Publicly, the impact of teachers unions on education is a hotly debated topic. Opponents posit that teachers unions constrain the ability of public officials to implement policy change, raise the cost of providing quality education, and divert funds from students. Since the 1970s, stimulated both by the 1966 Coleman Report and the increase in teacher unionization, social scientists have been interested in the role of teachers unions in the production of education. There has been substantial research on the role of teachers unions, both in economics[1] and education literature. This research explores the multitude of effects teachers unions may have including wage impacts, changes to overall expenditures, changes in the quality of teachers in a district, changes to the components of collective bargaining agreements, effects on student achievement, and class size effects.

In this paper, we conduct a meta-analysis of the oldest and largest subgroup of teachers union research, the differential wage impacts of unions. Beyond the fact that this literature is large enough to be credibly analyzed with meta-analytic techniques, teachers' wages account for approximately 60 percent of current expenditures in public schools.[2] Our objective is to synthesize the research to date and report any general conclusions. We are interested in the magnitude of union wage impacts to establish whether union wage effects are of primary importance.

Another method used to summarize the literature on teachers union wage impacts is to conduct a standard literature review. Freeman (1986), Lewis (1986), and Ehrenberg and Schwarz(1986) use this approach. Lewis finds that studies estimate a teachers union wage gap in the range of -1% to 21%.[3] From prior research, he suggests that increasing these measures by 3% would give a more accurate estimate of the union effect because fringe benefits are not included in the measure of compensation used in the included studies. This summary of results does not consider the accuracy of the estimated wage gaps and it is not clear how statistically insignificant results are treated. Our paper furthers the review of the literature by including more recent studies, studies with both

[1] The wage impacts of labor unions has been a topic of general interest in labor economics since the 1960s. See H. Gregg Lewis (1986) for a review of the early literature on union wage impacts.

[2] U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), "National Public Education Financial Survey," selected years 2000–01, 2005–06, 2009–10, and 2010–11. See Digest of Education Statistics 2013, Table 236.60.

[3] See Table 14, p. S295.

significant and insignificant wage effects, and combining results across studies to increase statistical power. In an effort to appropriately interpret the combined evidence on the teachers union wage gap, literature reviews often spend substantial effort evaluating the empirical methods of the included studies and making a judgment about how much weight we should assign to their results. The value of meta-analytic techniques is that we can specify moderators that capture differences in the studies' empirical methods and comment on whether these differences produce significantly different results.

I Collection of Studies

The studies were collected between February 11 and April 8, 2014. We performed a comprehensive search of the literature including both published and working papers. First, we identified a set of 136 possible studies by searching WebofScience, JSTOR, EconLit, GoogleScholar, the NBER Working Papers Series, researchers' curriculua vitae, the reference lists of relevant papers, and literature reviews. These studies were amassed using the following search terms in each database: teacher* union*, teacher* wages, collective bargaining, teacher* salary, and public sector unions.[4] We used three inclusion criteria for studies: 1) the study contains original empirical research , 2) the study contains a wage equation, 3) the sample includes both unionized and non-unionized districts, and 4) the study examines the U.S. teacher labor market.[5]

We created the inclusion criteria to mitigate "factual heterogeneity"[6], i.e. to be sure that we are examining the same phenomenon. From the original set of 136 studies only 19 studies meet our inclusion criteria. These studies yield 77 estimates of union wage effects. Table 1 provides summaries of all the papers included in our sample.

[4]A list of these studies is available upon request.

[5]We are aware of Dolton and Robson (1996), but have elected to exclude it because it utilizes data on England and Wales. The role of teachers unions may be substantially different in the United Kingdom than in the United States. In an earlier draft of the paper, we included their study. The results are unchanged and the study received very little weight in the analyses because of its small sample size. We also excluded Moore and Raisian (1987) because the paper did not include any sample sizes.

[6]This term is taken from Nelson and Kennedy (2009).

The first two criteria are self explanatory. The third criteria, however, necessitates discussion. Studies that include only unionized districts use covariate(s) of interest that capture union strength rather than union presence. While these studies tell us that unionization status does not conform well to a treatment and control paradigm, they do not show how unionized districts perform relative to non-unionized districts. Research on union strength is present in both older and newer papers. Ehrenberg and Chaykowski (1988) use data on 700 school districts in New York State represented by the American Federation of Teachers (AFT). Their covariates of interest are the presence of particular provisions in a district's collective bargaining contract. Strunk (2011) uses a more refined measure of union strength, i.e. "contract restrictiveness,"[7] to examine the effects of union strength on student achievement. She also includes only unionized districts. Although research on the heterogeneity of teachers unions is part of untangling their impact on public education, these papers are excluded from this analysis because they do not provide evidence of union wage effects. Without knowing that unionization status has monotonic effects on outcomes, progressing in the order of no union, weak union, strong union, these studies cannot be used to infer union effects.[8] Eberts and Stone (1985) aptly describe their research which includes only unionized districts by stating that the hypothesis they are testing "is affirmed for similar individuals who work for equally prosperous employers (and who, in a collective bargaining context, are members of equally strong unions)."[9]

A critical part of any analysis of estimates of the teacher union wage gap is an understanding that studies are measuring the wage gap in very different settings. A general theory of public sector unions, and/or teachers unions particularly, has been hard to pin down. The only paper we are aware of that proposes an explicit theoretical model is Babcock and Engberg (1997). Gregory and Borland (1999) lay out important factors to consider when examining the effect of a wage bargaining institution. Drawing on the classification of Maguire (1993) the three key factors appear to

[7]Strunk uses an item response framework to generate a measure of contract restrictiveness. Like the scoring of a standardized exam, she uses the sample of collective bargaining contracts to determine the percentile rank of a particular contract.

[8]Han (2012) has evidence that shows that the strength of the legal environment does not have monotonic effects on salaries. See Table 4, Column 1.

[9]See p. 279.

be: geographic scope, the form of the wage setting process, and the right of an organization to be the exclusive representative of a block of employees.[10]

At the most fundamental level, a union is interested in maximizing the well-being of its constituents. Teachers unions often negotiate for increases in teacher pay, reduced class sizes, better work environments, curriculum reforms, and the method of teacher evaluation. The empirical literature has not arrived at a general consensus on most bargaining outcomes. There is some convincing evidence that teachers unions reduce the likelihood that a pay-for-performance scheme will be implemented in the district (Goldhaber et al. (2008)). A union may have purely rent seeking goals or they may internalize student learning to some degree through a paternalistic view of their students. Even if a union is purely rent seeking, their behavior may have desirable consequences for the students affected. Findings on the productivity impacts of teacher's unions are still mixed. (See e.g. Allen (1986), Eberts and Stone (1985), Hoxby (1996), Milkman (1997), and Pantuosco and Ullrich (2010)) For instance, if unions successfully negotiate for reduced class sizes, this can have positive impacts on student learning and adult outcomes (Chetty et al. (2011)). Even increases in teacher salaries may cause districts to employ more qualified teachers or increase their dismissal of unsuccessful teachers before tenure binds (Han (2015)).

Part of the literature we survey attempts to identify the type of teachers that are served by the unions by examining the differing impacts on new and senior teachers earnings (Han (2011), Lipsky and Drotning(1973), West and Mykerezi (2011), Winters(2011), and Zwerling and Thomason (1995)). The consensus is that teachers unions increase the earnings of senior teachers, but not new hires. In the subsequent analysis, we check this result and gain new insight through the increased statistical power afforded by meta-analytic techniques.

One of the concerns when constructing a meta-analysis data set is the independence within and between study estimates. Meta-analysis practitioners are usually concerned about between study dependence when more than one study is produced by the same team of researchers or from the same data source. The first concern is not generally an issue in economics. The second source

[10] See section 3.3 of Gregory and Borland (1999) for a discussion of these factors.

of between study dependence is also not likely, given that each study's data set is compiled from multiple sources. There are 3 cases where a pair of studies use the same data source. The Census of Governments (COG) for 1972-1992 is a common source for wage and demographic data in Lovenheim (2009) and Hoxby (1996).[11] The 1980 Census of Population and Housing (CPH) is used in Kleiner and Petree (1988) and Tracy (1988). Both Baugh and Stone (1982) and Tracy (1988) gather data from the 1977 Current Population Survey (CPS). The selection of districts from the COG in Hoxby (1996) and Lovenheim (2009) is very different due to the unionization measure each employs. Lovenheim's study focuses on districts in three midwestern states. Hoxby's study is nationally representative of independent districts in the U.S. The CPH and CPS are both random samples of the U.S. population.

Many meta-analyses deal with within study dependence by selecting only one estimate from a study. Given the variety of data and measures of unionization utilized within some of these studies, we first select multiple estimates from a study based on the following criteria and then use robust variance estimation methods (RVE) to deal with any dependence between effect sizes[12] . We select estimates if the measure of unionization is significantly different[13] or the data source differs across specifications.[14] In Han (2012), we choose each measure of unionization and each estimate coming from a mutually exclusive legal environment[15]. The intention of this currently unpublished study is to examine the heterogeneity of union effects across legal environments. Therefore, Han (2012) contributes a large number of estimates to our analysis. We are sensitive to the weight provided to this study and conduct sensitivity analyses to check its impact on our results. In the majority of the papers sampled, multiple empirical specifications are presented that utilize the same data and measure of unionization. In this case, we begin by selecting the authors preferred specification. If the author does not state their preference, then we randomly select one of the estimates. We also

[11] Hoxby (1996) also uses it as source for her unionization measure.

[12] See Tanner-Smith and Tipton (2014) and Hedges et al (2010) for a discussion of this method

[13] For instance, Han (2012) uses four measures of unionization in separate empirical models; the presence of a collective bargaining agreement, the presence of a meet and confer agreement, union membership, and union density.

[14] For instance, the West and Mykerezi (2011) study contributes two estimates because they use the Teacher Rules, Roles and Rights data compiled by the National Council for Teacher Quality, as well as the Schools and Staffing Survey compiled by the Bureau of Labor Statistics.

[15] Legal environments range from states that prohibit collective bargaining to those that explicitly protect its use.

perform a sensitivity analysis by selecting only one estimate at random from a study to check the robustness of our results.

The RVE method mitigates any within study dependence by re-weighting the individual study estimates such that an estimate coming from a study that contributes multiple estimates will have less weight than a estimate that is the sole contribution of a study. The weights used are explained in Section 3.

II EMPIRICAL METHODS OF INCLUDED STUDIES

The studies included all contain a wage equation of the form

$$wage = \alpha + \beta union + X\gamma + \epsilon \tag{1}$$

where either the teacher, district or state is the unit of observation. $wage$ is a measure of a teacher's salary, $union$ is a measure of unionization, X is a vector of observables included to control for selection bias, and ϵ is the error. All of the papers justify their specification of (1) based on conceptual considerations and institutional knowledge of school districts. The papers vary in how they measure wages and unionization status, as well as the set of included covariates. As a result, the estimates of $\beta(\hat{\beta})$ cannot be directly compared. We use the standard meta-analytic technique of converting these into partial correlation coefficients to facilitate comparison [16]. The majority of studies use the district as the unit of observation. This is a reasonable choice since the union intervention occurs at the district level and district level data is readily available.

$wage$ is generally specified as the natural logarithm of a teacher's hourly wage to allow the coefficients on the measure of unionization to be read as percentage changes. Transforming wages by taking the natural logarithm has been found to fit the data well where returns to schooling are estimated.[17] 58 of the 78 estimates specify $wage$ as the natural logarithm of wages. A subgroup

[16] See Djankov and Murrell (2002) for a similar application

[17] See Lemieux (2006).

analysis shows no statistically significant difference in the mean effect size based on whether a specification uses a logarithmic transformation of the wage. Empirical models that specify the dependent variable as the natural logarithm of wages have an effect size of 0.033 compared to 0.025 for studies that do not transform wage data in this way.

Another important variation for our purposes is whether a study measures the average wage of teachers in the district (or state), the wage earned by new teachers, or the wage earned by experienced teachers. It is these type of differences in the measurement of $wage$ that lead us to follow the standard meta-analytic technique of calculating partial correlation coefficients to summarize the empirical results.

Partial correlation coefficients can readily be calculated from statistics reported in all empirical papers that use regression. One of the benefits of partial correlation coefficients is that they are unit-free and, therefore, allow the comparison of results across the heterogeneous specifications. A partial correlation coefficient is one type of effect size measure. Since we only utilize partial correlation coefficients, we use the term partial correlation coefficient and effect size interchangeably.

$union$ is also measured in different ways by researchers. There are generally three strategies for measuring unionization: 1) the presence of a collective bargaining agreement (CBA), 2) union membership/coverage, and 3) characteristics of the legal environment. The presence of a CBA is the most common measurement of unionization (accounting for 42.8% of the estimates surveyed). This is likely the result of a long history of similar specifications used to estimate union wage gaps in the private sector. Unionization of a district, however, does not necessarily result in a CBA. This distinction is unimportant if a large percentage of unionized districts have CBAs.

