Economics of Education Review 84 (2021) 102169

Contents lists available at ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

Are U.S. teacher salaries competitive? Accounting for geography and the retransformation bias in logarithmic regressions

McKinley L. Blackburn

Department of Economics, University of South Carolina, Columbia, SC 29208, United States

ARTICLE INFO

JEL codes:
J31
I21

Keywords:
Teacher pay
Wage differentials
Quasi-maximum likelihood estimation

ABSTRACT

Using data from the American Community Survey for 2012-2016, I estimate relative earnings differentials between teachers and observationally equivalent non-teachers. Two concerns at primary issue in the paper are adequately controlling for differing geographic locations of teachers and non-teachers, and addressing the bias that can arise in the use of logarithmic specifications of earnings regressions to estimate average wage differentials. I find that both issues are of relevance: while ignoring disparate location biases the differential away from zero, failing to account fully for differences in the distribution of earnings for teachers and non-teachers biases it towards zero. An analysis of data from the 2007-2011 American Community Survey suggests that that the magnitude of the teacher pay differential has increased since that time. Other suggested corrections, based on earlier research on differential time misreporting and benefit differences, lead to a smaller but still economically significant differential.

Concerns related to teacher compensation in the U.S. have been a constant in the evaluation of school performance. Such concerns received renewed attention in the 2018-2019 period, as several cities and states faced teacher strike activity. Pay was not the only motivating factor leading to strikes, but was an important concern in most cases. The public appears to be supportive of teachers' demands; a recent poll suggested that 73 percent of respondents would support public-school teachers striking for higher pay (Meckler, 2018).

The contention that teachers are relatively underpaid became an issue in the 2020 U.S. Presidential campaign, with top Democratic candidates supporting efforts to increase teacher pay (EdSource, 2020). Those supporting teacher raises often cite several studies that have used household survey data to compare the earnings of teachers to similar workers in non-teaching occupations. A particularly common source is a series of studies over time on teacher pay promoted by the Economic Policy Institute in Washington, D.C. (of which I focus on the recent study by Allegretto and Mishel, 2018). These studies point to a negative earnings differential for teachers compared to other college graduates, with the differential becoming increasingly large in magnitude over the years since 2000. By 2017, Allegretto and Mishel report an estimated weekly-wage gap of -19 percent for U.S. teachers. This series of studies has particularly been used by policy activists arguing for increases in teacher pay (as an example, see Benner et al., 2018). There has been a fair amount of criticism of the EPI studies as well – for example, see Aldeman (2019) and Biggs (2019).

Arguments for increased teacher pay have also cited evidence that the relative pay for teachers in the U.S. is considerably below that of other advanced economies (for example, see Strauss, 2017). The source most often cited is a series of roughly-500-page reports by the OECD (ironically titled *Education at a Glance*) that provide considerable cross-national evidence allowing comparisons of educational systems. For example, in OECD (2019) the U.S. was reported as paying primary-school teachers 45 percent less than similarly-educated non--teachers in the year 2017.[1] By comparison, the same differential across 20 OECD countries around that same time suggested teachers in those countries were paid only 15 percent less than the comparison group. The U.S. had the lowest relative teacher-pay across all countries included in the study.

The comparison of the estimated teacher-pay differentials between Allegretto and Mishel (2018) and OECD (2019) also serves to highlight a lack of agreement across studies as to the degree of perceived underpayment of teachers relative to similar workers in nonteaching occupations, as the OECD reports a differential more than twice as large in magnitude than that reported by Allegretto and Mishel. Even though both studies use Current Population Survey data, they use different salary reports to measure relative pay. The OECD study relies on annual

E-mail address: blackbrn@moore.sc.edu.

[1] Similar-sized differentials are reported at other teacher levels.

https://doi.org/10.1016/j.econedurev.2021.102169
Received 10 December 2020; Received in revised form 13 August 2021; Accepted 17 August 2021
Available online 2 September 2021

salary comparisons of teachers with nonteachers who are full-time, full-year college graduates, even though many teachers do not work the full year. Allegretto and Mishel focus on weekly wages reported in the Outgoing Rotations Group, that in theory could address the lower labor supply levels of teachers (though, as discussed later, criticisms are still relevant here). While this is likely to be a major source of the difference, it is also the case that Allegretto and Mishel incorporate an extensive set of controls for possible wage differences, something not incorporated in the simpler OECD comparisons.

Given this inconsistency in the magnitude of estimated teacher-pay differentials, it is relevant to consider factors that might cause an observed negative differential even though teachers have expected earnings close to those of comparable non-teachers. For one, the concern has been raised that work hours of teachers are more likely to be overstated, causing their weekly wage to be understated. Second, it has been noted that the salary differential ignores the fact that other sources of compensation (health insurance, retirement, for example) may be relatively higher for teachers than for nonteachers. Third, a related concern is that working conditions or the nature of the work may make the teaching occupation more attractive, suggesting the negative differential is primarily a compensating differential. Fourth, it has been suggested that a comparison of teachers to all college graduates is not relevant, if there is little possibility of teachers moving to high-paying alternative occupations. A fifth concern (related to the fourth) is the prevalence of education majors in the teaching occupation, an educational preparation that might leave them unqualified for many of these high-paying alternative occupations.

While I consider the importance of these explanations for the teacher-pay differential later in the paper, my primary contribution is to discuss two additional factors that are relevant to the size of the differential. One has to do with the fact that teachers are sorted across physical locations in a manner that is more representative of the general population than of other college graduates. For example, Moretti (2013) documents that college graduates are increasingly more concentrated in urban locations, where the demand for skill is higher (and where the cost of living is also higher). Previous research primarily relies on state-level controls (at best) in controlling for geography, but even locations in the same state may differ in their attractiveness to college-educated workers. In this paper, I use the American Community Survey (ACS) to estimate earnings differentials between public-school teachers and college-educated non-teachers in the private sector. Given the large scale of the ACS, I am able to control for geographic location at a more detailed PUMA level, with 982 distinct locations identified in the data I analyze. My analysis also includes estimates based on comparisons using exact matches between teachers and non-teachers working in the same PUMA.

A second concern addressed in this paper has to do with the way in which the nature of the distribution of earnings for teachers tends to differ from that of comparable non-teachers. Specifically, the dispersion of earnings tends to be lower for public-school teachers than for non-teachers, most likely due to the use of government pay scales in compensating teachers that tend to limit variation in pay. A common salary structure used by school districts is "step and lane," in which salaries are determined by experience (the "step") while educational qualifications can change the "lane" from which a teacher's payment may be determined. Salaries are also often set through explicit contracts set by collective bargaining (see Nittler, 2019). This naturally leads to a greater compression of salaries among public-school teachers, compared to a broader workforce of college graduates for whom pay may be set in a less-regimented manner.

The smaller variation in pay for teachers can actually lead to an overstatement of the relative pay of teachers compared to non-teachers when the focus is on traditional logarithmic earnings specifications (that is, the Mincer semi-log earnings equation). To see the issue, suppose that the distributions of earnings for teachers and non-teachers (after controls) have the same mean, but that the non-teacher distribution is more disperse. While the true earnings differential would be zero in this case, the concavity of the log function would tend to compress the earnings distribution above the mean relative to below the mean. This would lead to a log-wage comparison that suggests that earnings are lower on average for non-teachers compared to teachers, suggesting a positive differential in earnings for teachers. This is known as the retransformation problem, and the potential bias that can arise has been known in the health-economics literature at least since Manning (1998).

Problems with the retransformation of log earnings to estimate wage differentials have been noted in several previous wage studies. Blackburn (2007) provides evidence that the union/non-union wage differential is overstated when using log-wage regressions, related to the tendency for wage dispersion to be higher for non-union workers.[2] Hirsch and Schumacher (2012) show that earnings differentials for RN nurses (compared to other college-educated workers) are overstated when using a log-wage regression. Falk (2015) found a considerably smaller earnings differential for federal-government workers (compared to private-sector workers) once the issue was addressed, while a similar result is evident from the study of state-and-local government wages reported in Gittleman and Pierce (2012). As Falk emphasizes, comparisons of average wages are what is directly relevant when budget makers are concerned with the impact of changing compensation schemes on government outlays, so comparisons of expected wages (not log wages) is of most interest. These previous wage studies employ various alternative estimation methods for consistently estimating differentials in average earnings – adjusting log-wage coefficients, estimating wage equations directly (through quasi-maximum-likelihood), or matching on wages rather than log wages. All of these approaches are considered here in my analysis of teacher/non-teacher earnings differentials.

One additional aspect of the analysis in this paper should be highlighted. The results in this paper (and in the previous literature cited in the next section) should not be thought of as representing a causal effect on earnings of being a teacher – that is, the difference in earnings that current teachers may expect to face in choosing between a teaching job and a private-sector job. There may be many unobservable factors that could lead to the earnings of teachers being lower in the private sector than current non-teachers who are comparable based on observable characteristics. Given this, analyses referring to a "teacher wage penalty" are perhaps overly provocative, as the type of evidence used to support a teacher wage differential does not necessarily imply a penalty that any given teacher may be experiencing. Corcoran et al. (2004) present evidence suggesting that the cognitive ability of female teachers (as measured by their performance on tests) had been declining over time relative to comparable private-sector workers.[3] A decline in relative pay of teachers over time could just be a reflection of this decline in quality of the teaching workforce, for instance, as would be expected if labor markets are competitive with workers paid according to their relative productivity.[4]

If the estimates presented here (and in similar analyses) are not necessarily of causal parameters, are they interesting from a policy point of view? If the desire is to have a teaching force that is similar in quality to the average college graduate, then paying under-market wages may run counter to this goal. The federal government also has similar concerns, as reflected in their assessment of federal compensation in reports

[2] A similar result is found by Eren (2007) and Campolieti (2018) using matching estimators for the union wage differentials.

[3] Goldhaber and Walch (2014) provide evidence from SAT scores that the relative quality of teachers has more recently been improving. They focus on data from the recession years of 2008-2009, however, and note that it is not clear whether this was just an impact of the recession on teacher markets.

