Oligopsony and Collective Bargaining *

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

October 31, 2025

Abstract

Employers facing limited labor market competition may suppress wages below socially optimal levels. Unions can counteract this wage suppression through collective bargaining, 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 16 percent below the social optimum, while collective bargaining raises wages by 9 percent above the optimum. This average masks substantial district-level heterogeneity driven by variation in bargaining power. Twenty-seven percent of schools have negotiated salaries below the social optimum due to cross-district externalities, where high

*We thank participants at Chicago IO+, Yale, Penn, Princeton, Toronto, UT-Austin, KU Leuven, UNC, Wisconsin, Stanford, 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 thank Kenny Su, Sunmin Kim, Seohee Kim, and Kosuke Shimamoto for research assistance. We thank Michael Dinerstein, Steven Berry, David Berger, Gregor Jarosch, Paolo Somaini, and Ali Yurukoglu for comments. We have no financial interests to disclose.

†Duke University

‡Duke University and NBER

§Ohio State University

salaries at one school lead to hiring reductions, which increase labor supply in competing districts.

1 Introduction

A large and growing body of work measures employer monopsony power and considers its welfare implications on the labor market. With monopsony power, wages and employment are suppressed below the socially optimal level when firms face upward-sloping labor supply curves. 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 (where employers post wages) and evaluate how the settings compare to the socially efficient outcome. We use the model to assess the equilibrium effects of unionization as well as the spillover effects of unions on competing non-unionized employers. We also consider the welfare effects of alternative wage policies: minimum wages and state-level uniform wages.

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 school teachers, and at least in a fixed geographic area is a true monopsonist. Second, wages for school teachers 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. Third, in most school districts in the state, with the notable exception of most charter schools, wages

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

are negotiated through a collective agreement between each school district and the district's local chapter of the teacher union.

Given that many school districts compete for the same pool of teachers, we use an oligopsony model with hiring. In practice, true employer monopsonists are rare; oligopsony, therefore, is a more relevant market structure. The oligopsony model also allows us to capture spillover effects between school districts that lead to heterogeneous welfare effects from our policy counterfactuals.

Our collective bargaining model with oligopsony relies on 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) objective 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, i.e., the equilibrium point will be on the labor demand curve.

For labor supply, we estimate a mixed-logit labor supply model that depends on wages, commuting distance to each school district, other school characteristics, such as the share of students who qualify for free-lunch and charter status, and a random coefficient on the outside employment option that allows for positive correlation in teachers unobserved preferences for jobs in public education. Since our model incorporates constrained choice sets that we do not observe, a traditional estimation strategy that assumes workers can obtain employment at all possible districts would bias our estimates (Crawford et al., 2021). Instead, we estimate the preference parameters in our discrete choice labor supply model using the quit decision, since conditional on tenure, all teachers can either choose to keep their current job or "quit" and leave the public education sector. During estimation we adjust the density of the unobserved heterogeneity parameter to account for the fact that it differs for those that either quit or stay versus those that switch employers.

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 5.74, but varies between

4.4 and 6.[2] However, since teachers may only receive offers from a subset of schools in equilibrium, the average elasticity of labor supply conditional on these offers is lower at 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 rather than leaving the teaching profession.[3]

For collective bargaining, we use the Nash-in-Nash model commonly employed 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 bargaining power, which echoes previous work such as Grennan (2013).

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 $49,234 for the social planner, while predicting average wages of $45,536 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 $52,822 which is 16% higher than the oligopsonist solution and is also 7 percent above the social planner solution. However, the number of teachers falls from 113,237 to 107,292 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 oligopsony setting. To make sense of this result, notice that oligopsony 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

[2]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.

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

raise wages. However, when wages at a school district increase, the school will reduce hiring, which expands the pool of teachers that other school districts can hire.[4] Essentially, the reduction in hiring pushes out 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 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.

We also run counterfactuals looking at state-level wage setting, such as the practice in North Carolina of having a wage that applies for all teachers in the state. This policy combines a) replacing an oligopsony of school districts with a state monopsonist and b) an identical wage in each school district. While monopsony power alone would depress wages, the restriction to paying teachers the same wage everywhere limits this monopsony power. In a similar vein, we look at a minimum wage at the state level that could limit oligopsony power. Here, the large variation in the desirability of different school districts makes a uniform minimum wage less effective at implementing the socially efficient outcome.

Related Literature 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 unions raise wages, either by using cross-sectional comparisons, longitudinal data containing individuals that switch to or from union jobs, or quasi-experiments like close union elections. 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.

[4]We assume that these schools are in the labor demand binds case.

Despite the fact that 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 and they reach 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 there is 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 adding additional structure and an empirical model of union negotiations. The framework relates to an an older strand of labor and industrial relations theory surveyed in Farber (1986). As that literature preceded the development of many current structural methods in applied microeconomics, there are scant attempts to use this work for quantitative prediction with Azkarate-Askasua and Zerecero (2024) being a notable exception, as well as the work of

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). 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, but cost minimization does hold.

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). A useful survey is provided by Kline (2025).

Finally, there is a small literature trying 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 equilibrium effects of 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. We conclude in Section 7.

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 Chapter 25 of Robinson (1933), 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 hiring cost for a monopsonist is given by $MFC = w + L^S \frac{\partial w}{\partial L^S} = w\left(1 + (\epsilon^S)^{-1}\right)$, where ϵ^S is the elasticity of the labor supply curve. In order to hire additional workers the monopsonist needs to pay infra-marginal workers more. Thus, the marginal cost of hiring is always greater than the labor supply curve.

As such, the monopsonist will choose to hire at point B, the intersection of the labor demand curve and the marginal hiring cost curve, yielding L^0 workers and a lower wage w^0. This depression of wages and employment below the competitive level implies a deadweight loss from monopsony power, which is represented by the shaded triangle.

Notice that the size of the deadweight loss, as well as the depression of wages below the efficient level depends on the interaction between the elasticity of labor supply and the elasticity of labor demand. Indeed, if labor demand is completely inelastic, then there is neither deadweight loss, nor depression of wages below efficient levels, no matter the elasticity of labor supply.

Now suppose that workers can collectively bargain with the employer, which is captured in Panel b) of Figure 1. In this case, if the union's well-being increases more rapidly in wages than the employer, wages will be higher than B. In addition, if the union is willing to trade off less employment to obtain higher wages, wages could be bargained to w^b above w^*. 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 is chosen by the union or the employer. We follow the classic Dunlop (1944) approach used in Farber (1978), and assume that employers choose employment on the labor demand curve. [5] Then, labor will be chosen on the labor demand curve 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

[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, provides an example for the case 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 this point.

empirical methods to quantify the relative sizes of these deadweight losses.

Our assumption on employer choice of labor implies that collective bargaining may generate excess labor supply. We specify a rationing rule for labor in our model so that not all workers will get job offers in this case.

3 Institutional Background and Data

We study public school teachers for kindergarten to twelfth grade, who are employed in regular public schools or charter schools in the state of Pennsylvania from 2009 to 2023.[6]

3.1 Teachers

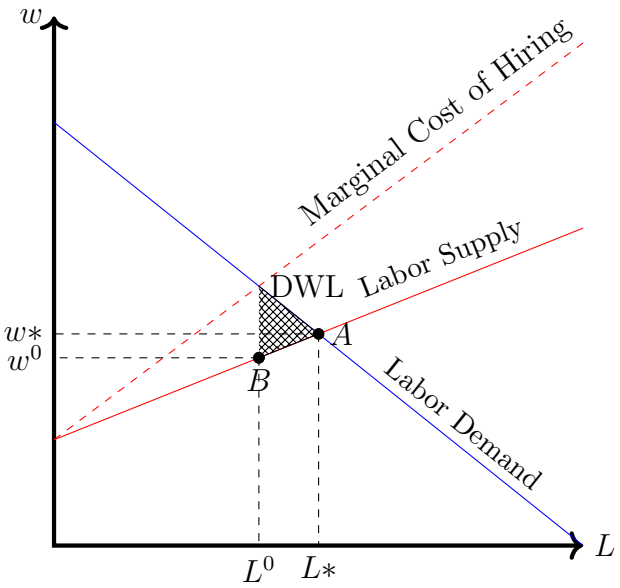
Teachers are licensed professionals hired by schools to teach children. After a 3-year period on the job, teachers become tenured, which is portable across all public school districts. As a practical matter, once teachers obtain tenure, they are very rarely fired, and indeed, are infrequently fired before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

Teachers in Pennsylvania are paid through a school district-specific 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). Experience transfers across school districts. An example of such a schedule is shown in Figure 2 for Abington. 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.[7] Second, uniform wage schedules reveal the wage a teacher would earn in an alternative job, which is crucial for modeling labor supply.

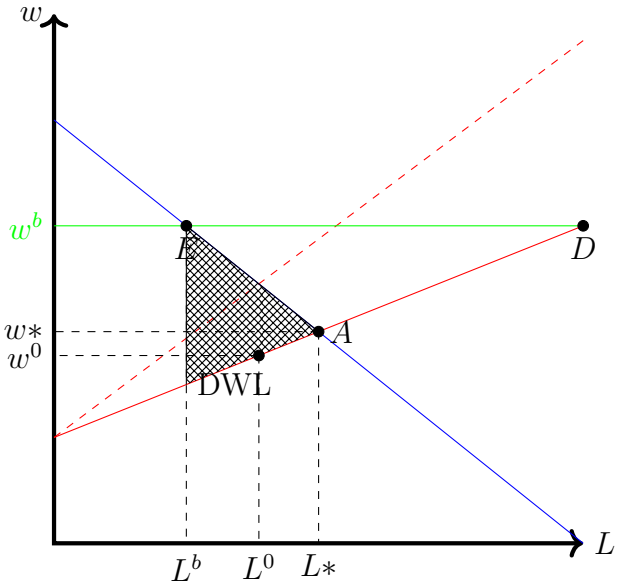
Teachers are also compensated with benefits, such as retirement and health insurance plans. These benefits represent about 30 percent of the cost of hiring a teacher in 2023, and are similar across school districts. For example, pension benefits are administered by the Pennsylvania Public School

[6] We have longitudinal panel ids for teachers from 2013-2023. Most of the analysis uses panel ids. We avoid using data from 2020-2023 to avoid complications related to Covid-19.

[7] More precisely this is the case in a context of complete information between workers and employers as any information asymmetries would reintroduce such distortions.



(a) Posted Wage Monopsonist



(b) Collective Bargaining

Figure 1: Welfare Analysis of Collective Bargaining and Monopsony Power

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.

3.1.1 Teacher Data

Our main data source is annual payroll data from Pennsylvania.[8] It contains the name, salary, experience, tenure, degree status, job description, demographic characteristics, and school district employer identification number and name 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 time-invariant employee identifiers, allowing us to track employees over time.

Second, we use data from InfoUSA that links names to home addresses for all Pennsylvania residents from 2010 to 2022.[9] 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 address to compute commuting times using the HERE georouting application,[10] for the school a teacher currently works at. For the collective bargaining model, we approximate the full school-teacher commute time matrix using a regression of our commute time data on as the crow flies distance.

Table 1 presents summary statistics of the public school teachers 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 peniten-

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

[9] 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.

[10] We used the stata georoute package which matches to the HERE API https://www.here.com/platform/geocoding.

	Mean	Standard Deviation	Observations
Charter Schools	0.07	0.26	113,224
Unionized	0.93	0.25	113,224
Percent Female	0.74	0.44	113,224
Percent White	0.95	0.23	113,224
Percent w/ Masters	0.56	0.50	113,224
Avg Years in Education	13.91	8.44	113,224
Avg. Commute Time	23.25	18.97	28,477
Average Wage	68,923.94	17,572.32	112,945
Exit Rate	0.04	0.20	113,224
Switching Rate	0.02	0.14	113,224

Notes: The summary statistics were computed on the 2017 population of public school teachers 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

tiaries or technical high schools, which make up about 5% of public school teachers. Public school districts hire about 90% of teachers relative to 10% at 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 teaching experience on average. They do not switch schools very often (only 2% per year), and indeed tend to exit teaching rather than switching schools. We find that teachers tend to exit either very early in their career or at the end when they retire. About half of exits are due to retirement. 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.

So that we can consider the effects on the extensive margin, we incorporate potential teachers who have chosen the outside option, not teaching. These teachers work outside the Pennsylvania public school system as private school teachers, unemployed teachers, or in different occupations. We estimate the number of potential teachers by using the number of newly licensed teachers and the number of teachers who quit in the previous four

years. We find that 22-36% of the teacher pool does not work in the public school system depending on the region. See Appendix A.1 for details on the construction of potential teachers.

3.2 Public Schools in Pennsylvania

There are two types of public schools in Pennsylvania, regular school districts and independently run charter schools.[11] In 2017, the state had 499 regular school districts and 165 charter schools in Pennsylvania, with over 92.5% of these 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 treat each school district as a separate employer.[12] Figure 14 shows the geography of school districts and intermediate units in Pennsylvania.