The 2011-12 School and Staffing Survey (SASS) reports that the percentage of public school districts in the United States with a collective bargaining agreement is 50.2%. The remaining half of districts either have no agreement (40%), a meet-and-confer agreement (8.4%), or some other, non-binding form of agreement (1.4%).[18] It would be helpful to have a measure of the percentage

[18] See Table 7. Moe (2011) reports the percentage of teachers covered by collective bargaining agreements and adjusts the SASS reported percentages because of changes in the wording of the questionnaire. Using his counts, 63% of teachers were covered by a CBA in 2008. (pg. 48)

of districts that are unionized in the United States, as well as the percentage of students educated in unionized districts. We are not aware of good nationwide measures of either of these. Data from the Census of Governments, such as the 1972-1992 data used by Hoxby (1996), is not available for recent years. Hoxby reports that in 1992 59% of districts in the U.S. (covering 69% of students) had meet-and-confer provisions in place, 52% (covering 63% of students) reported the use of collective bargaining as the form of negotiations, and 36% (covering 43% of students) met her definition of unionization (had a CBA, collective bargaining was the form of negotiations, and at least 50% of teachers belonged to the teachers' organization). This suggests that measuring unionization in other forms may be of considerable importance to understanding the link between unionization and teachers' wages.

Lovenheim (2009) presents evidence that in three states, Iowa, Indiana, and Minnesota, with duty-to-bargain laws 100% of districts that unionize successfully obtain a CBA. This, however, does not account for union activity in states with no collective bargaining law and states that prohibit collective bargaining. Moe (2011) shows that in Alabama, a state without a collective bargaining law, 84% of teachers report being members of the union. Estimates of the impact of unionization on wages in these environments are an important part of the population of interest.

Measuring unionization by membership or coverage captures unionization that is not represented by the presence of a CBA. These two measures of unionization are distinct from one another. To see this distinction, consider two states, state A and state B, each with 2 districts containing an equal number of teachers. If teacher membership for state A is 70% in one district and 40% in the other district, then the state will report 55% of its teachers are unionized. If teacher membership for state B is 58% in one district and 52% in the other district, then the state will also report 55% of teachers are unionized. Since teachers vote on whether to be unionized and a simple majority often determines the result, 50% and 100% of teachers are likely to be covered in state A and B, respectively. Of the 22 estimates that use these forms of unionization, only 2 measure the impact of union coverage. [19] As a result, no comparison can be made between how the estimates obtained

[19] Kasper (1970) and Kleiner and Petree (1988).

from these specifications differ. Data limitations likely account for the use of membership instead of coverage.

The final category of unionization measures uses characteristics of the legal environment to proxy for unionization. Stoddard (2005), for instance, creates a dummy equal to 1 if a district has a duty to meet, agency shops are permitted, or union shops are permitted; and 0 otherwise. Stoddard also uses an index of the favorability of the legal environment for union organizing. Gyourko and Tracy (1991) measure the impact of strong duty to bargain and weak duty to bargain laws. The distinction between legal environment and unionization is well established in the literature, see e.g. Moe (2011) and Ichinowski (1988). There are also studies that control for the impact of legal environment and then estimate the union wage impact parsed of this influence.

Finally, studies include different covariates in an attempt to obtain causal estimates of the union wage gap. These controls generally include characteristics of the teachers and the district. Teacher control variables such as education, experience, alternative wage, and gender are common. District controls include variables such as student socioeconomic status (SES), financial status of the district, and median house value.

The studies also use a range of empirical methods to estimate β. The majority of studies are cross-sectional. A few studies employ difference-in-difference or fixed effects specifications. There are substantive differences between estimates generated using within state variation in unionization, e.g. including state fixed effects, and those generated from across state variation. Studies that utilize across-state variation to identify the impact of unionization are comparing unionized districts in states like Massachusetts, i.e. states with strong laws supporting collective bargaining, to non-unionized districts in states like Virginia, where collective bargaining is disallowed. While some of these studies control for the legal environment (e.g. Duplantis et al. (1995), Freeman and Valletta (1988), Gyourko and Tracy (1991), and Han (2011)), the concern with these estimates is that unmeasured differences between states may bias the unionization estimates. It is not clear what direction this will bias estimates. If the omitted variable were solely legal environment, we would expect this to positively bias results.

Utilizing within state variation to identify the impact of unionization also creates challenges for establishing a causal estimate. Most states are either heavily union or non-union. In states where a large share of districts are unionized, within state variation means that the estimates compare the many unionized districts to the few (and possibly selected) non-unionized districts. Within state variation also makes the estimates more susceptible to the impact of threat effects. Winters (2011) utilizes a spatial model to explicitly control for threat effects. If threat effects are a significant driver of teachers wages, estimates utilizing within state variation that do not control for these effects may negatively bias the union impact.

Han (2011) is keenly aware of the tradeoff between these two levels of variation. Instead of utilizing within state variation her preferred empirical strategy utilizes variation within a legal environment. She then utilizes a mixed-effects model and propensity scores matching to contend with the endogeneity of unionization. We classify the majority of estimates from Han (2011) as utilizing across-state variation.[20] We, however, check the robustness of results by removing all Han estimates from the sample when we examine the role of across and within state variation.

Only three studies in our data set attempt an instrumental variables approach, Hoxby (1996), Kasper (1970), and Hirsch et al. (2011). This is not because researchers are unaware of the potential endogeneity of unionization, but rather because finding a credible instrument is difficult.[21] Hirsch et al. propose three sets of plausible instruments: a labor sentiment index for 1919 compiled from regulations and legislation that pertain to labor, an index created from the AFL-CIOs Committee on Political Education voting records for 1965-1975, and the 1964 state union density for the private sector. They ultimately reject the IV estimates because the stability of pro-labor sentiment makes it so these instruments may have a direct and current effect on wages.[22].

Given the degree of variation in empirical specification discussed in this section, it is not surprising that generalizations of the literature are difficult. The following meta-analytic approaches provide a first attempt to understand how these variations systematically contribute to the estimates

[20]There are a few estimates that utilize state fixed effects and we classify those as within state.
[21]For a thorough discussion of this difficulty, see Hirsch et al. (2011).
[22]See Hirsch et al. (2011) p. 8-9

of the union wage impact for teachers.[23]

III Standard Meta-Analysis

We are interested in the overall economic and statistical significance of union wage effects. Since the studies in our sample have many different specifications, we use the standard meta-analytic technique of converting the coefficients reported in the studies to partial correlation coefficients. The partial correlation coefficients for each included specification j is calculated as follows[24]

$$r_j = t_j / \sqrt{(t_j^2 + n_j)} \tag{2}$$

where t_j is the t-statistic for the unionization effect and n_j is the degrees of freedom in the specification. In our calculations, we have used sample size rather than degrees of freedom with the assumption that the size of the sample is large relative to the number of included covariates. We make this substitution because not all studies provide adequate information to calculate degrees of freedom.

Partial correlation coefficients are the simple correlation of the residuals from a regression of $wage$ on all included covariates other than $union$ and the residuals from a regression of $union$ on the same set of covariates. This statistic captures the explanatory power of the variable $union$ on $wage$ in a form that is very similar to the regression coefficients β, but without the issue of comparing different units due to the measurement of $wage$ and $union$. The sign r_j will be the same as the sign of β in the jth study. Further, the correlation of these residuals is not influenced by sample size. Therefore, we prefer partial correlation coefficients to a comparison of t-statistics across studies. Partial correlation coefficients have the desirable properties that they are unit-free and incorporate both magnitude and statistical significance. We note that the issue of selection bias is still a concern when aggregating the results. We will discuss the implications of selection bias in

[23] Jarrell and Stanley (1990) conduct a meta-analysis of the union wage gap generally.
[24] See Greene (2000) for a discussion of this formula.

the succeeding sections.

Before analyzing the overall effect size, we check for publication bias by plotting our effect size estimates against sample size. Publication bias occurs if journals are more likely to select an article for publication when the estimated coefficients are statistically significant. This may also occur if authors do not attempt to publish studies that find small or statistically insignificant effects. If publication bias is present, we would expect to see very few partial correlation coefficients near 0, particularly for studies with smaller sample sizes. Figure 1 reports the results of this analysis. The Figure does not show any evidence of publication bias. Therefore, we proceed to estimating the overall effect size for this sample of studies without concern that this may bias our estimate.

The modes by which unions find it effective and feasible to interact with districts is likely to differ depending on specific characteristics of the district and legal environment. For instance, unions in states that explicitly outlaw collective bargaining may be more likely to work toward improvements in teacher work conditions than increases in teacher pay. As a result of this, we employ a random effects model that does not impose the assumption that there is one true effect of unionization.[25] A random effects model requires only that the effects are drawn from the same underlying normal distribution. To be forthcoming with the evidence we have compiled, we report the results of a fixed effects model in the appendix, Figure A.1. The fixed effects model yields an overall partial correlation coefficient of 0.02. The fixed effects analysis weights smaller studies more than the random effects analysis. In section 4, we further investigate how unionization effects are shaped by the district and legal environment.

In a random effects model, the overall effect represents the mean of the true effects. The estimates are weighted by the inverse variance to account for within-study error and the between-study variance to account for the sampling from the population of true effect sizes. Between-study

[25] Borenstein et al. (2007) give the following example of when a random effects model is appropriate. We repeat the example at length because it is a direct fit to our use of this method.

> "[A]ssume that we are working with studies that assess the impact of an educational intervention. The magnitude of the impact might vary depending on the other resources available to children, the class size, the age, and other factors, which likely vary from study to study. We might not have assessed these covariates in each study. Indeed, we might not even know what covariates are related to the size of the effect...(p. 11)"

variance is calculated by subtracting the within-study variance from the observed total variance. The overall effect is calculated as follows:

$$\bar{r} = \sum_{j=1}^{J} \frac{1}{v_j} \mathrm{r_j} / \sum_{j=1}^{J} \frac{1}{v_j} \tag{3}$$

where v_j is the within-study variance plus the between-study variance and r_j is the particular correlation coefficient.[26]

Figure 2 reports the overall effect size and each study's contribution to it when we treat each partial correlation coefficient as independent, i.e. when we do not apply RVE procedures. The diamond represents the meta-analyzed measure of effect size for each study. It is centered around the effect size estimate and its length represents the 95% confidence interval. The vertical line is the line of no effect, i.e. an effect size equal to 0. The dashed line represents the overall partial correlation coefficient from the group of studies. Examining the location of study estimates to the dashed line is a visual representation of heterogeneity. The overall partial correlation coefficient is 0.03. The impact of unionization on wages is positive and significantly different from 0. The magnitude of the impact, however, is very small.

Since partial correlation coefficients do not convey economic significance, we use the overall correlation coefficient along with sample sizes and standard errors from papers in our sample to compute the resulting percentage change in teachers' wages. Figure 3 contains these values. Using all log-level specifications, we find that a 0.03 partial correlation coefficient on average generates a 4.81% increase in teacher's wages. The distribution is right skewed and two outliers in the upper tail (with a value of 26.21% and 35.68% for an estimate from Tracy (1988) and Baugh and Stone (1982), respectively) contribute to this mean impact being a poor representation of the typical finding. The median wage impact is 3.27% and 90% of values are less than 13%. This suggests a small wage impact of teacher's unions on wages. Given that the literature contains a welter of estimates about the size of these union impacts (ranging from no effect to nearly 20%), this meta-analytic result is of particular interest.This result also differs from the standard union wage gap of

[26]For further discussion of the random effects model, see Borenstein et al. (2007).

10 to 20% that is estimated for the private sector unions.[27]

In Figure 2 we note that the I^2 is very high (84.2%) and Han (2012) accounts for approximately 35% of our overall effect size. The I^2 statistic is the percentage of variation across estimates that is due to heterogeneity rather than to chance.[28]. We adjust the weighting of each study to account for potential clustering at the study level. On average 8 individual estimates are drawn from a study.

Figure 4 reports the result of this re-weighting. We account for any potential within study dependence by adjusting the standard errors. This is achieved by multiplying the standard errors by $\sqrt{1+l(b-1)}$ where b is the number of estimates from a particular study and l is the intraclass correlation within studies.[29] The average within group correlation is $l = 0.72695$.[30] This specification does not alter the overall effect size. The I^2 is much smaller at 54.0% and other studies are weighted more evenly with the Han study. We prefer this specification and report subsequent results with these adjusted errors. We also report a specification in the Appendix in Figure A.2 that sets intraclass correlation to 1, i.e. reflecting the extreme case where estimates from the same study are perfectly correlated. This does not change the results. As expected, the I^2 is slighlty smaller at 46%.

We check the sensitivity of the overall effect size to the inclusion of particular studies and estimates in several ways. First, we use a delete one and a trimming procedure. We then check our handling of the dependence of estimates within a study by drawing only one estimate from each study. The delete one procedure entails deleting the effect sizes from one study at a time and then recalculating the overall effect size. Figure A.3 in the Appendix contains the results of the delete one procedure that produces the largest change in effect size. Removing the Han study generates the largest change in overall effect size. Without the Han estimates, the overall effect

[27]See e.g. Jarrell and Stanley (1990). This difference may be the result of increased focus on fringe or non-pecuniary benefits by teachers unions.

