Year 2 Denver ProComp Evaluation Report:

Teacher Retention and Variability in Bonus Pay, 2001-02 through 2013-14

Allison Atteberry

Derek C. Briggs

Sarah LaCour

Charles Bibilos

Colorado Assessment Design Research and Evaluation (CADRE) Center

University of Colorado, Boulder

September 11, 2015

CADRE

School of Education
UNIVERSITY OF COLORADO BOULDER

Table of Contents

List of Figures

List of Tables

Focus Area 1: Teacher Retention in the ProComp Era

Overview of the Research Questions

Herein we present quantitative evidence about DPS's ability to retain more teachers in the district, with an emphasis on whether the kinds of teachers who are being retained have changed over time. In particular, we explore whether ProComp could have had an observable impact on trends in teacher retention since it first started (see discussion below about why it is difficult to determine unequivocally if these trends were *caused* by ProComp). A number of research questions were identified around the topic of teacher retention:

1. How has the ability of DPS to retain teachers in the district changed over time?
2. How do DPS's teacher retention rates compare to neighboring districts during the same time period?
3. How does the ability to retain more teachers to DPS differ by school type (e.g., elementary/middle/high school), or for schools serving historically disadvantaged populations (e.g., high minority, hard-to-serve, or turnaround schools)?
4. Have more teachers of higher performance (on MGPs) been retained in DPS during ProComp years?

We present figures that visually depict changes in overall rates of teacher retention over time, as well as among certain kinds of teachers and schools. We also use more technical approaches (i.e., regression-based statistical techniques) to estimate whether any observed changes in rates of retention are statistically significant. In other words, we estimate whether the data supports the hypothesis that an actual change in trend occurred, rather than non-systematic noise in trends from one year to the next.

Challenge of Estimating the Effect of ProComp on Teacher Retention

We first present simple descriptives on teacher retention outcomes over time. While we cannot say for sure that any patterns observed over time are definitively caused by the introduction of ProComp in Denver, we specifically explore whether the pre- and post-ProComp time periods have different retention patterns. If we do find that the district was able to retain more (and perhaps stronger) teachers in the period after ProComp began, that would provide first order evidence that ProComp may have played a role. However, while the onset of the first instantiation (2005-06) of ProComp preceded the economic recession, the full implementation of ProComp occurred simultaneous with this national event. Therefore it is difficult for us to determine whether ProComp itself or the recession of 2008-09 caused any observed changes in teachers' career movements. In more technical terms, we state that the analysis does not have a strong source of information about the counterfactual scenario—that is, what would have happened in DPS to rates of teacher retention from 2006 through the present if ProComp had not ever happened? This limits our ability to determine whether ProComp had a measurable effect on teacher retention.

One reasonable way to bring to bear some evidence about what Denver's teacher retention trends might have looked like in the absence of ProComp would be to examine retention trends in neighboring districts or the state at large. Though we do not have student- or teacher-level data from another location, we do have access to district-level data on teacher retention from the Colorado Department of Education. If teacher retention trends in Denver look quite different than those of neighboring districts or Colorado overall, it lends support to the argument that ProComp—a policy unique to Denver since 2006—can be linked to changes in teacher retention outcomes. While the district-level statistics on teacher retention from other locales certainly enhances our insight into whether Denver's teacher retention trends are unique in the post-ProComp period, we must remain concerned that what happened in other districts may not be appropriate for estimating what would have happened in Denver in the absence of ProComp during this time frame. Therefore, while the availability of teacher retention data from other districts is helpful for the retention analysis presented below, we continue to caution the reader against interpreting the findings below as the causal effect of ProComp on teacher retention.

Relevant Sample (and Necessary Limitations) for Retention Analysis

Whereas we are unable to determine whether teachers who appear in the first year of our dataset were new-hires (i.e., new arrivals) since we have no knowledge of their prior work history in DPS, we *are* able to determine whether teachers in the first year of data are retained (i.e., whether they leave the district in subsequent school years). On the other hand, we are unable to include the final year of the panel (2013-14), because we lack information about whether teachers who are present in that school year returned to the district the following year. Therefore, the retention analysis is limited to the 2001-02 through 2012-13 school years.

Methodology

We calculate an annual rate of teacher retention by counting the percentage of teachers in a given school year who appear in subsequent school year databases. For the sake of simplicity, we present the "retention rate" for each school year, which captures the percentage of teachers who were present in DPS in the previous year who reappear in the fall of the current school year. For instance, the Fall 2005 retention rate would be the percentage of teachers who were present in the 2004-05 school year who reappear in Fall 2005.

We divide the twelve-year longitudinal panel (from the 2001-02 school year through the 2012-13 school year) into three distinct segments: The first period begins with our first year of data in which we can observe teacher retention (Fall 2002), and ends with the last year in which ProComp was not in effect. The last school year included in this period was Fall 2005. ProComp was initiated in the middle of 2005-06, and thus teachers made decisions about leaving the district under ProComp for the first time in Fall 2006. Therefore we argue that teacher departures from the district were likely unaffected in Fall 2005 year, but potentially affected for the first time in the Fall 2006.

Our second time segment includes the Fall 2006 through Fall 2008 school years—that is, the years in which ProComp was in place but prior to its full implementation and also prior to the time in which it became mandatory for all new hires. We think that teachers employed in Denver during these three years were likely aware of the changes in compensation policy and thus could have been making their employment decisions based on ProComp. At the same time, since ProComp was quite new, not fully implemented, and optional, teacher retention may not have been as strongly affected in those three initial years.

The third and final time segment of interest was the period during which hires might be attracted to DPS because of the presence of the full implementation of ProComp. This includes retention in Fall 2009 and subsequent school years through the Fall 2013, which was the most recent opportunity to observe district retention.

Thus, our statistical model (shown below) separates time into those three time segments, and we use a linear function of time in each of those three periods:[1]

[1] We considered other models – a linear regression, a quadratic regression, and a cubic regression – in addition to our interrupted time/piece-wise model. Ultimately, however, we observed that the three periods were different enough to merit distinction in the model.

$$Retained_{py} = \beta_0 + \beta_1(Year_{py}) + \beta_2(Post05_{py}) + \beta_3(Year_{py} \times Post05_{py}) + \beta_4(Post08_{py}) + \beta_5(Year_{py} \times Post08_{py}) + \varepsilon_{py} \qquad \text{(Eq. 1)}$$

In Equation (1), the outcome variable is $Retained_{py}$, which is set equal to 1 if the given teacher (denoted by subscript p) returns to the district in the Fall of a given year (denoted by subscript y), and it is set equal to 0 if the given teacher does not appear in the district in the subsequent year(s). We predict variation in percentage of teachers who remain the district (i.e., percentage retained) as a function of time[2], captured by the variable $Year_{py}$. We interact the $Year_{py}$ variable with two post-period dummy variables, which allows us to estimate a separate retention trend in each period: $Post05_{py}$ is a dummy variable indicating whether or not the school year is 2005-06 or later (1), and $Post08_{py}$ is a dummy variable indicating whether or not the school year is 2008-09 or later. Each of these dummies are included in the model and interacted with the centered $Year_{py}$ variable. We estimate each coefficient (that is, each beta "β") using ordinary least squares regression techniques[3] and present results both in tables, as well as in figures (see *Appendix A: Teacher Retention Estimation Results*, for a complete description of the estimation results and their interpretation).

Findings

Overall Teacher Retention Rates across Study Years.

Figure 1 illustrates trends in the percentage of teachers who are retained in the district across the twelve years of the study, providing an overview of teacher retention trends before and

---

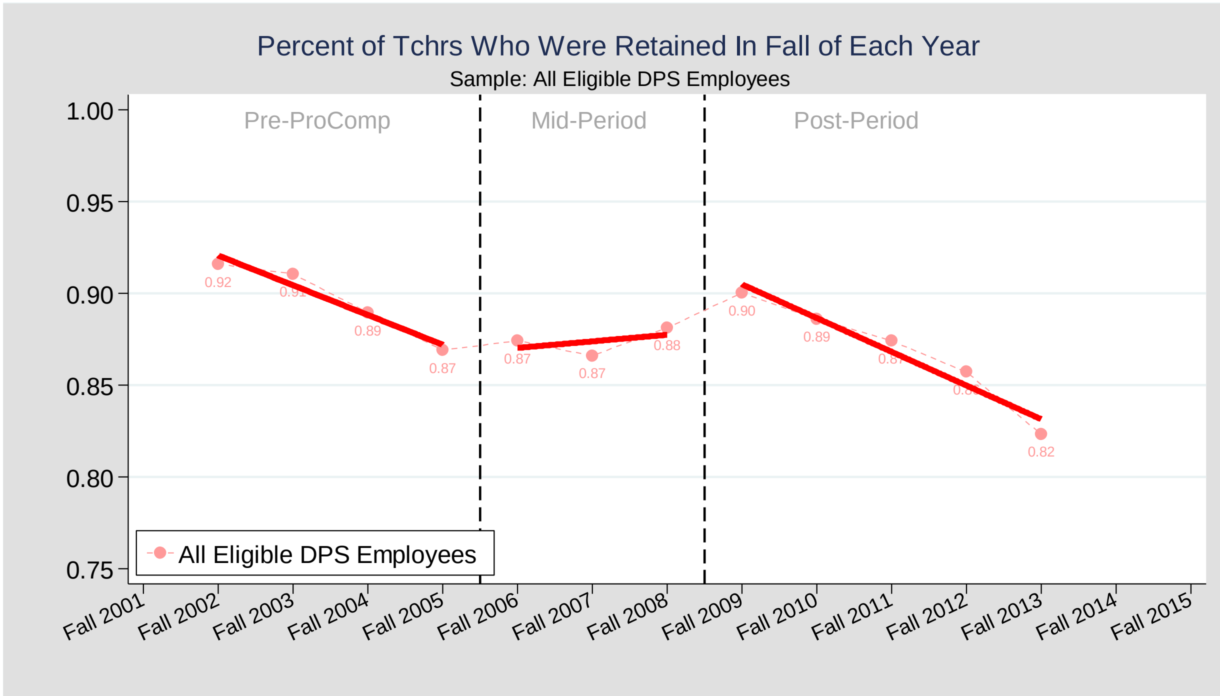
[2] We chose to center the year variable on the 2004-05 school year, so that $Year_{py}$ equals 0 in the 2004-05 school year—that is, the last year in which ProComp was not in effect (i.e., the last year of the first time segment. $Year_{py}$ equals 1 in the first year of ProComp implementation, equals 2 in the second year, equals 3 in the third year, etc.

[3] For ease of interpretation, we opted to present results from linear probability models, rather than logistic regression models. The patterns of statistical significance are largely consistent across the two approaches. Available upon request.

after the onset of ProComp. The data points, connected by a light, dashed line represent the observed “retention rates” in each school year. The retention rate is defined as the percentage of teachers present in the previous year who reappear in DPS in the given year. For instance, in Fall of 2002, 92 percent of teachers returned to the district. As described above, we also use linear regression models—see the foregoing equation and technical discussion—to estimate trends in retention rates separately in each of the three periods of interest. Because Figure 1 contains both the underlying data (raw retention rates in each year) and the more smoothed trend lines from the statistical model, compare how well the estimated trends lines from the model match up with the underlying data which tends to fluctuate somewhat from one year to the next.

As shown in Figure 1, retention rates had been falling in the initial pre-ProComp period. In Fall 2005, we estimate that 87.2 percent of teachers were new to the district. The annual change in retention rates during the pre-period (Fall 2005 and before) was negative: the rate of teacher retention decreased by an average of 1.6 percentage points each year. In Fall 2006, the first year following the onset of ProComp, the teacher retention rate was estimated to be 86.7 percent (a small, not statistically significant change from the previous year). In the mid-period, the annual trend in retention rates was a 0.4 percentage points increase each year—i.e., close to flat. The retention rate in Fall 2009 was 90.6 percent—a small jump up when compared to retention in the last year preceding ProComp’s implementation in Fall 2005 (87.2 percent). The annual change in retention between Fall 2008 and Fall 2013 is negative (a 1.8 percentage point decrease each year), and quite similar to the trend observed in the pre-ProComp period.

Figure 1. DPS Teacher Retention: District Retention Rates Across Three Time Segments (Pre, Mid-, and Post-ProComp Implementation)



Overall, retention rates decreased during both the first and third periods. The greatest downward trend in the rates of retention, on average, was during the most recent period. Interestingly, the rates of retention remained relatively constant (or perhaps increased slightly) during the middle period – during the initial iteration of ProComp. It is tempting to conclude that ProComp can be linked to *lower* rates of teacher retention in DPS since its full-scale implementation. However, we are unable to definitively ascribe this to ProComp itself because, absent an appropriate counterfactual, we cannot distinguish between effects of ProComp and those associated with the recession as well as any other policy efforts that were underway in DPS during the same time period.

One point of clarification is in order. In 2014, Dr. Eleanor Fulbeck published a peer-reviewed article on teacher mobility and financial incentives in the context of Denver's ProComp system. Similar to the current report, she explores labor market dynamics in relation to the onset of ProComp. Our findings for overall estimated retention trends differ from those estimated in Dr. Fulbeck's article[4], and we therefore take some time to delineate reasons why the findings do not match up.

Firstly, Fulbeck was more restrictive in her sample, limiting it to those whose job description titles included "teacher." Because we have multiple points of data on job descriptions, and given that not just direct instructional personnel are affected by ProComp, we elect to include all employees who were ProComp eligible in our analysis. Thus, we include personnel such as therapists and librarians that were not included in Fulbeck's analysis. We choose to do so because we believe the decisions to remain in the district for these individuals could be affected by ProComp. In addition, Fulbeck chose to exclude individuals who were terminated, retired, or let go due to a reduction in force, whereas we include those kinds of individuals in our sample and label them as not retained.[5]

Secondly, Fulbeck's outcome of interest includes both "departures" from a specific school, as well as the district as a whole. Our analysis is more focused on district retention of human capital, and we therefore do not consider movements across schools as a "loss" of a teacher that could potentially be attributed to the ProComp system. Thus, we do not label moves between schools as departures; we consider those teachers to have been retained by the district.

