, or pick the outside option. The school will offer jobs in sequence, and $\bar{\iota}(j)$ is the last worker that is offered a job.

Define the *best-response offer* from school j as $\mathcal{O}_j(\mathbf{w}, \mathcal{O}_{-j})$. Pick $\bar{\iota}_j(\mathbf{w})$ such that

$$L_j^D(w_j) = \sum_{\iota(j)=1}^{\bar{\iota}_j(\mathbf{w})} s_{\iota(j)j}(\mathbf{w}, (o_{\iota(j)j} = 1, \mathcal{O}_{\iota(j),-j})). \tag{21}$$

In practice, we use the following algorithm in Algorithm 1 below.

[37] We restrict offers to teachers within a 60-minute commute.

Algorithm 1 Offer Best-Response

procedure BEST-RESPONSE OFFER($\mathcal{O}_j(\mathbf{w}, \mathcal{O}_{-j})$)
  $\bar{\iota} = 0$
  $L_j^* = L^D(w_j)$.
  $T_j = 0$
  while $T_j \leq L_j^*$ or $\bar{\iota} \leq I$ **do**
    $\bar{\iota} = \bar{\iota} + 1$
    Update offer matrix $\mathcal{O}_{\bar{\iota},j} = 1$.
    $T_j = T_j + s_{\bar{\iota}(j),j}(\mathbf{w}, \mathcal{O})$
  end while
  return $\mathcal{O}_j$
end procedure

An equilibrium of this offer game is an offer matrix $\mathcal{O}^*$ such that offers are a best response, given the offers sent by other schools. That is $\mathcal{O}_j^* = \mathcal{O}_j(\mathbf{w}, \mathcal{O}_{-j}^*)$ for all j.

In practice our offer game algorithm does the following in Algorithm 2.

Algorithm 2 Offer Game

procedure OFFER GAME($\mathcal{O}(\mathbf{w})$)
  Start with initial offers sent to $\bar{\iota}_j = L^D(w_j)$ so that $\mathcal{O}_{ij} = 1$ for all j and for $i < \bar{\iota}_j$.
  $k = 0$ (iteration counter).
  while $|\mathcal{O}^k - \mathcal{O}^{k-1}| > 0$ **do**
    Update iteration counter $k = k + 1$
    for $j = 1, \cdots, J$ **do**
      Compute Best-Response for j: $\mathcal{O}_j' = \mathcal{O}_j(\mathbf{w}, \mathcal{O}_{-j}^k)$
      Update offer matrix for that school: $\mathcal{O}_j^k = \mathcal{O}_j'$
    end for
  end while
  return $\mathcal{O}$
end procedure

Marginal cost of Teacher	(1)	(2)
Average wage	1.486	1.420
	(0.030)	(0.037)
Constant	9022.611	5908.939
	(2527.594)	(2740.366)
Observations	531.0	5230.0

Notes: The table shows the regression of average wage on the marginal cost of a teacher. This regression allows us to estimate the fringe benefits from the wage. Column (1) shows the regression for just the year 2017. Column (2) shows the regression for all years with school and year fixed effects. Standard errors are clustered at the school district level.

Table 12: Fringe Benefits

C.4 Fringe Benefits

A school's cost of hiring a teacher is not just a wage, but also includes benefits. In our model, we set the cost of a teacher equal to wages plus a fringe benefit. In practice, we ran a regression of the average wage on the marginal cost of a teacher (expenditure on teachers divided by the number of teachers hired). Table 12

C.5 Planner Solution

The planner solution is simply the wage vector $\mathbf{w}$ such that labor supply and labor demand intersect. Define excess supply (excess demand) as $\xi_j(\mathbf{w}) = L_j^S(\mathbf{w}) - L_j^D(\mathbf{w})$. Thus, the planner solution $\mathbf{w}^*$ is such that $\xi_j(\mathbf{w}^*) = 0$ for all j.