[28]$I^2 = 100\% * \frac{Q-df}{Q}$. See Higgins et al. (2003) for a more detailed description.

[29]This produces very similar weights to those proposed in Hedges et al. (2010) to deal with correlated effects. The only difference is that the Hedges et al. weights assume that sample sizes are more similar within studies than across studies and use the same weight for all effect sizes coming from a particular study. We allow weights to vary within a study because this is not a reasonable assumption for our set of studies. See Han(2012), Lipsky and Drotning (1973) or Tracy(1988) for examples where sample sizes vary substantially within a study.

[30]This value is found by using STATAs loneway command.

size is 0.04. This effect size is statistically different from the overall effect size reported in Figure 3. This produces a mean and median percentage change of 6.414 and 4.366, respectively. A few of the studies we have included are unpublished manuscripts. We made the decision to include these papers to mitigate concerns about publication bias. Some of these manuscripts, however, receive a large weight in our analyses and the peer-review process should mitigate the bias of published estimates. As a result, we estimate the overall effect size for published studies. The overall effect size for published studies is 0.04 and therefore is statistically different from the result obtained by including all studies. The effect size is the same as the result obtained by excluding the Han study, largely because the Han study is unpublished and received significant weight in the original analysis.

We also check the sensitivity of our effect size to outliers by trimming the top and bottom 5%. Figure A.4 reports these results. The overall effect size is the same as the result reported in Figure 3. Figure A.5 reports the results of selecting only one estimate from each study, which can be viewed as a more conservative approach to handling potential dependence between estimates. The overall effect size is 0.05, but is not statistically different from the effect size reported in Figure 3.

The I^2 we find in our preferred specification, 54%, is still very large. Higgins et al. (2003) find that in a review of 509 meta-analyses about a quarter of meta-analyses have I^2 over 50%. These meta-analyses are predominately of medical studies. We are not aware of a similar accounting of I^2 values in economics or social science reviews, although it is likely that these would tend to have higher I^2 values. We view this large value as suggesting two possible sources of variation: 1)that the impacts of unionization will differ based on its form and the context in which it is applied, what we subsequently term 'true heterogeneity' and 2)that the empirical specification of different papers may deal with selection bias to different extents. We work on understanding this heterogeneity in the subsequent sub-group analysis.

It is also important to consider the role of selection bias when interpreting the overall effect sizes. The potential for endogenous assignment to unionization has been noted by other researchers and was discussed in the previous section. There is considerable variation in the amount of attention paid to mitigating bias in β across studies.

As a first look at understanding the magnitude and sign of the bias that may be present, we repeat the overall effect size analysis for only those studies that can be classified as using "quasi-experimental" empirical methods. We classify studies as having the best empirical methods if they have a quasi-experimental research design. There are only four papers that we identify as utilizing the best empirical methods. These are Han (2011), Hoxby (1996), Lovenheim (2009), and Winters (2011). We do not include estimates from Hirsch et al. because the author's question the large magnitude of their results in the instrumental variables specification. We present a random effects model including all papers with a best empirical method designation in Figure 5. Using only the best empirical methods estimates yields a mean partial correlation coefficient of 0.02.[31] The effect size is not statistically different from the results of the meta-analysis including all studies. Removing Han (2012) and repeating the analysis yields effect sizes more similar to the overall effect size reported above. The average effect size for best empirical methods without Han is 0.031. In the following sub-group analysis, we take a more detailed look at empirical specification to try to evaluate best practices.

The studies reviewed in this analysis, therefore, suggest that the economic significance of the teacher union impact on wages is modest, typically between 2-4.5%. We are tentative about concluding that these results are causal effects and discuss our concerns in the succeeding sections. The results are robust to a variety of specifications and tell us about the collective wisdom of research to date. Knowing that the wage impacts are modest should help to inform districts' and states' view of unions.

[31]This corresponds to an average increase of 3.21% and a median increase of 2.18%.

IV Sub-group analysis: Explaining the variation in union wage estimates

Since previous literature reviews, as well as our own evaluation of the literature, suggest that there is substantial variation in the union wage impact across studies, we try to explain this variation by identifying appropriate moderator variables. This analysis examines two types of variation that occur in the estimates: 1) the impact of selection bias and 2) the true heterogeneity of union effects. We first estimate mean partial correlation coefficients for subsamples defined by a few important moderator variables and compare across subsamples. To gain an understanding of the magnitude of each moderator's effect on the union wage impact and their interaction with one another, we then use meta-regression techniques.

Table 2 contains the subgroup analysis for a limited set of these characteristics. We focus on six subgroup analyses: whether the empirical method uses within or across state variation, the level of observations, the date a study was published, the measure of unionization, the decade of data, and the wage measure. The first three subgroups focus on the role of selection bias in the estimates. The last two are directed toward understanding the role of true heterogeneity. The measure of unionization subgroup likely involves both forms of variation. These analyses, of course, are suggestive and do not represent causal effects of these characteristics on effect size. We, therefore, favor the meta-regression analysis and present the impact of these and other moderators through those results.

Many of the moderators we use catalogue whether an empirical specification includes particular, relevant controls. We collect information on whether a specification includes controls for characteristics of the teacher and the district that may reduce selection bias, as well as information about the sample used for estimation. The moderators related to teacher characteristics identify whether a specification includes a control for experience, education, alternative wages that teachers may consider, and gender. Moderators that characterize the district environment are whether a specification includes controls for financial status, student-teacher ratios, the socioeconomic status of students, and median house values. The legal environment controls are 1) whether a study controls for variation in legal environment across states, 2) does not control for legal environment, or 3) uses the variation in legal environment as the measure of unionization. Other moderators include whether the specification utilizes within or across state variation in unionization, the unit of observation, the type of wage measured, and the quality of the journal in which the result was published.

Measures of teacher experience, education, gender, and alternative wages are all likely to be correlated with the unionization status of the district. Historical accounts of teacher unionization, such as Murphy (1992), provide anecdotal evidence that more educated and experienced teachers are more likely to be unionized. Murphy also discusses the importance of teacher gender in the formation of unions. Female teachers were less likely to unionize than their male counterparts. Further, male and female teachers initially did not have similar goals and this may have prevented early attempts to unionize. Research on the causes of unionization is not well developed. There are studies, such as Freeman (1986) and Ichinowski(1988), that show that changes in the legal rights of public sector unions increased unionization rather than appearing in response to its emergence.

We expect that not controlling for a teacher's level of education will create upward biased estimates of the union wage effect. More educated teachers are more likely to join unions and also more likely to earn a higher wage. The subsequent meta-regression provides evidence of such bias. We also expect similar upward bias in studies that do not control for teacher experience or the alternative wage.

The next set of moderators is determined by whether a specification includes particular controls for district characteristics. We collect information on whether the researcher has controlled for district financial status, legal environment, students' socioeconomic status, and median house values. We classify a specification as controlling for financial status if a measure such as district revenues, district per capita income, or debt service per pupil is included. A study is categorized as controlling for legal environment if the researcher either uses a sample with a homogeneous legal environment or includes dummy variables for differing legal environments (including state fixed effects). Studies often consider whether a state allows or prohibits collective bargaining for teachers, as well as the legal status of agency shops. Moe (2011) provides a useful classification of legal environments. He partitions states into 4 mutually exclusive categories: 1) states that have collective bargaining laws and allow agency shops; 2) states that have collective bargaining laws, but do not allow agency shops; 3) states that do not have collective bargaining laws; and 4) states that prohibit collective bargaining. Our original plan was to group studies according to this classification. This, however, is not possible since many studies utilize national data sets. We instead split the papers by whether the studies use within or across state variation.

Further research that examines the union wage gap in moderate and weak legal environment may help to sort out the heterogeneity of union wage impacts.[32] The meta-regression will allow us to parse the impact of this moderator from measuring unionization with a CBA and the quality of the empirical strategy.

The dependent variable teachers wages is specified as either the average wage of teachers, the wage for new teachers, or the wage for more experienced teachers. A well-established result from the literature on teachers unions is that unions increase the wages of senior teachers, but not the wages of new teachers. The subgroup analysis confirms that unions have a smaller effect on new teachers than senior teachers with partial correlation coefficients of 0.024 and 0.077, respectively. The subgroup analysis, however, additionally suggests that unions do have a positive and statistically significant impact on new teachers' wages. The effect size for new teachers' wages is not statistically different from the effect size when average wages are the measure.

We also classified the estimates by their measure of unionization. We divide the measures of unionization into the following mutually exclusive categories: 1) unionization measured by presence of CBA[33], 2) unionization measured as the number or percent of teachers who are members of the union or represented by the union, but have not necessarily established a CBA, and 3) unionization measured by variations in the legal environment, such as whether a district has the duty to bargain or meet-and-confer. The specifications that measure unionization by CBA have a higher

[32] This is the intention of Han(2012).

[33] We also included estimates where the measure is the coverage for a contract in this category.

overall effect size (r=0.058) than either the membership (r=0.025) or legal subgroup (r=0.014). The difference between the CBA and membership effect sizes is statistically significant. The difference between the membership and legal effect sizes is, however, not statistically significant. These effect sizes yield an average wage impact for unionization measured by CBA of 9.59% with a typical wage effect of 6.33%. In contrast measuring unionization by membership yields an average wage impact for unionization of 4.01% and a median impact of 2.73%.

Two potential explanations for the difference between these subgroup effect sizes are that these measures of unionization are capturing different forms of teachers unions or that there is attenuation bias due to greater measurement error in the membership and legal measures. The latter possibility was discussed in Hoxby (1996). The first explanation is plausible given that not all teachers unions have official CBAs with their districts. An open question is whether unionized districts with CBAs are more likely to increase teacher's wages than unionized districts that do not have them.

When the sample is stratified by the unit of observation, we find that the average effect size is smaller when the unit of observation is the individual teacher (r=0.022) than when it is the district (r=0.055). This result may reflect that teacher characteristics that increase wages are positively correlated with unionization. Teacher-level samples allow researchers to better control for these characteristics and provide a more accurate estimate of the role of teachers unions. We are able to test directly for the importance of these controls and any remaining effect of a teacher-level sample in the meta-regressions.

When we study the role of across-state versus within-state variation in union status, we find that utilizing our full set of estimates in a random effects model yields average effect sizes for the within-state group that are twice the size of the estimates for the across-state group, i.e. r=0.054 and r=0.027, respectively. These effect sizes equate to a mean percent increase in wages of 8.66% and 4.33%. We believe this result is important and check its robustness in the following three ways: 1) drop studies that address concerns generally raised about across-state variation, 2) drop unpublished studies, and 3) select one estimate from each study.

First, we drop the Han and Winters estimates because these empirical specifications address concerns about unobservables better than other studies relying on across-state variation. [34] Without these studies, we find that the average effect size for the within-state variation group is still double that of the across-state variation studies.[35] The difference, however, is not statistically significant. The I^2 for the across-state variation group is 54.3% with and 75.5% without the inclusion of their studies. [36]

Removing the unpublished studies from the analysis, confirms the results that the effect size for the within-state group is twice as large as the across-state group. Studies that utilize within-state variation would find on average that unions increase teacher's wages by 6.86%. In contrast, studies that utilize across-state variation would find a smaller increase of 3.45%. The difference is statistically significant in a one-tailed test at the 5%-level. The unpublished studies removed from the analysis are Han (2012), Hirsch et al. (2011), and Tracy (1988). Finally, the results from a select one method show that within-state variation produces effect sizes that are more than twice as large as the those produced by across-state variation, i.e. the average percentage increase in wages is 9.78% and 3.68% for within-state and across-state estimates, respectively. The difference is marginally statistically significant at the 7%-level. The select one method yields larger I^2 for each subgroup.

We also investigate the possibility that union wage effects change over time. We group studies according to the decade of data used. We exclude three studies from the initial analysis because the data utilized spans decades. The excluded studies are Kleiner and Petree (1988), Hoxby (1996), and Lovenheim (2009). The studies surveyed use data from years ranging from 1967 to 2008. The studies utilizing data from 1967-1969 are Lipsky and Drotning (1973) and Kasper (1970). We in-

[34] Winters shows that spatial correlation in errors is likely and addresses it with an inverse-distance weighting matrix. This empirical strategy addresses the concern over differences in state laws without using state fixed effects. Han groups districts into 4 mutually exclusive categories that explain the public sector labor laws. This allows her to control for the legal environment without utilizing either state fixed effects or confining analysis to particular states.

[35] The corresponding mean percentage wage increase in wages are 10.18% (median=7.05%) and 5.62% (median=3.89%).