[4] However, there is evidence of monopsony power in teacher markets (for example, see Ransom and Sims, 2010), so that any negative earnings differential could also be a reflection of the limited employment possibilities for teachers tied to a particular location.

such as Congressional Budget Office (2017), where the question addressed is how the compensation of federal civilian employees compares to that of observationally-equivalent workers in the private sector. Should teacher pay be set at a level competitive with a similar private-sector worker? While early evidence did not point to increasing teacher pay leading to improved student performance, Britton and Propper (2016) provide more recent evidence that paying below-market salaries to U.K. teachers reduced student achievement. Using Texas data, Hendricks (2014) finds that teacher experience increases as the base level of pay increases, which he argues should also lead to improved student performance. Increasing the pay of teachers to be closer to the average of comparable college graduates could also be warranted, if it attracts a new type of individual to the teaching profession. As noted by Hanushek et al. (2019), higher ability teachers do seem to correlate with better-performing students, leading to the suggestion that one reason why U.S. students do poorly relative to those in most other advanced countries is that the teacher wage differential is more negative in the U. S. Ultimately this becomes a normative question related to the relative valuation of educational quality and other goods. But it should be useful in this assessment to know where the current pay of teachers sits relative to the labor force from which teachers might be drawn.

My initial results from the ACS for the 2012-2016 period suggest a large negative earnings differential – roughly -20 percent – for teachers when compared to observationally-similar non-teachers. The magnitude of this differential is lessened by about 5 percentage points when a more precise measure of geographic location than state of residence is incorporated in the model. On the other hand, the magnitude of the differential is increased by around 6 percentage points when the bias in log-wage regression coefficients is addressed. The earnings differential is considerably larger in magnitude for men than for women, though the above biases work similarly for both genders. A similar analysis using the 2007-2011 ACS is consistent with the magnitude of the differential increasing (by about 5 percentage points) since that period. Several other suggested adjustments are considered in section 5 of the paper (see section 5.6 for a summary table presenting the size of these adjustments), suggesting that the best estimate of the differential could be around -10 percent lower wages for teachers compared to non-teachers.

1. Previous research on teacher pay

The report by Allegretto and Mishel (2018) is characteristic of a series of studies that have been publicized by the Economic Policy Institute to provide information on trends in the pay of teachers in the U.S. These reports compare over time the pay of public-school teachers with that of other college graduates in non-teaching occupations in the private sector. The teacher wage differentials are calculated using weekly wages from the Outgoing Rotation Group (ORG) sample from the monthly Current Population Survey (CPS). Estimates are adjusted for characteristics differences between teachers and non-teachers using log-wage regressions. The only geographic control is region of the country. They find a persistent "teacher wage penalty" of roughly 5 percent from 1979 to the mid-1990s, and then a persistent drop after the mid-1990s. Allegreto and Mishel (2018) also note that the downward trend in teacher differentials in wages is particularly striking for women, as the differential goes from essentially zero in the 1980s to roughly -15 percent in recent years. By comparison, the teacher differential in pay for men has always been negative and large (equaling -27 percent in 2017). In an earlier discussion of issues related to the design of teacher compensation, Hanushek (2007) noted that pay for female teachers was typically above the median for all women in the 1940s and 1950s, only falling below median in the 1990s (teacher pay for most men has been below overall male medians since the 1950s). Results from Corcoran et al. (2004) suggest that the decline in relative pay for female teachers has occurred along with a fall in relative ability (reflected in test scores intended to measure cognitive ability), perhaps due to the opening of other career opportunities for women who were historically crowded into the teaching profession.[5]

The possibility that measured teacher differentials in pay might be explained by geographic-location differences has been examined by Stoddard (2005). Cost-of-living adjustments have often been seen as important for policymakers and researchers in evaluating the extent to which teachers might be underpaid. She notes that attempts to measure available cost-of-living indices rely primarily on rent differences that in themselves work against correcting for compensating differentials associated with amenities and job opportunities. She argues that a relative comparison of teachers with comparable private-sector workers in the same geographic area appears to sufficiently account for these differentials (based on the idea that non-teacher state effects are similar to teacher state effects in wage regressions). Her analysis uses the 1980 and 1990 U.S. Census to estimate linear wage-level regressions (that is, the level of wages modeled as a linear function of characteristics). A similar finding is reported in Rickman et al. (2017) using log-wage regressions with ACS data from 2009-2011, with geography also measured at the state level. Both studies conclude that relative comparisons of teacher and non-teacher pay within states fully account for amenity differences across states.

Taylor (2008) also considers the importance of geography in her analysis of relative teacher pay in the 2000 U.S. Census. While she estimates teacher differentials from log-wage models with state effects, she also considers controlling for the Public Use Microdata Area (PUMA) of the worker. This geographic classification allows for identification of roughly 800 areas in the 5-percent microdata sample of the U.S. Census she uses. Taylor finds that PUMA controls reduce the estimated pay disparity between teachers and other college-educated workers by roughly 3 percentage points compared to only controlling for state effects. A pay gap of -8 percent still remains, however, even after PUMA controls.[6]

One criticism that arises in attempts to measure teacher differentials in pay has to do with measuring the relative labor supply of teachers and non-teachers. Pordgursky and Tongrut (2006) argue that measurement error in the ORG samples makes the use of its weekly earnings variable problematic – in particular for teachers, who tend to work less than the typical full-time, full-year worker. West (2014) argues that teachers are particularly prone to overstating their usual work hours (per week), based on comparing the usual CPS reports on hours worked to measures derived from diary data in the American Time Use Survey supplement to the CPS for 2003-2010. She finds evidence of a positive teacher differential in hourly wages once diary measures of hours are incorporated in log-wage regressions. This is the only study I am aware of using data from after the 1990s that does not find evidence of a negative pay difference.

A summary of the size of the estimated teacher wage differentials reported in these previous studies is provided in Table 1.[7] All studies report a differential estimated by comparing public-school teachers to college-graduate non-teachers. The OECD studies provide the largest

[5] The nature of their analysis, however, does not allow for a consideration of the extent to which the pay differential at a point in time might be accounted for by cognitive-ability differences.

[6] She also estimates models in which the comparison is only to a set of 16 non-teaching occupations considered most comparable to teaching, and finds only a -4 percent gap after PUMA controls.

[7] While Stoddard (2005) provides estimates of separate wage regressions for teachers and non-teachers, she does not provide sufficient information to calculate a differential. Rickman et al. (2017) also do not provide estimated teacher differentials.

Table 1.
Estimates from previous studies of the teacher/non-teacher earnings differential.

Study	Data	Model/Estimation Approach	Year	Estimated Diff.
Allegretto and Mishel (2018)	CPS ORG, various years	log-wage OLS regression/reported weekly earnings	1994 2004 2010 2017	-0.018 -0.108 -0.121 -0.171
OECD (2019) and earlier years	March CPS, and admin. data for teachers in 1995	average earnings comparisons of teachers and college-educated non-teachers/annual earnings	1995 2010 2017	-0.380 -0.300 -0.450
Podgursky and Tongrut (2006)	CPS ORG	median regression on log wages, using weekly earnings	2002	-0.075 (w) -0.270 (m)
Taylor (2008)	2000 U.S. Census	log-wage regression, using hourly earnings	1999	-0.084
West (2014)	American Time Use Survey	log-wage regression, using weekly earnings calculated from diary hours	2003-2010	0.013

Notes: Reported coefficients are from authors' preferred specifications (if noted). Estimates are for all workers, except for Podgursky and Tongrut where only separate estimates for men and women are provided. Differentials refer to salary compensation only, and do not include other sources of compensation.

differentials, but they are estimated with annual earnings and use no controls. Most of the other estimated differentials are from the 2000s, and suggest around 10 percent lower earnings for teachers. Allegretto and Mishel (2018) have numbers for more recent years, suggesting a large increase in the earnings differential since the 2000s.[8]

In my estimation of wage differentials, I consider how the level of location controls affects estimates of teacher differentials in pay. However, unlike in previous studies, I also examine the importance of the estimation method for the typical empirical model of wages used in estimating these differentials. My equations are nominally focused on annual earnings determination, but with controls for weeks and hours worked the results are best interpreted as relevant to the determination of hourly wages. I also consider the sensitivity of estimates to the mismeasurement concerns raised in Podgursky and Tongrut (2006) and West (2014).

2. Data

My analysis primarily uses public use microdata from the American Community Survey (ACS) for the years 2012-2016. The ACS collects data similar in nature to the old long-form data from the decennial U.S. Census (used in Stoddard, 2005, and Taylor, 2008). Conducted by the Census Bureau, the ACS attempts to survey roughly 3.5 million households each year (over the timespan analyzed here). Questions are asked about every person in the household, from which I use demographic information on schooling, age, gender, race, ethnicity, citizenship status, marital status, and geographic location. Information on work status includes earnings, weeks worked, and usual hours per week (all from the previous 12 months), along with occupation and employment sector.

I focus only on individuals who report having a college education or more, as this is the usual expected credential for teachers. Only public-sector teachers are included in the teacher sample, while the comparison group includes only private-sector workers. The teacher sample includes all teaching-related occupations identified in the ACS, including education administrators.[9] Given bias issues that imputation causes for regression parameter estimates when covariates are not used for forming imputations (see Hirsch and Schumacher, 2004), I exclude all observations with imputed earnings, labor supply, or occupation information. I also exclude non-black, non-white individuals, and non-citizens. I include only individuals who reported working at least 27 weeks in the previous year, who had average weekly earnings of at least $200 (in 2016) dollars), and who were between the ages of 23 and 54 (inclusive) at the time of the survey.

There are several ways in which to measure the location of individuals in the ACS. State of residence is possible (all 50 states, plus D. C.), but a more detailed identification is provided by the Public Use Microdata Area (PUMA) associated with the residential location of the household for the individual. As defined by the Census Bureau, each PUMA is located only in one state, and must have a population (at the last Census) of at least 100,000 individuals. PUMA definitions generally follow county lines, though larger counties can be broken across more than one PUMA. Categorization of the U.S. is complete, so that all individuals are located in one (and only one) PUMA. With 2,351 PUMAs identified in the ACS for the years I work with, the average PUMA population is roughly 135,000. The ACS does not identify metropolitan statistical areas (MSAs) in the public-use data, but does provided a crosswalk between the 2010 PUMA definitions and MSA definitions as used by the OMB ("MET2013").[10] This is only a rough translation, however, as not all of a PUMA is in an identified MSA (a PUMA is identified as in an MSA if a majority of its population is in that MSA). There are 382 MSAs identified by this crosswalk, and of course part of the population is characterized as outside of any MSA.