School districts get their funding from federal, state, and local sources. Local funding across districts varies substantially because it is usually sourced from local property taxes. In addition, school size has large variation, 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.[13] This contains budget data by school district and year. We complement this with information from the National Center for Education Statistics (henceforth NCES) from the common core of data.[14] 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 hire about 11% of Pennsylvania's teachers.[15] Due to lack of data, they are excluded from our model. In our model,

[11] There are also penitentiary schools and technical career schools, but we do not consider them in our analysis. They make up 1.75% of public school teachers in our data.

[12] Intermediate Units hire about 3.2% of public school teachers. These teachers tend to be music teachers, PE teachers, special needs teachers, that likely travel between smaller school districts. We do not consider them in our analysis

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

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

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

teachers who work for a private school have chosen the outside option.

Charter schools are generally much smaller than regular school districts with one-third of the students–824 students vs 3,129 students on average. About 67% of these charter schools are in urban areas, such as Philadelphia, while only 16% of regular school districts are in urban areas. School districts hire about 215 school teachers on average, while charter schools only hire about 50 teachers. In addition, charter schools have a higher student-teacher ratio of 15.88 students for each teacher compared to the school district's 14.01. Charter schools also tend to have poorer students as measured by the percentage of students that are eligible for the government's free lunch program. The summary statistics for private schools show that private schools are small and should not have a large impact on our results.

The lower panel of Table 2 contains summary statistics of school expenditures. Since school districts are much larger, they spend much more–$60.01 million compared to $11.46 million for charters. They also spend more per student–$18,230 vs $14,030 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 $112,316 on average, which includes salary and benefits. About 2/3 of the cost of a teacher is wages and salary.

Out of 678 districts and charter schools in 2017, we restrict our model to 486/499 school districts and 121/179 charter schools, which account for 97.5% of teachers working in the 678 schools.[16]

3.3 Collective Bargaining

Teachers at every regular school district are represented by a local chapter of a union, the Pennsylvania State Education Association.[18] The union

[16] We drop 9 cyber charter schools, 38 schools with less than 20 teachers, and 24 schools due to missing data. Twenty-two of the missing data schools have missing wage data due to our wage computation method. In order to avoid compositional variation in teachers, we compute school wages by taking the median salary of all full-time teachers at the school with a bachelor degree, 5-7 years of experience, and salaries above $30,000. The 22 schools that were dropped because they lack a teacher that matches these constraints, hired only 1.1% of all teachers and are small with a median value of 33 teachers.

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

	Public			Private
	Full	School Dist	Charter	
Basic Facts				
Charter Schools	0.25	0.00	1.00	0.00
	(0.43)	(0.00)	(0.00)	(0.00)
Urban	0.29	0.16	0.67	0.42
	(0.45)	(0.37)	(0.47)	(0.49)
Num of Schools	4.34	5.45	1.00	
	(9.12)	(10.29)	(0.00)	()
Num of Students	2,555.98	3,128.70	823.93	258.99
	(5,480.98)	(6,183.51)	(1,150.67)	(221.33)
Num of Teachers	173.74	214.83	49.47	25.93
	(323.50)	(362.70)	(54.17)	(23.00)
Num of Other Staff	30.08	35.70	13.09	
	(87.18)	(99.57)	(15.17)	()
Student-Teacher Ratio	14.47	14.01	15.88	10.37
	(2.72)	(1.62)	(4.39)	(4.39)
Percent Students on Free Lunch[17]	0.51	0.45	0.80	
	(0.27)	(0.22)	(0.26)	()
Unionized	0.76	1.00	0.04	0.00
	(0.43)	(0.00)	(0.19)	(0.00)
Expenditures				
Total Expenditures (mil)	48.05	60.01	11.46	
	(150.56)	(171.51)	(16.71)	()
Expenditure per Student (thou)	17.20	18.23	14.03	
	(4.81)	(4.35)	(4.77)	()
Percent Funding from Local Sources	0.61	0.52	0.90	
	(0.23)	(0.18)	(0.06)	()
Expenditure on Teachers (mil)	21.27	26.74	4.52	
	(39.26)	(43.77)	(5.09)	()
Teacher Share of Exp	0.30	0.30	0.30	
	(0.05)	(0.04)	(0.09)	()
Avg Teacher Cost	112,316.23	119,496.80	90,334.01	
	(24,180.63)	(17,923.08)	(27,444.61)	()
Avg. Salary Share of Teacher Exp	0.63	0.62	0.68	
	(0.04)	(0.02)	(0.06)	()
Obs	664	499	165	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

and schools negotiate a collective agreement every 3-5 years that incorporates teacher salary, retirement, health care, working hours, vacation, and termination rules, among other topics. Charter schools rarely negotiate with teacher unions; in 2016, only 14 charter schools were unionized.[19] In our model, we assume that the local union chapter negotiates individually with each school district and does not consider how negotiations at one school will affect 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. Each school district pays teachers according to a schedule 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 stipulated in the contracts. In the model, we assume that each school pays a single wage to all teachers that is determined as the median wage of all full-time teachers with a bachelor degree and 5-7 years of experience working at that school.

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 average salaries of $\$37,444$ in Turkeyfoot Valley to $\$97,791$ for Lower Merion. Some of this salary dispersion can be explained by compositional differences in the workforce, such as having teachers with more experience and more education. However, controlling for such differences only shrinks the overall variance in salary dispersion by one-third, indicating large differences in salaries across employers unrelated to worker attributes.

[19]The National Allicance for Public Charter Schools collects this data: https://publiccharters.org/wp-content/uploads/2023/01/Unionized-Charter-Schools-2016-17_0.pdf.

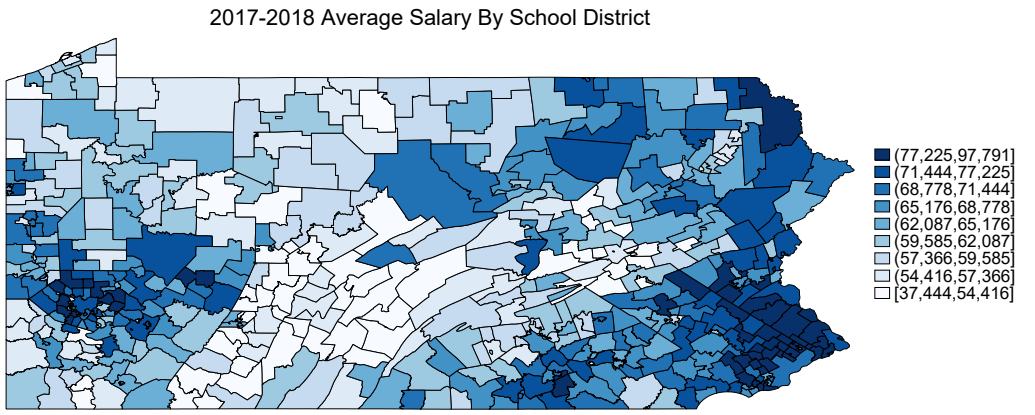
APPENDIX A SALARY SCHEDULE

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



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

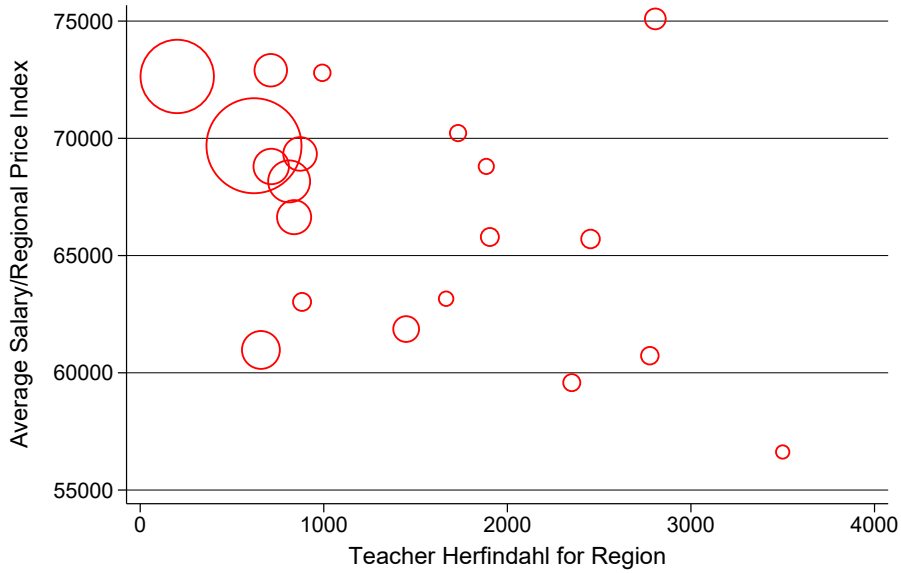
Interestingly, much of the salary dispersion in Pennsylvania is not explained exclusively by local labor markets, as locations such as Lower Merion and Philadelphia border each other and yet have average salaries that differ by over $\$20,000$ per year. We believe that differences in wages can be explained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

We define local labor markets as Metropolitan Statistical Areas from the Bureau of Labor Statistics.[20]. We drop rural school districts that are not part of an MSA. We also assume that teachers cannot find employment in multiple MSAs. 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 that 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 relation-

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



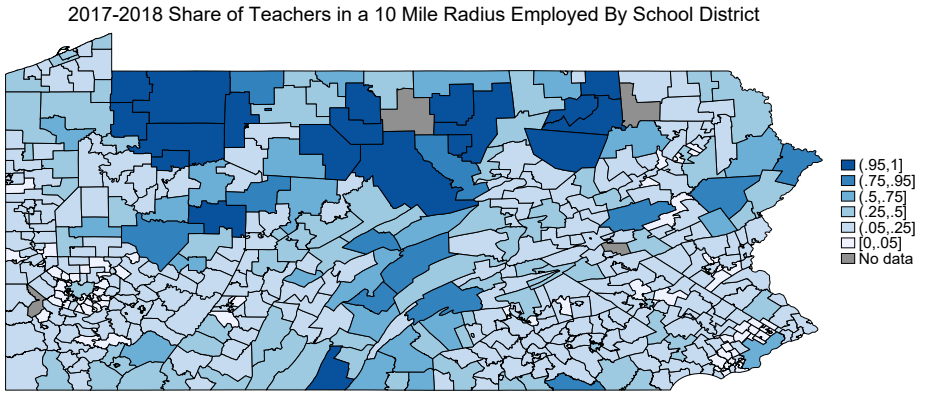
Notes: Each circle in the figure represents one of the 20 primary statistical areas in Pennsylvania. The Herfindahl is the sum of squared share of each employers teachers in the overall market, scaled by 10,000.

Figure 4: Regional HHI and Wages

ship 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).[21] 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 presented at the school district level. Notice the large variation in the share of teachers within a ten mile radius that work for this school district —

[21] 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.



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

essentially zero in the suburban districts in Pittsburg, and over 95 percent in certain rural school districts.

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.[22]

A one percent increase in the number of school districts within 10 miles 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.[23] 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

[22] See Footnote 21.

[23] The National Allicance for Public Charter Schools collects this data: https://publiccharters.org/wp-content/uploads/2023/01/Unionized-Charter-Schools-2016-17_0.pdf.

	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

school districts. Table 4 shows that the charter 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, it is still difficult to interpret the difference in salary as the causal effect of unions on salaries. For example, 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 non-unionized 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 use to estimate the model. The

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.

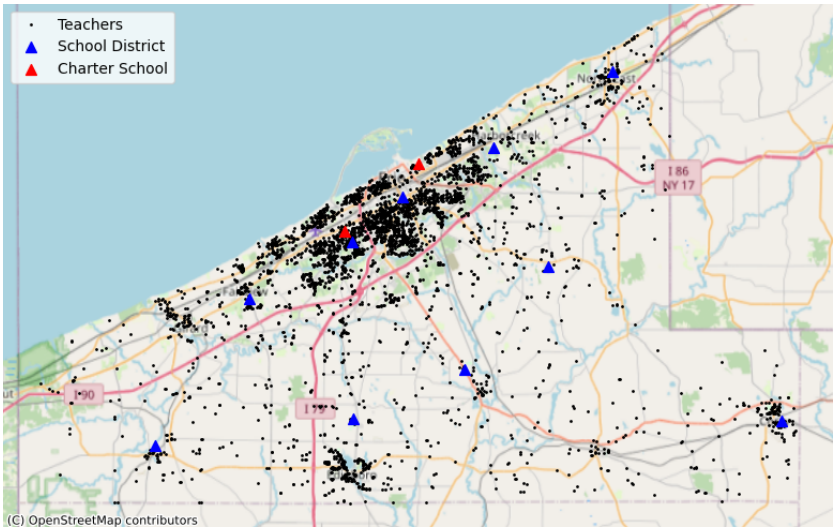
figure shows the location of teachers and school districts in Erie County in northwest Pennsylvania. Small black dots indicate the location of teachers,[24] while the blue triangles represent regular school districts and the red circles are the two charter schools.