[36] Han (2012) accounts for approximately 21% of estimates in the across-state group. Winters accounts for another 13% of the estimates. The heavy weights on these two studies are because their empirical strategies allow them to exploit larger data sets.

clude these studies and all other estimates coming from 1970s data in the 1970s category[37]. There are a few studies with data that spans the decade cutoff. Lentz (1998) uses data from 1989-1990. Winters (2011) and Hirsch et al. (2011) both use data from 1999-2000. We classify these studies according to the end date. The results suggest strong impacts of unionization in the early period of union activity, the 1970s, with an average effect size of 0.12. This would generate an average union impact of 19.24% and median impact of approximately 13%. The union impacts fall off in the 1980s and remain at those levels through the present decade. The average effect sizes for the 1980s, 1990s, and 2000s are 0.023, 0.035, and 0.033. These effect sizes are not statistically different and mimic the results of the reported overall effect sizes. The union wage impacts for this period range from 2.5-4%.[38]

To contrast true heterogeneity in union wage effects with possible selection bias, we analyze the role of publication date on partial correlation coefficients. Our thinking is that there may be trends in the specification of models and methods of analysis, as well as potential evolution toward better empirical tools. Many researchers implicitly consider newer studies to be more credible than older studies. We try to provide some empirical evidence to comment on this belief through a subgroup analysis and subsequent metaregression.

The classification of studies by publication date is generally straightforward. We choose to continue including unpublished studies in this analysis and then check the robustness of the analysis by excluding these studies. For unpublished studies, we assign the year of the most recent draft as the date of publication. We also choose to include the Han and Winters studies in the 2000s subgroup. The last draft of Han's study was written in 2012 and the Winters' paper was published in 2011.

There is a distinct break in the magnitude of the decade subgroup effect size between the 1980s and 1990s. The 1970s and 1980s have effect sizes of 0.069 and 0.066, respectively. These effect sizes correspond to a mean wage increase of approximately 11% and a median impact of approximately 7.5%. The 1990s and 2000s have effect sizes that are of much smaller magnitude

[37]The results are robust to excluding estimates from the studies using data from late 1960s.

[38]The median wage impacts for 1980s, 1990s, and 2000s are 2.51%, 3.82%, and 3.60%, respectively.

than those estimated during the early period, i.e. the effect sizes are 0.038 and 0.024, respectively. The 1990s effect size corresponds to a median wage impact of 4.15%. Similarly, the median wage impact for studies published in the 2000s is 2.62%. The effect size for studies published in the 1990s is not statistically different from either the 1970s or 1980s effect sizes. The 2000s effect size, however, is statistically distinct from both.

Since the Han (2012) study accounts for many estimates for the last decade , we also exclude the Han study and reexamine the results. Without Han, the effect size for studies published during the 2000s is 0.031 and the corresponding median wage impact is 3.38%. This result is marginally statistically distinct from the 1980s and 1990s. We also check this result by excluding unpublished studies from the sample. The results are not statistically different from those reported above.[39]

In many of these subgroup analyses, we still find a high I^2. This suggests that there is substantial factual or methodological heterogeneity remaining. The high I^2 statistic for CBA may be an indicator of the importance of particular provisions in a CBA contract.[40] We interpret these results as suggestive of the idea that factual heterogeneity exists. It is very likely that causal union effects exhibit a large degree of heterogeneity and that studies capture different types of unions due to data limitations and empirical specification. Additionally, variation in study quality may be an important determinant of the effect size.

IV.1 Metaregression: Understanding the interaction of study moderators

To parse the interaction of the moderator variables, we implement meta-regression techniques. The general equation we estimate is

$$effectsize = \beta_0 + district\beta_1 + teacher\beta_2 + sample\beta_3 + u \qquad (4)$$

[39] It is important to note that more recent studies receive larger weight in the overall meta-analysis conducted in the previous section because improvements in data availability increased sample sizes.

[40] There is a developing body of research examining the role of individual provisions, see e.g. Strunk (2011), Cowen and Fowles (2013), and Strunk and Grissom (2010).

where $effectsize$ is the partial correlation coefficient, $district$ is a set of moderators that identify the presence of district-level controls, $teacher$ is a set of moderators that identify the presence of controls related to teacher characteristics, and $sample$ is a set of moderators that characterizes the sample used by the researcher. We begin with a parsimonious specification of (4) that includes key moderators discussed in the previous section. For $district$, we include whether a specification includes controls for median value of houses and the SES of students. For $teacher$, we include whether a specification has a control for teacher education, teacher experience, teacher gender, and the alternative wage. We also include dummy variables to capture possible variation in union effects by decade.

In Table 3, we report the results of this analysis. The first column is a baseline specification that estimates the coefficients by ordinary least squares (OLS) and uses Huber-White errors to account for heteroskedasticity. Column 2 presents the results from a similar specification with the addition of a control for journal quality. Since we know the effect size estimates contain significant heteroskedasticity, we estimate similar specifications by weighted least squares (WLS) using the robust variation estimation weights. We prefer the WLS specification because it places greater weight on effect sizes that are estimated more precisely.[41] These results are reported in columns 3 and 4.

In the meta-analysis literature, there is an ongoing discussion about whether the quality of empirical specifications and methods greatly influences the results. Glass (1976) noted that "It is an empirical question whether relatively poorly designed studies give results significantly at variance with those of the best designed studies..." We directly test the role of empirical specification by checking for the presence of common controls, as well as measures of quality. In our WLS specifications, we find that the inclusion of teacher level controls significantly affects the estimated wage impact. The OLS specification does not confirm this finding.[42] Using only the OLS specification would lead one to believe that the model specification with regards to teacher

[41]Tests of model fit suggests that a linear model would appropriately fit the data. The results, however, of the OLS and WLS specifications are significantly different.

[42]Joint hypothesis tests of the significance of teacher education, teacher experience, teacher gender, and the alternative wage do not allow us to reject their joint insignificance.

controls is irrelevant. We strongly prefer the WLS specification and therefore focus on these results in subsequent discussion.

In addition to teacher, district, and sample characteristics, we also examine another dimension of quality by utilizing a measure of journal quality. We use the Source Normalized Impact Per Paper (SNIP) to rank the quality of the journals in which each study is published.[43] This ranking is, of course, unavailable for National Bureau of Economic Research (NBER) working papers, NBER book chapters, and other unpublished manuscripts. For NBER working papers and NBER book chapters, we infer that the ”journal” ranking is above average. The analysis controlling for journal quality excludes any other unpublished manuscripts.[44]

Controlling for the journal quality significantly decreases the estimated union wage impact in the WLS specification. A paper being published in a journal with an above average SNIP ranking is likely to find smaller effect sizes. The coefficient on the SNIP ranking being above average is -0.061. This corresponds to a median decrease of 6.66% in the union impact. This provides evidence that papers published in better journals find smaller wage impacts. It is not clear whether the smaller estimated impact is the result of paper quality or priors of early readers of the paper. Another notable effect of including this control is that the coefficient on the dummies for studies written in the 1980s and 1990s is of larger magnitude and significant. Testing the equality of the coefficients on the decade dummies shows that on average the estimates obtained in 1980s or 1990s are statistically different from those obtained in the 2000s. This specification suggests that union impacts were largest in the early period of teachers unionism, teachers began to unionize in significant numbers between the mid 1960s and mid 1970s, and then fell off substantially in the 1980s and continued to decline in 2000s. The difference between union effects in the 1980s, 1990s, and 2000s is not statistically different in column 4. The result that data from the 2000s generates significantly smaller union wage impacts than 1970s, however, is robust across all specifications.

[43] We also replicated the analysis using the Impact Per Publication (IPP) and SCImago Journal Rank (SJR). All three measures provide the same ranking of the journals. Both the SNIP and SJR measures are comparable across fields (www.journalmetrics.com).

[44] Han (2012), Hirsch et al. (2011) and Lentz (1998) are excluded in specifications containing this control. Lentz is published in a journal, but a SNIP ranking is not available.

The specifications in columns 1-3 do not find statistically distinct differences between the 1970s and the 1980s or 1990s. We view this result as showing that factual heterogeneity in teacher union wage impacts exists and that union wage impacts have decreased over time. [45]

In column 4, the sign of the coefficients on teacher control moderators are as expected. Controlling for teacher experience decreases the effect size by 0.048. There are two possible explanations for this finding. This may suggest that more experienced teachers are more likely to unionize or that more experienced and educated teachers are attracted to unionized districts due to higher salaries and other potential benefits. The sign on this moderator is consistent across specifications. Controlling for a teacher's education also significantly decreases effect size in the specifications where we have parsed the impact of study quality. In column 4, controlling for a teacher's level of education decreases the effect size by 0.080. This means that when a study attempts to mitigate bias by conditioning only on observables, including controls for teacher's education is particularly important. Studies that control for a teacher's gender also generate decreased effect sizes. Controlling for a teacher's gender decreases the effect size by 0.037. Given that female teachers on average earn less than male teachers (due to differences in credentialing and the labor markets they participate in), this may reflect that districts with more female teachers are less likely to unionize.

All four specifications confirm the commonly held belief that teachers unions increase the wages of senior teachers more than new teachers.[46] Changing the weighting of effect sizes makes the biggest difference in the magnitude of the effect. Effect sizes that are estimated more precisely (or that are the sole contribution of a study) show smaller impacts on senior teachers' wages. In column 4, the impact of measuring senior teachers wages is a 0.069 increase in effect size. A particularly notable result is that specifications which measure the impact on the average wage find smaller effect sizes than those that measure new teachers' wages. Columns 2 and 4, which control for the quality of the empirical specification, show statistically significant decreases in the effect

[45]We investigated the possibility of spatial dependence of estimates by date of publication. Our thinking was that there may be trends in the specification of empirical models unmeasured by our moderators that could account for the differences we see over time. We did not find any evidence of significant spatial correlation for papers and neighbors published in the surrounding 3 years.

[46]The excluded category in the regressions is that the dependent variable measures new teachers' wages.

size.

We did not expect to find an greater union-wage increase for new teachers relative to average teachers. After some reflection and review of relevant literature, we suggest a few interpretations. First, we split our thinking about this relationship into three concepts: the level effect (union-non union wage gap), dispersion effect, and returns to experience. Our result suggests that teachers unions either bargain for higher wages for new teachers or increase the qualifications of new teachers such that the average new teacher is paid better. The positive coefficient on new teachers is consistent with a positive union wage gap for new teachers coupled with lower returns to experience for teachers with experience less than the average teacher. It seems likely that the result reflects that diminishing returns to experience are more pronounced in non-unionized districts.

Measuring unionization as the presence of a collective bargaining agreement also appears to increase the effect size. This result is present in both WLS specifications and of larger magnitude in the specification that controls for quality. Measuring unionization as presence of CBA yields a 0.034 increase in the effect size when compared to proxying for unionization through legal status. In column 4, we see that the impact on effect size is not statistically different from measuring unionization as membership or coverage once we have included the quality index. Column 1 and 2 present a similar relationship between these two measures of unionization.

Finally, all four specifications show that controlling for student SES decreases the effect size in all specifications. The WLS specifications show that this control decreases effect sizes by 0.058 and 0.100 in columns 3 and 4, respectively. This result may be picking up the difference between average teachers in districts with high and low SES students. There is evidence that better teachers migrate to teach high SES students.[47] The difference in effect size is larger in the specification that controls for journal quality. Although the difference is statistically significant, this may be the result of few papers in good journals not controlling for the student SES.

[47] See e.g. Goldhaber, Choi and Cramer (2007), Clotfelter, Ladd and Vigdor (2005), and Lankford, et al. (2002).

Conclusion

The literature examining the impact of teachers unions on education is very large and diverse. Meta-analytic techniques allow us to better understand both the overall effect of unions on wages and the reasons behind differences in estimates from these studies. Our results suggest that it is important to take the context of union studies into account when examining their overall impact, and that the effects of unions on wages are shaped by both the district and legal environment being studied.

A key finding of this study is that the average wage impact estimated by the included papers is modest, around 2-4.5%. Our findings also suggest that the quality of an empirical strategy significantly affects the size of the estimated impact. Controlling for teacher experience, education, and gender all reduce the estimated wage impact. We also find that teachers union wage impacts have varied over time. The largest impacts appear to be following the rapid expansion of teacher unionism in the 1970s. Finally, we gain new insight into the goals of teachers unions by using the increased statistical power of meta-analytic techniques to show that unions increase the wages of new teachers and not just senior teachers.

REFERENCES

Allen, Steven G. (1986). “Unionization and Productivity in Office Building and School Construction.” Industrial and Labor Relations Review, 39: 187–201.

Babcock, L. C., and J. B. Engberg. A Dynamic Model of Public Sector Employer Response to Unionization. Journal of Labor Research 18, no. 2 (SPR 1997): 26586. doi:10.1007/s12122-997-1039-y.

Baugh, W.H., and J.A. Stone. “Teachers, Unions, and Wages in the 1970s - Unionism Now Pays.” Industrial and Labor Relations Review 35, no. 3 (1982): 368-76. doi:10.2307/2522816.