The ACS also provides information on the location of an individual's primary place of work. These "place-of-work" PUMAs are generally aggregations of existing PUMA definitions, so fewer are defined in the data. Like residential PUMAs, place-of-work PUMAs are defined to be in one state only. In my sample construction, I exclude all individuals who are identified as working outside the 50 states (or D.C.), so that I end up with 982 place-of-work PUMAs. As location of work is arguably more important than location of residence to whether pay is competitive, the place-of-work PUMA is the PUMA definition I use in the later analysis.[11]

Descriptive statistics for the teacher and non-teacher samples are provided in Table 2. Not surprisingly, teachers are much more likely to be female than non-teachers, and slightly more likely to be married. They are also half as likely to be foreign-born. Perhaps, the biggest difference is in educational attainment, as teachers are more than twice as likely to have a master's degree, but less likely to have a professional or doctoral degree. Average hours of work per week are similar across the two samples, but teachers are much more likely to have weeks worked in the 40-47 or 27-39 ranges.[12]

The bottom rows of Table 2 report average earnings statistics for teachers and non-teachers. In the raw numbers, teachers earn roughly $28,000 less per year than non-teachers, which suggests a -34 percent differential with respect to non-teacher earnings. A rough estimate of

[8] The reported differentials are differentials with respect to salary compensation. Allegretto and Mishel (2018) do provide estimates on other forms of compensation which are discussed later in section V.A of the paper. Podgursky and Torgut (2006) provide a limited set of differentials estimated using the National Compensation Survey, but it is difficult to construct an overall differential from these.

[9] The particular occupations included in the teacher sample (with their 2020 ACS OCC code values) are preschool and kindergarten teachers (2300), elementary and middle school teachers (2310), secondary school teachers (2320), special education teachers (2330), other teachers and instructors (2340), and education administrators (230).

[10] At the time of writing, this information was available on the IPUMS website of the Census Bureau.

[11] The Census Bureau does not provide a crosswalk between place-of-work PUMAs and MSA, so the MSA definitions I use later are defined on residential location, as would be typical of MSA variables provided in most other data sets.

[12] Weeks worked are reported only in these discrete ranges in the ACS for the years studied.

Table 2.
Descriptive statistics for public-sector teachers and private-sector non-teachers.

	Non-Teachers	*Teachers*
Education		
Masters	0.230	0.551
Professional	0.059	0.034
Doctorate	0.037	0.017
Age	38.9 (9.1)	39.7 (8.6)
Foreign Born	0.052	0.028
U.S. Citizen	0.931	0.959
Hispanic	0.073	0.080
Black	0.070	0.061
Other Race	0.017	0.017
Married	0.641	0.724
Widowed, Divorced, Sep.	0.091	0.096
Female	0.495	0.736
Lives in an MSA	0.942	0.890
Lives in a Large MSA	0.642	0.519
Labor Supply		
Average Hours Per Week	43.2 (9.6)	43.9 (8.5)
Weeks Worked:		
48-49	0.017	0.022
40-47	0.037	0.151
27-39	0.025	0.058
Annual Earnings	84157.2 (78790.8)	55748.7 (25010.9)
Weekly Earnings	1666.1 (1551.4)	1145.9 (517.5)
Log of Annual Earnings	11.060 (0.729)	10.840 (0.436)
N	978,688	112,294

Note: Numbers in parentheses are standard deviations. All variables are dummy variables, except age, hours per week, and earnings variables. All earnings variables are in 2016 dollars. Sample includes only individuals with at least a bachelors degree, who worked 27 weeks or more in the previous year.

average weekly earnings across the two subsamples suggests a -31 percent differential.[13] In comparison, the average log earnings for teachers is only 0.22 log points below that of non-teachers, suggesting a -20 percent differential. It is also notable that the variation in earnings, and log-earnings, for teachers is considerably below that of the non-teacher sample. This difference in variances is potentially important to the next section's discussion of issues regarding estimating teacher differentials with regression models.

3. Estimating wage differentials with regression models

To illustrate the issues associated with the use of log-wage regressions, it is helpful to clearly define the wage differential. For individual i, the expectation of earnings (w) can be written as $E(w_i|T_i, X_i)$, where T is a dummy variable indicating teaching occupation status, and X is a set of other covariates. The teacher wage differential for individual i would typically be defined as

$$\Delta_i = \frac{E(w_i|T_i = 1, X_i) - E(w_i|T_i = 0, X_i)}{E(w_i|T_i = 0, X_i)} = \frac{E(w_i|T_i = 1, X_i)}{E(w_i|T_i = 0, X_i)} - 1. \tag{1}$$

The calculation of the wage differentials presented in the previous section follows this formula, assuming no covariates and identical expectations across individuals.

Most studies that model earnings determination use a linear conditional mean for log wages, that is

$$E[\log(w_i)|T_i, X_i] = \beta_1 T_i + \beta_2' X_i.$$

This does not allow us to evaluate $\Delta_i$as defined in (1), because the expected value of log earnings has no direct relationship to the expected value of earnings. However, the log-earnings regression assumption can be stated equivalently as the assumption that

$$\log(w_i) = \beta_1 T_i + \beta_2' X_i + u_i \tag{2}$$

where the error term u has $E(u_i|T_i, X_i) = 0$. It follows that

$$E(w_i|T_i, X_i) = e^{\beta_1 T_i + \beta_2' X_i} E(e^{u_i}|T_i, X_i).$$

If the expectation of e^u does not depend on T (given X), we have $\Delta_i = e^{\beta_1}$, so that exponentiating the estimated log-wage coefficient on T provides a consistent estimate of the differential for all i.[14]

The retransformation problem becomes relevant to wage differentials when the conditional expectation of e^u does depend on T. One approach is to add assumptions to the log-wage regression model so as to allow for a consistent estimator of Δ_i. For example, if $u_i$is normally distributed with error variance

$$\sigma_i^2 = \lambda_1 T_i + \lambda_2' X_i + e_i \tag{3}$$

then

$$\Delta_i = e^{\beta_1 + (\lambda_1/2)}. \tag{4}$$

This follows from the properties of a log-normal distribution, as, given u_i is normal,

$$E(e^{u_i}|T_i, X_i) = e^{\sigma_i^2/2}.$$

Using only the log-wage regression coefficient β_1 to estimate $\Delta_i$would provide a biased and inconsistent estimator in this case. However, consistent estimators for both β_1 and λ_1 would allow the constant $\Delta_i$to be estimated consistently. While better than ignoring the problem, this approach is still somewhat unattractive, in that it requires a normality assumption for the error terms along with the functional form assumption inherent in equation (3).[15] However, it does highlight that a difference in the error variance between teachers and non-teachers should be a concern in evaluating log-wage regression results.

A preferred approach is to estimate the exponential regression directly.[16] In this case, the exponential regression can be written:

$$w_i = e^{\gamma_1 T_i + \gamma_2' X_i} + v_i \tag{5}$$

where $E(v_i|T_i, X_i) = 0$. This function can be estimated consistently as a generalized-linear model (see McCullagh and Nelder, 1989). Several estimators are possible. For instance, denoting the right-hand-side variables as $\dot{x}_i = [T_i \quad x_i]'$, the standard Poisson estimator solves the equation

$$\sum_{i=}^{n} \hat{v}_i \dot{x}_i = 0,$$

where $\hat{v}_i = w_i - \hat{w}_i = w_i - e^{\widehat{\gamma_1 T_i + \gamma_2 X_i}}$ is the residual formed after choosing estimates. Normality for v leads to the nonlinear least-squares

[13] This is rough because weeks worked is only in ranges; for this calculation, I used the midpoint of the range as the estimated weeks for that particular observation.

[14] Note that the assumption that u has zero conditional expectations does not provide the expectation of e^u.

[15] Duan (1983) has suggested a nonparametric approach to estimating $E(e^u)$, but this approach does not allow the expectation to vary with covariates conditional on other covariates. Ai and Norton (2000) adjusts Duan's estimator for differences in estimated error variances, but this correction is motivated by results assuming normality. Note that other models for the error variance could be considered (such as multiplicative heteroskedasticity), but the linear is the most convenient.

[16] Suggestions for estimating the models in this paper using Stata are provided in an online appendix.

(NLS) estimator, which solves:

$$\sum_{i=}^{n} \hat{v}_i \hat{w}_i \dot{x}_i = 0$$

while assuming a gamma distribution provides an estimator that solves

$$\sum_{i=}^{n} \frac{\hat{v}_i \dot{x}_i}{\hat{w}_i} = 0.$$

All these estimators are quasi-maximum-likelihood (QML) estimators, and are consistent under the assumption that the conditional mean of v is 0 (for example, see Wooldridge, 2010, pp 502-514). QML estimators are solutions to a maximization problem for a likelihood function, but consistency does not require that the likelihood function characterize the true distribution of the data.[17] Even though all of these QML estimators should be consistent under the conditional mean assumption, the first-order conditions make it clear that each estimator will be more sensitive to different parts of the distribution of $\hat{w}_i$; the NLS weights observations with larger expected wages more, and the gamma less, compared to the Poisson estimator. In fact, it is useful to calculate several QML estimators, as observed differences across the estimators suggest a misspecification issue with the conditional mean assumption. Note that with the QML estimators, we have $\Delta_i = e^{\gamma_1}$ without any additional adjustment.

In the following section, I estimate the log differential using the standard log-wage approach $(\hat{\beta}_1)$ and the adjusted log-wage approach $(\hat{\beta}_1 + (\hat{\lambda}_1/2))$, as well as estimating γ_1 by the three QML estimators. In the results, I report the estimates of the log of the differential effect (for example, $\hat{\gamma}_1$ rather than $\widehat{e^{\gamma_1}}$), as that is what is directly estimated. Tests based directly on $\hat{\gamma}_1$ are more appropriate than tests using the differential formula, as test results with the latter can depend on whether the coefficient is estimated as a teacher effect or a non-teacher effect.[18]

4. Results

4.1. Log-wage regression results

The first reported model is based on the direct estimation of the log-wage regression (equation 2) by OLS. This is the conventional approach used in most studies of teacher pay effects (and other wage determinants). The initial results are for a pooled sample of men and women, and are reported in the first row of results in Table 3. The different specifications allow for an assessment of the importance of adding controls for geography, and the degree of these controls, to the estimated teacher-pay coefficients. The raw log differential between teachers and non-teachers is estimated to be roughly -0.22 log points (as in Table 2), while adding additional demographic and labor-supply controls in column (2) slightly increases the magnitude of the estimated differential by almost 0.02 log points.

The last four columns of Table 3 show how the estimated teacher coefficients vary with changes in the controls for geography. The addition of state dummies for residence (available in most data sets) diminishes the estimated log differential's magnitude by only 0.01 log points. By comparison, controlling for residential MSA has a much more important effect, causing the magnitude of the differential to fall by 0.03 log points relative to column (2) with no controls for geography.[19] Addition of state dummies when MSA effects are already incorporated has essentially no impact on the estimated teacher coefficient. The last column shows estimated effects when PUMA effects (based on place of work) are instead included to control for geography – compared to results without geography controls, the estimated coefficient is about 0.07 log points smaller in magnitude.

Table 3.
OLS estimates of teacher effects from log earnings regression model.