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 a Nash-in-Nash bargaining game. Note that the Nash bargaining model nests the posted price solution; 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 accounts for the entire life-cycle of a teacher's career would be challenging to incorporate into the labor supply. Second, negotiations happen simultaneously.[25]

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

[25]Collective agreements are renegotiated every 4-5 years and are not negotiated simultaneously. Accommodating this feature would introduce dynamics into the bargaining



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

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 receives utility u_{ij} from working in a school district indexed by $j = 1, \cdots, J$:

$$\begin{aligned} u_{ij} &= \delta_{ij}(\nu_i) + \epsilon_{ij}; \\ \delta_{ij}(\nu_i) &= \psi\omega_j + x_j'\beta - \tau d_{ij} + \sigma\nu_i, \end{aligned}$$

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

problem which are very complex. Lee and Fong (2013) discuss the issues associated with it. Collard-Wexler et al. (2019) shows a model where agreements can occur sequentially, but in equilibrium, will occur simultaneously.

students eligible for free lunch and an indicator for whether the employer is a charter school. The term ψr determines the mean value of the outside option.

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 the school district j.[26] 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 $\mathcal{N}(0,1)$ random variable that is independently distributed across teachers. The ν_i term is common across all public teaching jobs, such that teachers in a public education job will switch more often to another public education job given a salary reduction rather than switching to the outside option.[27]

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 an employment option outside of Pennsylvania public schools. 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:

$$s_{ij}(\nu_i) = \frac{\exp(\delta_{ij}(\nu_i))o_{ij}}{\exp(\psi r) + \sum_k \exp(\delta_{ik}(\nu_i))o_{ik}}. \tag{1}$$

We integrate over ν_i to get:

$$s_{ij} = \int_\nu s_{ij}(\nu)\phi(\nu)d\nu. \tag{2}$$

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:

[26] The weighted average distance of the centroid to any school within a district is 1.58 miles. No school is more than 20 miles away from the centroid.

[27] 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. The logit model by itself would assume that switching schools and leaving education have the same horizontal differentiation.

$$T_j = \sum_{i \in \mathbf{O}_j} s_{ij}, \tag{3}$$

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 given by equation (2) is a very common mixed logit. A complicating factor is that we do not observe job offer sets, which in our model are incomplete due to rationing. In this case, erroneously assuming that all workers receive all offers would bias estimates of the marginal utility of job attributes towards zero.[28] Instead, we leverage the fact that conditional on ν, Independence of Irrelevant Alternatives (IIA) holds. We assume that teachers always have the option of either keeping their current job (stay) or leaving the public education sector (quit).[29] Conditional on $\nu_i = \nu$, which is the realization of the idiosyncratic preference shock common to all public education jobs for individual i, the likelihood of quitting or staying in one's current job is given by:

$$\begin{aligned}
\Pr(\text{quit}|\text{quit or stay}) &= \frac{\Pr(\text{quit})}{\Pr(\text{quit or stay})} \\
&= \frac{\exp(\psi r)}{\exp(\psi r) + \sum_k \exp(\delta_{ik}(\nu_i)) o_{ik}} \times \frac{\exp(\psi r) + \sum_k \exp(\delta_{ik}(\nu_i)) o_{ik}}{\exp(\psi r) + \exp(\delta_{ij}(\nu_i))} \\
&= \frac{\exp(\psi r)}{\exp(\psi r) + \exp(\delta_{ij}(\nu_i))}.
\end{aligned} \tag{4}$$

This result follows immediately from logit-driven IIA as discussed by McFadden (1978) for estimation with a subset of choices. We will call $s_{iq}(\nu, \phi, \beta, \tau, \sigma) \equiv \Pr(\text{quit}|\text{quit or stay})$.

[28] In most cases, except in the case of matching mechanisms, the entire set of offers is not 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 to teachers to fill a position.

[29] There are very few involuntary terminations in the teaching profession, and this is true not only after teachers receive tenure following a 3 year probationary period, but even before tenure is granted.

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

$$\mathcal{L}(\nu, \phi, \beta, \tau, \sigma) = \sum_i \log[1(q_i = 1) \int_\nu s_{iq}(\nu, \phi, \beta, \tau, \sigma)\phi(\nu|\text{quit or stay})d\nu + (1(q_i = 0)(1 - \int_\nu s_{iq}(\nu, \phi, \beta, \tau, \sigma)\phi(\nu|\text{quit or stay})d\nu)], \tag{5}$$

If there were no teachers that switch districts, the density of ν conditional on quit or stay would be standard normal (e.g. $\mathcal{N}(0, 1)$) and thus Equation 5 would be a mixed binary logit. When job switching rates are low, as in our application, misspecification bias from ignoring the selection problem is small. We propose an algorithm to correct for the selection bias and use monte carlo experiments to validate both that the selection bias is small when switching is low and that our selection correction procedure is accurate.[30]

5.3.1 Selected Job Movers

When some teachers switch jobs, $\phi(\nu)$ no longer equals $\phi(\nu|\text{quit or stay})$ due to a selection issue: teachers who quit teaching tend to have a lower value of working in public education, i.e., ν. The sample of teachers who quit or stay has a greater proportion of quitters than the full sample of teachers which includes job movers as shown in Figure 13 on page 73 in the appendix. Therefore, the conditional distribution puts more weight on lower values of ν. We develop a method to correct for this selection by estimating the conditional distribution of ν:

$$f(\nu|\text{stay or quit}, \mathbf{O}_i) = f(\nu) \underbrace{\frac{\Big(\Pr(quit|\nu_i, \mathbf{O}_i) + \Pr(stay|\nu_i, \mathbf{O}_i)\Big)}{\Big(\Pr(quit|\mathbf{O}_i) + \Pr(stay|\mathbf{O}_i)\Big)}}_{\text{weight}}. \tag{6}$$

Notice that when the number of job movers is small, the weight in Equation 6 will be close to 1, so that the estimation bias from job movers is small. This intuition is confirmed by the Monte-Carlos we run in section D in the appendix, which show in Table 21 that accounting for selection has a negligible impact on the estimated coefficients given a job move rate of 3% which

[30] More details on the monte carlo experiments are in appendix D.

is close to the 2% share in our own data (see Table 1). If a larger share of workers moved, then our coefficients would be biased. For these cases, which lie outside of the empirical context of this paper, we have developed a selection correction algorithm that we discuss below. In short, we use the structure of our model to estimate each piece of the weight in Equation 6.

First, we estimate the propensity score for quitting and staying denoted as $\Pr(quit|\mathbf{O}_i)$ and $\Pr(stay|\mathbf{O}_i)$ using a non-parametric logit regression. The probability of quitting is the product of the quit choice conditional on the offer set times the likelihood of this offer set, i.e. $\Pr(quit|X,Z) = \int_{\mathbf{O}_i} \Pr(quit|\mathbf{O}_i, X) \Pr(\mathbf{O}_i|Z)$. In the propensity score regression, we include both covariates X such as commute time and wages that affect the decision to quit given the offer set and covariates Z that affect the likelihood that a teacher receives a given offer set (which are discussed in more detail in section 5.5 discussing labor demand), such as school budgets, distance to the school district, and the number of other teachers located near that school district. For example, schools that have lower wages and higher free lunch share will send out more offers to account for the fact that teachers are less likely to accept offers.

In the context of our model, presented in section 5.5, the offer matrix of a teacher is a deterministic function, so the propensity score is expressed more simply by $\Pr(quit|X,Z) = \Pr(quit|\mathbf{O}_i(Z,X))$, and we estimate the propensity score as $\Pr(quit|\mathbf{O}_i)$ and $\Pr(stay|\mathbf{O}_i)$. We can obtain $\Pr(move|\nu_i, \mathbf{O}_i)$ using the fact that the events are mutually exclusive and the probabilities must sum to one.

Second, we tackle the probabilities conditional on ν_i and $\mathbf{O}_i$ and focus on the probability of quitting:

$$\Pr(quit|\nu, \mathbf{O}_i) = \frac{1}{1 + \exp(\sigma\nu)\gamma_i}, \tag{7}$$

where $\gamma_i \equiv \sum_k \exp(\delta_{ik}) o_{ik}$. Given σ, we can recover γ_i from $\Pr(quit|\mathbf{O}_i)$ since $\Pr(quit|\mathbf{O}_i)$ is strictly decreasing in γ_i, $\Pr(quit|\mathbf{O}_i)$ is known (estimated), and $f(\nu)$ is a normal density.

$$\Pr(quit|\mathbf{O}_i) = \int_\nu \Pr(quit|\nu, \mathbf{O}_i) f(\nu) d\nu = \int_\nu \frac{1}{1 + \exp(\sigma\nu)\gamma_i} f(\nu) d\nu \tag{8}$$

Third, we estimate $\Pr(stay|\nu, \mathbf{O}_i)$ and exploit the fact that movers and stayers have the same distribution of ν, which follows directly from IIA when

teachers do not choose the outside good. The relative odds of taking any two teaching jobs j and k is independent of ν:

$$\frac{\Pr(k|\nu_i, \mathbf{O}_i)}{\Pr(j|\nu_i, \mathbf{O}_i)} = \frac{\exp(\delta_{ik} + \sigma\nu_i)}{\exp(\delta_{ij} + \sigma\nu_i)} = \frac{\exp(\delta_{ik})}{\exp(\delta_{ij})}.$$

Each teacher either stays, quits, or moves jobs, so we can express $\Pr(stay|\nu_i, \mathbf{O}_i)$ in terms of $\Pr(quit|\nu_i, \mathbf{O}_i)$,$\Pr(move|\mathbf{O}_i)$, and $\Pr(stay|\mathbf{O}_i)$ all of which are known:[31]

$$\Pr(stay|\nu_i, \mathbf{O}_i) = \Big(1 - \Pr(quit|\nu_i, \mathbf{O_i})\Big)\Big(1 + \frac{\Pr(move|\mathbf{O_i})}{\Pr(stay|\mathbf{O_i})}\Big)^{-1} \tag{9}$$

Thus, we can construct the term labeled Weight in equation 6 and obtain the density of unobservable heterogeneity conditional on not switching jobs.

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 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 5 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 where $\sigma = 0$ and the model collapses to a binary logit. Each coefficient estimate is significant at the .05 level and has an intuitive sign.

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 (NEED TO UPDATE THESE). 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 (MRS) between salary and commute time of \$65.50 per hour. The second column corresponds to an intermediate value of σ. Columns 1-3 show that the implied MRS between commute time and salary is almost the same for different values of σ. Column 4 shows that

[31] We use $1 = \Pr(quit|\nu_i, \mathbf{O}_i) + \Pr(stay|\nu_i, \mathbf{O}_i) + \Pr(move|\nu_i, \mathbf{O}_i)$, Bayes Theorem $\Pr(move|\nu_i, \mathbf{O}_i) = \frac{Pr(move|\mathbf{O}_i)\Pr(\nu|move, \mathbf{O}_i)}{\Pr(\nu)}$, and that $\Pr(\nu|move, \mathbf{O}_i) = \Pr(\nu|stay, \mathbf{O}_i)$.

the parameter estimates and MRS are robust to dropping teachers without tenure, which is obtained after three years of experience. Column 5 shows that the parameters become blunted as we expand the sample to include teachers with more than 15 but less than 25 years of experience, likely because retirement becomes more salient in determining the decision to leave education. Despite this, the marginal effects are roughly the same as in column 3. Column 6 adds indicators for each of the MSA's in Pennsylvania along with an indicator for schools not in an MSA.

5.4.1 Labor Supply Elasticities

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:

$$\begin{aligned}\epsilon_{jk} &= \frac{\partial log(T_j)}{\partial log(\omega_k)} = \frac{\omega_k}{T_j}\sum_{i\in N}\int_\nu \frac{\partial s_{ij}(\nu)}{\partial \omega_k}\phi(\nu)d\nu \\ &= \begin{cases} \frac{\omega_k}{T_j}\sum_{i\in N}\int_\nu \psi s_{ij}(\nu)(1-s_{ij}(\nu))\phi(\nu)d\nu & \text{if } k=j, \\ \frac{\omega_k}{T_j}\sum_{i\in N}\int_\nu -\psi s_{ij}(\nu)s_{ik}(\nu)\phi(\nu)d\nu & \text{if } k\neq j \end{cases}\end{aligned} \tag{10}$$

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 2-3.[32] Table 6 illustrates the elasticity matrix for some of the schools in Erie and Crawford County, including charter schools. The entry in column j and row k is equal to ϵ_{jk}. Notice that the cross-wage elasticities between school districts are heavily influenced by distance, such that schools that are farther away tend to have smaller cross-wage elasticities in magnitude. For example, compare the cross-elasticities between S-ERI and S-MIL to those between S-ERI and S-CON which are much further apart.[33] Teachers can also choose to switch to the outside option. 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

[32] 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.

[33] Figure 15 shows the locations of the schools.