As a consequence of these differences in samples and in our conception of a teacher departure, we arrive at noticeably different conclusions than those presented in the Fulbeck article.

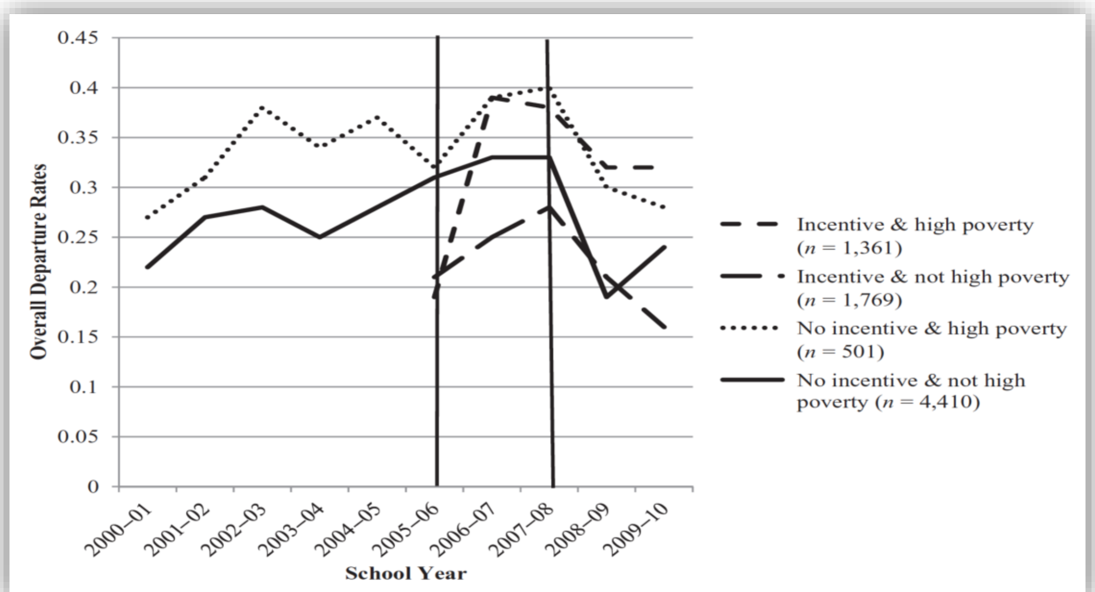
---

[4] Fulbeck, E. S. (2014). Teacher Mobility and Financial Incentives A Descriptive Analysis of Denver's ProComp. *Educational Evaluation and Policy Analysis*,*36*(1), 67-82.

[5] We did not have information on the reasons teachers did or did not remain in the district. However, even if given the opportunity to exclude these individuals, we believe that would have been a different story, in that it would not have shown as clearly the kinds of individuals who were leaving the district, on the whole. We feel it is important for our purposes to explore whether higher quality teachers are more likely to remain in the district, on average.

Fulbeck, in looking at departures rather than retention, as indicated by her original Figure 1 on page 72 of the publication (reproduced below), saw a decrease in the number of teachers leaving high poverty schools in the 2007-08 to 2009-10 school year period as well as a decrease in incentivized teachers leaving non-high poverty schools. The only increase in departures was among non-incentivized teachers leaving non-high poverty schools. We do not observe these same decreases in departures (or, in our case, a corresponding increase in retention). This discrepancy likely arises because we are counting a different phenomenon (departing from the district rather than a given school) and using a less restrictive different sample (including personnel not engaged in direct classroom instruction and those whose departure was other than voluntary). Furthermore, our data panel includes three additional years of data. Together, these factors likely account for the discrepancy between her findings and our own.

Fulbeck 2014: "FIGURE 1. **Teachers' overall departure rates, 2000–2001 to 2009–2010."**

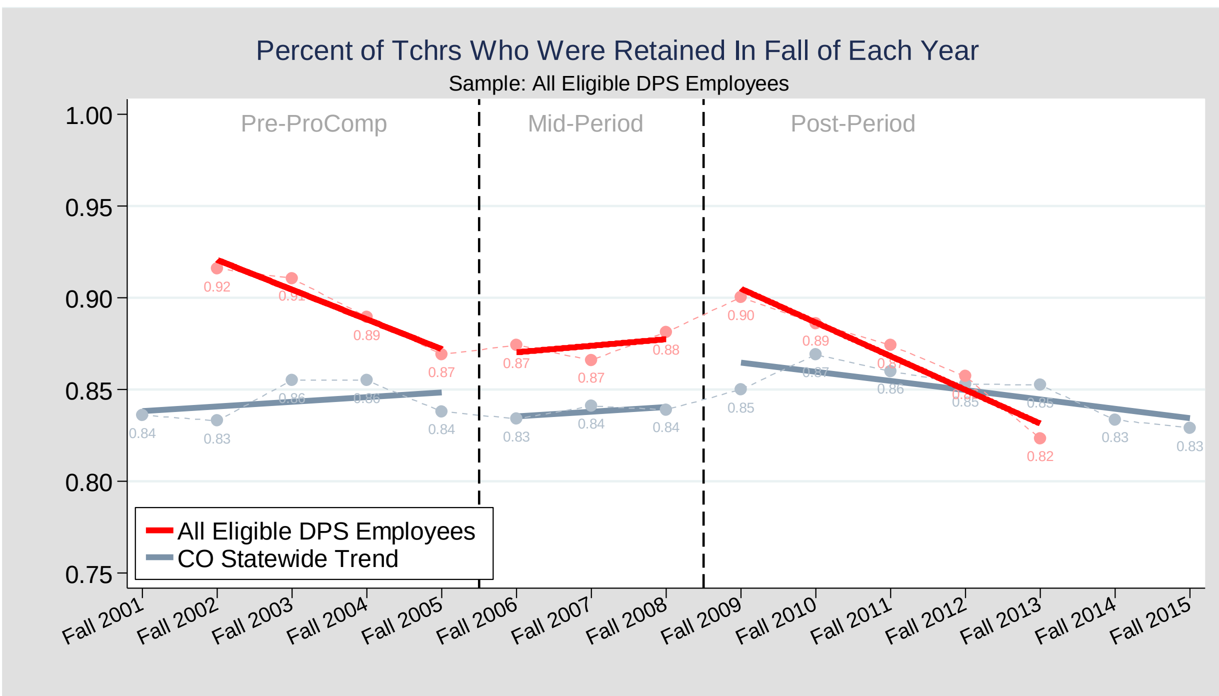


Source: [1] Fulbeck, E. S. (2014). Teacher Mobility and Financial Incentives a Descriptive Analysis of Denver's ProComp. *Educational Evaluation and Policy Analysis,36*(1), page 72.

Retention Rates over Time across the State and in Neighboring School Districts.

As stated above, it is difficult to interpret the patterns in teacher retention outcomes in relation to ProComp given that Denver—like all school districts—experienced the economic recession in 2008-09 and the following three to four years. Therefore it is difficult to make sense of the teacher retention trends apparent in Denver without a larger context for teacher labor markets. In Figure 2 below, we present information from the Colorado Department of Education on teacher retention rates throughout Colorado in order to provide greater context for the Denver trends.

Figure 2. Teacher Retention Rates, Denver vs. Colorado Statewide (Fall 2001 - Fall 2015)



Source for Statewide Data: http://co.chalkbeat.org/2015/05/28/more-colorado-teachers-left-their-school-districts-last-year/#.Ve9fb_lVhuA

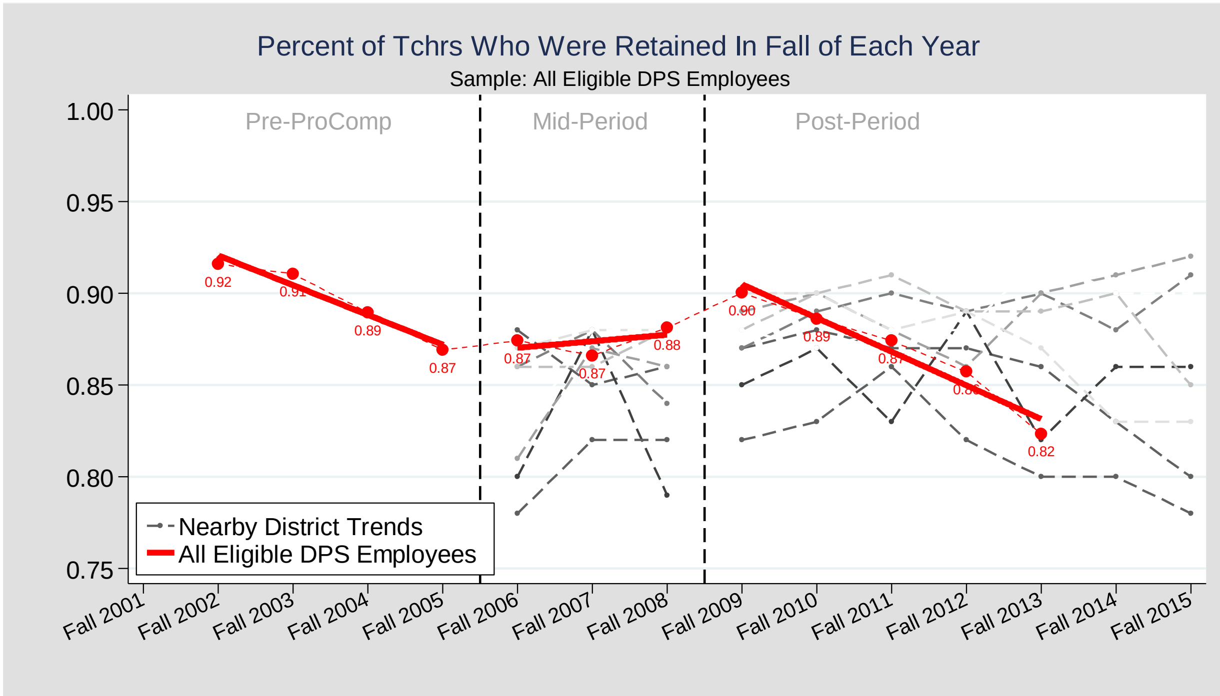
Figure 2 reveals that teacher retention trends in Denver mirror, to some extent, the patterns evident throughout the state: In the three "mid-period" years after ProComp was introduced, but before its full implementation in 2008-09, teacher retention was relatively constant across

Colorado. Beginning in 2009 (and more dramatically in 2010), the statewide retention rates increase by several percentage points. The temporary uptick in teacher retention throughout Colorado likely reflects the economic downturn associated with the recession. This was a point in time when workers in almost all sectors were unlikely to voluntarily leave their jobs. In the most recent years, teacher retention rates have decreased once more, leading to an overall downward trend in teacher retention throughout Colorado from 2008-09 through the most recent school year. Taken together, this suggests that one should not necessarily conclude that the downward trend apparent in Denver was a consequence of the introduction of ProComp. It seems more likely that Denver experienced the same kinds of economic challenges that are visible across the state during this time frame. That said, we also see that the retention rates for Denver have fallen *more precipitously* since Fall 2009 than have statewide averages. This may be an area of concern for DPS: It does appears that the implementation of ProComp does not coincide with an increasing rate of retention.

In Figure 3 below, we also present teacher retention rates in seven school districts geographically proximate to Denver Public Schools. The Colorado Department of Education publishes data on personnel turnover rates, by district and publishes those reports on their website.[6] We selected the following eight districts due to their geographic proximity to DPS: Adams 12 , Aurora Public Schools, Boulder Valley, Brighton School District, Cherry Creek School District, Douglas County School District, Jeffco Public Schools, and Littleton Public Schools. Teacher retention data at the district level is currently only available since 2005-06, which prevents us from comparing these districts' pre-ProComp retention trends to those observed pre-ProComp in Denver.

[6] Source: http://www.cde.state.co.us/cdereval/staffcurrent

Figure 3. Teacher Retention Rates in Eight Neighboring Districts since 2005-06



Source for Statewide Data: http://co.chalkbeat.org/2015/05/28/more-colorado-teachers-left-their-school-districts-last-year/#.Ve9fb_lVhuA

Figure 3 shows substantial variability in overall teacher retention in the local teacher labor market around the Denver Metro Area. The takeaway from Figure 3 is that many neighboring districts possess quite volatile retention rates over these years: some are trending upwards, some trend downwards, and some fluctuate from year-to-year and overall remain flat. However Denver exhibits clearer patterns in both the mid-period (2006-2008) and the full implementation post-period (2009+). In Table 1 below, we also estimate the average annual change in trends during these two periods separately for each district, allowing the reader to make comparisons between Denver and specific nearby districts. An example is instructive: Table 1 reports that, in Aurora, from the mid-period from 2006 to 2008, the retention rate increased on average by 2 percentage points each year. However in the post-period from 2009-2015, the retention rate in Aurora *dropped* on average by 0.9 percentage points each year.