Borenstein, Michael, Larry Hedges, and Hannah Rothstein. 2007. “Meta-Analysis: Fixed vs. random effects.” Mimeo. accessed online at https://www.meta-analysis.com/downloads/Meta-analysis%20fixed%20effect%20vs%20random%20effects.pdf.

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmaneul Saez, Diane Whitmore Schanzenbach, and Danny Yagan. “How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star”, Quarterly Journal of Economics 126, no. 4 (2011): 1593–1660.

Clotfelter, Charles T. Helen F. Ladd and Jacob Vigdor (2005). “Who Teaches Whom? Race and the Distribution of Novice Teachers,” Economics of Education Review 24(4):377–392.

Cowen, Joshua M. “Teacher Unions and Teacher Compensation: New Evidence for the Impact of Bargaining.” Journal of Education Finance 35, no. 2 (2009): 172-93. doi:10.1353/jef.0.0010.

Cowen, Joshua M., and Jacob Fowles. 2013. Same Contract, Different Day? An Analysis of Teacher Bargaining Agreements in Louisville Since 1979. Teachers College Record 115 (5): 1-30.

Djankov, Simeon, and Peter Murrell. 2002. “Enterprise restructuring in transition: A quantitative survey.” Journal of Economic Literature 40 (September 2002): 739-792.

Dolton, P., and M. Robson. “Trade Union Concentration and the Determination of Wages: The Case of Teachers in England and Wales.” British Journal of Industrial Relations 34, no. 4 (December 1996): 539-55.

Duplantis, M., T. Chandler, and T. Geske. “The Growth and Impact of Teachers Unions in States with Collective-Bargaining Legislation.” Economics of Education Review 14, no. 2 (June 1995): 167-78. doi:10.1016/0272-7757(95)90396-P.

Eberts, Randall W., and Joe A. Stone. ”‘Communications Wages, Fringe Benefits, and Working Conditions: An Analysis of Compensating Differentials.”’ Southern Economic Journal 52, no. 1 (July 1985): 274-280.

Ehrenberg, Ronald and Richard Chaykowski. 1988. “On Estimating the Effects of Increased Aid on Education,” in When Public Sector Workers Unionize, ed. by Richard B. Freeman, and Casey Ichniowski. The University of Chicago Press, Chicago.

Ehrenberg, Ronald and Joshua Schwarz. “Public Sector Labor Markets.” Handbook of Labor Economics, edited by Orley Ashenfelter and Richard Layard. North-Holland, 1986.

Freeman, Richard B. “Unionism Comes to the Public Sector.” Journal of Economic Literature 24, no. 1 (1986): 41-86.

Freeman, Richard B., and Robert G. Valletta. 1988. "The Effects of Public Sector Labor Laws on Labor Market Institutions and Outcomes," in When Public Sector Workers Unionize, ed. by Richard B. Freeman, and Casey Ichniowski. The University of Chicago Press, Chicago.

lass, G.V. "Primary, Secondary, and Meta-Analysis of of Research." Educational Researcher 5, no. 1 (1976): 3-8.

oldhaber, D., DeArmond, M., Player, D., and Choi, H. J. "Why do so few public school districts use merit pay?" Journal of Education Finance, 33, no. 3 (2008): 262289.

Goldhaber, Dan, Hyung-Jai Choi, and Lauren Cramer (2007). "A Descriptive Analysis of the Distribution of NBPTS-Certified Teachers in North Carolina," Economics of Education Review 26(2):160–172.

Greene, William. 2000. Econometric Analysis. Upper Saddle River, NJ: Prentice Hall.

regory, R and S.R. Borland. 1999. "Recent Developments in Public Sector Labor Markets," in Handbook of Labor Economics, edited by Orley Ashenfelter and David Card. Elsevier, North Holland.

Gyourko, Joseph, and Joseph Tracy. 1991. "Public Sector Bargaining and the Local Budgetary Process." In Ronald G. Ehrenberg (ed.), Research in Labor Economics, vol. 12, pp. 117–36. JAI Press Inc.

Han, Eunice. 2015, October. "The Myth of Unions' Overprotection of Bad Teachers: Evidence from the District Teacher Matched Panel Data on Teacher Turnover." Job Market Paper.

Han, Eunice. 2012, September. "The Impact of Teachers Unions on Teachers Well-being under Different Legal Environments: Evidence from Districts and Teachers Matched Data." Working paper.

Hanushek, Eric A., and Steven G. Rivkin. 1997. "Understanding the Twentieth-Century Growth in U.S. School Spending." Journal Of Human Resources 32, no. 1: 35–68. EconLit, EBSCOhost (accessed July 23, 2014).

Hedges, L.V., Elizabeth Tipton, and M.C. Johnson. 2010. "Robust variance estimation in meta-regression with dependent effect size estimates." Research Synthesis Methods. 1(1): 3965. Erratum in 1(2): 164–165.

Higgins, Julian P.T., Simon G. Thompson, Jonathan J. Deeks, and Douglas G. Altman. "Measuring Inconsistency in Meta-Analyses." BMJ, September 6, 2003, 327, no. 7414 (2003): 557-60.

Hirsch, Barry T., David A. Macpherson, and John V. Winters. 2011. "Teachers Salaries, State Collective Bargaining Laws, and Union Coverage." Working paper.

Hoxby, C. M. "How Teachers Unions Affect Education Production." Quarterly Journal of Economics 111, no. 3 (August 1996): 671-718. doi:10.2307/2946669.

Ichniowski, Casey. 1988. "Public Sector Growth and Bargaining Laws: A Proportional Hazards Approach with Time-Varying Treatments," in When Public Sector Workers Unionize, ed. by Richard B. Freeman, and Casey Ichniowski. The University of Chicago Press, Chicago.

Jarrell, Stephen B., and T. D. Stanley. "A Meta-Analysis of the Union-Nonunion Wage Gap." Industrial & Labor Relations Review 44, no. 1 (October 1990): 54-67.

Kasper, H. "The Effects of Collective Bargaining on Public School Teachers' Salaries." Industrial and Labor Relations Review 24, no. 1 (October 1970): 57–72.

Kish, Leslie. 1965. Survey Sampling. New York, NY: John Wiley and Sons, p. 162.

Kleiner, Morris and D. L. Petree. 1988. "Unionism and Licensing of Public School Teachers: Impact on Wages and Educational Output," in When Public Sector Workers Unionize, ed. by Richard B. Freeman, and Casey Ichniowski. The University of Chicago Press, Chicago.

Lankford, Hamilton, Loeb, Susanna, and Wyckoff, James (2002)."Teacher Sorting and the Plight of Urban Schools. A Descriptive Analysis," Educational Evaluation and Policy Analysis, 24(1), 37-62.

Lemieux, Thomas. The Mincer Equation Thirty Years After Schooling, Experience, and Earnings. In Jacob Mincer A Pioneer of Modern Labor Economics, edited by Shoshana Grossbard, 127–145. Springer US, 2006.

Lentz, C. "The Effects of Collective Bargaining on Teacher Compensation: Lessons from Illinois." Journal of Collective Negotiations in the Public Sector 27, no. 2 (1998): 93-106.

Lewis, H. Gregg "Chapter 20 Union Relative Wage Effects." In Handbook of Labor Economics, edited by Orley C. Ashenfelter and Richard Layard, 2:1139-81. Elsevier, 1986. http://www.sciencedirect.com/science/article/pii/S1573446386020102.

Lipsky, David B. and John E. Drotning. "The Influence of Collective Bargaining on Teachers' Salaries in New York State." Industrial and Labor Relations Review 27, no. 1 (1973): 18–35.

Lovenheim, Michael F. "The Effect of Teachers Unions on Education Production: Evidence from Union Election Certifications in Three Midwestern States." Journal of Labor Economics 27, no. 4 (October 2009): 525-87.

Milkman, M. 1997. "Teachers Unions, Productivity, and Minority Student Achievement." Journal of Labor Research, 30(1), 137–150.

Moe, Terry M. 2011. Special Interest: Teachers Unions and Americas Public Schools. Washington, D.C: Brookings Institution Press.

Moore, William and John Raisian. "Union-Nonunion Wage Differentials in the Public Administration, Educational and Private Sectors:1970-1983." The Review of Economics and Statistics 69, no. 4 (November 1987): 608-16.

Murphy, Marjorie. Blackboard Unions: The AFT and the NEA, 1900–1980. Ithaca, N.Y.: Cornell University Press, 1992.

Nelson, Jon P., and Peter E. Kennedy. The Use (and Abuse) of Meta-Analysis in Environmental and Natural Resource Economics: An Assessment. Environmental and Resource Economics 42, no. 3 (March 1, 2009): 345–77. doi:10.1007/s10640-008-9253-5.

Pantuosco, L.J., and Ulrich, L.D. 2010. "The impact of teachers unions on state-level productivity". Journal of Education Finance, 35(3), 276–294.

Retsinas, Joan. "Teachers: Bargaining for Control." American Educational Research Journal 19, no. 3 (Autumn, 1982): 353–72.

Stoddard, Christiana. "Adjusting teachers salaries for the cost of living: The effect on salary comparisons and policy conclusions." Economics of Education Review 24, (2005): 323–39.

Strunk, Katherine. "Are Teachers' Unions Really to Blame? Collective Bargaining Agreements and their Relationship with District Resource Allocation and Student Performance in California." Education Finance and Policy 6 (2011): 354-98.

Strunk, Katharine O., and Jason A. Grissom. 2010. Do Strong Unions Shape District Policies? Collective Bargaining, Teacher Contract Restrictiveness, and the Political Power of Teachers Unions. Educational Evaluation and Policy Analysis 32 (3): 389406. doi:10.3102/0162373710376665.

Tanner-Smith, Emily E., and Elizabeth Tipton. 2014, “Robust variance estimation with dependent effect sizes: Practical considerations including software tutorial in Stata and SPSS.” Research Synthesis Methods 5 (1): 13-30.

Tracy, Joseph. 1988. “Comparisons between public and private sector union wage differentials: Does the legal environment matter?” NBER Working Paper No. 2755.

U.S. Department of Education, National Center for Education Statistics, Common Core of Data (CCD), ”National Public Education Financial Survey,” in the Digest of Education Statistics 2013.

U.S. Department of Education, National Center for Education Statistics, 2011-12 School and Staffing Survey (SASS).

West, Kristine Lamm, and Elton Mykerezi. “Teachers Unions and Compensation: The Impact of Collective Bargaining on Salary Schedules and Performance Pay Schemes.” Economics of Education Review 30, no. 1 (February 2011): 99-108. doi:10.1016/j.econedurev.2010.07.007.

Winters, John V. “Teacher Salaries and Teacher Unions: A Spatial Econometric Approach.” Industrial & Labor Relations Review 64, no. 4 (July 2011): 747-64.

Zwerling, H.L., and T. Thomason. “Collective-Bargaining and the Determinants of Teachers Salaries.” Journal of Labor Research 16, no. 4 (Fall 1995): 467-84. doi:10.1007/BF02685721.

TABLE 1:
Summary of Included Studies

Author(Year) and Journal	Summary of Study
Baugh and Stone (1982). Industrial and Labor Relations Review.	**Data Source:** Current Population Survey 1974-75 and 1977-78, national sample of school teachers. **Unit of observation:** Teacher. **Type of Econometric Model(s):** (a) First difference (b) OLS cross-sectional. **Measure(s) of dependent variable:** (a) Log (hourly wage 1974 / hourly wage 1975) (b) Log (hourly wage 1977 / hourly wage 1978) (c) Log (hourly wage). **Measure(s) of unionization:** union member. **Estimates obtained from this study:** 2
Cowen (2009). Journal of Education Finance	**Data Source:** SASS 1999-2000 and Common Core of Data, 2005-2006, districts in 14 states (includes only with more than 10% of districts estimated as bargaining or non-bargaining). **Unit of observation:** District. **Type of Econometric Model(s):** (a) OLS cross-sectional (b) state fixed effects. **Measure(s) of dependent variable:** ln(total expenditures paid as teachers salaries) **Measure(s) of unionization:** collective bargaining. **Estimates obtained from this study:** 1
Duplantis et al. (1995). Economics of Education Review.	**Data Source:** Several, including: Survey of superintendents, Bureau of the Census, and Dept. of Labor, 1992, 88 districts in 11 states. **Unit of observation:** District. **Type of Econometric Model(s):** OLS, cross-sectional. **Measure(s) of dependent variable:** Ln (average teachers salary) **Measure(s) of unionization:** existence of CBA. **Estimates obtained from this study:** 1
Freeman & Valletta (1988). NBER book chapter.	**Data Source:** CPS, 1984, nationally representative sample. **Unit of observation:** Teacher. **Type of Econometric Model(s):** OLS, cross-sectional. **Measure(s) of dependent variable:** ln(hourly wage) **Measure(s) of unionization:** (a) Legal index (b) collective bargaining agreement. **Estimates obtained from this study:** 2
Gyourko & Tracy (1991). Research in Labor Economics.	**Data Source:** Census of Population, 1980, nationally representative sample in 131 cities. **Unit of observation:** Teacher. **Type of Econometric Model(s):** OLS, cross-sectional. **Measure(s) of dependent variable:** ln(weekly wage) **Measure(s) of unionization:** (a) strong duty-to-bargain law (b) weak duty to bargain (c) percent organized. **Estimates obtained from this study:** 3