Table 5: MLE Estimates of the Quit Decision

Variable	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)
Real Salary	-0.229	-0.394	-0.696	-0.562	-0.270	-0.587	-0.531
(10k)	(0.019)	(0.033)	(0.059)	(0.072)	(0.044)	(0.069)	(0.109)
	[-0.0093]	[-0.0090]	[-0.0077]	[-0.0058]	[-0.0028]	[-0.0062]	[-0.0007]
Commute Time	0.009	0.015	0.028	0.026	0.018	0.020	0.025
(Minutes)	(0.001)	(0.003)	(0.005)	(0.005)	(0.004)	(0.006)	(0.007)
	[0.0004]	[0.0004]	[0.0003]	[0.0003]	[0.0002]	[0.0002]	[0.0000]
Free Lunch Share	0.583	1.038	1.815	1.284	1.882	1.431	1.399
	(0.162)	(0.294)	(0.544)	(0.524)	(0.443)	(0.374)	(0.797)
	[0.0238]	[0.0236]	[0.0200]	[0.0133]	[0.0197]	[0.0152]	[0.0019]
Charter	0.735	1.404	2.550	2.905	2.646	2.480	4.458
	(0.111)	(0.199)	(0.359)	(0.346)	(0.342)	(0.393)	(0.573)
	[0.0300]	[0.0319]	[0.0282]	[0.0301]	[0.0277]	[0.0263]	[0.0062]
σ	0	4.425	8.85	8.85	8.85	8.85	8.85
$Pr(quit)$	0.050	0.050	0.050	0.044	0.047	0.050	0.044
Experience	< 15 *years*	< 15 *years*	< 15 *years*	< 15 *years* > 3 *years*	< 25 *years*	< 15 *years* > 3 *years*	< 15 *years* > 3 *years*
Observations	66,745	66,745	66,745	54,962	110,299	54,962	46,467

Notes: The table reports maximum likelihood estimates of mixed binary logit models for different values of the random coefficient σ. The likelihood was computed using Gauss-Hermite quadrature with 30 nodes and estimated on full-time teachers from the 2013-14 through 2018-19 academic years. Columns 1-3 restrict the sample to teachers with less than 15 years experience. Column 5 conditions on teachers with less than 25 years of experience. Columns 4, 6 and 7 restrict the sample to teachers with more than 3 but less than 15 years of experience. All specifications include year indicators. Column 6 adjusts the quadrature weights to account for selection. Column 7 contains indicators for MSA's. Standard errors clustered by school district are in parentheses. Average marginal effects are in brackets.

	S-CON	S-ERI	S-MIL	S-PEN	C-ROB	C-MON	Out Opt
All Offers							
S-CON	2.804	−0.115	−0.135	−0.261	−0.105	−0.133	−0.033
S-ERI	−0.03	2.411	−0.054	−0.037	−0.058	−0.055	−0.008
S-MIL	−0.232	−0.353	2.537	−0.252	−0.355	−0.35	−0.055
S-PEN	−0.321	−0.17	−0.18	2.592	−0.163	−0.178	−0.041
C-ROB	−0.0	−0.001	−0.001	−0.001	2.28	−0.001	−0.0
C-MON	−0.001	−0.001	−0.001	−0.001	−0.001	1.923	−0.0
Out Opt	−1.093	−1.054	−1.055	−1.083	−1.054	−1.054	2.305

Notes: The table shows own- and cross-elasticities of a subset of schools Erie and Crawford 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 15 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 and Crawford County

job offers each teacher has. When jobs are rationed in some districts, these relationships can break down as highlighted in Section 5.8.

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, such as 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 flexible. Imposing the budget constraint $B_j = X_j + w_j T_j$ yields the problem:

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

$$s.t. \qquad T_j(O_j, \mathcal{O}_{-j}) = \sum_{i=1}^{O_j} s_{ij}(\mathcal{O}_{-j}),$$

$$O_j \leq N, \tag{11}$$

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.[34]

If $O_j < N$, 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:

$$\frac{\frac{\partial W}{\partial T}}{\frac{\partial W}{\partial x}} = \frac{\gamma}{1-\gamma} \cdot \left(\frac{T_j}{X_j}\right)^{\rho-1} = w_j. \tag{12}$$

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

$$T_j(w_j) = \frac{w_j^{\frac{1}{1-\rho}} B_j}{(\frac{\gamma}{1-\gamma})^{\frac{1}{\rho-1}} + w_j^{\frac{\rho}{\rho-1}}} \tag{13}$$

When $O_j < N$, 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.

When $O_j = N$ the district cannot hire as many teachers as it would like at the prevailing wage. In this case the benefit of hiring another teacher exceeds the cost:

$$\frac{\frac{\partial W}{\partial T}}{\frac{\partial W}{\partial x}} = \frac{\gamma}{1-\gamma} \cdot \left(\frac{T_j}{X_j}\right)^{\rho-1} > w_j \tag{14}$$

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.[35]

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

[35] 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 log-transforming Equation 13:

$$\begin{aligned} \log(w_j T_j) - \log(X_j) &= \frac{1}{\rho - 1}\left[-log(\frac{\gamma}{1-\gamma}) + \rho \log(w_j)\right] \\ &= \gamma_1 + \rho_1 \log(w_j) \end{aligned} \tag{15}$$

Table 7 presents estimates constructed by fitting Equation 15 to the data using OLS. The independent variable, labeled teacher marginal factor cost, is total spending on teachers over the number of teachers hired, i.e., wages and benefits spent on the average teacher. We exclude observations from non-unionized charter schools, since they 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, Column 3 includes both year and region fixed effects, and Column 4 includes year and school district fixed effects. In each case, the coefficient on the teacher marginal factor cost is significantly different from zero, so that Cobb-Douglas preferences are rejected. 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 unionized 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.

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.

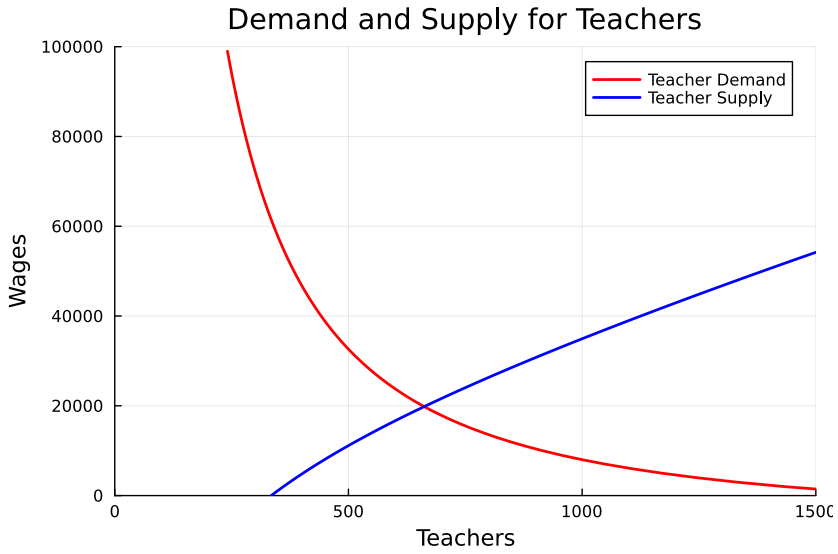
5.6 Wage Negotiation

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

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 remove non-unionized charter schools in the regression, since they 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



Notes: The graph shows the labor supply and demand curves for Millcreek Township School District (S-MIL) in Erie County, Pennsylvania. The demand curve comes directly from Equation 13. The labor supply curve is determined by keeping wages and offers fixed at all other schools, while allowing S-MIL to give offers out to everyone. Then we sum up how many teachers are willing to work at S-MIL at each wage. Wages at other schools were kept fixed at the wages seen in the data.

Figure 7: Labor Supply and Labor Demand

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 salaries w_{-j}:

$$\begin{aligned}\mathcal{N}(w_j|w_{-j}) &= [W(T_j, B_j - w_j T_j))]^{\alpha_b(j)} \left[T_j(w_j - r_o)\right]^{1-\alpha_b(j)} \\ &= [W(T^*(w_j, w_{-j}), B_j - w_j T^*(w_j, w_{-j}))]^{\alpha_b(j)} \times \\ &\quad [T^*(w_j, w_{-j})(w_j - r_o)]^{1-\alpha_b(j)},\end{aligned} \tag{16}$$

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 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 $T_j w_j$.[36][37] In the event of disagreement, the union payoff is $T_j r_0$, where we set $r_0 = 30,000$.[38]

The first-order condition of the logarithm of Equation 16 with respect to w_j is:

$$\frac{\partial \log \mathcal{N}_j}{\partial w_j} = \frac{\alpha_b(j)\frac{\partial W_j(X_j,T_j)}{\partial w_j}}{W_j(X_j, T_j)} + \frac{(1-\alpha_b(j))}{T_j}\frac{\partial T_j}{\partial w_j} + \frac{(1-\alpha_b(j))}{w_j - r_0} = 0. \tag{17}$$

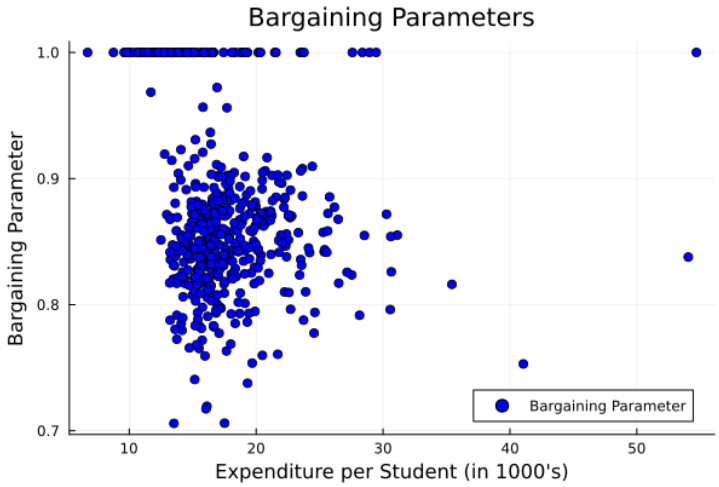
The derivative, $\frac{\partial T_j}{\partial w_j}$, depends on whether labor demand or labor supply binds. Rearranging this equation allows the bargaining parameter to be expressed as a function of data and our estimated labor demand and labor supply parameters:

$$\alpha_b(j) = \frac{W_j(X_j, T_j)(T_j + (w_j - r_0)\frac{\partial T_j}{\partial w_j})}{(W_j(X_j, T_j))(T_j + (w_j - r_0)\frac{\partial T_j}{\partial w_j}) - (T_j \frac{\partial W_j}{\partial X_j} + \frac{\partial T_j}{\partial \omega_j}\left[\frac{\partial W_j}{\partial T_j} - \omega_j \frac{\partial W_j}{\partial X_j}\right])T_j(\omega_j - r_o)}. \tag{18}$$

[36] 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.

[37] 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.

[38] 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 expenditure per student 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

Thus, we can recover each bargaining power parameter, $\alpha_b(j)$. We assume non-unionized schools 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 per student on the horizontal axis. The unweighted correlation between budget and bargaining parameters is 0.027 and the weighted correlation is 0.56 when weighting by students. The lack of strong correlation indicates that wealthy schools are not significantly stronger bargainers than 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 directly from the data using regressions and first order conditions. The remaining two parameters, σ and r, which appear in the labor supply model, are estimated via indirect inference. 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

Parameter	Estimate	Standard Error
σ	9.4	1.75
r	258,000	25,310.52

Notes: Standard Errors computed by bootstrapping the estimates of labor supply in Table 5 and labor demand in Table 7 and reestimating the coefficients that minimize the criterion function (equation 19) for 50 bootstrap resamples.

Table 8: Indirect Inference Estimates

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

$$Q(\sigma, r) = \left(\mathbb{M}(\sigma, r) - \hat{\mathbb{M}}\right)' \mathbb{W} \left(\mathbb{M}(\sigma, r) - \hat{\mathbb{M}}\right), \tag{19}$$

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. We compute standard errors by bootstrapping the moment matrix $\hat{\mathbb{M}}$ at the school district level.

For computational convenience, we restrict our procedure to Erie and Crawford Counties in Pennsylvania.

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 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.[39] 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,non-unionized charter schools become closer substitutes to regular school districts under restricted offers, since they send offers to all teachers. 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 11 presents average salary and total number of teachers hired in the data and under the model of collective bargaining for the entire state of Pennsylvania. Our model predicts average wages of $ 52,822 in contrast to average wages of $ 53,449 in the data. Likewise, our model predicts that 107,292 teachers will be hired, which is close to the 104,748 that were actually hired in 2017. 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 fit is substantially better for districts on the labor demand curve, with a correlation coefficient of 0.86, versus schools on the labor supply curve, where this same correlation is 0.23. When labor demand binds, the Nash bargaining weights absorb much of the wage variation. However, we lack this structure for schools where labor supply binds. In addition, the model struggles to explain why districts with large free lunch shares do not

[39]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.