Table 1. Rate of Change in Retention Rates in Two Periods (2006-2008, 2009+), by District

	Average Annual Change	
DISTRICT	2006-2008 (mid-period)	2009-present (post-period)
Adams	-1.0	-1.1
Aurora	2.0	-0.9
Boulder	3.0	0.6
Brighton	-0.5	0.0
CherryCreek	-1.0	0.4
Douglas	0.5	-1.3
Jeffco	1.0	-0.4
Littleton	2.5	0.5
Denver	0.4	-1.8

We can see in Table 1 that some districts had decreasing rates of retention in the mid-period (Adams, Brighton, Cherry Creek), while others had rapidly increasing rates of retention (e.g., Aurora, Boulder, Littleton). Retention rates in Denver remained relatively flat during the mid-period, increasing only about half a percentage point each year (similar to Douglas). During the post-period (which coincided with the Great Recession), many districts experienced falling retention rates (Adams, Aurora, Douglas, and Denver), several experienced relatively stable retention (Boulder, Brighton, Cherry Creek, Jeffco, and Littleton), and none of these 8 districts experienced an annual increase even above 1 percentage point per year. This is likely emblematic of the overall labor market trends during this time. It is notable however, that annual decreases in teacher retention rates were steepest in Denver during the post-period since ProComp was fully implemented.

Retention Rates over Time, by School Characteristics.

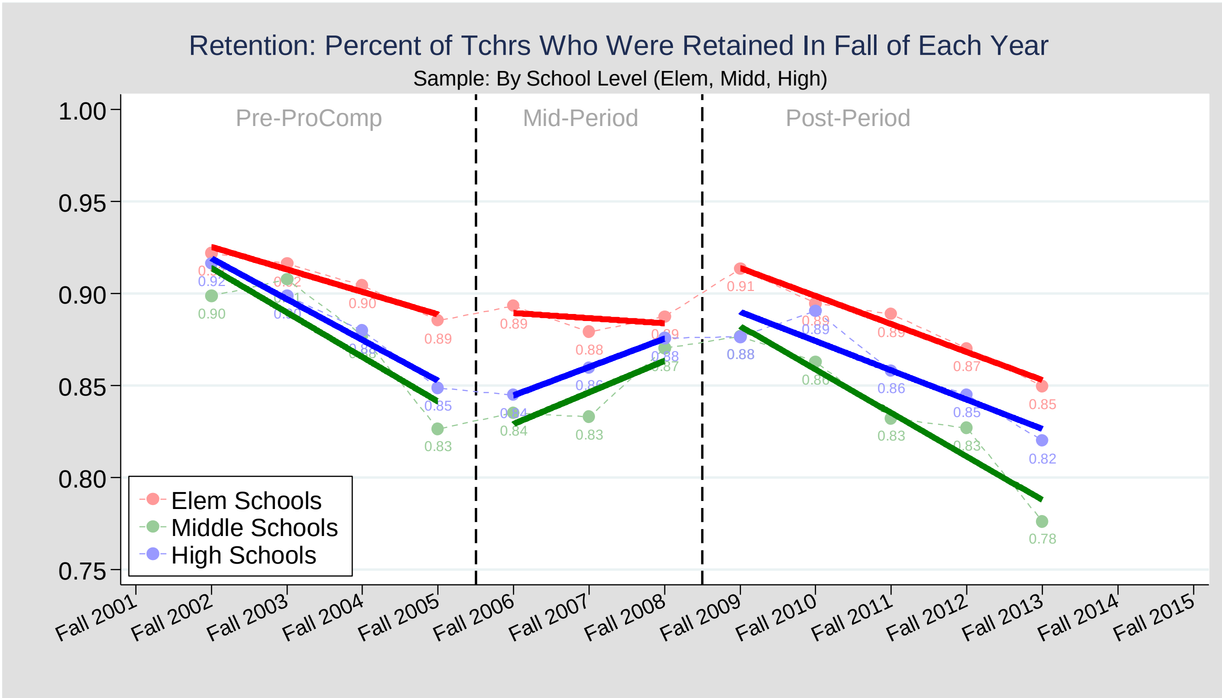
We next explore whether certain kinds of schools exhibit teacher retention trends that differ from the overall average. Specifically, we compare rates of teacher retention across elementary, middle, and high schools. We also examine whether schools that are either designated as "high minority" (greater than 75% non-white), "high poverty" (greater than 75% free-lunch eligible), "hard to staff", or "turnaround" experience atypical teacher retention rates.[7]

Figure 4 suggests that teacher retention has been the greatest challenge in Denver middle schools (green line), followed by high schools (blue line), and finally teacher retention is least problematic in Denver's elementary schools (red line). For instance, the Fall 2005 retention rate in elementary schools was estimated to be 88.9 percent, followed by 84.1 in middle schools, and 85.3 percent in high schools. That gap is even larger in Fall 2006, the first year after ProComp began (retentions rates are 89.3, 81.2, and 83.0 percent in elementary, middle, and high schools, respectively). That gap appears to close somewhat by Fall 2008, however by Fall 2013, the gaps in retention rates across school levels have reemerged.

While middle school teachers exhibit the lowest retention rates relative to their elementary- and high-school counterparts, it appears that the trends have been quite similar over time across all groups: In other words, teachers in all school levels exhibited decreasing rates of teacher retention prior to the initial onset of ProComp in 2005-06, followed by a three-year period of increased retention, and then a marked drop in retention rates at the onset of the Great Recession.

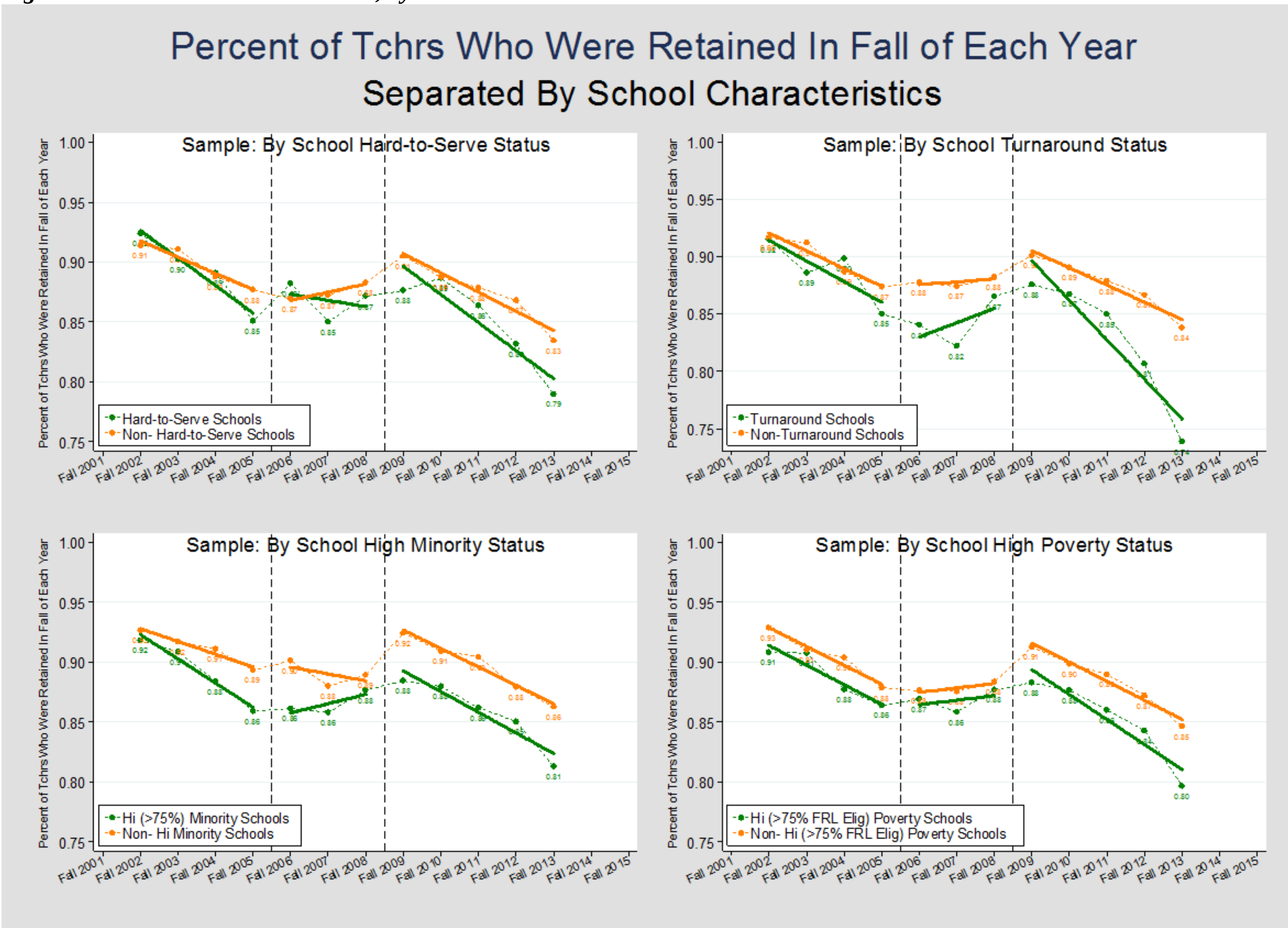
[7] A school was considered "turnaround" if it was ever in turnaround status during the period of study. A school was considered a "hard to serve" school if it was ever designated as such on the historic lists of Hard to Serve schools listed on DPS's ProComp website (http://denverprocomp.dpsk12.org/history). A school was designated high minority if it ever served a population of students in a given year that was greater than 75 percent non-white during the study. A school's grade level (elementary, middle, or high) was determined by its most recent grade configuration.

Figure 4. Denver Teacher Retention Rates over Time, by School Level (Elementary, Middle, High School)



We also examine whether schools that are either designated as "high minority" (greater than 75% non-white), "high poverty" (greater than 75% free-lunch eligible), "hard to staff", or "turnaround" experience atypical teacher retention rates. Again we see that the overall pattern of teacher retention rates from 2001-02 through 2012-13 are quite similar across the four school types of interest in Figure 5 (below). It is interesting to note that, prior to the onset of ProComp in Fall 2006, retention rates were not so different between these four kinds of schools and their non-disadvantaged counterparts. For instance, the retention rates in turnaround vs. non-turnaround schools was quite similar between 2002 and 2005. However, in the post-period from 2009 through 2013, the differences in retention rates have widened. In 2013 specifically, turnaround schools experienced a low retention rate of only about 75 percent, whereas non-turnaround school about 84 percent of their teachers. Indeed, it does appear that turnaround schools have experienced an especially steep decrease in teacher retention in the time since Fall 2009.

Figure 5. Retention Rates over Time, by School Characteristics



Teacher Retention Rates over Time, by Teacher Performance (MGP's).

Measuring a given teacher's performance and impact on students is a complex issue. A well-established literature examines the predictive validity of statistical measures of teacher effectiveness (e.g., value-added measures such as median growth percentiles, a.k.a. MGP's). This literature generally suggests that there is some useful signal amongst the noise but the measures are imprecise for individual teachers (Braun, Chudowsky, & Koenig, 2010; Briggs, 2012; Donaldson & Papay, 2014; Kane, McCaffrey, Miller, & Staiger, 2013; McCaffrey, 2012; McCaffrey, Sass, Lockwood, & Mihaly, 2009). In this report, we use MGPs as one example of a measure of teaching performance. We do so not because MGPs capture all aspects of teaching that are important, nor because any district proposes to use MGPs in isolation. In fact, DPS bases personnel decisions on measures of teaching effectiveness combine multiple sources of information as part of its LEAP system. . We focus on MGPs in the report as an imperfect proxy for teaching effectiveness that is being used in DPS as part of its LEAP system. As data from LEAP become more widely available it will become possible to use multiple measures to characterize teachers for the same sort of analysis.

By their nature, the use of MGPs restricts our sample to those teaching in tested grades and subject areas, and with a minimum number of students. In any given year, typically 20 to 30 percent of teachers are linked to students in tested subjects and grades. Relatedly, teachers sometimes do not have an MGP in every school year as their assignments to subject- and grade-levels change. In addition, MGPs are quite noisy measures of teachers in any single year, with the amount of noise inversely related to the number of student growth percentiles available to compute the MGP (McCaffrey, Sass, Lockwood, & Mihaly, 2009). We chose to use all available MGP data for each teacher during the study timeline to characterize his or her teaching performance. We therefore calculate an average MGP for every teacher across all subject and school years in which that teacher is present. Of the 13,520 unique teachers that appear in our dataset, 4,019 (approximately 28.7 percent) have at least one MGP score and can be grouped using MGP scores. We categorize the set of 4,109 teachers for whom we can estimate an overall average MGP score (across years and subjects) into a top, middle, and bottom group of the performance distribution. Specifically, a teacher was designated a "top" performer if her overall MGP score was greater than or equal to 60, a "middle" performer with an average MGP score between 40 and 59, and a "bottom" performer with an average MGP score below 40. These are the same thresholds used to categorize teacher MGPs in LEAP.

Using the overall average MGP score has the advantage of being as inclusive as possible, allowing any teacher who has ever had at least one MGP score to be part of the analysis. On the other hand, one disadvantage is that these averaged MGP scores may be more or less imprecise across teachers, depending on how many MGP scores are available. For instance, it is possible that a given teacher could have MGP scores every year in both subjects—for this hypothetical teacher, the overall MGP score reflects a long history of performance across many school years. At the same time, another teacher in the sample might only have an MGP from one school year. This second hypothetical teacher's overall average MGP score is more prone to instability and error (i.e., more likely to be extreme) simply because there is less data available about how student test scores grew while assigned to her. We were particularly concerned that there would be a systematic, negative relationship between a teacher's probability of being grouped into the highest or lowest categories and the number of scores that went into his/her overall average MGP score. [8]

[8] We were concerned that teachers with fewer years of MGP data would be more likely to have overall average MGP scores in the top or bottom portions of the distribution, simply because their MGP's are more noisy and thus more

These differences across teachers in the imprecision of the MGP's as a measure of their overall effectiveness are inherent to the analysis below.