	(1)	(2)	(3)	(4)	(5)	(6)
Log Earnings Regression Dummy Variable Coefficient						
	-0.220	-0.235	-0.224	-0.204	-0.203	-0.169
	(0.015)	(0.013)	(0.011)	(0.009)	(0.009)	(0.008)
Squared Residual Dummy Variable Coefficient						
	-0.342	-0.147	-0.150	-0.135	-0.135	-0.134
	(0.011)	(0.007)	(0.004)	(0.003)	(0.003)	(0.003)
Estimate for log(Δ), assuming normal errors						
	-0.391	-0.308	-0.299	-0.272	-0.270	-0.235
	(0.020)	(0.016)	(0.012)	(0.010)	(0.010)	(0.008)
Demographic and Labor Supply Controls		x	x	x	x	x
State Dummies			x		x	
MSA Dummies				x	x	
PUMA Dummies						x

Note: All regressions are estimated on a sample of 1,046,003 observations. Demographic and labor supply controls include three education-level dummies, four year dummies, three weeks-worked dummies, quartics in age and usual weekly hours, and dummies for foreign born, citizen status, hispanic, black, gender, and dummies for married and widowed/divorced/separated also interacted with gender. PUMA is defined on the individuals place of work. There are 382 MSAs, and 982 PUMAs, identified in the data. Standard errors allow for clustering of errors at the PUMA level. For the estimated effect on log E(w|X), standard errors are calculated allowing cross-equation correlation in errors for each PUMA.

As argued in section III, expected log-wage differences may not be informative about the differential in expected wages (as defined in equation 1) if the error term is not independent of the covariates. Evidence of heteroskedasticity in the error terms related to the covariates, for example, would imply a lack of independence. The second row of results in Table 3 suggests that heteroskedasticity of this type does exist with respect to teacher occupational status. The reported OLS coefficient estimates are for the teacher coefficient in equation (3), where the squared residuals from the log-wage regression are themselves regressed on the teacher dummy and other covariates.[20] In all specifications in the table, there is a clear suggestion that the residual variance is lower for teachers compared to non-teachers. The third row of results provides an estimate of the exponent in equation (4), consistent if the error terms are normal and have a variance that follows equation (3).[21] It is clear that the correction for non-independence is substantial in all cases, making

[17] For example, OLS estimators applied to a linear model are QML estimators – they are ML if the error term in the model is normal, but are known to be consistent (under weaker assumptions) when the errors are nonnormal. With QML estimators, it is important that sandwich versions of standard errors be used, given the likely use of a misspecified likelihood function (see White, 1982).

[18] This is because inference using the delta-formula calculation of a standard error for $\widehat{e^{\gamma}}$ will have different outcomes based on whether the regression is specified with a teacher dummy or with a non-teacher dummy.

[19] For estimation of the teacher coefficient, the most important MSA-related controls would appear to be for "large" MSAs (over 1 million in population), and for residences outside MSA. If dummies only for those two characteristics are used in the model, the estimated coefficient (std. error) is -0.205 (0.011). As many data sets (such as the CPS) would not have PUMA indicators for workplace, but would have MSA indicators, it appears more appropriate to incorporate MSA controls in regressions than state or region.

[20] Note that the test in row (2) is from a Breusch-Pagan regression, which has somewhat more generality for testing for heteroskedastiticy than assuming that the variance is a linear function of the covariates (see Greene, 2012, p. 276).

[21] Standard errors for the estimates in row 3 are calculated allowing for cross-equation correlation in the log-wage and heteroskedasticity equation errors within clusters (where clusters are defined on the basis of the place-of-work PUMA.)

the estimated differential larger in magnitude. While its importance declines as the number of covariates are expanded, it still provides an estimate of the log differential equal to -0.24, compared to -0.17 from the simple log-wage regression (both with PUMA controls).

4.2. Quasi-maximum likelihood results

One limitation of the "corrected results" in Table 3 is that their consistency relies on the assumptions of normality in the error term in equation (2) and the functional form assumption for heteroskedasticity in the error expressed in equation (3). As discussed in section II, QML estimation of the underlying parameters relaxes these assumptions by directly estimating equation (5), where the level of earnings variable is an expressed as an exponential function of covariates. In Table 4, I report results using these alternative estimators, for specifications that use either MSA or place-of-work PUMA in controlling for geography differences between teachers and non-teachers.[22] For all three estimators – NLS, Poisson, or gamma – the use of PUMA rather than MSA dummies reduces the magnitude of the negative teacher (log) differential by about 4 log points. However, the choice of estimator is clearly important to the estimated size of the remaining differential. The NLS estimator provides the largest differential, while the gamma provides the smallest.

Two other aspects of the QML results in Table 4 should be noted. First, the gamma estimators are quite close to the corrected estimates in row (3) of Table 3. The fact that the gamma results would be similar to the log-wage results would not be surprising if the error terms in the log-wage regression were homoskedastic – Firth (1988) noted the similarity between these two estimators when the error term in equation (2) is homoskedastic.[23] The fact that the heteroskedasticity correction with log-wage estimates provides results similar to the gamma may suggest that, in these data, the assumptions necessary for this correction to be appropriate may come close to holding. Second, the difference across estimators does suggest that the regression model should be more closely examined, as all estimators should be consistent if the model is correctly specified. This suggests that a more careful consideration of the specification of the regression equation should be undertaken.

4.3. Interactive specifications

The specifications estimated in tables 3 and 4 may be overly restrictive in a number of ways. First, results are combined for men and women, when it is likely that many wage determinants (including teacher status) have different effects on wages across genders. Second, it is a natural extension to allow the wage-determinant coefficients to vary between teachers and non-teachers. I consider this latter extension, first, and then consider how results vary as I use separate samples for men and women.

As noted by Wooldridge (2010, pp.919-920), a parameter interpretable as an average treatment effect for the treated (ATT) can be estimated for linear and nonlinear models through a standard interactive specification. In particular, I extend the log-wage model in equation (2) to

$$\log(w_i) = \beta_1 T_i + \beta_2' X_i + \beta_3' T_i \left(X_i - \overline{X}_T\right) + \eta_i + u_i \tag{6}$$

where $\overline{X}_T$ is the sample means of the covariates among teachers, and $\eta_i$is a PUMA effect (which I do not interact with the teacher dummy). Using deviations from teacher means implies that β_1 can be interpreted as:

$$\beta_1 = E\left[\log(w)\middle|T=1, X=\overline{X}_T\right] - E\left[\log(w)\middle|T=0, X=\overline{X}_T\right],$$

that is, the difference in expected log wages between teachers and non-teachers for an individual with characteristics equal to those of the average teacher. There is still the retransformation problem, but that can be handled by either assuming u in equation (6) is normally distributed (and estimating a variance function) or by directly estimating the exponential regression by QML. In the latter, we now estimate the equation

$$w_i = e^{\gamma_1 T_i + \gamma_2' X_i + \gamma_3' T_i \left(X_i - \overline{X}_T\right) + \eta_i} + v_i$$

which, if correctly specified, allows the use of $\widehat{\Delta}_i = e^{\widehat{\gamma_1}}$ to provide a consistent estimate of the teacher differential for an individual with average teacher characteristics (that is, Δ_i when $X_i = \overline{X}_T$).[24]

Table 4.
QML estimates of teacher effects on log E(w|X).

	NLS		Poisson		Gamma	
	(1)	(2)	(1)	(2)	(1)	(2)
	-0.454	-0.407	-0.358	-0.316	-0.273	-0.237
	(0.017)	(0.012)	(0.013)	(0.010)	(0.010)	(0.008)
MSA Dummies	x		x		x	
PUMA Dummies		x		x		x

Note: See note to Table 3 for sample size and additional controls (all estimated models include all demographic and labor supply controls). Standard errors are calculated allowing for error clustering at the PUMA level.

Table 5.
Estimates of average teacher effects from models with teacher interactions.

Unadjusted Log Earnings Effect	Adjusted Effect from Log Earnings Model	NLS	Poisson	Gamma
Estimated Effect at Teacher Means for Covariates				
-0.159	-0.241	-0.253	-0.249	-0.246
(0.007)	(0.007)	(0.010)	(0.008)	(0.008)
Estimated Effect at Overall Means for Covariates				
-0.210	-0.290	-0.298	-0.294	-0.296
(0.007)	(0.007)	(0.010)	(0.008)	(0.007)
Estimated Effect at Teacher Means, Women Only				
-0.136	-0.202	-0.216	-0.206	-0.208
(0.007)	(0.007)	(0.009)	(0.008)	(0.007)
Estimated Effect at Teacher Means, Men Only				
-0.282	-0.375	-0.389	-0.386	-0.380
(0.008)	(0.008)	(0.011)	(0.009)	(0.008)

Note: See Table 3 for additional controls. All models include PUMA dummies; all controls are interacted with teacher dummies except PUMA dummies. The estimates for men are from models estimated with the sample of only men (n=500,804), with covariate means estimated only for men. Estimates for women (n=545,199) are constructed in a similar fashion. All estimates are the teacher effect on log E(w|X), except the first column is for E(log(w)|X).

[22] In results not reported here, I also estimated models that controlled for residential PUMA (rather than place-of-work). A substantially larger number of residential PUMAs are available (2,378, compared to 982 place-of-work PUMAs), but the impact of using residential PUMA on the teacher coefficient (and estimated log differential) is very similar to using place-of-work PUMA (estimated coefficients differ by about 0.003 log points).

[23] His primary conclusion was the relative efficiency (under reciprocal misspecification) was higher for the gamma, but that efficiency loss was small in either case.

[24] There is an issue in developing a standard error for $\widehat{\gamma}_1$, as the estimated models use sample means rather than population means. In my calculation of standard errors, I ignore sampling variation in the sample means; given the sample sizes being used here, this issue is highly unlikely to be important to any inferences drawn.

Estimates of models with teacher interactions evaluated at teacher means are reported in the first row of results in Table 5.[25] The first two columns are from log-wage models – one unadjusted and one adjusted for heteroskedasticity in the error term (using an interactive model as well). The results are similar to the non-interactive results in column (6) of Table 3, with the corrected result in Table 5 less than 0.01 log points higher in magnitude than in Table 3. Where the interactions are more important is in the comparison of QML estimators – unlike the large differences across estimators shown in Table 4, the interactive models provide results that are now quite similar. For example, the NLS estimate is less than 0.01 log points above the gamma estimate (with the Poisson in between). This similarity suggests that the interactive specification seems to largely handle the misspecification problem evident from the difference in estimates reported in Table 4. The QML estimates suggest a teacher log-differential of roughly -0.25 – considerably higher in magnitude than the uncorrected log differential of -0.16.