In practice, we define the vector function $\xi(\mathbf{w}) \equiv \begin{bmatrix} \xi_1(\mathbf{w}) \\ \xi_2(\mathbf{w}) \\ \dots \\ \xi_J(\mathbf{w}) \end{bmatrix}$, and look for the zero of this vector function using a root finding algorithm.

C.6 Posted Prices Solution

For the posted price solution, each school district sets its marginal rate of technical substitution for teachers equal to its marginal factor cost:

$$\frac{\frac{\partial W}{\partial T}}{\frac{\partial W}{\partial x}} = \frac{\gamma}{1-\gamma} \cdot \left(\frac{T_j}{X_j}\right)^{\rho-1} = w_j(1+\frac{1}{\epsilon_j}). \tag{22}$$

We can define $\xi_j^{PO}(\mathbf{w}) = MRS_j - w_j(1+\frac{1}{\epsilon_j})$. Thus, the posted solution $\mathbf{w}^*$ is such that $\xi_j^{PO}(\mathbf{w}^*) = 0$ for all j.

In practice, we use a best-response algorithm. Schools make wage decisions holding other school wages fixed. We iterate until we reach convergence where $\xi_j^{PO}(\mathbf{w}^*) = 0$ for all j.

C.7 Nash-in-Nash Bargaining

For the Nash-in-Nash Bargaining solution, we first define how to find the wage that maximizes the Nash-Product. Find $w_j = w_j^{NB}(\mathbf{w}_{-j})$ which maximizes $\mathcal{N}_j(w_j^{NB}|w_{-j}^k)$.

Take the first-order condition of the Nash Product given by $\frac{\partial \mathcal{N}_j(w_j'|w_{-j}^k)}{\partial w_j} = 0$. Define the FOC $\xi^{NB}(w_j|\mathbf{w}_{-j}) \equiv= \frac{\partial \mathcal{N}}{\partial w_j}$, and we find $w_j^{NB}(\mathbf{w}_{-j})$ such that $\xi^{NB}(w_j^{NB}|\mathbf{w}_{-j}) = 0$.

A Nash-in-Nash equilibrium is defined as a vector of wages $\mathbf{w}^{NN}$ such that all schools $j \in \mathcal{J}$ are choosing wages that maximize their Nash Product given the wages of other schools, that is $w_j^{NN} = w_j^{NB}(\mathbf{w}_{-j}^{NN})$ for all $j \in \mathcal{J}$.

Again, in practice we use the following algorithm 3

Algorithm 3 Nash-in-Nash Bargaining

procedure NASH-IN-NASH BARGAIN($\mathbf{w}$)
    Start with inital wage guess $\mathbf{w}^0 = \{w_j^0\}_{j=1}^J$.
    Iteration Counter $k = 0$.
    while $|\mathbf{w}^k - \mathbf{w}^{k-1}| > \epsilon$ **do**
        for $j = 1, \cdots, J$ **do**
            Find w_j' such that $\frac{\partial \mathcal{N}_j(w_j'|w_{-j}^k)}{\partial w_j} = 0$.   ▷ This is a root finding exercise in one dimension.
            Update $w_j^k = w_j'$
        end for
    end while
    return w
end procedure