While there are downsides to using a teacher's *average* MGP across all available subjects and school years (we refer to this as "Approach A"), we believe it is the best approach given then data limitations. However, we also present sensitivity analyses to assess whether our findings are sensitive to the number of unique MGPs being used to place them into effectiveness categories.

We were provided with student-level standardized test score data from 2004-05 onward, allowing us to estimate teacher level MGP's in 2005-06 and beyond. This lack of student test score data from prior school years imposes an important limitation on the teacher retention analysis that follows: We would like to examine patterns of differential attrition from DPS in the period prior to the onset of ProComp to see if there is any relationship between teacher performance and teacher retention in the absence of the pay-for-performance system. Unfortunately, because we have no MGP data prior to 2005-06, we have no information about the performance of departing teachers before ProComp began.

Nonetheless, we compare the rate of teacher retention from Denver in the ProComp implementation periods (both during its first three years, as well as after the full implementation in 2008-09). This allows us to examine whether trends in the retention of teachers who have high, middle, or low performance on MGP's are different after the introduction of ProComp. If we find that high-performing teachers are more likely to be retained under the ProComp policy—and concurrently low-performing teachers are more likely to leave—then this would be a pattern that

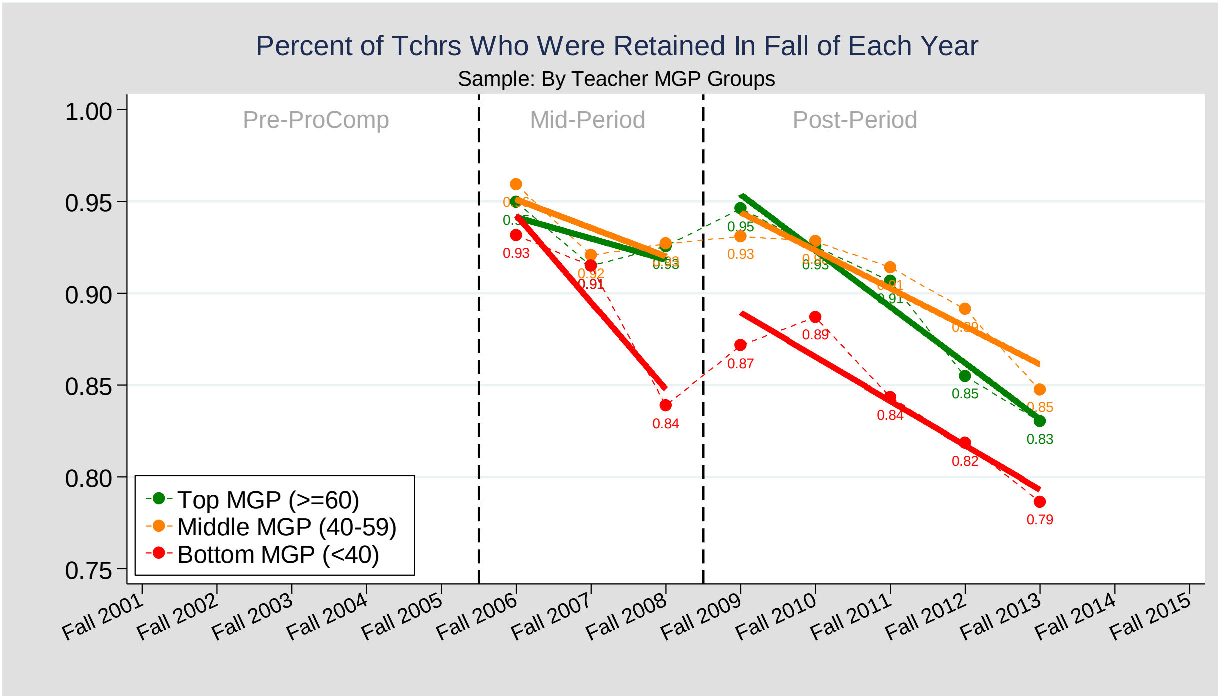
---

likely to be extreme. In order to test this, we ran three logistic regression models: (1) Predict the probability that a teacher's overall average MGP (across all subjects and years) puts her in the highest performing group, as a function of the number of MGP scores he/she possess, (2) the same for the predicting assignment to the middle-performing group, and (3) the same for predicting assignment to the lowest-performing group. The results (available upon request) indicate that number of MGP scores a teacher possesses is not a statistically significant predictor of whether he/she is allocated to the high, middle, or low-performing group. This alleviates our concern that the quality groupings were driven by differences in the amount of MGP data teachers possess.

is consistent with the theory of action underlying pay-for-performance systems: That is, teachers with the strongest performance and thus greater earning potential are encouraged to remain in the district while other, perhaps weaker, teachers are not rewarded by the compensation system and thus are discouraged from remaining. As a result, it is hypothesized that the overall quality of the teacher workforce would increase through this selection process. In the analysis that follows, we examine whether the data is consistent with that pattern in the time period since ProComp began in Denver.

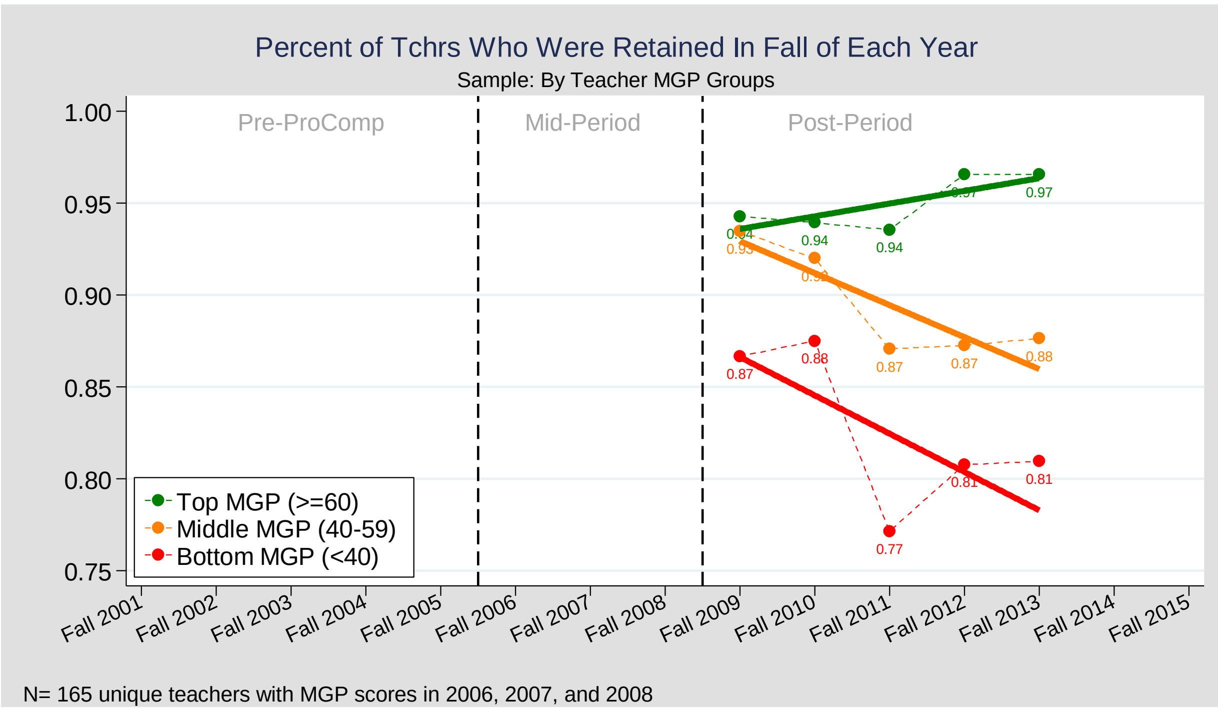
Figure 6 depicts the percentage of teachers that are retained by DPS each year, separately for the highest (>=60), middle (40-59), and lowest (<40) groupings of overall teacher MGP scores. Again, recall that the sample is now necessarily limited to the subset of teachers who taught in classrooms (i.e., subjects and grades) in which end-of-year standardized tests are administered, and we cannot observe the MGP-based performance of teachers who departed prior to 2005-06 (N = 4,109 unique teachers). We find that teachers in the lowest third of the performance distribution were the least likely to be retained, relative to teachers in the middle- or top- third of the performance distribution. Moreover, the trend over time appears to be steepest for the lowest-performing teachers. It is important to acknowledge that all three groups of teachers exhibited a downward trend in retention during the economic recession.

Figure 6. Retention Rates Since Fall 2006, Separately for Top, Middle, and Bottom-Performing Teachers.



While we believe that using the overall average MGP score for each teacher ("Approach A") is the best approach despite the drawbacks mentioned above, we also considered a number of alternative approaches to using MGP scores to group teachers into high- , middle-, and low-performing categories *other* than Approach A described above. For instance, as a sensitively check, we limited the sample to a group of teachers who possessed MGP scores in 2005-06, 2006-07, and 2008-09 ("Approach B"). The idea here is to characterize teacher effectiveness based on the mean of those exact years and then explore whether these groups exhibited differential rates of attrition in 2010 and beyond. Our objective with Approach B was to balance the need to have some minimum number of MGP scores (to address imprecision), with the need to use a comparable set of MGP scores so that categorizations are unrelated to the number of scores a teacher possesses. A major limitation of Approach B is that only 165 teachers possess MGP scores in 2006, 2007, *and* 2008—only 4 percent of all teachers who ever have an MGP score. This group is quite small and it is reasonable to be concerned that the retention patterns among this select group do not mirror retention patterns in DPS as a whole. However, the advantage of Approach B is that we eliminate the concern that variation in number of MGP scores is biasing the results.

Figure 7. Retention Rates Since Fall 2006, Separately for Top, Middle, and Bottom-Performing Teachers. ("Approach B" to Grouping Teachers)



Results in Figure 7 differ qualitatively when we limit the analytic sample to teachers with MGP scores in 2006, 2007, and 2008. In Figure 7, we see that retention rates of top performers are actually *increasing* in the ProComp period, while retention rates among middle- and bottom-performing teachers are steeply declining. On the one hand, this pattern is quite consistent with a pay-for-performance theory of action. On the other hand, these patterns are generated from only a handful (165) of the 4,109 teachers who are ever linked to students during this period. The main takeaway here should likely be that results are sensitive to how one uses MGP scores to measure teaching effectiveness.[9]

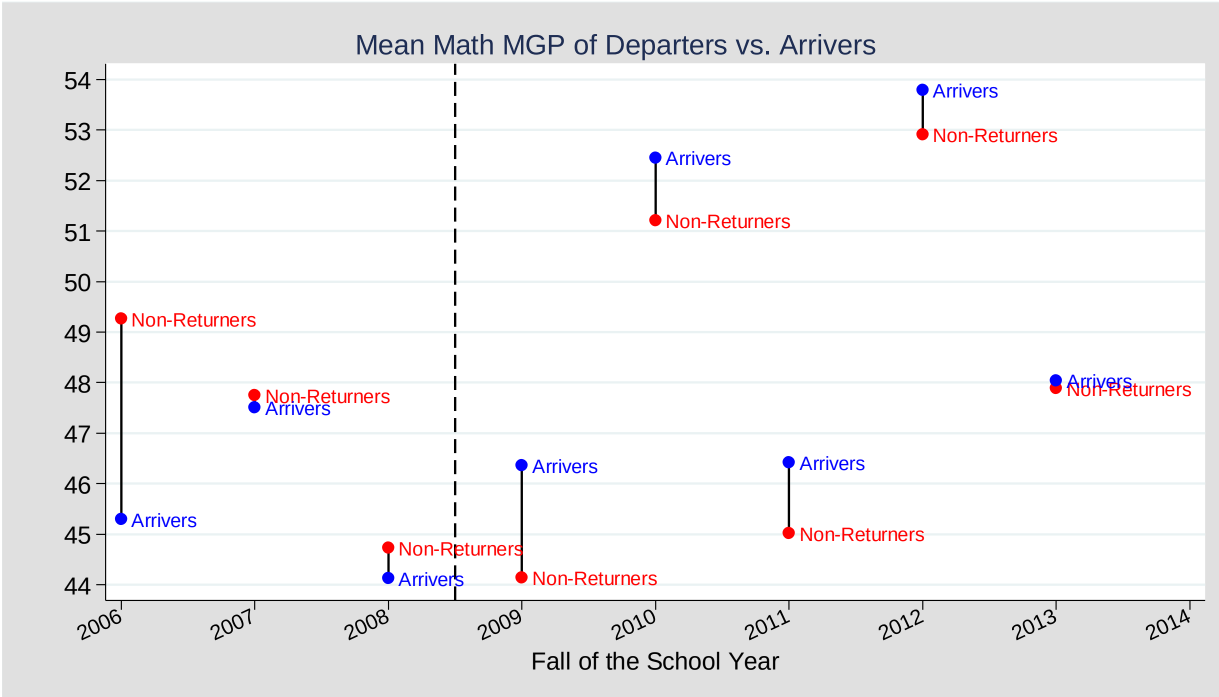
Again, we caution against definitively concluding that ProComp caused low-performing teachers to exit the system at higher rates and/or caused high-performing teachers to remain. An alternative explanation, for instance, is that teachers have always self-selected out of the profession based on their effectiveness, and we would have observed the same pattern prior to ProComp. A fruitful avenue for further exploration would be to estimate MGP scores for teachers who left Denver prior to ProComp to see if teacher retention was as strongly related to teacher performance before ProComp began.