Continued on next page

TABLE 1 – *Continued from previous page*

Author (Year) Journal	Summary of Study
Han (2012). Job Market Paper.	**Data Source:** SASS and School District Finance Survey 2007-2008, National (includes roughly 1/3 of all public school districts.) **Unit of observation:** Several, including Teacher and District. **Type of Econometric Model(s):** (a) Weighed OLS (b) Clustered, mixed effects. **Measure(s) of dependent variable:** (a) log(base salary) (b) log(max salary) **Measure(s) of unionization:** (a) union member (b) union density (c) collective bargaining (d) meet and confer. **Estimates obtained from this study:** 19
Hirsch et al. (2011). Unpublished draft.	**Data Source:** (a) CPS 2000-2009 (b) SASS 1999-2000. **Unit of observation:** Teacher. **Type of Econometric Model(s):** OLS, cross-sectional. **Measure(s) of dependent variable:** (a) ln(hourly earnings) (b) ln(salary) **Measure(s) of unionization:** (a) collective bargaining agreement (b) CB law index. **Estimates obtained from this study:** 4
Hoxby (1996). Quarterly Journal of Economics.	**Data Source:** District data from Census of Governments (1972,1982,1992), Unionization Measure from Census of Governments, NEA Negotiating Agreement Provisions, and Perry and Wildman, Demographic Data and high school dropouts from Census, NBER Public Sector Collective Bargaining Law Data Set, 1972, 1982, 1992, national 95% of independent school districts in the USA. **Unit of observation:** District. **Type of Econometric Model(s):** (a) Cross-section (b) First difference (c) Diff-in-diff (d) IV diff-in-diff **Measure(s) of dependent variable:** log(average teacher salary/1000 in current dollars) **Measure(s) of unionization:** collective bargaining exists, a contractual agreement exists, and 50% of teachers unionized. **Estimates obtained from this study:** 1
Kasper (1970). Industrial and Labor Relations Review.	**Data Source:** Several including unpublished reports of the NEA/AFT and personal mail survey. State level data, 50 states plus DC, 1966-67, 1967-68. **Unit of observation:** State. **Type of Econometric Model(s):** (a) OLS, cross-sectional (b) 2SLS. **Measure(s) of dependent variable:** (a) average statewide teacher salary (b) arithmetic mean of teacher salaries 1966-67 and 1967-68 (c) ratio of 1967-68 teacher salary to 1967 average police entrance salary **Measure(s) of unionization:** (a) Proportion of teachers represented by an organization (b) Proportion of school districts which had representation (c) Proportion of state teachers covered by formal CBAs (d) Proportion of teachers represented by NEA (d) Proportion of teachers represented by AFT **Estimates obtained from this study:** 2

Continued on next page

TABLE 1 – *Continued from previous page*

Author (Year) Journal	Summary of Study
Kleiner & Petree (1988). Chapter in NBER book.	**Data Source:** State-level sample, data from Census and author collected information on teachers union membership and licensing laws, 1972-82, 50 states. **Unit of observation:** State. **Type of Econometric Model(s):** (a) OLS, cross-sectional (b) Fixed effects. **Measure(s) of dependent variable:** log(average teacher wages) **Measure(s) of unionization:** (a) Percent members (b) Percent covered by contracts **Estimates obtained from this study:** 2
Lentz (1998). Journal of Collective Negotiations in the Public Sector.	**Data Source:** District level data in Illinois from Illinois State Board of Education and School District Data Book, 1989-1990, for (a) Illinois (b) Chicago Metro Area (c) Rural and Suburban Illinois. **Unit of observation:** District. **Type of Econometric Model(s):** OLS, cross-sectional **Measure(s) of dependent variable:** salary plus hospitalization and life insurance for teacher and families **Measure(s) of unionization:** Existence of CBA **Estimates obtained from this study:** 1
Lipsky & Drotning (1973). Industrial and Labor Relations Review.	**Data Source:** Hand-collected data on New York State, 1968-69, New York State, 441 districts with contracts and 255 without. **Unit of observation:** District. **Type of Econometric Model(s):** OLS, cross-sectional **Measure(s) of dependent variable:** (a) salary paid to 1st year teacher with a BA (b) salary paid to a teacher with 7 years experience and BA+30 hours (c) salary paid to a teacher with 11 years experience and BA+60 credit hours (d) mean salary **Measure(s) of unionization:** Existence of CBA **Estimates obtained from this study:** 8
Lovenheim (2009). Journal of Labor Economics.	**Data Source:** Hand-collected teachers' union certification dates for Iowa, Indiana, and Minnesota, Census of Governments, Census, 1972-1991. **Unit of observation:** District. **Type of Econometric Model(s):** (a) Diff-in-diff (b) Fixed effects **Measure(s) of dependent variable:** (a) ln(real average monthly salary) (b) log(average teacher salary/1000 in current dollars) **Measure(s) of unionization:** (a) state has a duty to bargain law (the treatment group is states without duty to bargain) (b) union certification election **Estimates obtained from this study:** 4

Continued on next page

TABLE 1 – *Continued from previous page*

Author (Year) Journal	Summary of Study
Retsinas (1982). American Educational Research Journal.	**Data Source:** 37 school districts that constitute Rhode Island. Data from RI Association of School Committee, RI Dept of Education, Moody's rating, RI Dept of Community Affairs, RI Dept of Elderly affairs 1973-74, 1974-75, 1977-78. **Unit of observation:** District. **Type of Econometric Model(s):** OLS, cross-sectional **Measure(s) of dependent variable:** Salary index **Measure(s) of unionization:** Number of members **Estimates obtained from this study:** 3
Stoddard (2005). Economics of Education Review.	**Data Source:** 5% public use microdata sample 1980 and 1990 US Census, nationally representative sample. **Unit of observation:** Teacher. **Type of Econometric Model(s):** OLS, cross-sectional **Measure(s) of dependent variable:** Yearly wage **Measure(s) of unionization:** (a) dummy=1 if administration has duty to meet, agency shops are permitted, or union shops are permitted (b) teacher union index that ranks legal environment **Estimates obtained from this study:** 4
Tracy (1988). NBER working paper.	**Data Source:** Varies, CPS, Census, 1977, 1980. **Unit of observation:** Teacher. **Type of Econometric Model(s):** (a) OLS (b) GLS **Measure(s) of dependent variable:** ln(wage) **Measure(s) of unionization:** (a) Meet and confer (b) Duty to bargain, no strikes or arbitration (c) Duty to bargain, access to strikes or arbitration **Estimates obtained from this study:** 6
West & Mykerezi (2011). Economics of Education Review.	**Data Source:** Varies, Teacher Rules, Roles and Rights (TR3) compiled by National Council for Teacher Quality, SASS, 2006-07, National. **Unit of observation:** District. **Type of Econometric Model(s):** OLS, cross-sectional **Measure(s) of dependent variable:** ln(starting wage) **Measure(s) of unionization:** collective bargaining **Estimates obtained from this study:** 2

Continued on next page

TABLE 1 – *Continued from previous page*

Author (Year) Journal	**Summary of Study**
Winters (2011). Industrial and Labor Relations Review.	**Data Source:** Schools and Staffing Survey, School District Demographic System, Bureau of Labor Statistics, 1999-2000, 48 contiguous states. **Unit of observation:** District. **Type of Econometric Model(s):** (a) OLS, cross-sectional (b) Spatial model **Measure(s) of dependent variable:** (a) log(base salary for 20 yrs experience and a master's degree) (b) log(base salary for no teaching experience and a bachelor's degree) **Measure(s) of unionization:** (a) collective bargaining (b) meet and confer (c) share of districts in the state with collective bargaining agreement (d) state union membership **Estimates obtained from this study:** 8
Zwerling and Thomason (1995). Journal of Labor Research.	**Data Source:** Main data source is a national sample of districts from 1984, Administrator-Teacher Survey of the National Logitudinal Survey: High School and Beyond, 1984, nationally representative, 186 schools w/ unions and 77 schools w/o unions. **Unit of observation:** School. **Type of Econometric Model(s):** OLS, cross-sectional **Measure(s) of dependent variable:** (a) ln(highest salary in school) (b) ln(lowest salary in school) **Measure(s) of unionization:** (a) collective bargaining (b) union density at state level **Estimates obtained from this study:** 5

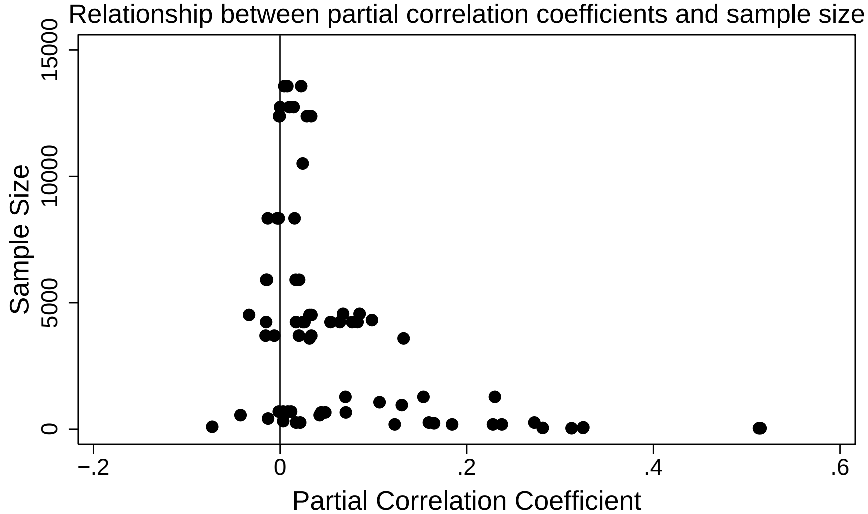
TABLE 2:
Comparing the Size of Union Wage Effects Across Specifications
Subgroup Analysis (weighting for within study dependence)

Partial correlation coefficients and test statistics derived from 78 estimates from 20 studies.

Specification Characteristic	Number of Estimates	Effect Size	95% Confidence Interval	Z Statistic	I^2 Statistic
Unit of Observation					
Teacher	36	0.022	[0.012,0.032]	4.47	55.8
District	33	0.055	[0.033,0.077]	4.94	36
State	4	0.063	[-0.067,0.194]	0.95	51.6
Test statistic for difference	*-2.557,-0.121*				
Wage Measure					
Average	35	0.031	[0.020,0.043]	5.28	68.7
New Teacher	28	0.024	[0.008,0.040]	2.9	0
Senior Teacher	8	0.077	[0.041,0.112]	4.27	0
Test statistic for difference	*0.7,-2.754*				
Unionization Measure					
Collective Bargaining Agreement	33	0.058	[0.037,0.079]	5.39	51.5
MembershipCoverage	22	0.025	[0.001,0.048]	2.07	19.7
Legal	22	0.014	[0.006,0.022]	3.53	30.9
Test statistic for difference	*2.119,0.859*				
Within vs Across State Variation					
Within State	29	0.054	[0.028,0.079]	4.17	48.3
Across State	48	0.027	[0.017,0.037]	5.26	54.3
Test statistic for difference	*1.940*				
Decade of Data					
Data from 1970s	29	0.120	[0.063,0.178]	4.10	24.7
Data from 1980s	48	0.023	[0.007,0.038]	2.84	67.3
Data from 1990s	4	0.035	[0.003,0.067]	2.12	86.6
Data from 2000s	48	0.033	[0.020,0.047]	4.74	29.2
Test statistic for difference	*3.129,1.09,0.2*				
Decade Published					
Published during 1970s	10	0.069	[-0.009,0.147]	1.74	0
Published during 1980s	15	0.066	[0.026,0.106]	3.22	73.8
Published during 1990s	10	0.038	[0.009,0.068]	2.53	68.8
Published during 2000s	42	0.024	[0.015,0.033]	5.17	34.8
Test statistic for difference	*0.067, 1.12, 0.886*				

Figure 1

Relationship between partial correlation coefficients and sample size



Note: This graph provides no evidence of publication bias. 6 estimates with sample sizes greater than 20,000 have been excluded from this graph. These estimates were excluded so that the scaling of the y−axis would allow the reader to better see the data. The excluded estimates have a mean partial correlation coefficient of 0.014 and a range of 0.004 to 0.040.