	S-CON	S-ERI	S-MIL	S-PEN	C-ROB	C-MON	Out Opt
All Offers							
S-CON	2.804	−0.115	−0.135	−0.261	−0.105	−0.133	−0.033
S-ERI	−0.03	2.411	−0.054	−0.037	−0.058	−0.055	−0.008
S-MIL	−0.232	−0.353	2.537	−0.252	−0.355	−0.35	−0.055
S-PEN	−0.321	−0.17	−0.18	2.592	−0.163	−0.178	−0.041
C-ROB	−0.0	−0.001	−0.001	−0.001	2.28	−0.001	−0.0
C-MON	−0.001	−0.001	−0.001	−0.001	−0.001	1.923	−0.0
Out Opt	−1.093	−1.054	−1.055	−1.083	−1.054	−1.054	2.305
Constrained Offers							
S-CON	0.634	0.0	0.0	−0.011	−0.004	−0.01	−0.019
S-ERI	0.0	0.664	−0.313	−0.005	−0.973	−0.93	−0.087
S-MIL	0.0	−0.167	0.648	0.0	−0.086	−0.098	−0.047
S-PEN	−0.015	−0.001	0.0	0.85	−0.016	−0.018	−0.025
C-ROB	−0.001	−0.032	−0.005	−0.002	1.957	−0.318	−0.004
C-MON	−0.002	−0.045	−0.009	−0.003	−0.464	1.432	−0.006
Out Opt	−1.256	−1.229	−1.257	−1.25	−1.826	−1.846	1.579

Notes: The table shows own- and cross-elasticities of a subset of schools in Erie and Crawford County. S refers to a school district, while C refers to a charter school. See the map in Figure 15 for the location of the schools. Each cell 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. The top table shows the elasticities when schools give out all offers, i.e., they do not constrain their hiring. In the bottom table, school districts can limit hiring by sending out fewer offers. Thus, changes in wage can have zero effect on teachers, because those teachers were not getting offers to begin with.

Table 9: Unconstrained and Constrained Elasticities: Erie and Crawford County

	Labor Demand Binds	Labor Supply Binds
% Hired with Emergency Permit	0.4%	4.0%
School Districts	419	106

Notes: The table shows the percentage of a school's long-term teachers hired using emergency permits, which are used to proxy teacher shortage level. Our model predicts that labor demand binds in 419 school districts and charter schools, and labor supply binds in 105. Schools where labor supply binds hire a greater percentage (6.2%) of their long-term teachers using emergency permits than schools where labor demand binds (1.1%).

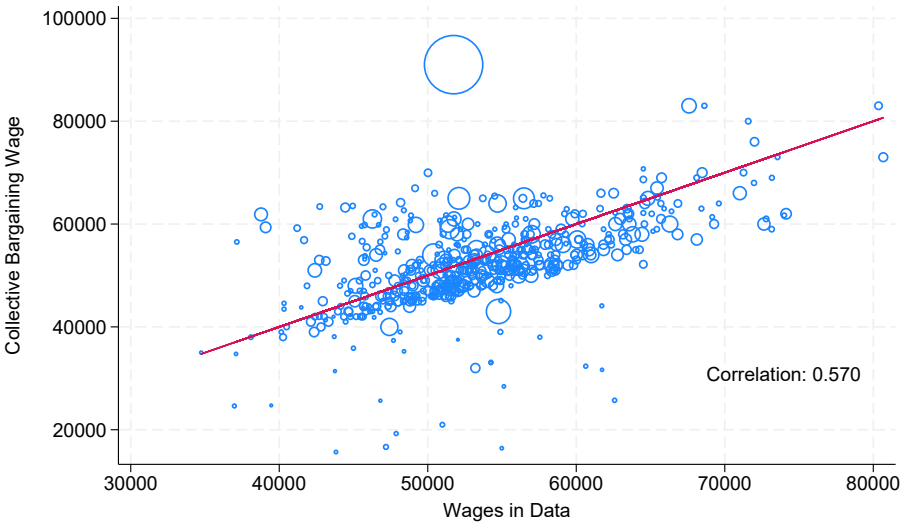
Table 10: 2017 Percent Teachers hired with Emergency Permit

pay better.[40]

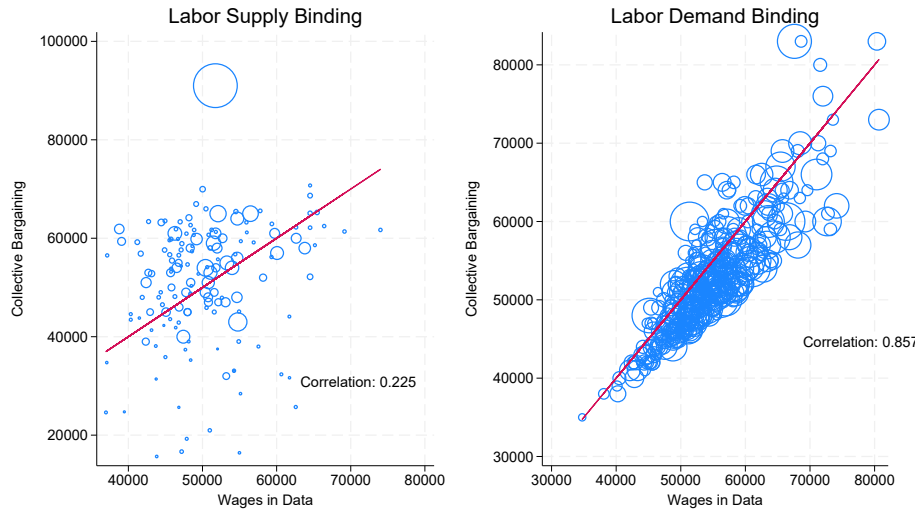
As the number of teachers hired by a district is strongly correlated with budgets, the model does a better job of explaining the number of teachers hired. The correlation between the model's prediction and actual teachers hired is essentially 1. More interesting is that our model predicts an average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

Schools on the labor supply curve face shortages; they would like to hire more teachers at the prevailing wage, but cannot do so. Thus, we can use teacher shortage data as a means to validate the model's prediction of whether labor supply or labor demand binds. We used data on emergency permits for long-term positions for the year 2017, and discuss this data in more detail in appendix A.3. Emergency permits for long-term positions can be used as a proxy for teacher shortages since they indicate long-term positions that were filled by unqualified teachers due to a lack of supply. Schools where labor supply binds hire 6 times more teachers using emergency permits than schools where labor demand binds, indicating that our measurement of binding labor supply captures much of the variation in the data of the ability of schools to hire additional teachers.

[40] An attentive reader would note that the largest district in our data, which happens to be the Philadelphia school district, over-predicts wages. One of the issues with such a large school district, both in terms of students and geography is that the travel distance from teacher house to centroid of the school district may over-predict true commuting times.



(a) Predicted and Actual Wages



(b) Predicted and Actual Wages

Figure 9: Model Fit

6.2 Welfare of Collective Bargaining and Oligopsony

Our first counterfactuals address the challenge posed in section 2: we compute the social planner solution and compare it to the collective bargaining setting. We also compare the planner and collective bargaining setting to the posted wage environment, which determines the equilibrium under pure oligposony. Thus, we can understand how oligopsony and collective bargaining individually distort the market. We will often refer to the comparison of collective bargaining to oligopsony as a "no-union" counterfactual, since it removes the collective bargaining function of labor unions. Such policies have been implemented recently in the market for school teachers in the United States, such as Wisconsin passing Act 10 in 2011 which effectively restricted collective bargaining for Wisconsin teachers.

In the oligopsony and the social planner outcome, all teachers receive offers from every school; otherwise, schools and the social planner leave welfare on the table.[41],[42] This simplifies the computation of equilibrium considerably compared to the collective bargaining solution. In addition, the oligopsony solution is nested by Nash-in-Nash, specifically for $\alpha_j^b = 1$ for all j.

The second through fourth rows of Table 11 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 average wages of $ 50,459, which is above the oligopsony model prediction of wages of $ 42,536, or a 16% decrease

[41] 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, 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.

[42] Collective bargaining involves job rationing, while posted wages do not, nor does the social planner solution. Thus workers benefit from additional potential jobs in their choice set, which is the analogue of the product variety effect but for the labor market. We do not attempt to quantify the welfare benefit from the expansion of the choice set in our nested logit model, mainly because it can be difficult to obtain quantitatively sensible estimates, that is, estimates that are not implausibly large. This is a long-standing issue in the literature.

in wages due to oligopsony power. Associated with this decrease in wages is a decrease in the number of teachers from 121,875 to 118,590, or a two percent fall. As teacher supply is fairly inelastic in terms of substitution between teaching and the outside job rather than between schools, a 16% drop in wages only results in a 2% 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 $\epsilon_S = 2.8$, which would imply a markdown in the monopsonist model in section 2 of 27 percent. However, wages are not depressed by nearly 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 the supply and demand side. In addition, the wage distortion due to oligopsony power is fairly uniform across school districts in the data.

The collective bargaining model predicts wages of $ 55,100 which is above both the planner and the posted solutions. Collective bargaining is strong enough to eliminate the oligopsony distortion on wages, but it introduces a new distortion because wages are too high. Given a 23% increase in wage, the collective bargaining model leads to 7% fewer teachers being employed, reducing teachers from 121,000 in the planner solution to 110,656 hired teachers. This is because under collective bargaining rationing of job offers occurs in 390 out of 525 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 posted — or no-union — wages on the 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 teacher unions increases wages. Thus, eliminating unions should decrease wages, since teacher union bargaining power drops. 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 non-unionized 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 non-unionized 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 at rival school districts.[43]

Model	Average Wage	Teachers Hired
Data	54,190	104,472
Collective Bargaining	55,100	110,656
Planner	50,459	121,875
Posted	42,526	118,590

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.

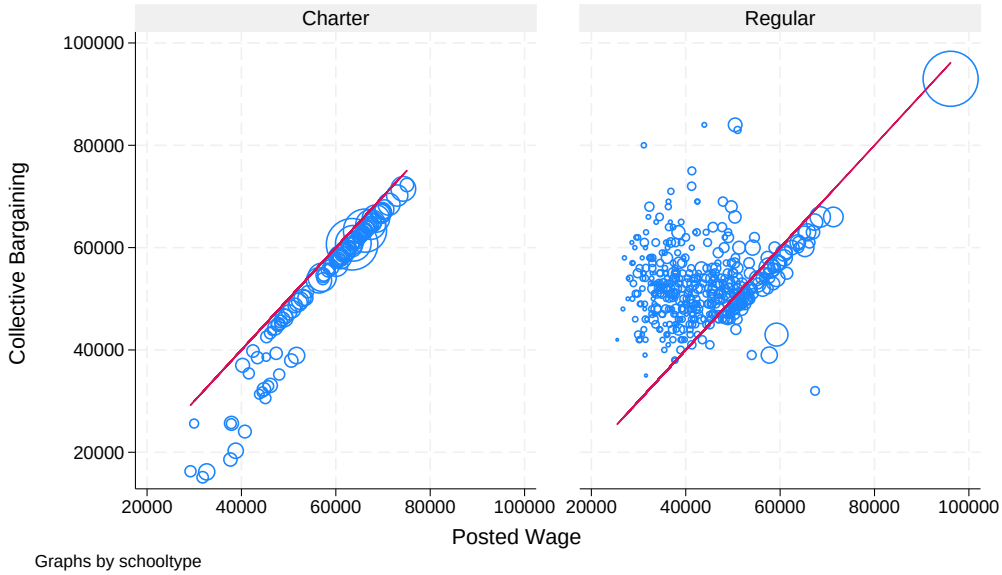
Table 11: Welfare, Collective Bargaining and Oligopsony Distortion

Collective bargaining substantially changes the distribution of wages. The posted wage equilibrium and planner problem 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 school level factors that induce teachers to quit in Table 5 and wages. Instead, the collective bargaining model rationalizes this pattern by assigning high bargaining power to schools that pay relatively low wages.

6.3 State Level Wage Setting

Many states determine teacher compensation by setting a flat wage schedule for the entire state, rather than having local school districts decide on teacher

[43]In a oligopsony with logit teacher preferences, wages are strategic complements, that is, raising the wage at one school district raises wages at rival school districts. This follows from, for instance, Caplin and Nalebuff (1991). Thus, it is likely (but not strictly a theorem in a BLP context), that oligopsony with a mixed logit also yields strategic complementarity. In contrast, if all schools were on the labor demand curve, then wages are likely (but again this is not a theorem given the random coefficients on distance) to be strategic substitutes. If a school raises wages, then that school will hire fewer teachers. This expands labor supply at rival schools leading to lower 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

compensation. For example, North Carolina paid a base salary of $41,870 in 2023-2024 for all teachers in the public school system with a bachelor's degree and three years of experience, including charter schools, along with some adjustments for the local cost of living.[44] Statewide wage posting clearly enhances the monopsony power of schools because it internalizes spillovers between employers. We find, however, that state-level wage schedules generate inequality between schools, since many schools struggle to hire enough teachers when they cannot differentiate themselves with higher wages.

To assess the effect of statewide wage posting, we endow the state agency with an objective function that aggregates the welfare of individual school districts.[45] We consider a state regulator that sets a state-level wage $\bar{w}$, and allocates budgets B_j to each school district to minimize total expenditures ($\sum_j B_j$), subject to the constraint that school districts are no worse off than under an initial wage w_j^0 and budget B_j^0. That is, budgets are adjusted to keep each district's welfare constant at its initial level, consistent with the concept of compensating variation. This can be thought of as a Pareto constraint at the school district level when setting state wages $\bar{w}$. Section C.10 in the appendix provides further details on our method.

We initially examine a simplified market that clearly shows the mechanisms at play when setting a state-level wage. We then show state-level wage results for the State of Pennsylvania. In particular, a state wage: a) allows the state to internalize wage externalities between school districts and thereby depresses wages, but b) when significant school heterogeneity exists, an equalized state wage cannot provide compensating wage variation, which leads to inefficiency. We begin by analyzing a market with identical schools.