Compare the Quality of Those Who Leave to Those Who Arrive

We now further probe the notion that ProComp may discourage relatively low-performers to depart while simultaneously attracting high-performers to the district. If these two mechanisms were successful, one would expect to see that the average MGP score of the teachers who are not retained (i.e., who do not return in the fall) are lower than the average MGP score of the teachers who arrive new to the district in the fall. We begin with math, and calculate the average MGP score of each departing teacher in the year immediately preceding his or her retention. We compare that to the average math MGP score of each new-to-district teacher in his or her first year in DPS. These means are depicted in Figure 8 (math) and Figure 9 (reading). The results are also reported in Table 2 (math and reading together).

[9] In *Appendix D: Using MGP Scores to Group Teachers into Performance Categories*, we report results according to a third approach to grouping teachers into performance categories based on MGP scores.

Figure 8. Comparison of Math MGP Scores of Non-Returning Teachers to the Math MGP Teachers Who Arrive in the Following Year



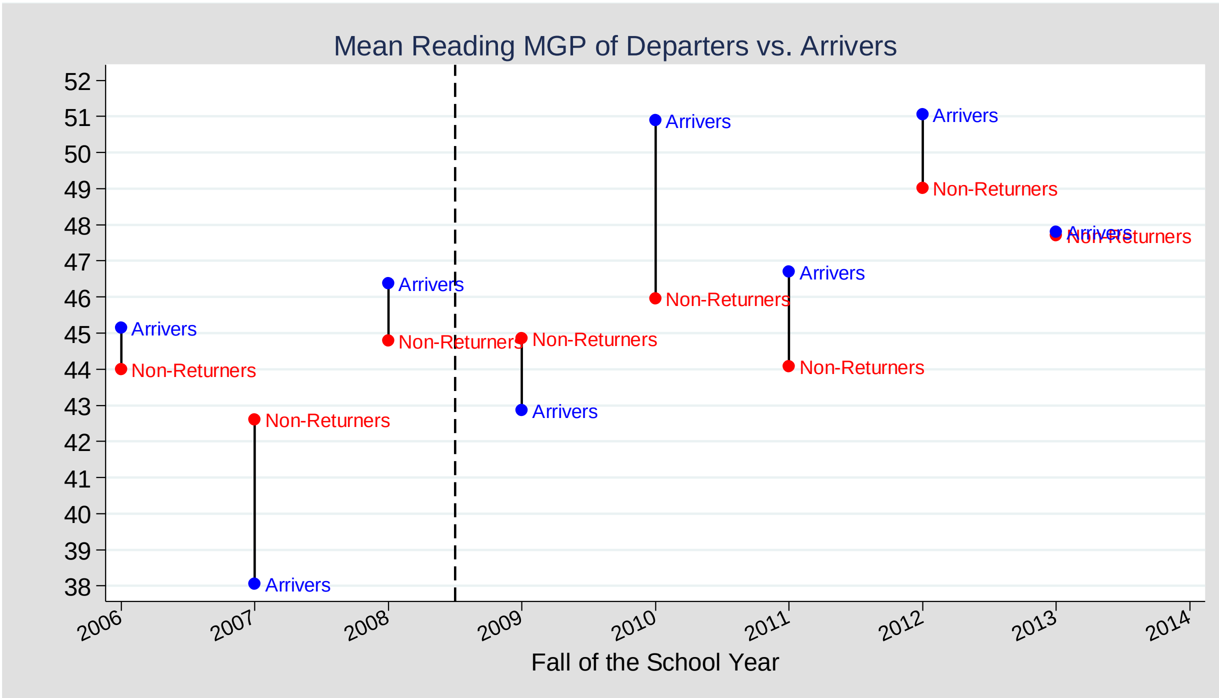
There is a great deal of variability in the difference between the MGP scores of non-returners and arrivers across the nine school years shown in Figure 8. For instance, the teachers who did not return to DPS in Fall 2006 had an average math MGP score of 49.3, however the arriving teachers only had a mean math MGP score of 45.3. In that particular year, then, the attrition/retention cycle resulted in the new teachers being of lower measured quality than the teachers they replaced. In 2007 and 2008, the pattern holds—MGP scores of the new teachers are lower, on average, than their departing predecessors—though the gap is not as large. However starting in 2009, this pattern flips, and the quality of arrivers is higher than that of the departers. Though the difference between the two group means is not always large (e.g., in 2013), it is interesting that this more productive pattern emerges in the year following the full implementation of ProComp in 2008-09.

Table 2. Comparison of MGP Scores of Non-Returning Teachers to the MGP Teachers Who Arrive in the Following Year, by Subject

	Math			Reading		
	Non-Returners	Arrivers	Difference	Non-Returners	Arrivers	Difference
2006	49.3	45.3	4.0	44.0	45.1	-1.1
2007	47.8	47.5	0.2	42.6	38.0	4.6
2008	44.7	44.1	0.6	44.8	46.4	-1.6
2009	44.1	46.4	-2.2	44.9	42.9	2.0
2010	51.2	52.5	-1.2	46.0	50.9	-4.9
2011	45.0	46.4	-1.4	44.1	46.7	-2.6
2012	52.9	53.8	-0.9	49.0	51.1	-2.0
2013	47.9	48.0	-0.1	47.7	47.8	-0.1

For ELA teachers, the pattern is less consistent (Figure 9). In 2007, the arriving teachers have a lower group mean reading MGP score than the teachers who do not return, but in all other years, arriving ELA teachers have a higher group mean MGP score than the non-returners (with the exception of Fall 2013, in which the two group means are almost equal). The reading patterns have a less clear relationship with the full implementation of ProComp than did the math patterns. That said, the tendency for entering ELA teachers to outperform leaving teachers is a positive sign for the district overall.

Figure 9. Comparison of Reading MGP Scores of Non-Returning Teachers to the Reading MGP Teachers Who Arrive in the Following Year



Summary of Teacher Retention Findings

We have presented figures that visually depict changes in overall rates of teacher retention over time, as well as among certain kinds of teachers and schools. Overall trends in teacher retention in DPS were summarized in Figure 1 and Figure 2. When reflecting on the reported trends in teacher retention in DPS, it is tempting to conclude that the initiation of ProComp temporarily caused fewer teachers to leave DPS: Prior to the onset of ProComp, teacher retention was trending downward, and then we see the trend “flatten out” in the first three years of ProComp. However we also see that, beginning just after ProComp was fully implemented in 2008-09, DPS retention rates began to decline again. As discussed above, it is difficult to separate the powerful impact of the recession on teacher labor markets from the potential impact of the introduction of ProComp on teachers’ decisions to remain in the district.

We also explored whether teachers who are in the top third of the MGP performance distribution (an average MGP of 60 or above) are retained in DPS at higher rates than teachers in the bottom group (below 40). We do find that, during the ProComp years, teachers in the bottom third have a 5 percentage point *higher* probability of leaving DPS than teachers in the top third (see Figure 6). This is potentially consistent with the hypothesis that ProComp differentially retains more effective teachers, it is also possible that these top-third MGP teachers would have remained in the District at higher rates even in the absence of ProComp.

Focus Area 2: Teacher Salary Variability due to ProComp

Overview of the Research Questions

A major focus of work from the first year of the internal evaluation was the nature of salary variability and the distribution of bonuses under the ProComp system. A new set of research questions emerged from this work, concerning the ways in which teachers perceive the nature of the incentive system and the year-to-year variation in their own salaries. In the Year 2 Work Scope Plan, we proposed to meet with teacher focus groups to discuss their perceptions of ProComp and/or conduct a survey with a random sample of DPS teachers, and several of the research questions required the more qualitative insights provided by those data collection sources. Ultimately, however, the District requested that we not conduct either the focus groups or the survey, which limited the number of the questions we could answer in the Year 2 Report. Nonetheless, we were able to quantify teachers' "experience" of ProComp (e.g. the variation in incentives and total compensation received by a given teacher) and examine the association between this experience and retention rates. Below we list the set of questions related to teacher perceptions of salary variability and indicate which we are still able to answer once limited to the administrative datasets: The original research questions included:

1. *What is the average dollar amount that individual teacher compensation (ProComp incentive pay) varies from year to year? What is the average variance from year to year for the district's highest performing teachers?*

2. *Is there evidence that pay variance from year to year causes high-performing DPS teachers to leave or consider leaving?*

Methodology

The ProComp system includes 10 distinct financial incentives, divided into four broad categories: Knowledge and Skills, Student Growth, Market Incentives, and Comprehensive

Professional Evaluations. The dollar amounts awarded for each of the 10 ProComp incentives are based on a pre-negotiated percentage of an overall index, which has ranged from $33,300 in the 2005-2006 school year to $38,118 in the 2013-2014 school year.[10]

Payment for meeting the various ProComp incentives comes in the form of either a one-time bonus or a permanent, "base-building" salary increase, depending on the incentive and the teacher's years of experience. Of the ten incentives currently offered by DPS, six are offered only as a one-time bonus, two are offered only as base-building salary increases, and two are applied as either a bonus or a salary increase, depending on the circumstances. In addition, three of the ten bonuses are awarded at the school-level (High Growth, Top Performing, and Hard to Serve), while the rest are awarded based on teacher-level criteria. We can therefore calculate the annual pay, in dollars, that each teacher receives through ProComp, separately for one-time and base-building incentives.[11] (See *Appendix E: Overview of ProComp Incentives* for a complete description of the frequency with which teachers received bonuses and their dollar amounts.)

Findings

Figure 10 below illustrates variation in the receipt of ProComp incentives across teachers in each school year since 2005-06. For our purposes, a base building incentive was "earned" in

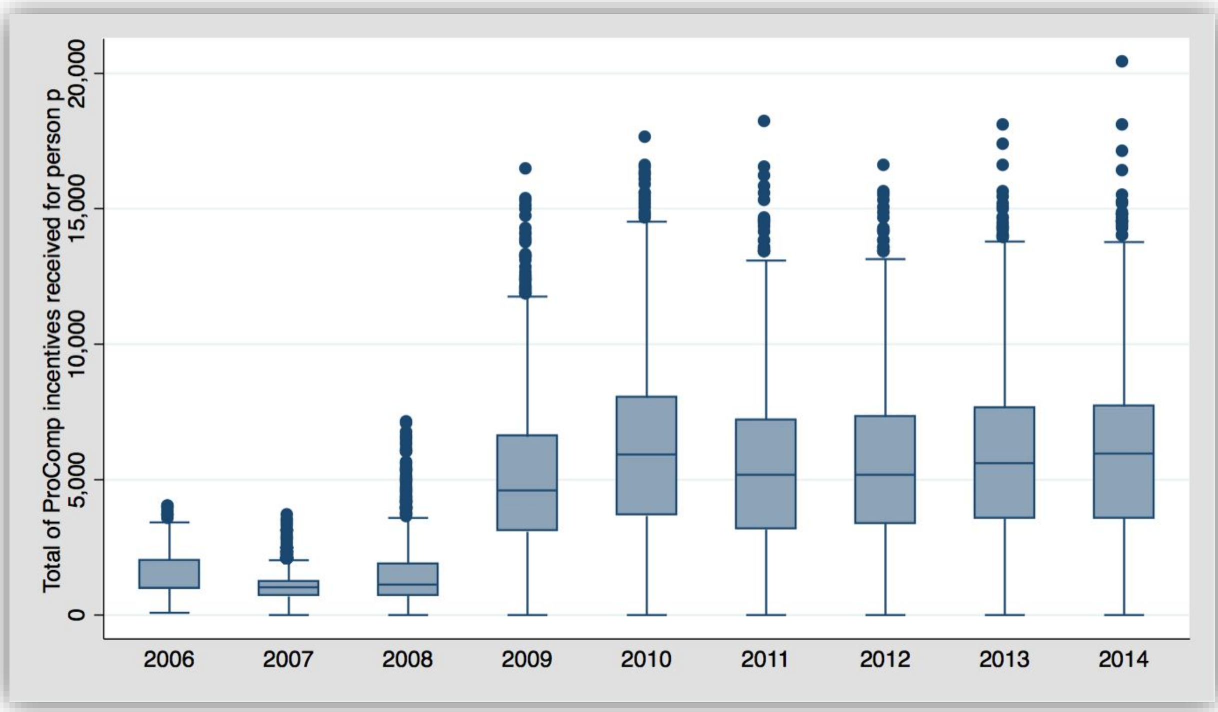
---

[10] Although information regarding the amount of each incentive is available via the ProComp website (http://denverprocomp.dpsk12.org/history) for the 2014-15 school year, in this report we only address those years for which we had data (i.e. 2005-06 through 2013-14).

[11] As an initial matter, we established the variation in incentives earned and total compensation across teachers and years. To do this, we limited the analytic sample to all ProComp participants for which we had incentive data. A review of the data showed several extreme individual incentive values in excess of $7,000: There were 55 such teacher-year observations for the Advanced Degree bonus, 7 of the CPE Base Building incentive, and 40 of the PDU Base Building incentive. Because these values were more than double the scheduled value of any of the individual incentives, and, given the relatively small number of them, we determined that they were likely due to a data entry error. We elected to drop these values because, despite the relatively small number of them in the sample, they were so far from what was likely the true value that they would have a noticeable effect on the mean. A complete discussion of the remaining discrepancies between scheduled and entered values of incentives can be found in Appendix B: Outlier Incentive Data.

the first year in which it was received. Thus, while such an incentive would continue to appear in the total compensation for a given teacher in future years, it would not appear in the succeeding years' earned incentives. The graph and whisker plot illustrates range of values by depicting outlier incentive amounts using a "dot", the 25-to-75 percentile range, and median incentive pay in each school year. The greater the spread of each box-and-whisker plot in each year, the greater the variability in incentive pay across teachers. Figure 10 indicates that the range of incentive earnings has increased over time. As the box-and-whiskers above the 2005-06 and 2006-07 school year markers indicate, even the highest outliers earned less than $5,000 in total incentives during those years. However, by the 2009-10 school year, the average value of incentives earned by an individual teacher was stabilizing above the $5,000 mark. Also during this period, the interquartile range (i.e. the difference in incentive earnings between someone in the 25th percentile and someone in the 75th percentile) seemed relatively consistent. By 2013-14, a small handful of teachers appear to be earning more than $15,000 in incentives in a year. However, not every participant earned incentive compensation as indicated by the minimum value bar consistently tracking across the $0 line on the y-axis.