The second row of results in Table 5 are from a different specification that deviates covariates from means for the overall sample, rather than for teacher-specific means. Note that this alternative specification doesn't change any of the values of the coefficients other than the coefficient on the teacher dummy. If the interactive model is a correct specification, these could be interpreted as average treatment effects, rather than average treatment effects on the treated. The results are much larger in magnitude than the results at teacher means – suggesting that teachers have covariate values that tend to be associated with lower teacher differentials (compared to the overall population). It is again the case that the corrected log-wage/QML estimates are quite similar – and considerably above the uncorrected estimates.

Previous studies have also noticed a difference across genders in the level of teacher pay, or in teacher-pay differentials (see Allegretto and Mishel, 2018; Fox, et al., 2019).[26] As such, an additional interaction to consider is with gender. I also estimated separate models for men and women, with teacher interactions of all covariates (except PUMA effects) within each model. The bottom two rows of Table 5 report estimated log differentials (evaluated at gender-specific teacher means for the covariates) for women and for men. The results are consistent with Allegretto and Mishel, in that the teacher differential is estimated to be substantially larger in magnitude for men than for women – almost twice as large in most cases. Otherwise, results are largely similar to those using the pooled sample: (1) ignoring heteroskedasticity in the wage equation biases downward the magnitude of the differential; (2) adjusting the log model provides similar estimated differentials to the gamma estimator; and (3) the results across the QML estimators are similar for either given gender.

4.4. Matching estimates

While the rough similarity of the QML estimators is consistent with an appropriate specification when teacher interactions are included, it could still be the case that the underlying linear specification is too restrictive in modeling how wages are associated with the covariates. One possible alternative is to consider matching estimators for the teacher/non-teacher wage difference. As before, the focus is on the average treatment effect for the treated (teachers). Matching algorithms provide an estimate of the wage difference, rather than a wage differential, though I discuss how the estimated difference can be used to construct a relevant differential below. In particular, matching can be used to estimate:

$$d_i = E(w(1)|T_i = 1, X_i) - E(w(0)|T_i = 1, X_i)$$

where w(1) is the wage as a teacher and w(0) as a non-teacher. The first expectation can be directly estimated from the sample of teachers, so the challenge is providing an estimate of $E(w(0)|T_i = 1, X_i)$, the expected wage for teachers if they were not actually teachers.

As in Abadie and Imbens (2011), I use a nearest neighbor match (with replacement) for each sampled teacher to identify a non-teacher with characteristics as similar as possible to the teacher. I use the Mahalanobis metric to minimize the difference in the covariate vector between each teacher and the matched non-teacher. Given the importance placed on geographic location in this study, I only consider matched non-teachers who are in the same place-of-work PUMA as the teacher. This approach generally does not lead to exact matches on all covariates, so a wage regression estimated with the non-teacher sample is used to adjust the non-teacher's earnings to better reflect what would be expected if the covariates were the same.[27] Abadie and Imbens argue that this bias-adjustment term can provide a consistent estimator for the treatment effect even if the regression used for adjustment is misspecified.

I used nearest-neighbor matching to provide estimates of teacher earning differences across the pooled sample, and then separately for women and men. In general, only a small number of observations could not be matched (given the exact matching on PUMA).[28] As an assessment of the extent to which matching creates a more comparable sample across teachers and non-teachers, I present in Table 6 the standardized differences in covariate means between teachers and non-teachers for both (1) the overall sample of teachers and non-teachers (the "raw" difference) and (2) the sample of teachers and matched non-teachers ("matched" difference). There are obvious raw differences in several variables – education, and weeks worked in particular. Matching re-

Table 6.
Standardized differences in earnings determinants, before and after matching.

Earnings Determinant	All *Raw*	*Matched*	Women Only *Raw*	*Matched*	Men Only *Raw*	*Matched*
M.A.	0.697	0.111	0.659	0.125	0.744	0.170
Prof. Degree	-0.118	0.015	-0.118	0.016	-0.111	0.023
Ph.D.	-0.123	0.004	-0.145	0.003	-0.059	0.009
Native	-0.123	0.008	-0.119	0.010	-0.138	0.015
Black	-0.035	0.015	-0.097	0.018	0.030	0.021
Other Race	-0.002	0.003	-0.020	0.004	0.025	0.007
Hispanic	0.025	0.011	0.007	0.014	0.040	0.017
Age	0.097	0.052	0.141	0.062	0.068	0.044
Work Hours	0.077	0.200	0.284	0.256	-0.015	0.115
Married	0.180	-0.017	0.251	-0.029	0.146	-0.002
Wid./Div./Sep.	0.017	0.029	-0.031	0.038	-0.003	0.031
Citizen	0.124	-0.008	0.117	-0.013	0.140	-0.017
Female	0.511	0.045	–	–	–	–
Weeks Worked						
48-49	0.035	0.003	0.009	0.009	0.078	0.014
40-47	0.399	0.015	0.373	0.030	0.394	0.020
27-39	0.168	0.005	0.155	0.010	0.140	0.012

Note: All matches are exact based on the PUMA of the respondent. Four year dummies were also used in the matching.

[25] As I wish to directly compare the wages of teachers to non-teachers in the same geographic location, I do not include PUMA/teacher interactions in these models.

[26] Fox, et al. also show evidence that annual earnings from the ACS overstates the degree of gender difference in teacher pay, as men are more likely to work second/summer jobs than women. Controlling for weeks worked may handle this concern somewhat in the current analysis.

[27] Let w_{NT} be the wage of the matched non-teacher who has characteristics X_{NT}. Then $E(w(0)|T_i = 1, X_i)$in equation (6) is estimated as $w_{NT} + b'_{NT}(X_i - X_{NT})$ where b_{NT} is the OLS coefficient-vector estimate from a regression of the wage on X using the matched non-teaching sample.

[28] Robust standard errors suggested by Abadie and Imbens (2011) are used; these require two matches for each teacher, though only the closest match is used in constructing the counterfactual non-teaching expectation.

duces considerably the differences in most variables; for the overall sample, one worrisome difference is that the teacher sample is still more likely to have an M.A. than the non-teacher sample (which is more likely to have an undergraduate degree). Another concern may be the matching on work hours, especially for women (though, as with education, we end up with a difference that unadjusted would lead to an upward biased estimate of the earnings difference between teachers and non-teachers).

The estimated average treatment effects for teachers are reported in Table 7. For comparison purposes, panel A reports the raw difference in average earnings between teachers and non-teachers, and the estimated teacher pay differential (that is, $(\overline{w}_T - \overline{w}_{NT})/\overline{w}_{NT}$). The matched earnings difference with exact PUMA matches is reported in panel B, and all estimated differences are highly statistically significant. The estimated earnings difference is considerably larger in magnitude for men (-$33,091) than for women (-$16,152). Reported below these numbers is the implied estimated differential, and this is also larger in magnitude for men (-0.29) than for women (-0.21).[29] Using MSA rather than PUMA as the geographic control (again forcing exact matches on geography) leads to even higher differences in earnings (reported in panel C) – as before, it is useful to use finer geographic distinctions in estimating teacher wage differentials. Of more importance is the examination of earnings levels rather than earnings in logs: matching estimates using log earnings as the outcome variable suggest considerably smaller differentials than those estimated in levels.[30]

The last panel in Table 7 reports selected coefficients from Table 5 in exponentiated form. These are estimates from a modelling of treatment effects through an interactive specification. The matched log earnings differentials are similar to the differentials suggested by the unadjusted log earnings regression model (column 1 in Table 5), with the biggest difference for men. Matching leads to a higher magnitude for the differentials than the QML estimators in Table 5 for all workers and for women only, but a lower estimated differential for men (the gamma is reported in Table 7). These differences are not large in magnitude, however (roughly 2 percentage points in general).

Table 7.
Matching estimators of earnings differences: average treatment effect for teachers.

	All	Women Only	Men Only
A. Unmatched Differences			
Earnings Levels	-28409.0	-13508.1	-38800.9
	(111.5)	(117.1)	(213.7)
Unmatched Differential	-0.338	-0.202	-0.384
B. Matching with Exact PUMA Matches			
Difference in Earnings Levels after Bias Adjustment	-20713.3	-16151.9	-33091.0
	(209.7)	(210.3)	(496.3)
Matched Differential	-0.237	-0.207	-0.293
C. Matching with Exact MSA Matches			
Difference in Earnings Levels after Bias Adjustment	-22321.0	-17689.2	-35230.9
	(208.2)	(217.3)	(483.9)
Matched Differential	-0.249	-0.220	-0.303
D. Matching with Exact PUMA Matches, Log Earnings			
Difference in Log Earnings after Bias Adjustment	-0.163	-0.135	-0.246
	(0.002)	(0.003)	(0.004)
Matched Differential	-0.150	-0.126	-0.218
E. Differentials from Estimates in Table 5			
Unadjusted Log Earnings Differential	-0.147	-0.127	-0.246
QML Gamma Differential	-0.218	-0.188	-0.316

Note: Matching is nearest neighbor based on the Mahalanobis distance using characteristics listed in Table 6 (with exact match on either PUMA or state), with a bias adjustment based on the same characteristics. Standard errors are robust to heteroskedasticity. Differentials for the top 3 panels are calculated using the treatment effect divided by the average non-teacher earnings prediction for the matched teacher sample. Differentials from Table 5 estimates use the formula $e^{\widehat{\beta}_{teach}} - 1$.

[29] Let the estimated average treatment effect for teachers be $\widehat{d}$. An estimated teacher differential from the matching estimators is calculated using $\widehat{\Delta} = \widehat{d} / (\overline{w}_T - \widehat{d})$, where $\overline{w}_T$ is the average earnings for teachers who are in the sample with matches.

[30] Here, the matched differential is calculated as $e^{\widehat{d}} - 1$, where $\widehat{d}$ is now the estimated difference in log earnings. This exponentiated coefficient estimate formuala is also used to form the unadjusted differentials reported in panel E.

5. Additional estimations and other issues

5.1. Benefits differences

While the comparisons in this paper focus on the wage compensation of workers, there may be other factors associated with the teaching occupation that make it more or less attractive than work in the private sector. For example, Allegretto and Mishel (2018) note that non-wage benefits tend to be higher for teachers than for non-teachers, and that adjusting for nonwage compensation would lower the differential by as much as 6.5 percentage points in 2015. The ACS does not provide sufficient information to allow such a comparison, but it does provide information on whether or not the individual has health insurance through their current or former employer or union. Linear probability models in which a dummy variable for employment-provided health insurance is the dependent variable are reported in panel A of Table 8, and the results are consistent with a roughly 6 percentage-point increase in the probability of health insurance for teachers compared to non-teachers (or about a 7 percent higher probability, relative to the 90 percent average insurance coverage rate of non-teachers). These results might suggest

Table 8.
Estimated teacher effects with alternative outcomes and specifications.