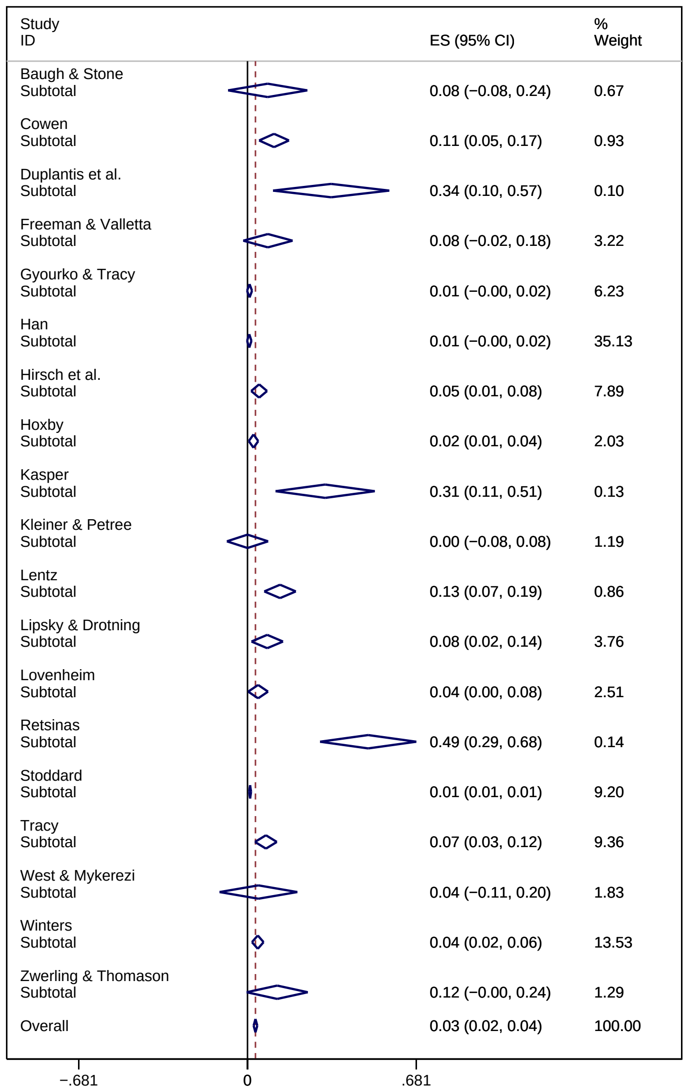
TABLE 3:
The Impact of Study Moderators on Effect Size

	(1) OLS β / SE	(2) OLS β / SE	(3) WLS β / SE	(4) WLS β / SE
Controls for teacher's experience	−0.036	−0.064	0.007	−0.013
	(0.040)	(0.061)	(0.014)	(0.020)
Controls for teacher's education	−0.026	−0.164*	0.047*	−0.130*
	(0.068)	(0.095)	(0.025)	(0.075)
Controls for alternative wage	−0.064	−0.081	−0.060***	−0.067*
	(0.052)	(0.056)	(0.020)	(0.036)
Controls for teacher gender	−0.040	0.092	−0.084***	0.159**
	(0.048)	(0.105)	(0.019)	(0.070)
Controls for student SES	−0.120***	−0.158***	−0.058***	−0.100***
	(0.025)	(0.026)	(0.012)	(0.032)
Controls for median home value	−0.084	0.047	−0.022	−0.010
	(0.057)	(0.086)	(0.027)	(0.045)
Unionization varies within state	−0.091**	−0.064	−0.103***	−0.104***
	(0.043)	(0.067)	(0.023)	(0.038)
Salary measure is average salary	−0.096**	−0.185***	−0.031**	−0.074
	(0.045)	(0.061)	(0.015)	(0.049)
Salary measure is for senior teachers	0.067	0.057	0.055**	0.057**
	(0.047)	(0.049)	(0.025)	(0.025)
Unionization measured as CBA	−0.036	−0.034	0.030***	0.041*
	(0.030)	(0.043)	(0.010)	(0.021)
Unionization measured as membership or coverage	0.019	−0.016	0.027*	0.006
	(0.028)	(0.053)	(0.016)	(0.029)
Data from 1980s	−0.042	−0.045	−0.004	−0.086***
	(0.040)	(0.057)	(0.021)	(0.032)
Data from 1990s	0.034	0.052	−0.003	−0.088***
	(0.048)	(0.062)	(0.021)	(0.032)
Data from 2000s	−0.103**	−0.313***	−0.060**	−0.155*
	(0.045)	(0.093)	(0.027)	(0.080)
Journal has above average SNIP ranking		−0.054		−0.061*
		(0.061)		(0.034)
Constant	0.393***	0.540***	0.191***	0.323***
	(0.088)	(0.093)	(0.042)	(0.108)
Observations	74	50	74	50
R^2	0.546	0.655		
Goodness-of-Fit				

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses.

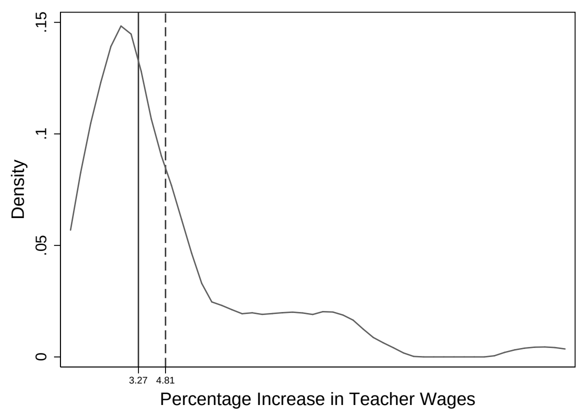
Note: The excluded category for salary measure is new teacher salaries. Both weighted least squares specifications use inverse-variance weights calculated from the adjusted standard errors. We compute adjusted standard errors by multiplying the standard errors by $sqrt[1 + l(b - 1)]$ where b is the number of estimates from a particular study and l is the intraclass correlation within studies.

FIGURE 2
Forest plot of partial correlation coefficients for all studies



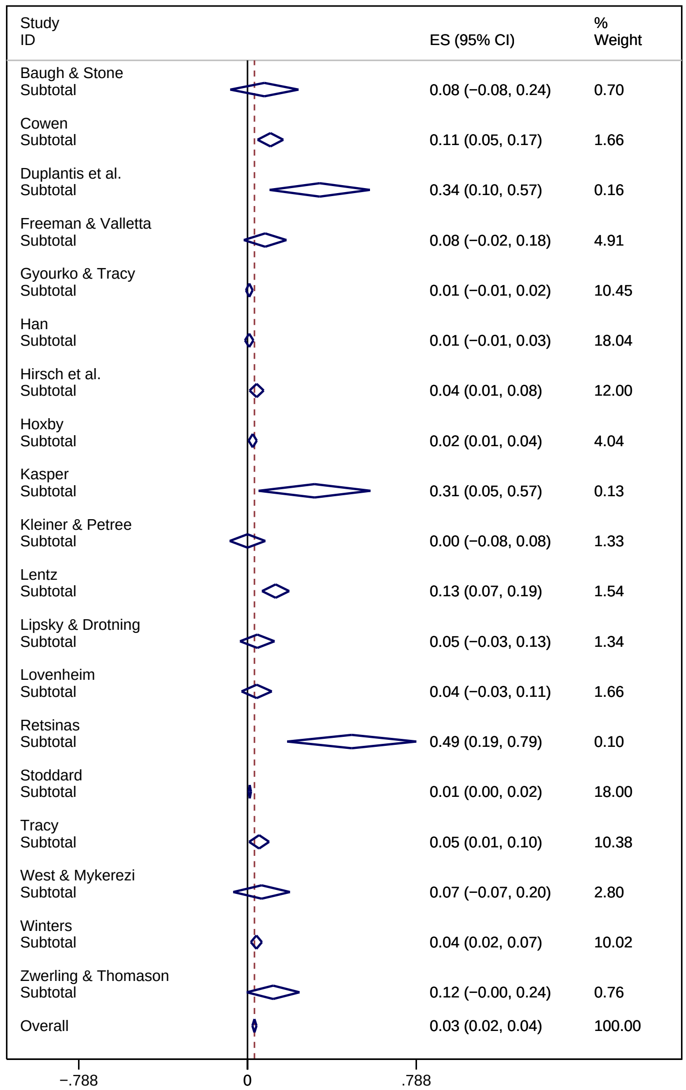
A random-effects model is utilized to deal with the fact that there is unlikely to be one true effect size generated by unionization. The effect sizes are weighted by their inverse-variances. Each effect size is assumed to be independent. See the text for an explanation of how we selected studies to mitigate within study dependence. The overall I^2 is 84.2%.

FIGURE 3
Economic Significance of Wage Impacts



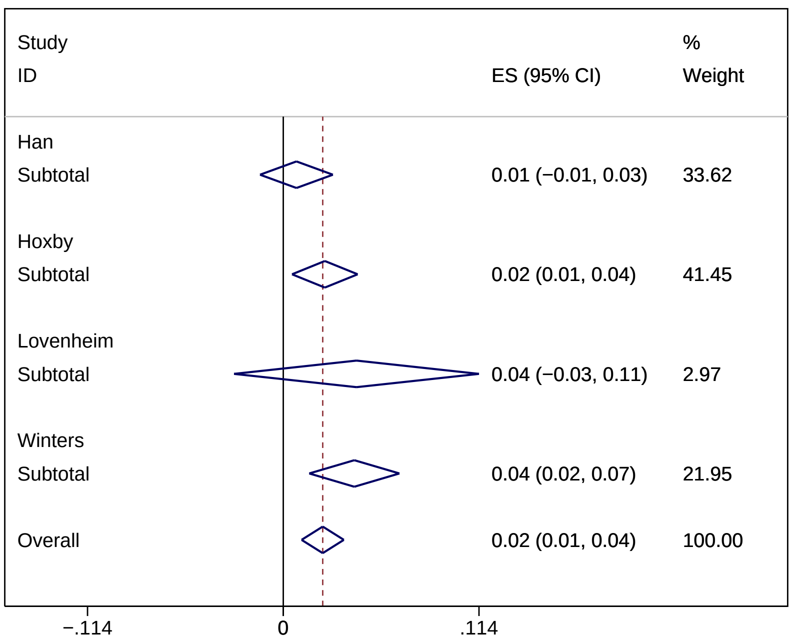
A random-effects model is utilized to deal with the fact that there is unlikely to be one true effect size generated by unionization. The economic impacts in this figure are found by reversing the partial correlation transformation for log-linear specifications, coeff*100=0.03*standard error*sqrt(sample size)

FIGURE 4
Forest plot of partial correlation coefficients for all studies weighted for within study dependence



A random-effects model is utilized to deal with the fact that their is unlikely to be one true effect size generated by unionization. The effect sizes are weighted by the inverse-variance method. To deal with the possibility of intrastudy dependence we compute adjusted standard errors by multiplying the standard errors by $sqrt[1 + l(b - 1)]$ where b is the number of estimates from a particular study and l is the intraclass correlation within studies. This method comes from Kish(1965), and is the same as the method proposed in Hedges et al. (2010). The overall I^2 is 54.0%.

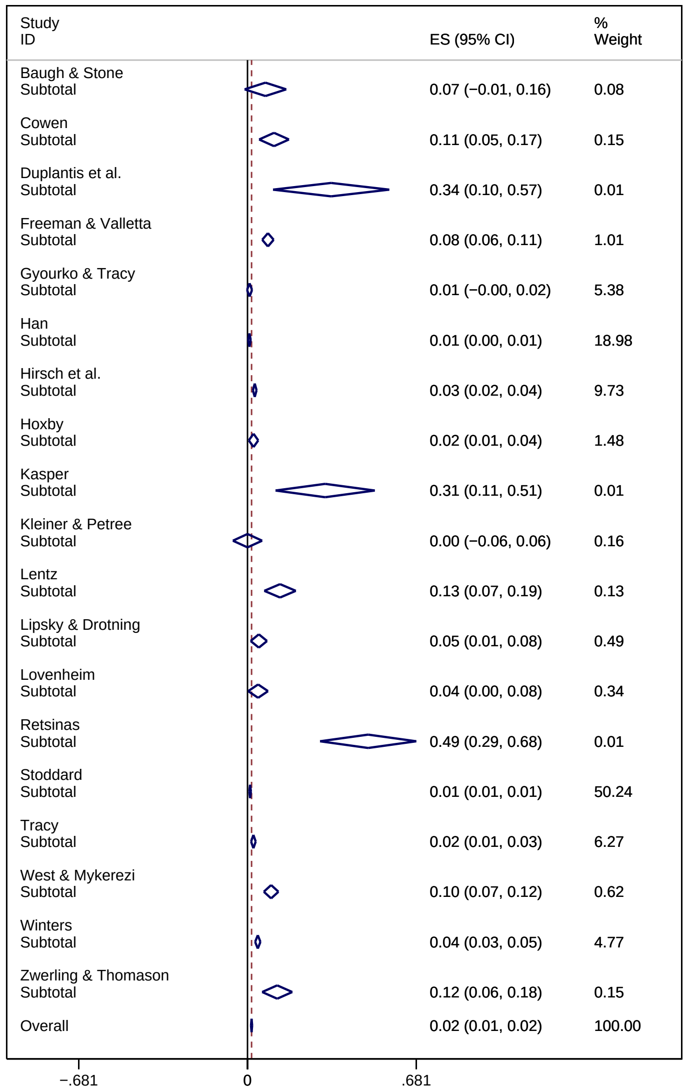
FIGURE 5
Forest plot of partial correlation coefficients for good empirical methods



A random-effects model is utilized to deal with the fact that there is unlikely to be one true effect size generated by unionization. The effect sizes are weighted by their inverse-variances. See the text for an explanation of how we mitigate within study dependence. The overall I^2 is 0%.