Figure 10. Box and Whisker Plot: Variation in ProComp Incentive Earnings per Teacher, Across Study Years



When we considered whether or not these annual incentive totals were attributable to one-time bonuses or base-building incentives, we found that one-time bonuses showed a similar marked increase in their average between the 2007-08 school year and the 2008-09 school year (Figure 11). Also, similar to the trend in total incentive earnings, the average annual sum of one-time bonuses stabilized just shy of the $5,000 mark beginning with the 2009-10 school year and the 25th percentile stabilized around $2,500 beginning in the 2008-09 school year. Figure 11 also indicates that some teachers in each year earned no one-time bonuses. By contrast, the average annual sum of base-building incentives remained more volatile throughout the period of study, as did the inter-quartile range and overall range of earnings, as shown in Figure 12.

Figure 11. Box and Whisker Plot: Variation in ProComp One-Time Bonuses Earnings per Teacher, Across Study Years

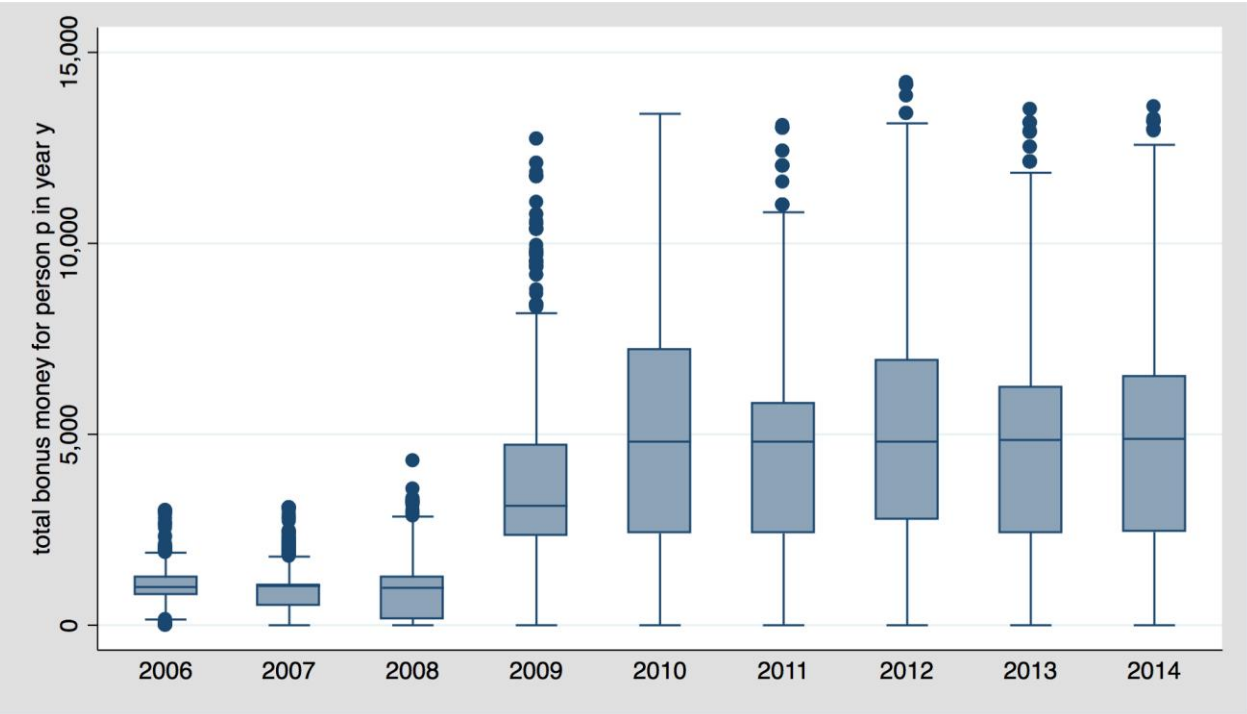
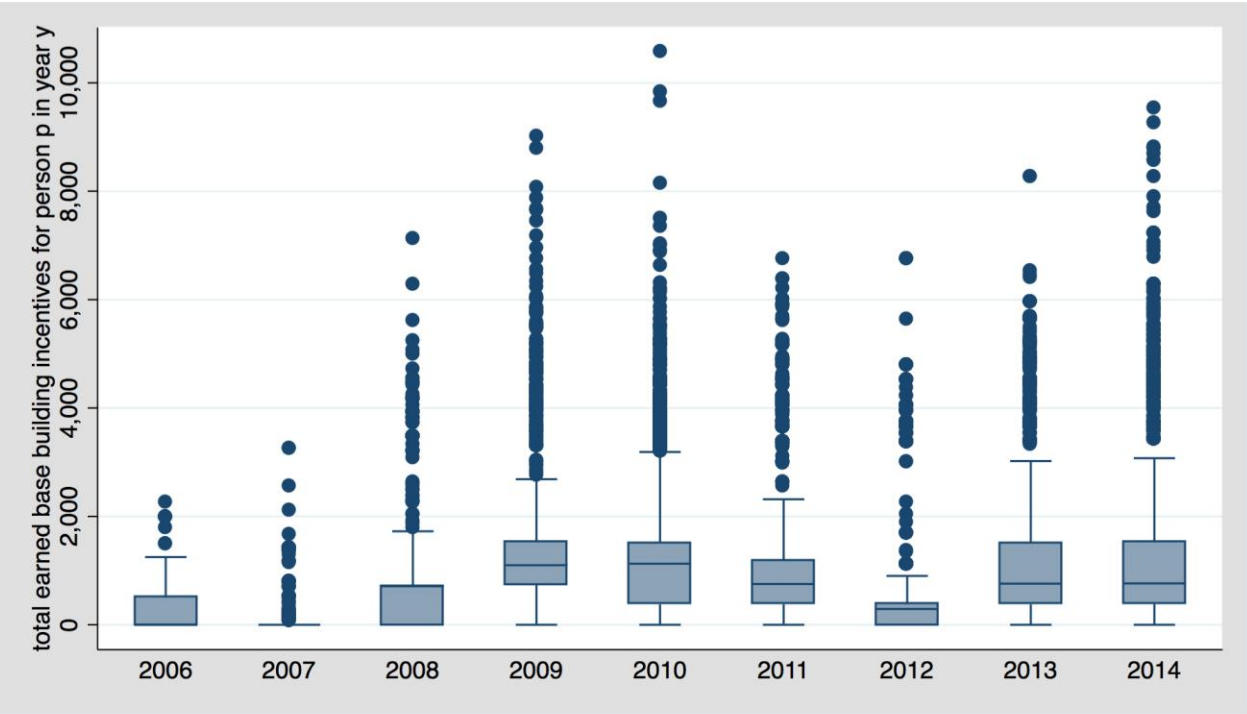


Figure 12. Box and Whisker Plot: Variation in ProComp Base-Building Incentive Earnings per Teacher, Across Study Years

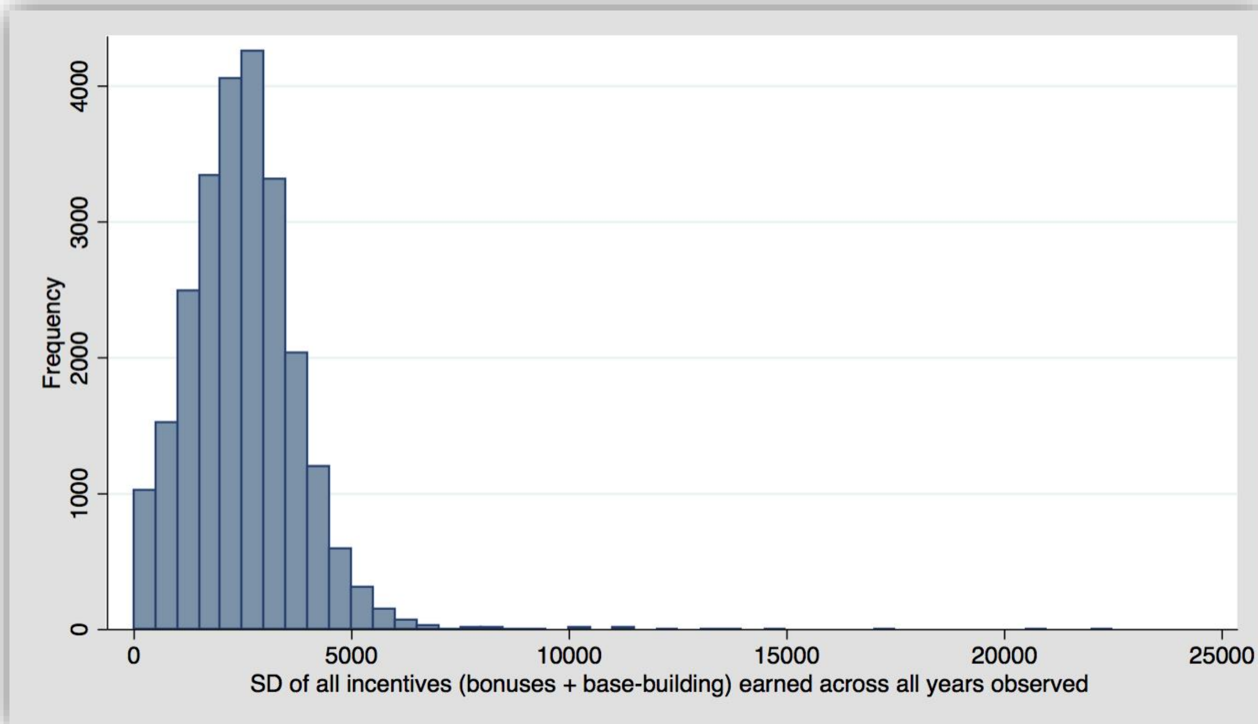


We also sought to simply characterize the average size of the fluctuation in incentive pay each teacher experienced during the study years. To do so, we calculated the standard deviation of incentive pay for each teacher. The standard deviation is the average amount (in dollars) that the given teacher's incentive pay changes from one year to the next. If, for instance, a teacher had a standard deviation in incentive pay of $0, she would have received the exact same amount, in dollars, of incentive pay in every year. On the other hand, a teacher with a standard deviation of $2000 would expect that her incentive pay could fluctuate, on average, by $2000 every year. Teachers with large standard deviations have experienced a great deal of variation in incentive pay during their time in Denver, while teachers with small standard deviations have experienced relatively stable incentive pay from one year to the next.

Figure 13 presents the distribution of each teacher's standard deviation of incentive pay during the study year. The x-axis represents the standard deviation of incentive pay (in $500 increments), and the y-axis represents the numbers of teachers with the given amount of variation in their pay. For instance, approximately 1,200 teachers have a standard deviation of ProComp pay of $500 or less. Another 1,100 teachers have a standard deviation of incentive pay between $4,000 and $4,500—a significant amount of variability from one year to the next. As the figure indicates, although some teachers experienced standard deviations of incentive pay in excess of $6,000, most had standard deviations in incentive pay of between $1,500 and $4,000.[12]

[12] We were concerned that teachers who were newer to Denver—and thus had participated in ProComp—would have greater standard deviations simply because they had fewer years of ProComp incentive pay data. We therefore conducted a sensitivity analysis to ensure that the number of years each teacher participated in ProComp was unrelated with the standard deviation in pay. Results from that analysis are presented *Appendix C: Is Variation in ProComp Payments Related to Number of Years Observed*.

Figure 13. Histogram of Teachers' Standard Deviation of Incentive Pay across Study Years.



Overall, Figure 13 indicates a wide range of experiences across teachers in terms of teachers' variation in ProComp payments. One potential concern that arose from the Year 1 Focus Groups was that the variability in compensation that results from a pay-for-performance system can be frustrating. It is possible that the instability in pay—though it may arise because teachers are being paid *more* in some years—may be frustrating and may induce teachers to leave for a district with a more stable salary schedule. We explore this question in Focal Area 3.

Summary Variability in Teacher ProComp Payments

A major focus of work from the first year of the internal evaluation was the nature of incentive pay variability and the distribution of bonuses under the ProComp system. A new set of research questions emerged from this work, concerning the ways in which teachers perceive the nature of the incentive system and the year-to-year variation in their own salaries. Without conducing teacher surveys or focus groups, it was not possible to characterize teachers' qualitative perceptions of ProComp payments, their fairness, their relation to effort, and the role of bonus pay in terms of remaining at the district. Nonetheless, we were able to quantify teachers' "experience" of ProComp (e.g. the variation in incentives and total compensation received by a given teacher) and examine the association between this experience and retention rates.

We find that the median size of teacher ProComp payments have increased since the program began in terms of both one-time bonuses (Figure 11) and base-building incentive (Figure 12). In particular, we observe a sizable jump in the size of incentives earned starting in the 2008-09 school year, as ProComp entered its full implementation phase. Moreover, the total size of ProComp payments varies more across teachers in each school year than was originally the case in the early years of ProComp. As variability across teachers widens, teachers may become more aware that some teachers are bigger "winners" than others under a pay-for-performance system like ProComp.