Specification	All Workers *No Controls*	All Workers *Full Controls*	Women Only *Full Controls*	Men Only *Full Controls*
A. Linear Probability Model for Health Insurance from Employer ($\overline{Y} = 0.90$)				
Teacher	0.061	0.059	0.051	0.075
	(0.002)	(0.002)	(0.002)	(0.003)
B. Linear Model for Commuting Time to Work (in minutes) ($\overline{Y} = 28.2$)				
Teacher	-7.55	-5.53	-5.49	-5.62
	(0.62)	(0.16)	(0.17)	(0.20)
C. Wage Models with Select Non-Teaching Occupations				
Teacher	-0.222	-0.219	-0.219	-0.193
	(0.018)	(0.008)	(0.007)	(0.011)
D. Wage Models with Education-Major Controls				
Teacher	-0.391	-0.292	-0.260	-0.374
	(0.021)	(0.009)	(0.009)	(0.010)
Teacher*(Education Major)	0.378	0.324	0.319	0.320
	(0.013)	(0.006)	(0.007)	(0.010)
Education Major	-0.436	-0.271	-0.264	-0.273
	(0.013)	(0.004)	(0.005)	(0.007)
E. Wage Models with State-Level Wage-Bargaining Law Interactions				
Teacher	-0.362	-0.214	-0.176	-0.313
	(0.022)	(0.008)	(0.008)	(0.009)
Teacher*(No Statues on Wage Bargaining)	-0.187	-0.103	-0.105	-0.086
	(0.019)	(0.028)	(0.028)	(0.031)
Teacher*(Wage Bargaining Illegal)	-0.160	-0.114	-0.107	-0.119
	(0.017)	(0.016)	(0.015)	(0.020)
F. Wage Models with Adjusted Hours for Over-Reporting				
Teacher	-0.412	-0.184	-0.145	-0.283
	(0.021)	(0.007)	(0.007)	(0.008)

Note: For panels A and B, “full controls” include all controls noted in Table 3, including PUMA dummies. For remaining panels, estimations incorporate teacher-dummy interactions (as in Table 5) deviated from teacher-level means (except for the education-major/teacher interaction in panel D). Panels C through F are gamma QML estimates. See text for “non-teaching occupations” included in panel C. Standard errors are calculated allowing for error clustering at the PUMA level.

that a percentage-point downward adjustment in the implied teacher/non-teacher wage differential is appropriate.[31] This comparison is limited, however, as it doesn't address the potential for health insurance provided to teachers to be more generous than that provided to non-teachers.

The benefits comparisons reported in Allegretto and Mishel (2018) are based on reports on employer costs devoted to non-wage compensation as reported in Bureau of Labor Statistics (2021). Taking averages over the starting and ending points of the 2012-2016 period examined here, the reported difference in benefits suggests a benefit cost of $2.62 per hour more for teachers than non-teachers.[32] Based on average annual earnings of non-teachers (reported in Table 2), this benefit difference would be roughly 5 percent of total compensation for a full-time year-round worker.[33] This might also suggest a rough correction adding about 5 percentage points to the estimated teacher wage differential reported in this paper (making the teacher pay differential smaller in magnitude). This is "rough" at least partly because there is no ability to control for any important differences (including geography).

An additional caveat arises in valuing the teacher/non-teacher benefit differential, in that a substantial portion of the benefit difference is in retirement benefits. Fitzpatrick (2015) used data on a voluntary retirement contribution among public-school employees in Illinois to estimate the value of retirement benefits to these employees. Her results suggested that retirement benefits were valued by employees at about 20 percent of the cost to the employer. Given that retirement benefits represent almost 40 percent of the total benefits received by teachers, adjusting the compensation value to teachers by this amount would reduce the benefit advantage to about 3.5 percent.[34]

5.2. *Compensating differentials*

Another possible advantage to the teaching occupation is that work locations can often be located more closely to the individual's residential location, given that schools are spread throughout communities so as to be close to the students' homes. The ACS does collect information on travel time to work for the primary job, which I use as a dependent variable in a linear model with a teacher dummy and other characteristics. The results are reported in panel B of Table 8, and are consistent with a shorter commute for teachers compared to non-teachers, with an estimated 6-minute shorter time-to-work value (or a roughly 20 percent lower commute time compared to the average time of 29 minutes for non-teachers). The value of this shorter commute time, however, would likely account for only a small part of the estimated differential – using average hourly wages of teachers would suggest that this shorter commute is worth about $1300 per year – or about 5 percent of the raw difference in annual earnings between non-teachers and teachers.[35]

A related concern has to do with whether all non-teaching occupations for college graduates are comparable in skill expectations and work environment to the teaching occupation. Following Taylor (2008), I reduced the non-teacher sample to a set of 16 occupations considered to be more comparable to the teaching occupation, and then re-estimated the wage models (with teacher interactions).[36] Gamma QML estimates of these models are reported in panel C of Table 8, and do suggest a slightly smaller (in magnitude) overall teacher wage differential when using this alternative control group (a 20 percent differential, compared to 22 percent in the comparable Table 5 results). It is interesting that all of this reduction has to do with a substantial decline in the estimated differential for men (from a 32 percent differential in Table 5 to an 18 percent differential with this alternative control sample). In fact, in these estimations there no longer appears to a difference in the differential between men and women.

Another concern might be that teachers tend to have an educational background that focuses on the teaching occupation, and that this particular skill may not be valued highly in the general labor market. For the years studied here, the ACS does include a "field of degree" variable for those with a bachelor's degree, and this allows me to identify individuals who state their primary major as being in one of several education-related fields.[37] I then re-estimate the earnings models of Table 5 (with teacher interactions evaluated at teacher means), also including an education major dummy and an interaction between teacher and education major. This latter effect is not deviated from means, so the teacher main effect reflects the coefficient for teachers who are not education majors, and the education-major effect is for education majors who are not teachers. The results are reported in panel D of Table 8, and lead to an interesting set of conclusions. Having an education major is associated with a lower earnings level, and being a teacher is associated with lower earnings as well, but teachers who were education majors actually have slightly higher earnings than teachers who had some other major. With all controls included, teachers without an education degree have an estimated differential of approximately -25 percent when compared to non-teachers without an education degree, while teachers with an education degree have a differential of -21 percent when compared to non-teachers without an education degree. (This difference in estimated differentials is statistically significant at any conventional level). In any case, it does not appear that the negative teacher earnings differential is simply a reflection of lower earnings for individuals with an education major.

There may be other dimensions in which the teaching profession is seen differently by teachers when compared to possible work in non-teaching occupations. For example, there may be an intrinsic value to being a teacher, to the extent that teachers value the influence they are able to have on their students' development. This would be hard to measure directly, but it could explain why teachers are willing to work at a job where they may be paid less than they would expect if they held a non-teaching job.

5.3. *Teacher unions*

Collective bargaining in the teacher labor market is more prevalent than among college-educated non-teachers – the numbers for collective bargaining reported in Hirsch et al. (2012) are consistent with 73

[31] This is based on employer costs for health insurance being about 10 percent of total compensation on average (see Bureau of Labor Statistics, 2021).

[32] I omit "required legal benefits" (mostly tax contributions), assuming these are not valued highly by workers. Required legal benefits are slightly higher for non-teachers, given their higher average pay.

[33] This calculation uses the fact that total benefit costs are $12.98/hour for non-teachers, which is added to an hourly earnings of $38.57 for non-teachers implied by the averages reported in Table 2. The annual benefit difference of $2.62 is 5.1 percent of the total hourly compensation suggested for teachers.

[34] The other two major sources of benefits reported in the data are insurance (which is mostly health insurance) and paid leave (for which non-teachers have an advantage). The difference in retirement-benefit hourly costs ($3.63) is actually larger than the overall benefit difference of $2.62 – if one were to value this difference at only 20 percent of its value, the benefit differential would be reduced essentially to 0.

[35] Average hourly earnings for teachers in the ACS is roughly $26, and the shorter commute would save about an hour per week for teachers. Over 50 weeks in the year, this would be worth about $1300, about 1.5 percent of the average non-teacher salary.

[36] The particular occupations are: accountants and auditors; insurance underwriters; human resources, training, and labor relations specialists; inspectors, testers, sorters, samplers, and weighers; agricultural inspectors; architects, except naval; conservation scientists, and foresters; registered nurses; occupational therapists; physical therapists; archivists, curators, and museum technicians; clergy; technical writers; editors; news analysts, reporters, and correspondents; and computer programmers.

[37] Roughly 60 percent of teachers are classified with education as their primary degree.

percent of teachers represented by a collective bargaining agreement, for instance. The meta-analysis of Merkle and Phillips (2017), however, suggests a mild effect of teacher unions on wages (a premium between 2 and 4.5 percent), and the higher prevalence of unions for teachers compared to non-teachers implies that heavier unionization would work against explaining the negative estimated teacher earnings differential. The ACS does not collect information on union representation of individual workers, so it is not possible to estimate earnings models with individual union controls. However, it is possible to consider whether the strength of state-level collective-bargaining laws for teachers affects the teacher earnings differential.

To measure collective-bargaining strength, I use the measures provided in Sanes and Schmitt (2014). For the purposes of wage bargaining, they identify states in which wage bargaining for teachers is explicitly illegal (there are five such states), and states in which there is no statue allowing wage bargaining (seven states). All other states have wage bargaining for teachers as legal. To estimate the relevance of these laws to estimated teacher earnings differentials, I re-estimated equations like in Table 5, but with the teacher effect interacted with a dummy for bargaining being legal, and with a dummy for bargaining being without statutory defense. The results are reported in panel E of Table 8, and are consistent with the teacher coefficient being considerably larger in magnitude (by about 0.1 log points) in states without protections for wage bargaining. It does not seem to matter whether it is explicitly illegal or just not protected (the two interactions' coefficients are not statistically significantly different), and there is still a considerable negative teacher earnings differential in states with legal protection for wage bargaining.[38]

5.4. Labor supply differences between teachers and non-teachers

Previous research has suggested that teacher-pay comparisons may be biased towards understating relative teachers' earnings due to possible misreporting of work hours in standard population surveys. Podgursky and Tongrut (2006) criticize the use of directly-reported weekly wages in surveys like the Current Population Survey, as they suggest that teachers who work part-year are prone nonetheless to estimate weekly wages as annual earnings divided by 52 weeks. The current study uses reports on annual earnings as the pay variable, but controls for the part-year status of teachers with weeks-worked controls – if weeks worked are reportedly accurately, then controls for the weeks-worked categories should address potential labor-supply differences through the year.[39] As reflected in Table 2, teachers are considerably less likely to report working full year than non-teachers.