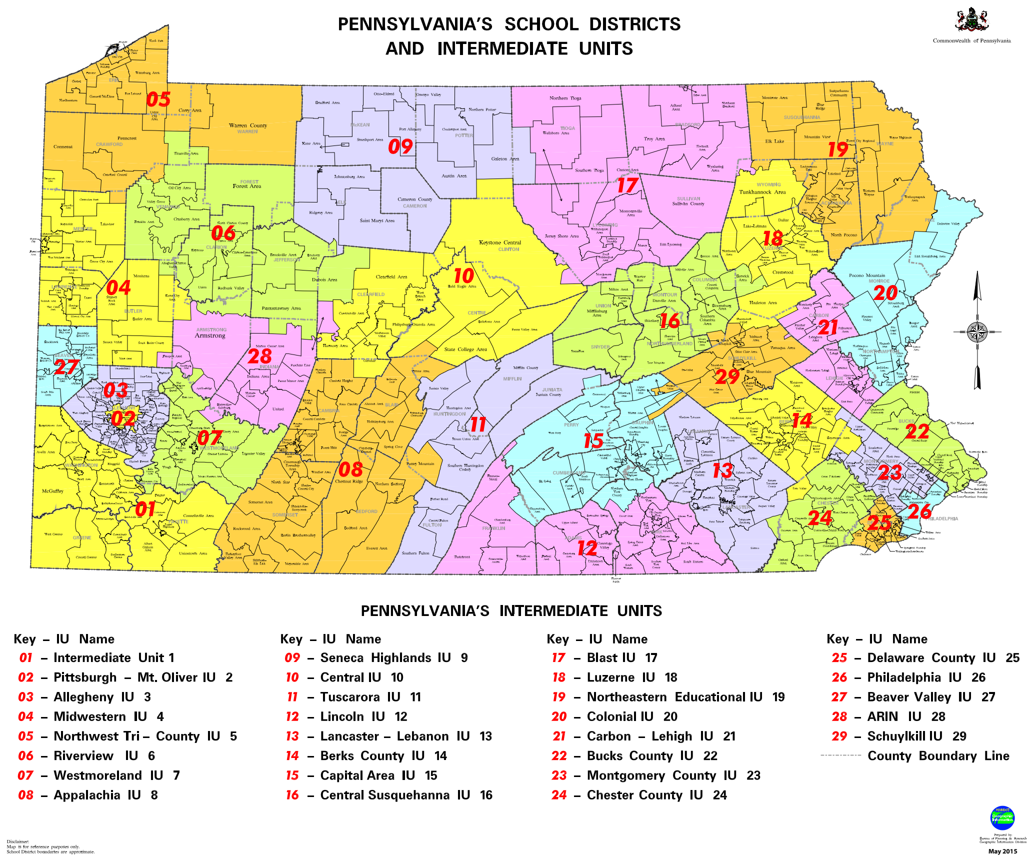


Figure 11: Map of School Districts in Pennsylvania

C.8 Parallel Implementation

Given there are 525 schools and 113,242 teachers, we need to find a way to speed up the computation of the Nash-in-Nash equilibrium, which otherwise can take over a month to compute. To do this we use the julia distributed.jl package along with SlurmClusterManagers.jl to run the for loops in parallel across 512 nodes on the Duke DCC compute cluster.

D Additional Tables and Figures

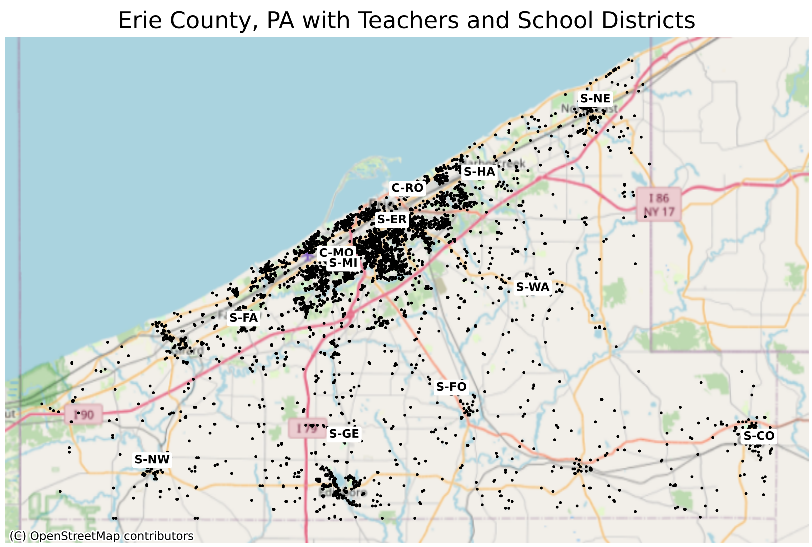


Figure 12: School District Locations in Erie County, Pennsylvania

Notes: The map shows Erie County along with teacher locations and school locations. The schools are labeled with the first two letters of their name.