In addition, we also document that individuals are likely to experience large swings in their ProComp payments from one year to the next (Figure 13). The median teacher's standard deviation of ProComp payment is approximately $2,700. This suggests that teachers may find it difficult to anticipate how much their total compensation will vary from one year to the next. That said, high-performing teachers have higher probabilities of receiving bonuses related to student achievement growth, and teachers who have received ExEx in a prior year have a high probability of receiving it again. This finding ameliorates some of the concern that the inherent uncertainty that comes along with a bonus-based compensation policy will differentially affect high-performing teachers.

Focus Area 3: Salary Variability and Retention

Overview of the Research Questions

Given the foregoing findings on increasing variability in teacher compensation, and the District's goal of retaining high quality teachers, we next turn our attention to exploring whether that apparent increased instability in teacher pay is associated with the departure of high-performing teachers. While it seems straightforward that most employees would appreciate having the opportunity to receive bonuses above and beyond their base pay, a bonus system like ProComp is also inherently unpredictable. Teachers will not know precisely how much income they should expect in any particular school year. During the Year 1 teacher focus groups, a theme that came up several times was that this compensation instability could be vexing. Moreover, some teachers voiced the opinion that the relative satisfaction of receiving their ProComp bonuses could be undermined by the challenge of not being able to anticipate one's total pay amount. The concern then arises that the dissatisfaction with that instability could lead some teachers to seek employment in a different district. In this section, we conduct some exploratory analyses to examine whether teachers who experience more instability in their pay from year-to-year are, indeed, more likely to leave Denver.

Methodology

We estimate the probability that a teacher leaves at the end of a given year as a function of the amount of variability in ProComp incentives she has experienced to date. Thus, our statistical model (shown below) uses a logistic regression to test the hypothesis that teachers with greater pay variability are more likely to leave DPS.

$$logit(LeftAfter_{py}) = \beta_0 + \beta_1\left(IncentiveVariability_{py}\right) \qquad \text{(Eq. 2)}$$

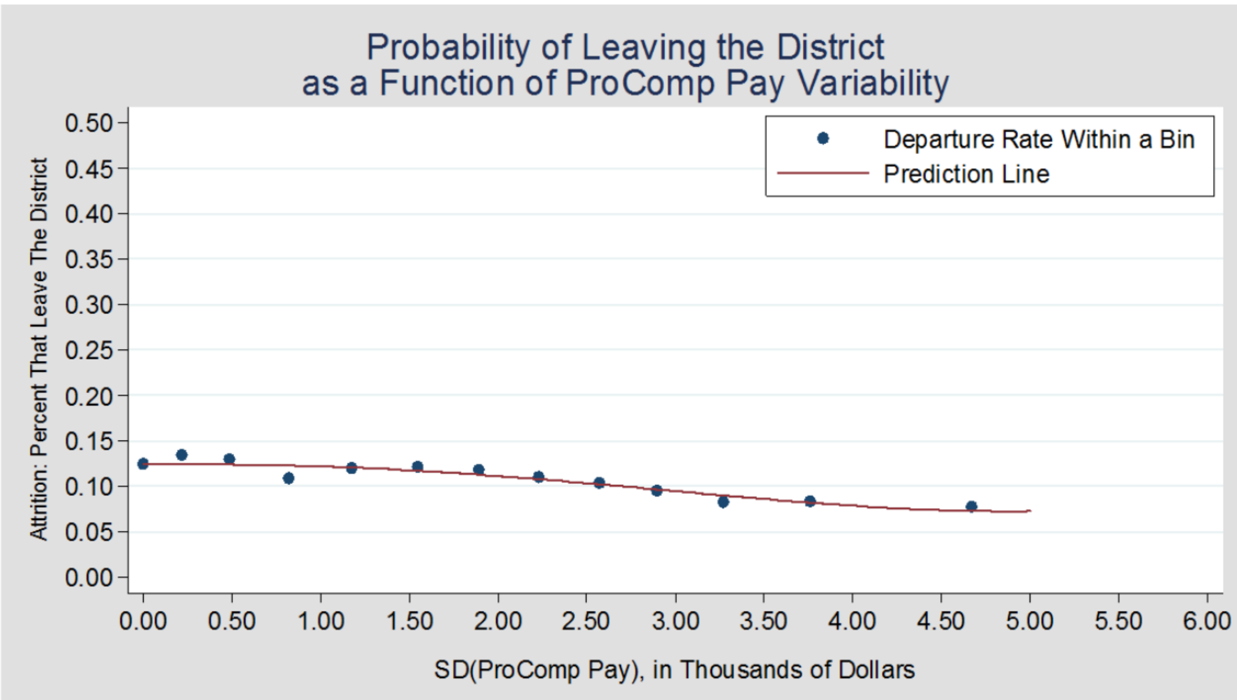
In Equation (2), the outcome, $LeftAfter_{py}$, is a binary outcome variable that is set equal to 1 if the teacher p leaves DPS at the end of school year y. To measure the extent to which the teacher has experienced instability in her pay due to ProComp, we calculate the standard deviation of incentive earnings experienced by a given teacher up to the given year. We use the standard deviation as a proxy for incentive variability, $IncentiveVariability_{py}$. Note that $IncentiveVariability_{py}$ is a time-varying characteristic of the teacher. Over time, she may experience more or less variability in her ProComp incentives. We estimate Equation (2) using logistic regression and examine whether the coefficient on $IncentiveVariability_{py}$ is statistically significant.

Figure 14 presents an illustration of the relationship between departure rates (y-axis) and variation in ProComp incentive earnings (i.e., the standard deviation of ProComp payments, in thousands of dollars) on the x-axis. The solid red line captures the direction of the overall relationship between these two factors.[13] As the graph shows, teachers with greater volatility in ProComp payments do *not* have higher probabilities of leaving the district. On the contrary, there appears to a slightly negative relationship between the two suggesting that greater variability is associated with lower probability of departure. Specifically, for every $1,000 increase in the standard deviation of teacher incentive earnings, we there was almost equal odds of departure (odds ratio = 0.89).[14] It is apparent then, that there is only a weak (negative) relationship between salary instability and attrition from the district. We can see from the figure that a teachers with a

[13] To construct the fitted curve in
Figure 14, we performed a logistic regression and clustered the standard errors at the teacher-level. We chose to cluster the standard errors because we have repeated observations of an individual teacher's decision to remain in the district every school year, and an individual teacher's decision to leave or stay in a given year is not entirely dependent of that same teacher's decision in a previous or subsequent year.
[14] This negative association is statistically significant at the .001 level. We can therefore reject the null hypothesis that there is no relationship between volatility in ProComp pay and teacher departures from the district. This finding holds when we also control for a teacher's MGP score in the given year, or when we include teacher fixed effects.

standard deviation of pay of $500 has an estimated probability of departure of about 12 percent, whereas a teacher with a standard deviation of incentive pay 6 times larger ($3,000) has an estimated probability of departure of 10 percent. While a relationship exists, it is weak and it runs opposite to the hypothesized concern. Ultimately, we do not find evidence that greater volatility in ProComp incentive pay is associated with teacher departures from the district.

Figure 14. Probability of Leaving the District as a Function of Variability in ProComp Payments.



Summary Variability and Retention

We find no evidence to support the claim that teachers who have experienced more variability in ProComp payments to date are more likely to leave the district (if anything, higher variability is associated with lower propensities to leave the district). We caution against making causal claims here—e.g., we are not able to definitively claim that teachers would have been more likely to remain in the district had they experienced greater ProComp payment fluctuations.

However, the first order evidence appears inconsistent with the concern that volatility in pay is associated with a level of frustration that causes high-performing teachers to leave the district.

Appendix A: Teacher Retention Estimation Results

In Equation (1)—written out again here for ease— we predict variation in the percentage of teachers who are retained in the district as a function of time over the last twelve study years.

$$Retained_{py} = \beta_0 + \beta_1(Year_{py}) + \beta_2(Post05_{py}) + \beta_3(Year_{py} \times Post05_{py}) + \beta_4(Post08_{py}) + \beta_5(Year_{py} \times Post08_{py}) + \varepsilon_{py} \quad \text{(Eq. 1)}$$

In Equation (1), the outcome variable is $Retained_{py}$, which is a dummy variable set equal to 1 if the given teacher p returns to the district in the Fall of year y, and it is set equal to 0 if the given teacher does not appear in the district in the subsequent year(s). We predict variation in percentage of teachers who remain the district (i.e., percentage retained) as a function of time, captured by the variable $Year_{py}$. We chose to center the year variable on the 2004-05 school year, so that $Year_{py}$ equals 0 in the 2004-05 school year—that is, the last year in which ProComp was not in effect (i.e., the last year of the first time segment. $Year_{py}$ equals 1 in the first year of ProComp implementation, equals 2 in the second year, equals 3 in the third year, etc. We do so for purely to simplify the estimation of statistical significance, and the centering has no substantive impact on the estimates. We interact the $Year_{py}$ variable with two post-period dummy variables, which allows us to estimate a separate retention trend in each period: $Post05_{py}$ is a dummy variable indicating whether or not the school year is 2005-06 or later (1), and $Post08_{py}$ is a dummy variable indicating whether or not the school year is 2008-09 or later. Each of these dummies are included in the model and interacted with the centered $Year_{py}$ variable. We estimate each coefficient (that is, each beta "β") using ordinary least squares regression and present estimation results in the following tables.

It is useful to walk through what each coefficient represents. The equation coefficients allow us to focus on the differences between periods: Specifically, β_0 captures the percentage of teachers who were retained in DPS in Fall 2005, and β_1 captures the linear trend in retention rates across school years in the pre-ProComp period. β_2 captures the jump up (or drop down) in retention rates from the year just before to just after the start of ProComp in 2005-06. If β_2 is statistically significant, it suggests that there may have been an immediate change in DPS retention rates in the first year ProComp started. β_3 captures whether the linear trend in retention rates in the mid-period (i.e., Fall 2006 through Fall 2008) were any different than the linear trend before ProComp. If it is not statistically significant, we cannot reject the hypothesis that trends in retention did not change between time segment 1 and 2. β_4 and β_5 follow the same interpretation as β_2 and β_3, but now for time segment 3 (i.e., full and mandatory implementation of ProComp in 2007-08 and after). β_4 is the estimated difference in annual retention rates between Fall 2005 and Fall 2009, if any exists. Finally, β_5 captures whether the trend in retention rates in the third time segment (i.e., Fall 2009 through Fall 2013) was any different than the linear trend before ProComp (prior to Fall 2006).

In Column (1) of Table 3 below, we present the estimated coefficients from the above teacher retention prediction model. Note that these estimates correspond directly to the function (red line) in Figure 1—a visual approach to presenting the exact same information. We begin by examining estimated coefficients pertaining to the pre-ProComp period: In Fall 2005, we estimate that 87.2 percent of teachers were new to the district. The annual change in retention rates during the pre-period (Fall 2005 and before) was negative: the rate of teacher retention decreased by an average of 1.6 percentage points each year.

We now turn to the results related to the mid-period (Fall 2006 – Fall 2008). In Fall 2006, the first year following the onset of ProComp, the teacher retention rate was 86.7 percent (a small, not statistically significant change from the previous year). The coefficient on "post", β_2 is estimated to be -0.005, which reflects the fact that retention rates dropped by 0.5 percentage points from the year just before to just after the start of ProComp in Fall 2006. In the mid-period, the annual trend in retention rates was 0.4 percentage points increase each year—i.e., close to flat. The coefficient on "timeXpost06", $\beta_3 = 0.020^{***}$ is statistically significant, which means that we can reject the hypothesis that trends in retention did not change between time segment 1 and 2. This mirrors what we can observe from Figure 1, in which we see an abrupt change in retention trends right around the onset of ProComp.

Finally, we explicate results related to the third time period, Fall 20098 and after. β_4 (the estimated coefficient on "post08") and β_5 (the estimated coefficient on "timeXpost08") follow the same interpretation as β_2 and β_3, but now for the post-period (i.e., full and mandatory implementation of ProComp). The retention rate in Fall 2008 was 90.6 percent—a small jump up when compared to retention in the last year preceding ProComp's implementation in Fall 2005 (87.2 percent). The annual change in retention between Fall 2008 and Fall 2013 is negative (a 1.8 percentage point decrease each year), and quite similar to the trend observed in the pre-ProComp period.

Note that the subsequent columns (2 through 10) Table 3 can be interpreted analogously to those described above. In each column, we limit the analytic sample to a relevant subset of the teacher population, and these correspond directly to the figures presented throughout the discussion of Focus Area 1. For instance, the results presented in Columns 2, 3, and 4 of Table 3 are reflected in Figure 4; results in Columns 5 through 8 are reflected in Figure 5; and the results in Columns 9, 10, and 11 are reflected in Figure 6 .