An additional bias could come from a tendency for teachers to misreport hours of work per week in survey data. West (2014) used the American Time Use Survey component of the CPS from 2003-2010 to study this issue, finding that teachers have a greater tendency than non-teachers to over-report actual hours worked (as determined by a 24-hour diary). Both teachers and non-teachers tend to over-report their hours in the CPS "usual hours" question – during the regular school-year months, non-teachers are estimated to over-report their hours by 2.8 hours per week, while teachers over-report by 4.8 hours (compared to the "restrictive" definition for diary work hours). The CPS "usual hours" question relates to the job held at the time of they survey, and it is difficult to know the implications for the retrospective question on "usual hours per week" in the preceding year that is asked in the ACS. In the CPS data, "usual hours" is similar for teachers in the school-year and summer months, and the biggest source of over-reporting of hours has to do with over-reporting by teachers in the summer (when their diary work hours fall). It would appear that some teachers report typical work hours through the year in those months, rather than the reduced hours they record as work-related in their diaries in those months. To the extent that the ACS weeks-worked question handles reduced work activity in the summer months, this may be less of an issue in the present analysis.

Over-reporting of work hours is still a potentially serious issue. To attempt a rough adjustment to the ACS data to assess this importance, I re-estimated the models reported in Table 5 (evaluated at the mean characteristics for teachers) adjusting reported "usual work hours" downward for each teacher and non-teacher according to the average over-reporting noted in the previous paragraph. The gamma QML results of this re-estimation are reported in panel F of Table 8, and show that the overall log differential would fall in magnitude by about 0.06 log points if this adjustment were appropriate.[40] The reduction is particularly significant for the male differential (a fall of roughly 0.11 log points). It is difficult to assess implications of this result for the general literature estimating teaching-related earnings differentials, as diary hours are not generally available in data sets that would be used. For the present case, the large differentials reported may be somewhat overstated, but there is still evidence of a significantly lower average pay for teachers compared to observationally comparable non-teachers.

5.5. Sampling error and the percentage-difference formula

A number of studies have pointed to a bias in the use of the exponentiated log-wage coefficient estimate on a dummy variable to estimate the true percentage differential. As noted in section III, the estimator $\widehat{\Delta} = e^{\widehat{\gamma}} - 1$provides a consistent estimator for $\Delta = e^{\gamma} - 1$ if $\widehat{\gamma}$ is consistent for γ, but $\widehat{\Delta}$ is not an unbiased estimator for Δ. Kennedy (1981) suggested the use of a simple corrected estimator

$$\widetilde{\Delta} = e^{\widehat{\gamma} - 0.5 * \widehat{V}(\widehat{\gamma})} - 1$$

which Giles (1982) has argued is approximately unbiased, especially in large samples (here, $\widehat{V}(\widehat{\gamma})$is the estimated variance of $\widehat{\gamma}$). This bias issue could potentially be relevant for all of the estimators presented (including the QML estimators). However, given the small standard errors for the coefficient estimates, the recommended adjustment by Kennedy would be quite minor. Even for the least precisely estimated coefficient in Table 5 (the NLS estimator for men) the adjustment would suggest that the implied differential of -0.322 should be lowered in magnitude by 0.000041 percentage points.

5.6. Summary of corrections to the raw log-wage differential

Several suggested corrections to a simple comparison of (log) wages between teachers and non-teachers have been suggested in this paper, so I provide Table 9 as a summary of the estimated impact of each of these corrections. The table begins with the exponentiated log-wage difference, providing a raw differential of -20 percent. The impact of sequentially changing the set of controls or estimated method is noted in

[38] Using a different classification of state collective bargaining laws, Brunner and Ju (2019) find a similar impact of mandatory collective-bargaining laws on teacher wages, also using the ACS data.

[39] Use of annual earnings does imply that individuals who work only part of the year as teachers (say, taking a non-teaching summer job) may still report being a full-year worker, with earnings reported including pay from the non-teaching job. This may be a limitation of using annual earnings, or it may be interpreted as a reflection of the average impact on one's annual earnings from having teaching as a primary occupation.

[40] An alternative would be to adjust work hours downward for average over-reporting in West (2014) for the full year (including summer months), noting this this could be an over-adjustment for workers in the ACS who do not consider themselves as working in those weeks by ACS (and so would not be in the CPS sample for those weeks). This alternative adjustment would lower the estimated log differential (std. error) to -0.128 (0.007), or a fall in magnitude of 0.11 log points.

Table 9.
Suggested adjustments to estimated log differentials.

	Table/Text Source	*Estimated Log Diff.*	*Δin Est. Diff.*
Raw Log Differential (Table 3, Spec. (1))		-0.198	
Adjustment			
Add Usual Controls (including State Dummies)	Table 3 (col. 3)		-0.003
Add PUMA dummies	Table 3 (col. 6)		+0.045
Exponential instead of Log Regression	Table 4 (gamma col. 2)		-0.056
Interactive Specification	Table 5 (row 1, gamma)		-0.007
Correction for Benefits Difference	text section 5.1		+0.05
Reduced Commuting Time	Table 8 (row B) and section 5.2.		+0.015
Hours Over-Reporting	Table 8 (row F)		+0.050
Log Differential After Suggested Corrections		-0.103	

Note: Reported corrections are for all workers combined; results using exponential regression (Tables 4, 5, and 8) are for models estimated by gamma QML.

the final column of the table (note that a negative change leads to a differential farther from zero).[41] The choice to directly estimate the exponential function has a slightly larger impact on the differential than the addition of more-detailed PUMA controls, but the resulting differential after these changes (leading to the "iterative specification" row) is only slightly larger in magnitude (-21 percent) than the original raw differential.

The table also provides three additional adjustments based on concerns related to benefits, commuting time, and hours mismeasurement. While these are more speculative, they do appear to substantially reduce the suggested teacher differential. When added together the result is an implied differential of roughly -10 percent. This could be even smaller in magnitude if one thought that the restriction to a reduced set of occupations were more appropriate (which lowers the size of the overall differential by about 2 percentage points), or larger if one thought the value of retirement benefits should be reduced below their cost to employers (raising the size by 1.5 percentage points or more). Nonetheless, there is still the suggestion that teachers are paid less than observationally-similar private-sector workers by at least 10 percent.

6. Results from the 2007-2011 ACS

Allegreto and Mishel (2018) have pointed to a tendency for the negative teacher-pay differential to have increased in magnitude over time. Reasonably large samples from the ACS only started in the mid-2000s, so it is not possible to perform a consistent analysis with the ACS for as long a period as Allegreto and Mishel do with the ORG samples (which go back to 1979). However, it is possible to consider whether the differential was smaller in magnitude in the five-year period prior to the 2012-2016 period analyzed above. A sample with the same restrictions was taken from the American Community Surveys from 2007 to 2011, and the models discussed above were also estimated with these data. A selected set of results are reported in Table 10.[42]

The comparison across specifications reported in Table 10 is similar to what was found with the 2012-2016 ACS. Simply using a log-earnings regression provides a considerably lower teacher differential (of -11 percent) compared to the estimated differential of -17 percent when log-earnings variance differences are taken into account. Estimation using interactive specifications (evaluated at mean characteristics for teachers) provides a very similar estimate, while matching estimates provide a slightly higher estimated differential of -20 percent. A roughly similar pattern of results is suggested for both genders when estimation is done separately for men and women.

The results are consistent with an increase in the magnitude of the differential from 2007-2011 to 2012-2016. For example, the gamma QML estimator with the interactive specification suggests an increase in the magnitude of the estimated differential from -17 percent to -22 percent for all workers, from -14 percent to -19 percent for women, and from -27 percent to -32 percent for men. A similarly-sized increase over this period was noted by Allegretto and Mishel (2018).

Table 10.
Teacher pay differential estimates, 2007-2011 American Community Survey.

	All	**Women Only**	**Men Only**
Raw Differential	-0.345	-0.213	-0.437
Log-Earnings Regression Based Estimates			
Log-Earnings Coefficient	-0.118	-0.083	-0.200
	(0.006)	(0.006)	(0.007)
	[-0.111]	[-0.080]	[-0.181]
Coefficient Corrected for Variance Differences	-0.180	-0.143	-0.277
	(0.006)	(0.006)	(0.007)
	[-0.165]	[-0.133]	[-0.242]
Interactive Specifications			
Corrected Log-Based Differential	-0.178	-0.140	-0.285
	(0.006)	(0.006)	(0.007)
	[-0.163]	[-0.131]	[-0.248]
QML – Normal	-0.186	-0.151	-0.317
	(0.008)	(0.007)	(0.009)
	[-0.170]	[-0.140]	[-0.272]
QML – Gamma	-0.183	-0.145	-0.291
	(0.006	(0.006)	(0.008)
	[-0.167]	[-0.135]	[-0.253]
Matching Estimates – Levels	-14887	-10987	-26376
	(178)	(173)	(418)
	[-0.197]	[-0.163]	[-0.265]

Note: Numbers in parentheses are standard errors. All specifications include place-of-work PUMA dummies (with exact matching on PUMA for the matching estimates), and the other controls noted in earlier tables. Numbers in brackets are the implied teacher pay differential based on the reported coefficient (or function of coefficients). Matching estimates use a linear regression bias-adjustment correction. The main sample size is 1,028,312 (523,694 female; 504,573 male).

7. Conclusion

I use ACS data for the years 2007-2016 to estimate wage models that allow for the comparison of the earnings of teachers to those of observationally-comparable non-teachers, and find estimates of an overall negative estimated differential in which teachers are paid roughly 20 percent less than non-teachers over this period (with the magnitude of the differential growing over time). While the ACS has its limitations, it has advantages that allow for particular consideration of two issues related to the estimation of teacher differentials. For one, its large sample size allows for the incorporation of very detailed controls for labor-market location – using place-of-work geographic indicators leads to a reduction in the magnitude of the overall differential by about 5 percentage points relative to a less-detailed set of geographic controls based on state of residence. The large sample size also allows for a precise consideration of issues related to the use of log earnings models to estimate differentials. In this case, the use of log earnings to estimate differentials in expected earnings for teachers (compared to non-teachers) leads to a downward bias in the magnitude of the estimated differential. More appropriate estimation methods increase the magnitude of the estimated differential by about 6 percentage points. The net result, then, of the consideration of these two issues is a slight increase in

[41] While the size of these estimated corrections does depend on the order in which the changes are made, in practice the sizes of the regression-method or geographic-control impacts are largely invariant to the order in which they are applied.