6.3.1 Identical School Districts

We consider a simplified market with 10 identical schools and 500 identical teachers and present the entire calibration in Table 19 in Appendix C.10. Both schools and teachers are in the same location. All schools have a budget of $750,000 and the same school-level characteristics, so that each school yields identical mean utility to teachers δ_{ij}. Table 12a shows that the posted

[44] https://www.dpi.nc.gov/documents/fbs/resources/salaryschedule2324hb259pdf/download?attachment Accessed October 1, 2025, Table 1.

[45] In the case of a private firm, it would be straightforward to simply consider joint profit maximization. Since school districts do not maximize profit, we must take a different approach.

wage would be $ 53,247 for each school. The state-level wage that minimizes the total spending of all school districts is substantially lower at $ 22,000. This would reduce the spending on schools by 7.2 percent, which is substantial given that Table 2 showed that schools spend a third of their budget on teachers. The state has a lot of monopsony power because most of the diversion of teachers when wages fall is to other school districts, rather than to the outside option, as was shown in the elasticity matrix in Table 6. Indeed, the number of teachers only falls by 1.7 % between the posted and state wage setting outcomes even though wages have dropped by almost 50%.

6.3.2 Nine Identical School Districts, One Charter School

Additional complications arise when we move to a setting with differentiated teachers and schools so that teachers have varying willingness to work at different schools. When a school is particularly unattractive, perhaps due to its remote location, charter status, or high free lunch share, a flat state wage precludes the school from offering higher wages to compensate for the lack of amenities leading to a precipitous drop in teachers hired. Returning to our simple example, we assign charter status to the last school, while the remaining nine schools are unchanged. The average posted wages would then be $ 53,549 (versus $53,247 before), but the wages of the charter school would be $ 68,320. If the state imposed a flat wage across all school districts, the wage that would minimize expenditures, again subject to the constraint that no school district is worse off, would be a state wage of $ 70,000 which is far above the level of posted wages.[46] In addition, the state would need to spend 6.3 percent *more* to implement such a policy relative to posted wages. Heterogeneity in school amenity values makes uniform wage policies difficult to implement, unless the state is willing to accept welfare losses at some schools to lower wages.

6.3.3 Pennsylvania School Districts

Schools in Pennsylvania have significant heterogeneity in amenity values, making it difficult for the state to exploit its monopsony power under the constraint of keeping all schools just as well off as before. We focus our analysis on the greater State College Region in Pennsylvania and present

[46] A high wage makes non-charter schools hire fewer teachers, increasing the residual labor supply from which the charter school can hire.

	Symmetric Schools	One Charter
Average Posted Wage	53,247	53,549
Charter Posted Wage	.	68,320
State Wage	22,000	70,000
Change in State Budget	-7.2%	6.3 %
Change in Number of Teachers	-2%	-17%

Notes: The table shows a simulated region with 10 school districts and 1000 teachers, all of which are in the same location. In column 1 all schools are regular school districts, while in column 2, one school is a charter school. State Wages denote the budget minimizing uniform state wage where the starting point is the posted wage equilibrium.

(a) State Wage Setting for Case Study Market

Average Posted Wage	34,441
State Wage	57,000
Change in State Budget	14.8%

Notes: The table shows the state college region. State Wages denote the budget minimizing uniform state wage where the starting point is the posted wage equilibrium.

(b) State Wage Setting for State College

results in Table 12b. We choose as a starting point the wages and hiring outcomes from the posted wage outcome. The optimal state wage would be set to $ 57,000 and would increase spending by 14.8%. Thus, while state wage setting could lower wages due to the ability to coordinate wages between districts, it is not effective in areas with significant heterogeneity across school districts. As such, there is little benefit, at least with the type of Pareto constraint we impose, in having uniform wage setting at the state level. Indeed, states that impose uniform wage schedules tend to allow schools with hiring difficulties to set wages higher, which could be one way to balance these incentives.[47]

6.4 Minimum Wage

In this counterfactual, we consider how a minimum wage for teachers across the state of Pennsylvania would affect the equilibrium allocation of teachers. In particular, we ban unions but allow the state to set a minimum wage for

[47] North Carolina, for example, allows schools to provide a hiring bonus if they are having trouble hiring.

all schools, which is the policy in 9 states.[48] We find that minimum wages can raise wages to the social planner level; however, they cannot achieve social planner hiring due to inefficiencies at the minimum wage constraint.[49]

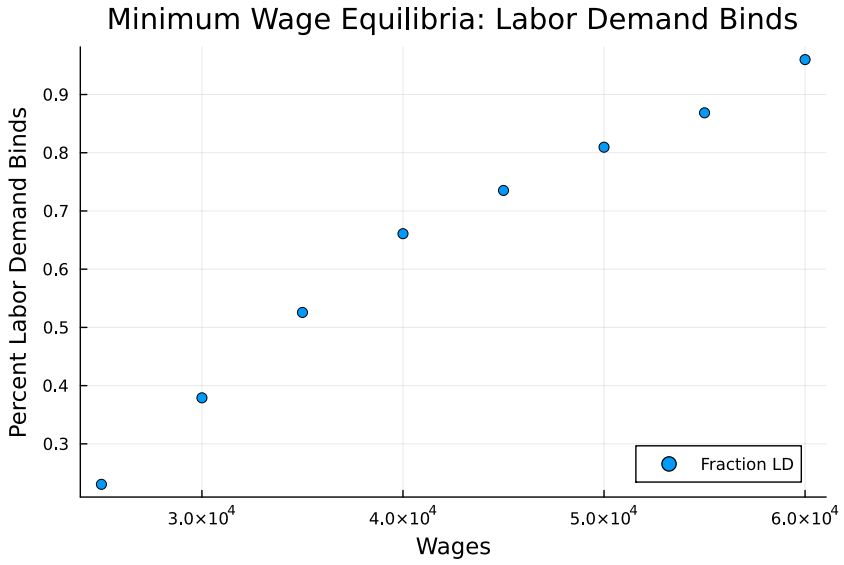
Consider a school that finds the minimum wage binding. Too many teachers want to work at the school, i.e., labor demand is below labor supply. Thus, the school will implement the rationing rule and offer fewer jobs than it would in the wage posting setting. Thus, wages increase, but rationing leads to lower equilibrium values of teachers hired. In our counterfactual, we allow schools where minimum wage binds to play the same offer game that they do with collective bargaining.

Figure 11 shows the equilibrium points of various minimum wage policies that range from \$25,000 to \$60,000 in \$5,000 increments. Panel 11a shows how changing the minimum wage on the horizontal axis increases the share of schools where labor demand binds. For instance, at a wage of \$40,000, 65 percent of schools have binding labor demand. Panel 11b presents the average wage and number of teachers hired for the state as the minimum wage increases. As minimum wages increase, so do average wages; for example, at a minimum wage of \$40,000, the average wage would be \$47,000. Equilibrium employment, meanwhile, first increases and then decreases in minimum wage as schools shift from hiring on the labor supply curve to the labor demand curve.

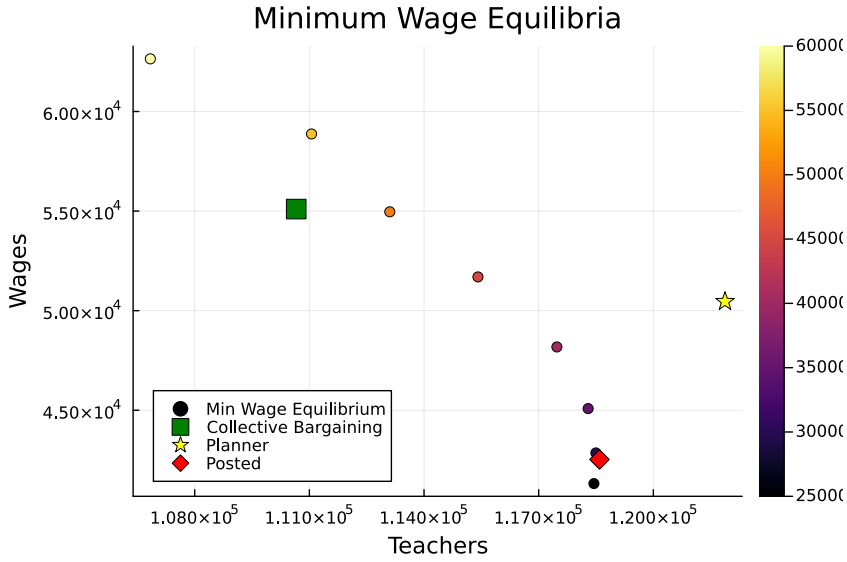
Importantly, minimum wages can increase wages to the social planner level, but employment is always substantially below the social planner outcome. This occurs because attractive schools, that could hire enough teachers at lower wages, respond to higher wages by rationing jobs.

[48] Source: https://www.nctq.org/research-insights/state-of-the-states-2022-teacher-compensation-strategies/.

[49] Of course, the social planner solution could be implemented if the state could set a unique minimum wage *for each school district*. The state would set the school level minimum wage $\underline{w}_j = w_j^*$ to the social efficient wage. The complication with setting minimum wages has to do with having a single wage floor for the entire state.



(a) Minimum Wages and Binding Labor Demand



(b) Minimum Wages Equilibria: Employment and Wages

Figure 11: Minimum Wage Counterfactual

Notes: The figure maps minimum wage policies to an equilibrium outcome. In particular, we show the equilibria for minimum wages from \$30,000 - \$ 60,000 in \$5,000 increments. The red diamond shows the posted equilibrium, the green square the collective bargaining equilibrium and the yellow star the planner outcome.

7 Conclusion

Many labor markets are characterized by a few dominant firms, that can be represented using a model of oligopsony power. This power 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.[50]

We find the potential for substantial monopsony power in the market for public school teachers in Pennsylvania, with wages under a pure wage posting arrangement 16 percent below the social planner solution on average. However, collective bargaining allows public teachers to have countervailing power that pushes their wages 23 percent higher on average than in the oligopsony scenario, which is even higher than wages under the social planner. Thus, we find that collective bargaining not only corrects the distortions from oligopsony, but introduces an upward distortion in wages. The effect of collective bargaining, however, depends crucially on bargaining power. Where bargaining power is too strong, collective bargaining adds 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 since it would need to be conditioned on which side of the market is binding.

Furthermore, we show that other policies such as statewide uniform wage setting, have the potential for depressing wages due to greater employer consolidation are offset by having to make sure relatively undesirable school

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

districts can retain teachers.

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

Azkarate-Askasua, Miren, and Miguel Zerecero. 2024. "Union and firm labor market power."Technical report, Mannheim.

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.

Caplin, Andrew, and Barry Nalebuff. 1991. "Aggregation and imperfect competition: On the existence of equilibrium." *Econometrica* 1 25–59.

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.

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.

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.

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.

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.

Kline, Patrick M. 2025. "Labor Market Monopsony: Fundamentals and Frontiers." (33467): .

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.

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, but not teaching. We refer to this group as the potential teachers. Instead, we use data on certifications and the number of teachers who leave the profession to estimate the number of non-teachers. We do our analysis by region since the exit rate is variable across regions. We find that the potential teachers make up between 22% to 36% of total teachers depending on the region. After calculating the number of total teachers in each region, 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.[51]

Teachers need an Instructional I certificate to practice as teachers.[52] Once received, it lasts for 6 years. Generally, teachers would apply for an Instructional II certificate soon after they receive their Instructional I certificate. The number of Instructional I certificates issued is a useful proxy for 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 3 years a teacher who does not teach will never teach and has exited the potential pool.

Next, we find the number of teachers who entered the teacher dataset for the first time and have years of experience equal to one. By subtracting the number of starting teachers from the number of newly licensed teachers, we

[51] Data were downloaded from https://www.pa.gov/agencies/education/data-and-reporting/school-staff/educator-preparation-and-certification/act-82-of-2018-report-on-educator-preparation-and-certification.

[52] Some teachers can get emergency permits to teach due to shortages. We ignore these permits.

Region	Actual	Certificate	Exiters	Total	Model
Philadelphia	36,107	3,752	16,673	56,532	32,605
Pittsburgh	19,717	2,049	4,562	26,328	16,487
Harrisburg	18,587	1,931	5,305	25,823	17,514
NW	11,194	1,163	2,362	14,719	9,422
NE	11,094	1,153	1,983	14,230	9,593
State College	7,265	755	1,711	9,731	6,292
Allentown	15,177	1,577	3,789	20,543	12,559
Total	119,141	12,380	36,385	167,906	104,472

Notes: The table shows the number of teachers and the potential teachers by region. Actual represents the number of teachers in each region in our teacher dataset. Certificates represents the number of teachers who have been certified but are not teaching. Exiters represents the number of teachers who have exited and have less than 20 years of experience in education. Total sums actual certificate and exiters to get the total number of qualified teachers in each region. Finally, the model has the number of teachers at in sample schools.

Table 13: Potential Teachers by Region

can get a value of the number of teachers in each year who are licensed, but not teaching. In 2017, this adds 12,380 potential teachers. We allocate these teachers proportionally across region since our certification data is not region specific.