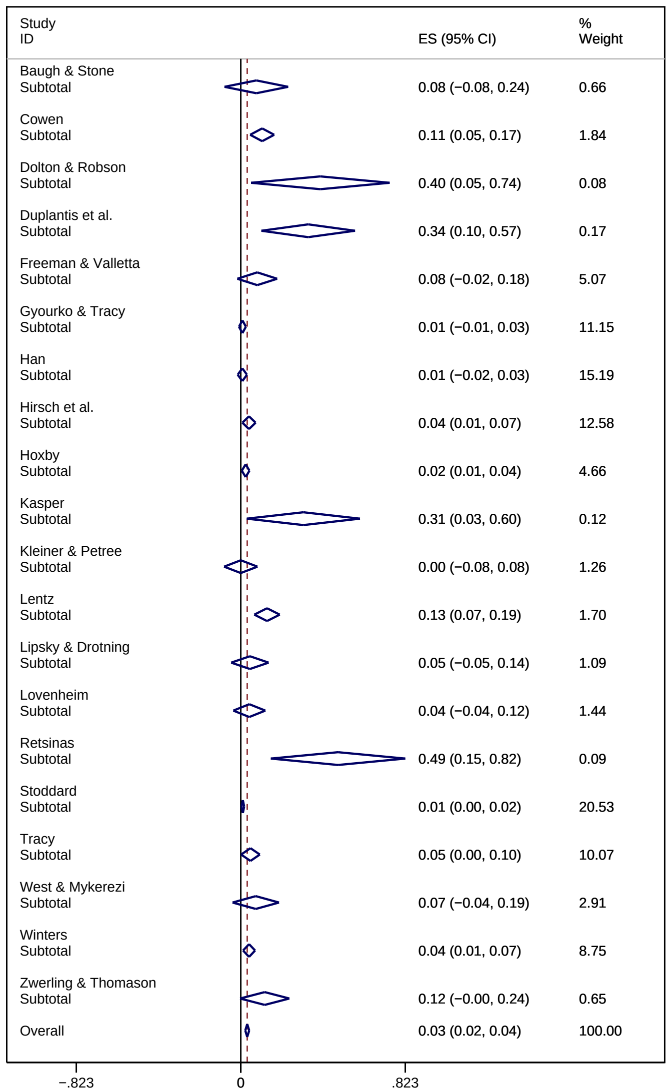
APPENDIX

FIGURE A.1
Forest plot of partial correlation coefficients for all studies



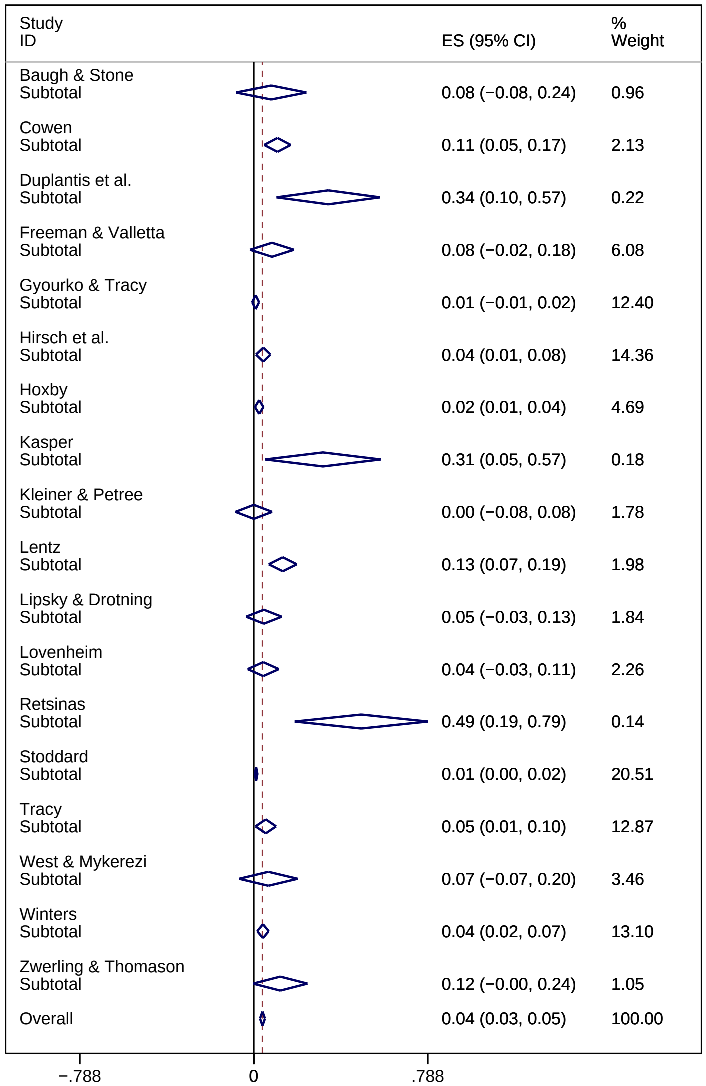
The table above is a fixed effects meta-analysis. Smaller studies receive less weight in this specification than in the one reported in Figure 3. The effect sizes are weighted by their inverse-variances. To deal with the possibility of intrastudy dependence we compute adjusted standard errors by multiplying the standard errors by $sqrt[1 + l(b - 1)]$ where b is the number of estimates from a particular study and l is the intraclass correlation within studies. This method comes from Kish(1965), and is the same as the method proposed in Hedges et al. (2010). The overall I^2 is 54.0%.

FIGURE A.2
Forest plot of partial correlation coefficients if within study estimates were perfectly correlated



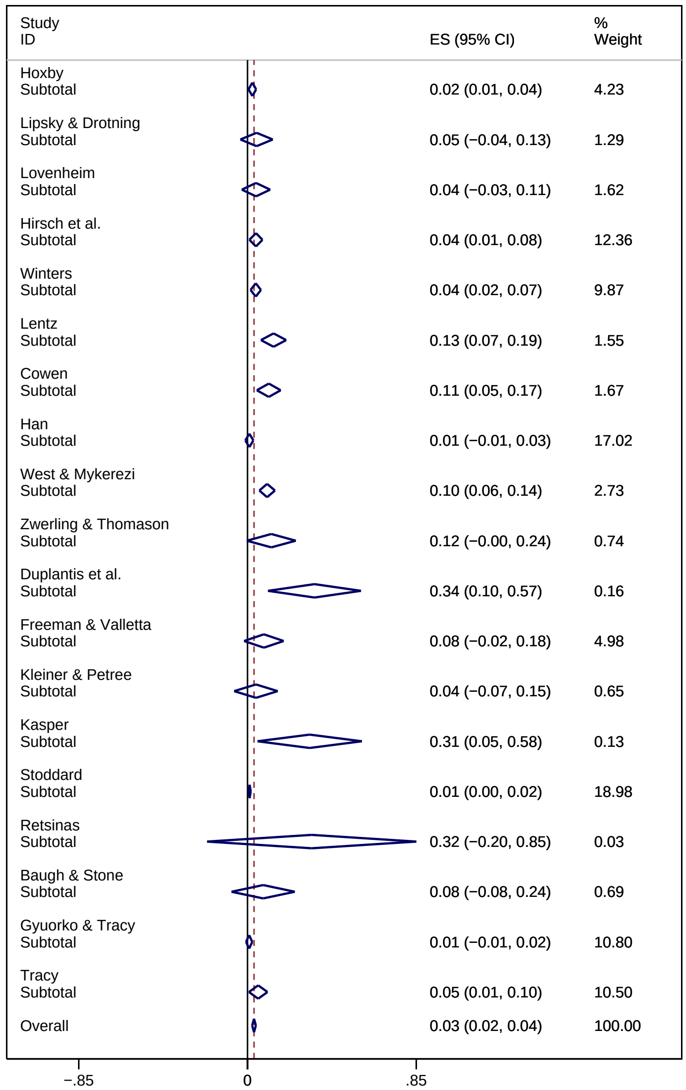
The table above estimates the overall effect size. This is a random effects meta-analysis. The effect sizes are weighted by their inverse-variances. To deal with the possibility of intrastudy dependence we compute adjusted standard errors by multiplying the standard errors by $sqrt[1 + l(b-1)]$ where b is the number of estimates from a particular study and l is the intraclass correlation within studies. In this table, l is set equal to 1. This method comes from Kish(1965), and is the same as the method proposed in Hedges et al. (2010). The overall I^2 is 46%.

FIGURE A.3
Forest plot of partial correlation coefficients without Han effect sizes



The table above estimates the overall effect size without the Han study. The same analysis has been performed for all studies in our sample. We include this one because it results in the largest change in the overall effect size. The other studies produce estimates that range from 0.029(deleting Freeman and Valletta(1988)) and 0.039 (deleting Hoxby(1996)). This is a random effects meta-analysis. The effect sizes are weighted by their inverse-variances. To deal with the possibility of intrastudy dependence we compute adjusted standard errors by multiplying the standard errors by $sqrt[1 + l(b - 1)]$ where b is the number of estimates from a particular study and l is the intraclass correlation within studies. This method comes from Kish(1965), and is the same as the method proposed in Hedges et al. (2010). The overall I^2 is 64.4%.

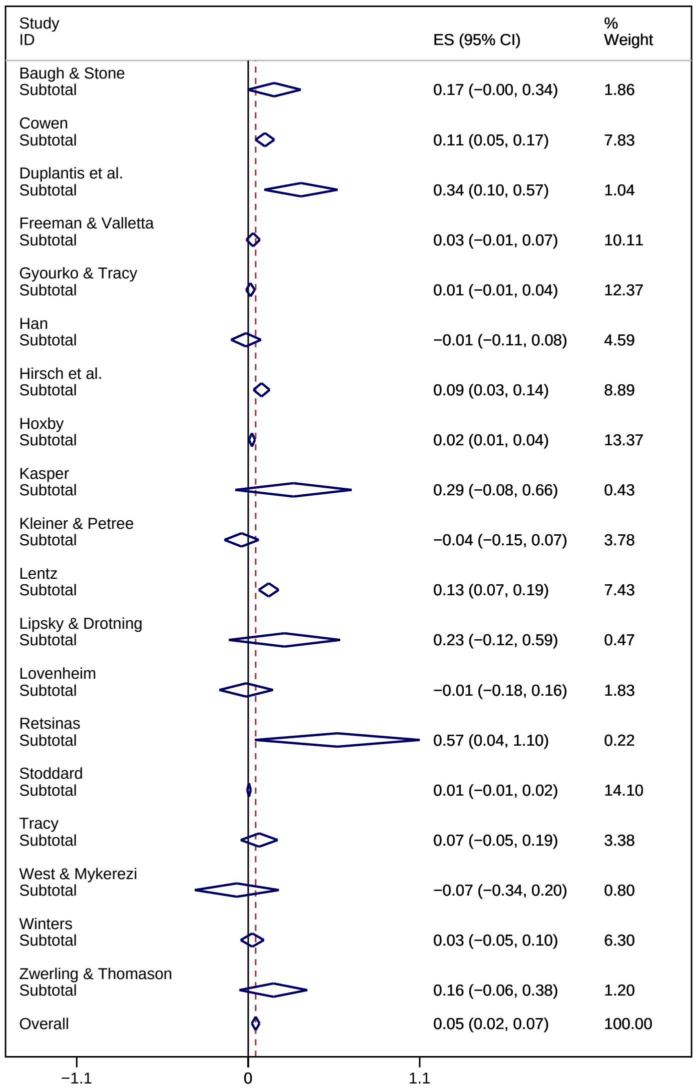
Figure A.4
Forest plot of partial correlation coefficients trimming top and bottom 5%



The table above estimates the overall effect size after trimming the top and bottom 5% of effect sizes. This is a random effects meta-analysis. The effect sizes are weighted by their inverse-variances. To deal with the possibility of intrastudy dependence we compute adjusted standard errors by multiplying the standard errors by $sqrt[1 + l(b-1)]$ where b is the number of estimates from a particular study and l is the intraclass correlation within studies. This method comes from Kish(1965), and is the same as the method proposed in Hedges et al. (2010). The overall I^2 is 53.3%.

Figure A.5

Forest plot of partial correlation coefficients when one estimate per study is selected



A random-effects model is utilized to deal with the fact that there is unlikely to be one true effect size generated by unionization. The effect sizes are weighted by their inverse-variances. Only one estimate is selected from a study in this specification. This is a standard method for addressing within study dependence. The overall I^2 is 68.2%.