[42] No results are reported in Table 8 that allow for the direct comparison of using MSA rather than PUMA controls. In practice, the impact of using the more detailed geographic controls is quite similar to results with the 2012-2016 data.

the magnitude of the estimated differential.

There are other issues in estimating wage differentials for which the ACS is less well-suited to address. For example, other evidence suggests that benefits may be more commonly provided for teachers, and my results on employer-provided health insurance are consistent with this. If benefits are valued by workers as equivalent to wage income, then the differential is likely overstated – the results reported in Allegretto and Mishel (2018), and my own calculations, are consistent with the negative differential being overstated by roughly 5 percentage points over the period I examine. West (2014) has reported concerns with CPS work-hour measures for teachers – a simple approach to addressing this concern suggests that the differential could be overstated by 5 percentage points (or more) if this is relevant to the ACS data. Addressing both of these concerns would likely lead to a much smaller differential than the 22 percent underpayment suggested in the ACS data. Nonetheless, my results would still be consistent with a differential on the order of -10 percent.

A consistent finding in the estimation of U.S teacher pay differentials is the tendency for the estimated differential between teachers and non-teachers to be considerably smaller in magnitude for women than for men. While there is variation depending on estimation method, the wage differential is generally larger in magnitude by roughly 10 percentage points for women.[43] This could be indicative of a lower degree of gender discrimination in the public sector – another way of stating the result is that the gender differential is smaller among public-school teachers than among observationally-similar private-sector workers. It is also relevant in this consideration that the one adjustment that has substantially different impacts on the estimated teacher differentials across gender is a restriction of the private-sector comparison group to only a subset of occupations considered most similar to teaching. This adjustment has only a small impact on the estimated differential for women, but makes the estimated male differential slightly smaller in magnitude than that for women. This is consistent with the suggestion of Blau and Khan (2017) that occupational access still plays a key role in explaining the gender wage gap.

Declaration of Competing Interest

None

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.econedurev.2021.102169.

References

Abadie, Alberto, & Imbens, Guido W. (2011). Bias-Corrected Matching Estimators for Average Treatment Effects. *Journal of Business and Economic Statistics, 29*(1), 1–11.

Ai, Chunrong, & Norton, Edward C. (2000). Standard Errors for the Retransformation Problem with Heteroscedasticity. *Journal of Health Economics, 19*(4), 697–718.

Aldeman, Chad. (2019). We Should Probably Stop Citing EPI's "Teacher Wage Gap" Data. *Eduwonk Oct. 7. Accessed at*. https://www.eduwonk.com/2019/10/we-should-probably-stop-citing-epis-teacher-wage-gap-data.html.

Allegretto, Sylvia, & Mishel, Lawrence (2018). *The Teacher Pay Penalty Has Hit a New HIgh: Trends in the Teacher Wage and Compensation Gaps Through 2017. Economic Policy Institute Report, September 5*.

Benner, Meg, Roth, Erin, Johnson, Stephenie, & Bahn, Kate (2018). How to Give Teachers a $10,000 Raise. *Center for American Progress* Accessed at https://www.americanprogress.org/issues/education-k-12/reports/2018/07/13/453102/give-teachers-10000-raise/.

Biggs, Andrew G. (2019). The Truth about Teacher Pay. *AEI*. Sept. 20. Accessed at https://www.aei.org/articles/the-truth-about-teacher-pay/.

Blackburn, McKinley L. (2007). Estimating Wage Differentials without Logarithms. *Labour Economics, 14*, 73–98.

Blau, Francine D., & Kahn, Lawrence M. (2017). The Gender Wage Gap: Extent, Trends, and Explanations. *Journal of Economic Literature, 55*(3), 789–865.

Britton, Jack, & Propper, Carol (2016). Teacher Pay and School Productivity: Exploiting Wage Regulation. *Journal of Public Economics, 133*, 75–89.

Bureau of Labor Statistics. (2021). *Employer Costs for Employee Compensation Historical Listing: National Compensation Costs, March 2004-March 2021*.

Brunner, Eric. J., & Ju, Andrew (2019). State Collective Bargaining Laws and Public-Sector Pay. *ILR Review, 72*(2), 480–508.

Campolieti, Michele (2018). Matching and Inverse Probability Weighting Estimates of the Union Wage Premium: Evidence from Canada, 1997-2014. *Industrial Relations, 57*(1), 101–130.

Congressional Budget Office. (2017). Comparing the Compensation of Federal and Private-Sector Employees, 2011 to 2015. *CBO*. April.

Corcoran, Sean P., Evans, William N., & Schwab, Robert M. (2004). Changing Labor Market Opportunities for Women and the Quality of Teachers, 1957-2000. *American Economic Review, 94*(2), 230–235.

Duan, Naihua (1983). Smearing Estimate: A Nonparametric Retransformation Method. *Journal of the American Statistical Association, 78*(3), 605–610.

EdSource. 2020. "On Education: What Democratic presidential candidates are promising" February 24. Accesed at https://edsource.org/2020/on-education-what-democratic-presidential-candidates-are-promising.

Eren, Ozkan (2007). Measuring the Union-Nonunion Wage Gap Using Propensity Score Matching. *Industrial Relations, 46*(4), 766–780.

Falk, Justin R. (2015). Comparing Federal and Private-Sector Wages without Logs. *Contemporary Economic Policy, 33*(1), 176–189.

Firth, David (1988). Multiplicative Errors: Log-Normal or Gamma? *Journal of the Royal Statistical Society: Series B, 50*(2), 266–268.

Fitzpatrick, Maria. (2015). How Much Are Public School Teachers Willing to Pay for Their Retirement Benefits? *American Economic Journal: Economic Policy, 7*(4), 165–188.

Fox, Daniel, Gmeiner, Michael, & Price, Joseph (2019). The Gender Gap in K-12 Educator Salaries. *Economics of Education Review, 68*, 23–26.

Giles, David E. (1982). The Interpretation of Dummy Variables in Semilogarithmic Equations. *Economics Letters, 10*(1-2), 77–79.

Goldhaber, Dan, & Walch, Joe (2014). Gains in Teacher Quality: Academic Capabilities of the U.S. Teaching Force are on the Rise. *Education Next, 14*(1).

Greene, William H. (2012). *Econometric Analysis*. Boston: Prentice-Hall. Seventh Edition.

Gittleman, Maury, & Pierce, Brooks (2012). Compensation for State and Local Government Workers. *Journal of Economic Perspectives, 26*(1), 217–242.

Hanushek, Eric A. (2007). The Single Salary Schedule and Other Issues of Teacher Pay. *Peabody Journal of Education, 82*(4), 574–586.

Hanushek, Eric A., Marc Piopiunik and Simon Wiederhold. 2019. "Do Smarter Teachers Make Smarter Students?" *Education Next*, Spring.

Hendricks, Matthew. (2014). Does it Pay to Pay Teachers More? Evidence from Texas. *Journal of Public Economics, 109*, 50–63.

Hirsch, Barry T., David A. Macpherson, and John V. Winters. 2012. "Teacher Salaries, State Collective Bargaining Laws, and Union Coverage." Unpublished manuscript.

Hirsch, Barry T., & Schumacher, Edward J. (2004). Match Bias in Wage Gap Estimates Due to Earnings Imputation. *Journal of Labor Economics, 22*(3), 689–722.

Hirsch, Barry T., & Schumacher, Edward J. (2012). Underpaid or Overpaid? Wage Analysis for Nurses Using Job and Worker Attributes. *Southern Economic Journal, 78*(4), 1096–1119.

Kennedy, Peter E. (1981). Estimation with Correctly Interpreted Dummy Variables in Semilogarithmic Equations. *American Economic Review, 71*(4), 801.

Manning, Willard G. (1998). The Logged Dependent Variable, Heteroscedasticity, and the Retransformation Problem. *Journal of Health Economics, 17*(3), 283–295.

McCullagh, Peter, & Nelder, John A. (1989). *Generalized Linear Models* (Second Edition). Boca Raton: Chapman and Hall.

Meckler, Laura (2018). 2 in 3 Americans Say Public School Teachers are Underpaid. *The Washington Post*. August 27.

Merkle, Jessica S., & Phillips, Michelle Andrea (2017). The Wage Impact of Teachers Unions: A Meta-Analysis. *Contemporary Economic Policy, 36*(1), 93–115.

Moretti, Enrico (2013). Real Wage Inequality. *American Economic Journal: Applied Economics, 5*(1), 65–103.

Nittler, Kency. 2019. "The Ins and Outs of Teacher Salaries." *National Council on Teacher Quality*, May 30. Accessed at https://www.nctq.org/blog/The-ins-and-outs-of-teacher-salaries .

OECD. (2019). *Education at a Glance 2019*. Paris: OECD Publishing.

Podgursky, Michael, & Tongrut, Ruttaya (2006). (Mis-)Measuring the Relative Pay of Public School Teachers. *Education Finance and Policy, 1*(4), 425–440.

Ransom, Michael R., & Sims, David P. (2010). Estimating the Firm's Labor Supply Curve in a 'New Monopsony' Framework: Schoolteachers in Missouri. *Journal of Labor Economics, 28*(2), 331–355.

Rickman, Dan S., Wang, Hongbo, & Winters, John V. (2017). Adjusting State Public School Teacher Salaries for Interstate Comparison. *Public Finance Review, 47*(1), 142–169.

Sanes, Milla and John Schmitt. 2014. "Regulation of Public Sector Collective Bargaining in the States." Center for Economic and Policy Research, Washington D.C.

Strauss, Valerie. (2017). Teachers in U.S. Paid Far Less than Similarly Educated Professional, Report Finds. *Washington Post*. September 14.

Taylor, Lori L. (2008). Comparing Teacher Salaries: Insights from the U.S. Census. *Economics of Education Review, 27*, 48–57.

Stoddard, Christiana (2005). Adjusting Teacher Salaries for the Cost of Living: The Effect on Salary Comparisons and Policy Conclusions. *Economics of Education Review, 24*, 323–339.

[43] Fox et al. (2019) provide evidence that the gender difference in actual teacher pay could be even larger, as ACS salary information for male teachers is more likely to include second-job incomes than is the case for female teachers.

West, Kristine L. (2014). New Measures of Teachers' Work Hours and Implications for Wage Comparisons. *Education Finance and Policy, 9*(3), 231–263.

White, Halbert (1982). Maximum Likelihood Estimation of Misspecified Models. *Econometrica, 50*(1), 1–25.

Wooldridge, Jeffrey M. (2010). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press. Second Edition.