Exiting Teachers: Next, we consider exiting teachers, i.e., teachers who have quit teaching or taken a gap year. To avoid counting retirees as part of the pool, we restrict our sample of exiting teachers to those with less than 20 years of experience in education. We assume that after 4 years out of the labor market, a teacher who permanently exits will not return to teaching.

For each year beginning in 2013, we calculate the number of teachers who permanently exit and have less than 20 years of experience. We also add teachers who have taken a gap year. Table 13 shows the number of teachers in our teacher dataset as well as the number of potential teachers that are added.

We believe our potential teacher numbers match well to other statistics. The number of Pennsylvania private school teachers in 2017 was 23,630, which is about half of our non-teachers.[53] In addition, the Bureau of Labor

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

Statistics says that the number of teachers in Pennsylvania is 152,980.[54]

A.1.2 Potential Teachers in the Model

For the structural model, we need to add potential teachers to our teacher data. In the model, teachers enter as addresses. We have 151,539 unique teacher addresses that were collected from InfoUSA. These addresses include any unique address where a teacher has lived between 2010-2021.

We sample with replacement from our teacher address database by region until we reach the total number of teachers desired, i.e., column "Total" in ??.

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 14 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.

A.3 Teacher Shortages

When the labor supply curve binds, school districts want to hire more teachers at the negotiated wage but not enough teachers want to work for them. Thus, we would expect school districts where labor supply binds to report shortages. We collected various shortage measures to evaluate the ability

[54] 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.

	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, not at all, or more than once to the InfoUSA data in 2017.

Table 14: 2017 Statistics Of Teachers Matched to InfoUSA

of the model to correctly identify school districts where labor supply binds. Using data from the 2017-2018 school year (which we used to calibrate the model), we show that the school districts where labor supply binds according to our model, are significantly more likely to have a shortage. [55]

Our preferred measure of shortages comes from emergency permit data, which can be used as a proxy for shortages.[56] Emergency permits allow teachers without the proper certifications to teach at schools. The Pennsylvania Department of Education issues these if "An emergency permit is issued by the Department upon the request of an employing public school entity (LEA) when a position has been advertised and no fully qualified and properly cer-

[55] Vacancy data from 2024-2025 exists, but it was not collected in earlier years. If we assume that shortages are stable over time, then the vacancy data also supports our model. Schools with labor supply binding are more likely to have vacancies than those without. Data available at: https://ceepablog.wordpress.com/2025/03/28/teacher-shortages-examining-pennsylvania-teacher-vacancies-in2023-24-and-2024-25/

[56] Accessed: Accessed September 12, 2025 https://www.pa.gov/agencies/education/data-and-reporting/school-staff/educator-preparation-and-certification.html#accordion-43b97ca7af-item-61eeb6dfb7

	LD Bind	LS Bind
Shortage	0.4%	4.0%
Number	419	106

Notes: The table determines the average percentage of teachers hired at a school using an emergency permit in 2017-18, which is a proxy for a teacher shortage (weighted by teachers). Our model predicted that 419 schools were bound by labor demand, whereas 106 schools had labor supply binding. The table indicates that schools where labor supply binds hired almost 8 times more teachers using emergency permits than those where are model predicted that labor demand binds.

Table 15: Percent Teachers hired with Emergency Permits

	LD Bind	LS Bind
Shortage	6.5%	28.6%
Number	418	7

Notes: Using 2017-2019 data from the US Department of Education, we characterize school districts into whether there is a teacher shortage area or not. The table shows that schools where labor supply binds are much more likely to be in a teacher shortage area, than those where labor demand binds. We did not include charter schools for this result.

Table 16: Teacher Shortage Areas

tificated applicant is available"[57]. They are used primarily at schools that struggle to hire teachers. The number of emergency permits, therefore, tells us how many substitutes and under-qualified teachers school districts have hired in each year. In particular, we focused on teachers hired for long-term positions either with or without educational obligations. The 2017-2018 data demonstrates that schools where labor supply binds are more likely to have shortages. They hire 6 times more teachers via emergency permits, than schools where labor demand binds.

Next, we looked at data on teacher shortage areas from the US Department of Education.[58] This dataset designates shortage areas at the school district level, intermediate unit level, and state level. We only use the school district level data, since the data does not have good coverage of charter schools. As shown in Table 18, 40% of school districts where labor supply binds are considered to be in a shortage area on average for the years 2017-2019, while only 5.8% of schools where labor demand binds have shortages.

[57] See https://www.pa.gov/agencies/education/programs-and-services/educators/certification/current-pa-educators/emergency-permits, Accessed September 12, 2025

[58] Available at: https://tsa.ed.gov/#/reports.

B Proofs and Derivations

B.1 Nash Bargaining Derivations

B.2 Union utility function: w*T

The problem is:

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

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:

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

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

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

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

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

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

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

$$\alpha_j = \frac{W_j(X_j,T_j)(T_j + (w_j - r)\frac{\partial T_j}{\partial w_j})}{(W_j(X_j,T_j))(T_j + (w_j - r)\frac{\partial T_j}{\partial w_j}) - (-k_1 T_j \frac{\partial W_j}{\partial X_j} + \frac{\partial T_j}{\partial \omega_j}\Big[\frac{\partial W_j}{\partial T_j} - (k_2 + k_1\omega_j)\frac{\partial W_j}{\partial X_j}\Big])} \tag{20}$$

B.3 Selection Correction for Labor Supply

Each teacher's utility from accepting a job offer in public education has a common component ν_i that follows a $N(0,1)$ distribution. We estimate preferences over job attributes by maximing the likeihood of quitting instead of keeping the current job, which is given by Equation **??**. In order to construct this likelihood function, we must integrate over the distribution of ν conditional on quitting, $q_i = 1$, or keeping the current job and not quitting, $q_i = 0$. However, the distribution of ν conditional on not switching jobs will not be $N(0,1)$, as individuals who switch jobs will tend to have higher values of ν_i than

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:

$$\begin{aligned} s_{ij} &= \int_\nu \frac{o_{ij} \exp(\psi\omega_j + x_j'\beta - \tau d_{ij} + \sigma\nu)}{\exp(r) + \sum_{k=1}^k o_{ik} \exp(\psi\omega_j + x_j'\beta - \tau d_{ij} + \sigma\nu)} \phi(\nu) d\nu \\ &\approx \sum_{n=1}^N w_n \frac{o_{ij} \exp(\psi\omega_j + x_j'\beta - \tau d_{ij} + \sigma\nu_n)}{\exp(r) + \sum_{k=1}^k o_{ik} \exp(\psi\omega_j + x_j'\beta - \tau d_{ij} + \sigma\nu_n)} \end{aligned} \tag{21}$$

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.[59]

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:

$$L_j^D(w_j) = \frac{w_j^{\frac{1}{\rho-1}} B_j}{\left(\frac{\gamma}{1-\gamma}\right)^{\frac{1}{\rho-1}} + w_j^{\frac{\rho}{\rho-1}}}. \tag{22}$$

The elasticity of labor demand is given by:

$$\frac{\partial log(L_j^D)}{\partial log(w_j)} = \frac{1}{\rho - 1} - \frac{\frac{\rho}{\rho-1} w^{\frac{\rho}{\rho-1}}}{(\frac{\gamma}{1-\gamma})^{\frac{1}{\rho-1}} + w^{\frac{\rho}{\rho-1}}}. \tag{23}$$

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

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

$$L_j^D(w_j) = \sum_{\iota(j)=1}^{\bar{\iota}_j(\mathbf{w})} s_{\iota(j)j}(\mathbf{w}, (o_{\iota(j)j} = 1, \mathcal{O}_{\iota(j),-j})). \tag{24}$$

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

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.

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 17

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: $\mathcal{O}'_j = \mathcal{O}_j(\mathbf{w}, \mathcal{O}^k_{-j})$
      Update offer matrix for that school: $\mathcal{O}^k_j = \mathcal{O}'_j$
    end for
  end while
  return $\mathcal{O}$
end procedure

District Cost Per 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 17: Fringe Benefits

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 $\xi(\mathbf{w}) \equiv \begin{bmatrix} \xi_1(\mathbf{w}) \\ \xi_2(\mathbf{w}) \\ \cdots \\ \xi_J(\mathbf{w}) \end{bmatrix}$, 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:

$$\frac{\frac{\partial W}{\partial T}}{\frac{\partial W}{\partial x}} = \frac{\gamma}{1-\gamma} \cdot \left(\frac{T_j}{X_j}\right)^{\rho-1} = w_j(1 + \frac{1}{\epsilon_j}). \tag{25}$$

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

C.8 Parallel Implementation

Given there are 525 schools and 113,242 teachers in teaching jobs, 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. This is done by making school level choices run in parallel for both planner, poster, and Nash Bargaining solutions, holding the choices of other schools implicitely fixed.

C.9 Partition Analysis

Even with parallel implementation, the Nash Bargaining computation would take a month to run. Instead, we run our analysis using a partition of Pennsylvania into separate labor markets. We choose seven regions composed of groups of counties. We pick county borders to minimize school district interactions across borders, that is borders where there were fewer school districts directly next to each other across the border. We wanted each region to have less than 100 schools to aid computation, the exception being the Philadelphia Area. Figure 12 shows how we partition Pennsylvania. Using this method, computing the Nash Bargaining solution goes from a month if we pick the market to be the State of Pennsylvania to 2 days with a seven

Penn Pl Wage	Partition Pl Wage	Penn Po Wage	Partition Po Wage
48,427.72	48,907.63	40,339.61	40,785.36
(20,013.99)	(20,028.02)	(20,203.36)	(20,242.74)

Notes: We ran the planner and posted counterfactuals for the whole state of Pennsylvania and also using the partition method. The results here show that the partition methodology produces similar results as the entire state of Pennsylvania.

Table 18: Partition Comparison

market partition.

As our counterfactuals for the Planner and Posted outcomes for the entire state only take a couple of hours, versus a month for the Nash Bargaining solution, we can compare how the partition performs relative to the whole state. We find that the wages are quite similar for the seven market partition versus the whole state analysis as shown in Table 18. We find that wages are somewhat different around border regions, as would be expected. However, these differences average out when we consider the whole state.

C.10 State-Level Wage Setting

To assess the effect of such a statewide wage posting, we need to endow the state agency with an objective function. In general, when looking at the effects of consolidation we would simply add up firm profits, but this is not possible in our setting since the school district objective functions are not in comparable units. Instead, we rely on compensating variation. Consider the indirect utility function $V_j(w_j, B_j) \equiv \max_{T_j \leq \sum_i s_{ij}(w_j)} W(w_j T_j, B_j - w_j T_j)$, or the maximum utility that a school district achieves given a wage w_j, budget B_j, and the labor supply constraint. Given a wage w_j^o and budget B_j^o, the compensating variation for movement to a wage w' would be the budget change $B' - B_j^o$ that yields the same indirect utility. That is, we want $B^{CV}(w')$ such that $V_j(w_j^o, B_j^o) = V_j(w', B^{CV})$. Note that because of externalities between school districts operating on competing offers, we are actually looking for a market equilibrium in terms of this compensating variation. Consider the vector of wages $\vec{w} = \{w_1, w_2, \cdots, w_J\}$ and budgets $\vec{B} = \{B_1, B_2, \cdots, B_J\}$. We find a vector of $\vec{B}^{CV} = \{B_1^{CV}, B_2^{CV}, \cdots, B_J^{CV}\}$ such that given the wage change to $\vec{w}'$ we have:

1. $V_j(\vec{w}, \vec{B}) \equiv \max_{T_j \leq \sum_i s_{ij}(w_j, w_{-j}, \mathbf{O}_{-j})} W(w_j T_j, B_j - w_j T_j)$, where the choice

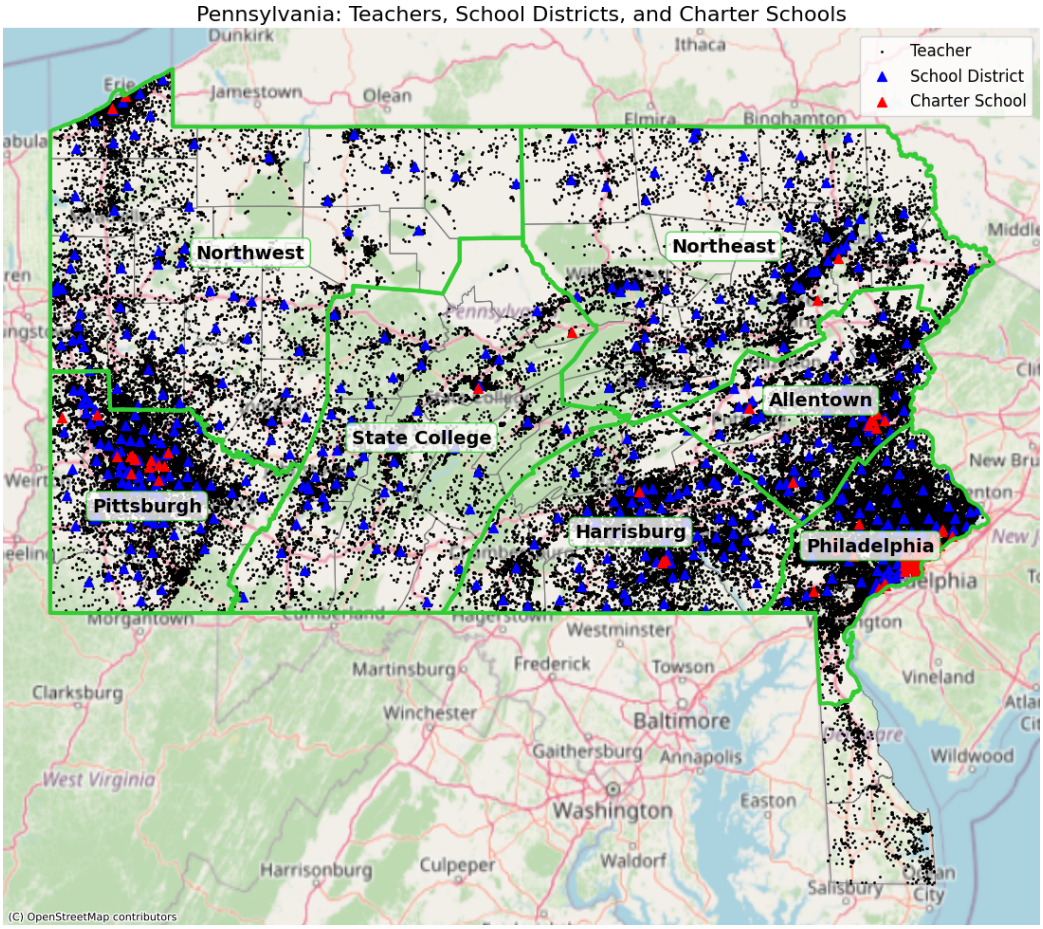


Figure 12: Pennsylvania Partition

Notes: The map depicts teacher locations and school locations in Pennsylvania. It also shows the seven regions that we used to partition Pennsylvania when we compute Nash Bargaining for the whole state.

Parameter	Value
Number of Schools	10
Number of Teachers	500
Budget per school	750,000
Mean δ_j for Regular School	10.0
Mean δ_j for Charter	9.0

Note: All other parameters are those used in the body of the paper.

Table 19: Parameters for Identical Schools

of T_j also defines an offer set $\mathbf{O}_j$. This is the equilibrium of the labor demand system in section 5.5 with a new set of wages and budgets.

2. $\vec{B}^{CV}$ such that $V_j(\vec{w}^o, \vec{B}^o) = V_j(\vec{w}', \vec{B}^{CV})$ for $j = 1, \cdots, J$.

The state budget is given by $\mathbf{B} = \sum_j B_j$. Suppose then that the state agency chooses a state wage $\bar{w}$ to minimize total expenditures subject to the constraint that no school is worse off under this allocation.[60] That is:

$$\min_{\bar{w}} \sum_j B_j^{CV}(\bar{w}). \tag{26}$$

C.11 Identical School District Simulation

For the simulation of identical school districts we use the following calibration:

D Monte-Carlo

In this section, we perform a Monte-Carlo to assess the the quitting regression used to estimate labor supply.

The data generating process is the following. First, offers o_{ij} are sent out using the following:

$$o_{ij} = 1(\gamma_0 + \gamma_1 w_j + \gamma_2 d_{ij} + \gamma_3 c_j + \gamma_4 f_j + \epsilon_{ij}^o > 0) \tag{27}$$

[60] Clearly this exercise depends on the exact starting point to compute compensating variation. We show the results with starting points at the real, planner, posted, and bargaining solutions

Offer Parameters	Low Switching	High Switching
γ_0	-1	20
γ_1	-0.00001	-0.00001
γ_2	-0.04	-0.04
Implied Move Rate	4%	80%
Utility Parameters		
β_0	3	3
β_1	0.00005	0.00005
β_2	-0.01	-0.01
β_3	-0.5	-0.5
β_4	-1.2	-1.2
σ	13	13

Table 20: Parameters for Monte-Carlo

where $\epsilon^o_{ij} \sim \mathcal{N}(0,1)$. This is simply a probit like model to generate offers, that may or may not be correlated with the covariates we are using. For one of the options, called the *stay* option, teachers receive an offer with certainty, thus, o_{ij} if j is teacher i's stay school. As well, all teachers receive the outside option with certainty. Thus, the stay and quit decisions are always in the offer set.

Second, we generate utilities for teachers given by:

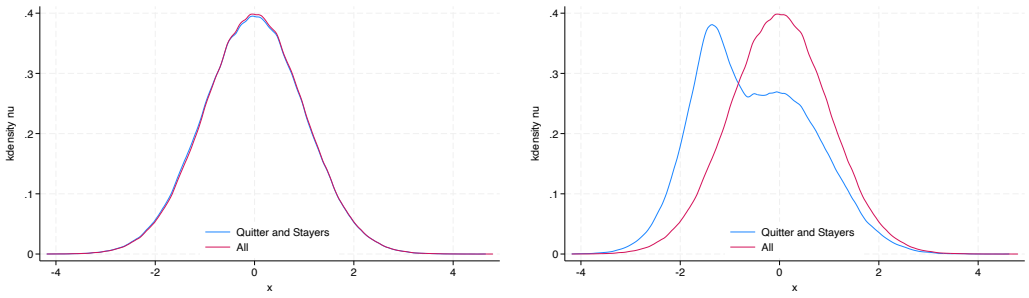
$$u_{ij} = \beta_0 + \beta_1 w_j + \beta_2 d_{ij} + \beta_3 c_j + \beta_4 f_j + \sigma\nu_i + \epsilon_{ij}, \tag{28}$$

where ϵ_{ij} are logit draws and $\nu_i \sim \mathcal{N}(0,1)$, c designates charter schools, f the free lunch share.

Third, agents choose the highest u_{ij} among the offers o_{ij}, for us to get $j^*_i \equiv js.t.u_{ij} \geq u_{ik} \forall k$.

We use the parameters in Table 20, where the utility parameters, i.e. the β's, are chosen to mimic our empirical results, while the offer parameters are more arbitrary, and are mainly targeted to get a specific move rate, either close to the data at around 3 percent, or a much higher rate around 70 or 80 percent to illustrate how higher move rates would cause different problems.

We show the distribution of ν, conditional on teachers choosing to stay in their current job or quit the teaching profession in Figure 13, plotted in blue, while the normal distribution, i.e. the unselected distribution, is plotted in



(a) ν distribution: 5% move rate (b) ν distribution: 80% move rate

Figure 13: Selection on ν

red. Panel a) shows the case with a 5 percent move rate, which matches the empirically observed move rate in our data, and Panel b) shows an 80 percent move rate. The distributions in panel a) are very hard to distinguish, and indeed, accounting for selection will make little difference in our estimates. In panel b, the selected distribution of ν is quite different from the unselected one, in particular, since we lose a lot of mass in the upper tail since we take out the movers who tend to have higher taste to work in public education.

Table 21 presents results from our Monte-Carlo. Column 1 shows the issue with a high move rate of 80 percent and $\sigma = 13$. Column 2 simply shows the case where $\sigma = 0$, where selection on ν is not an issue. Column 3 present the case where $\sigma = 13$, but there are no moves. Column 4 presents $\sigma = 8$ and a low move rate of 5 percent, which is above what we see in the data. Column 5 presents $\sigma = 13$ and a low move rate of 5 percent, which is above what we see in the data. Column 6 present our selection correction for the same specifications as column 5.

All columns except for column 1 correctly estimates the parameters as expected.

	True	(1)	(2)	(3)	(4)	(5)	(6)
Parameters							
σ		13	0	13	8	13	13
Move Rate		80%	80%	0%	4%	4%	4%
Selection Correction							X
Estimate							
Wage	0.50	0.26	0.50	0.52	0.49	0.51	0.52
		(0.60)	(0.10)	(0.17)	(0.11)	(0.17)	(0.07)
Distance	-1.00	-0.64	-0.97	-0.98	-0.95	-0.99	-1.03
		(2.25)	(0.34)	(0.72)	(0.45)	(0.69)	(0.29)
Charter	-0.50	-0.33	-0.50	-0.51	-0.48	-0.51	-0.49
		(0.84)	(0.13)	(0.19)	(0.13)	(0.19)	(0.09)
Free Lunch	-1.20	-0.57	-1.21	-1.26	-1.20	-1.25	-1.16
		(0.82)	(0.12)	(0.23)	(0.13)	(0.21)	(0.09)
Constant	3.00	23.47	2.98	3.61	3.45	3.44	4.09
		(2.51)	(0.44)	(0.73)	(0.49)	(0.72)	(0.29)
Replications		100	100	100	100	100	50

Table 21: Monte Carlo Results

Standard deviation in parenthesis.

E Additional Tables and Figures

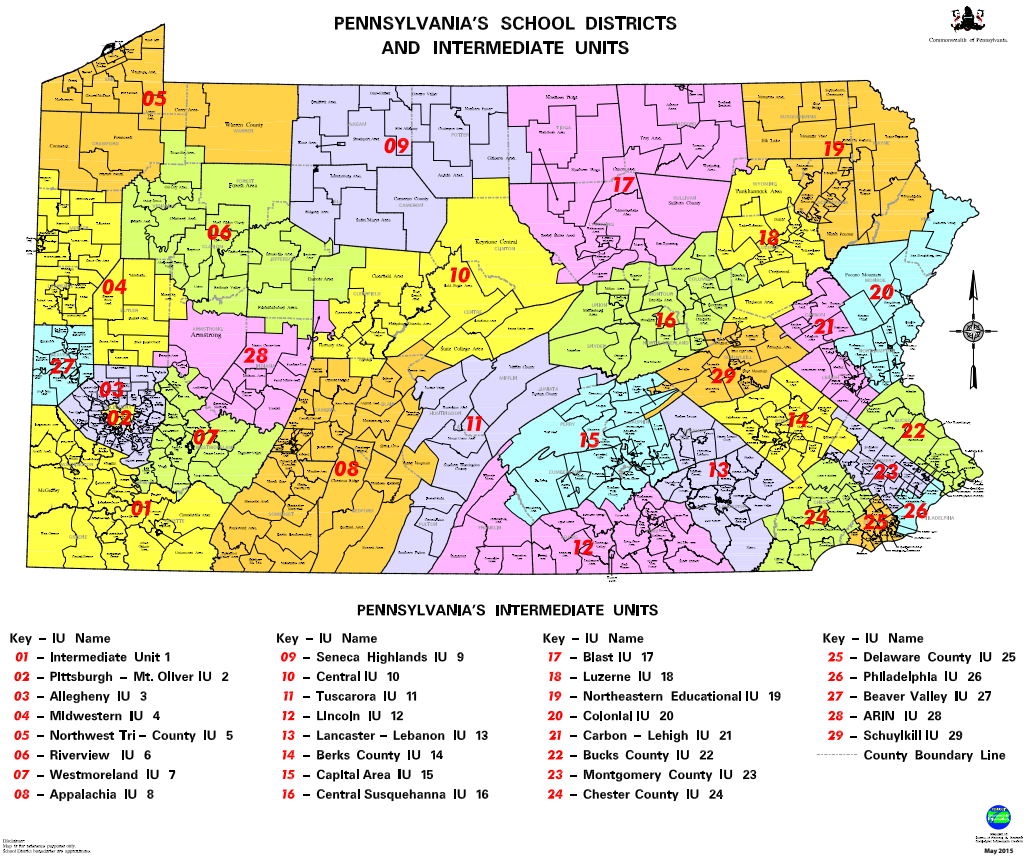


Figure 14: Map of School Districts in Pennsylvania

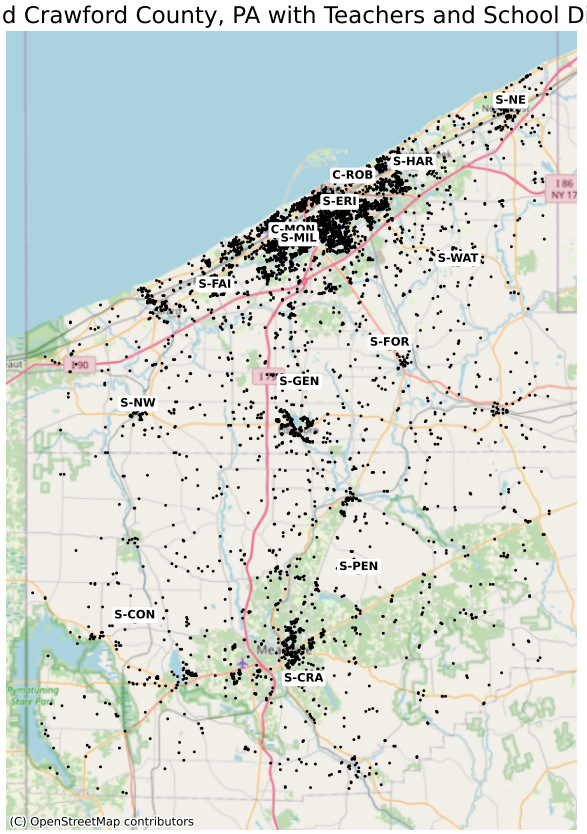


Figure 15: School District Locations in Erie and Crawford County, Pennsylvania

Notes: The map depicts teacher locations and school locations in Erie and Crawford county. The schools are labeled with the first three letters of their name.