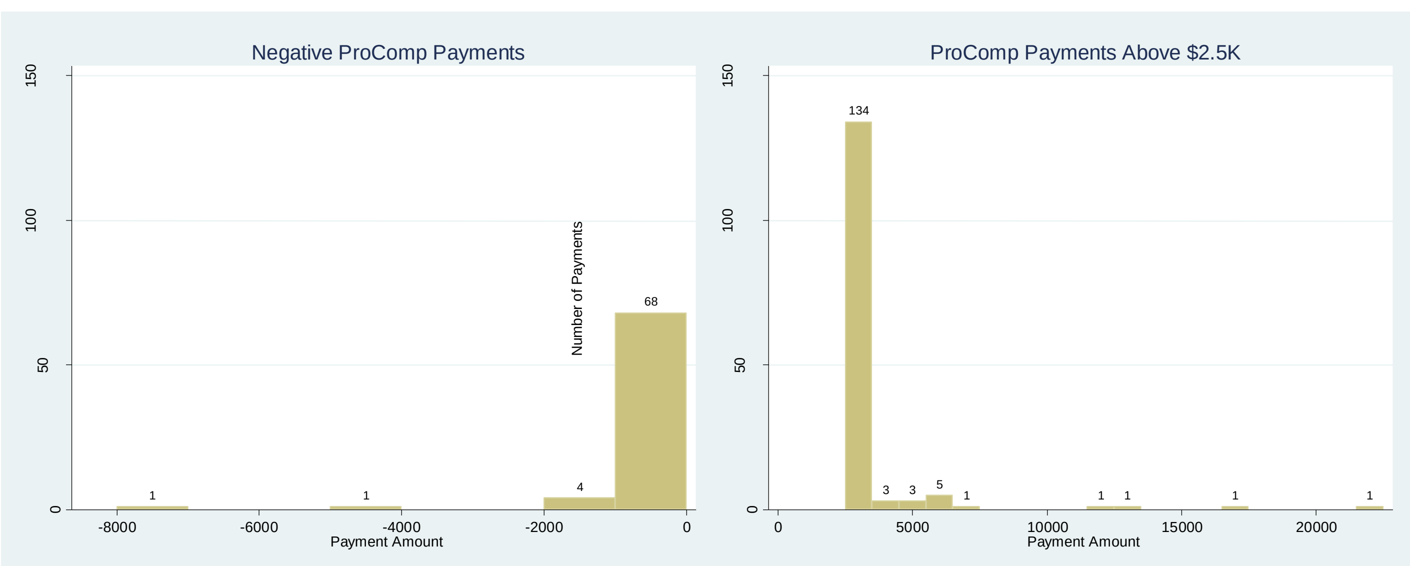
Table 3. Estimation Results for Teacher Retention, by Teacher and School Subgroups

	All Eligible Participants	Elem. Schools	Middle Schools	High Schools	Hard-To-Serve	Not Hard-To-Serve	High Minority	Not High Minority	High Poverty	Not High Poverty	Turnaround Schools	Non-Turnaround
time	-0.016 ***	-0.012 ***	-0.024 ***	-0.022 ***	-0.023 ***	-0.016 ***	-0.020 ***	-0.016 ***	-0.016 ***	-0.016 ***	-0.018 **	-0.016 ***
	(0.002)	(0.003)	(0.007)	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.007)	(0.002)
post	-0.005	0.004	-0.029	-0.023	0.021	-0.011	-0.012	-0.011	-0.004	-0.011	-0.043	-0.001
	(0.008)	(0.011)	(0.027)	(0.020)	(0.018)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.025)	(0.009)
post08	0.106 ***	0.086 ***	0.136 ***	0.101 ***	0.133 ***	0.098 ***	0.099 ***	0.098 ***	0.112 ***	0.098 ***	0.174 ***	0.091 ***
	(0.010)	(0.013)	(0.035)	(0.024)	(0.020)	(0.014)	(0.013)	(0.014)	(0.014)	(0.014)	(0.028)	(0.011)
timeXpost	0.020 ***	0.009	0.041 **	0.038 ***	0.018 *	0.019 **	0.028 ***	0.019 **	0.020 ***	0.019 **	0.031 *	0.018 ***
	(0.004)	(0.005)	(0.013)	(0.010)	(0.009)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.012)	(0.004)
timeXpost08	-0.002	-0.003	0.001	0.006	0.000	0.000	0.003	0.000	-0.004	0.000	-0.016 *	0.001
	(0.003)	(0.004)	(0.008)	(0.007)	(0.006)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.008)	(0.003)
constant	0.872 ***	0.889 ***	0.841 ***	0.853 ***	0.858 ***	0.882 ***	0.863 ***	0.882 ***	0.865 ***	0.882 ***	0.860 ***	0.874 ***
	(0.004)	(0.006)	(0.013)	(0.010)	(0.009)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.012)	(0.004)
Retention Rate in '05	0.872	0.889	0.841	0.853	0.858	0.882	0.863	0.882	0.865	0.882	0.860	0.874
Retention Rate in '06	0.867	0.893	0.812	0.830	0.879	0.871	0.851	0.871	0.861	0.871	0.817	0.873
Retention Rate in '09	0.906	0.915	0.885	0.89	0.899	0.916	0.894	0.916	0.897	0.916	0.898	0.905
Pre-Trend '01-'05	-0.016	-0.012	-0.024	-0.022	-0.023	-0.016	-0.020	-0.016	-0.016	-0.016	-0.018	-0.016
Mid-Trend '06 - '08	0.004	-0.003	0.017	0.016	-0.005	0.003	0.008	0.003	0.004	0.003	0.013	0.002
Post-Trend '09 - '13	-0.018	-0.015	-0.023	-0.016	-0.023	-0.016	-0.017	-0.016	-0.020	-0.016	-0.034	-0.015
R-squared	0.005	0.004	0.009	0.005	0.009	0.004	0.006	0.004	0.007	0.004	0.016	0.004
N	59825	29205	6049	10258	14736	25146	31728	25146	30004	25146	8208	46942
Covariates?	No	No	No	No	No	No	No	No	No	No	No	No

Appendix B: Outlier Incentive Data

A review of the data showed several extreme individual incentive values: We found some one-time, specific bonus awards that were listed at unrealistically high values, including a $25,842 PDU salary incentive that was awarded to Teacher ID XXXXXXXX [redacted] in 2009-2010. We also observe negative paycheck entries listed under ProComp. There are sometimes needs for negative adjustments for logistical reasons (we see about 75 of these), but a handful of those negative entries were greater than -$4,000 (and up to -$8,000). On the other hand, we see about 150 bonus-specific payments that are >$2,500 (and four that are greater than $10,000). Because these values were more than double the scheduled value of any of the individual incentives, and, given the relatively small number of them, we determined that they were likely due to scrivener error. We elected to drop these values because. Despite the relatively small number of them in the sample, they were so far from what was likely the true value that they would have a noticeable effect on the mean. See Figure 15 for a distribution of the frequency and size of these unusual ProComp payments, which we excluded from the analyses presented in this report.

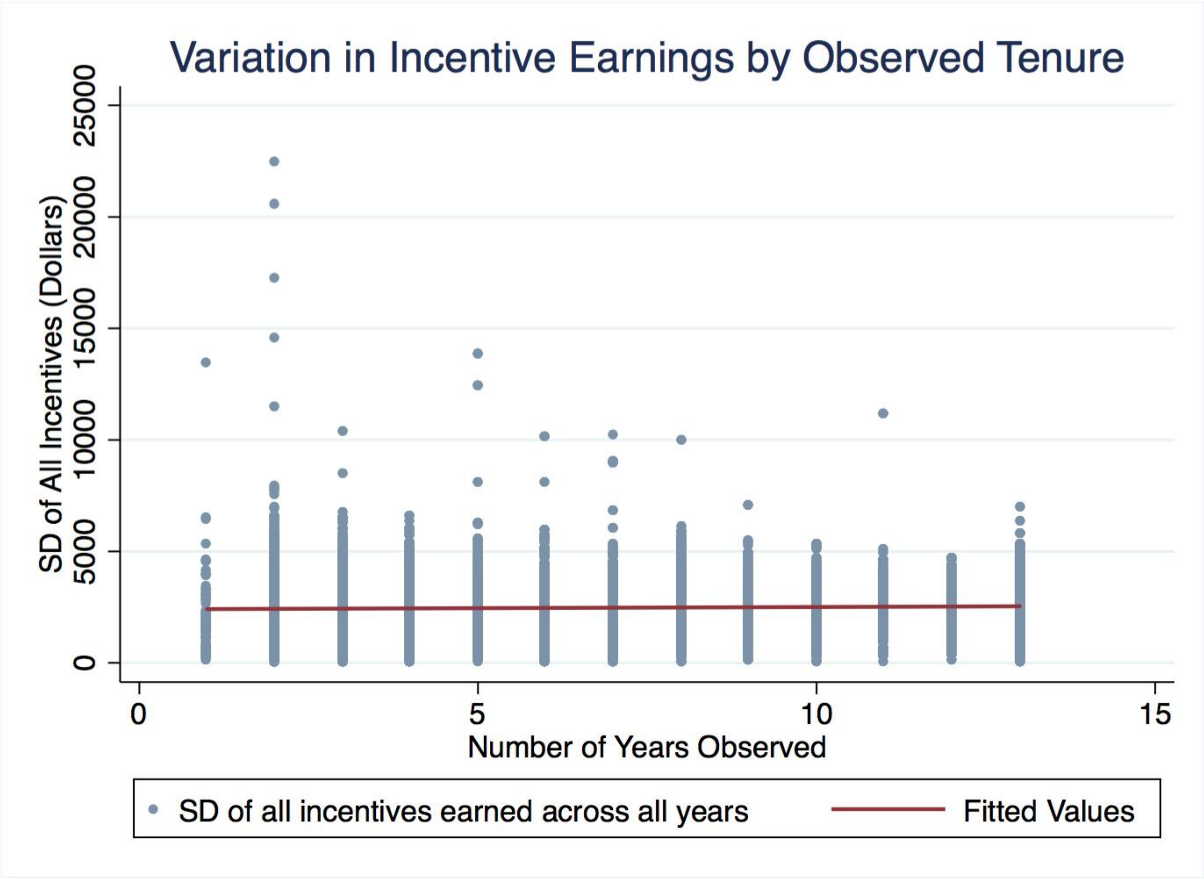
Figure 15. Frequency and Size of Unusual ProComp Payments (Negative or Above $2,500)



Appendix C: Is Variation in ProComp Payments Related to Number of Years Observed

In considering whether the length of a teacher's observed tenure with DPS we plotted the standard deviation of incentive earnings across all observed years for a given teacher against the number of years which that teacher was observed and fitted it with a line. As shown by Figure 5 below, there was an increase of around $10 on average for each additional year a teacher was observed. Though this number is quite small, as a practical matter, it was statistically significant at the 0.001 level. See Figure 16,Figure 16 below.

Figure 16. The Relationship between a Teacher's Variation in ProComp Bonuses (SD of Incentives) and the Number of Years the Teacher is Observed in the ProComp System.

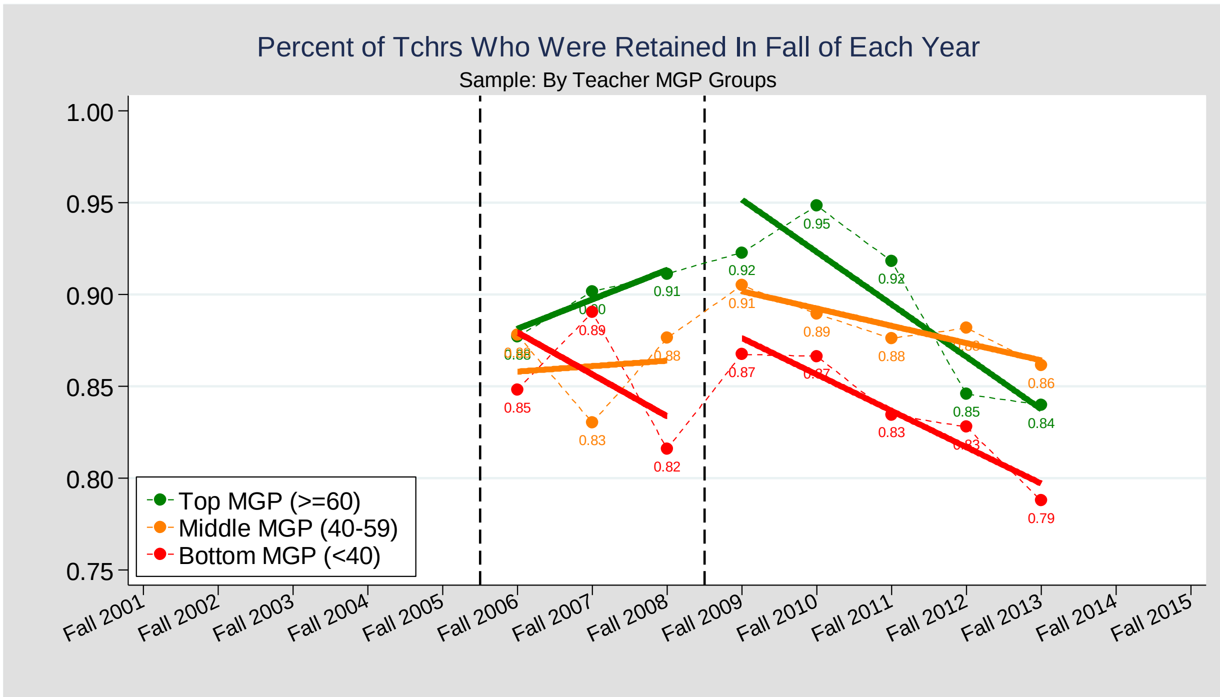


Appendix D: Using MGP Scores to Group Teachers into Performance Categories

In the paper, we follow the practices of DPS to use MGP scores to group teachers into three broad performance categories: high- (>=60 MGP), middle- (40-59 MGP), and low- (<40 MGP) performing categories. We chose to take advantage of all available MGP data for each teacher during the study timeline to characterize his or her teaching performance. We therefore calculate an average MGP for every teacher across all subject and school years in which she is present (we refer to this as "Approach A"). Of the 13,520 unique teachers that appear in our dataset, 4,019 (approximately 28.7 percent) have at least one MGP score and can be grouped using MGP scores. This approach has the advantage of being as inclusive as possible, allowing any teacher who has ever had at least one MGP score to be part of the analysis. On the other hand, one disadvantage is that these averaged MGP scores may be more or less imprecise across teachers, depending on how many MGP scores are available. We therefore also reported results using a second, more restrictive approach to using MGP's to group teachers ("Approach B")—see Figure 7.

We also considered a third approach ("Approach C"), in which we grouped teachers into the three categories simply based on their MGP score in the single, *current* year. This approach has two advantages: First, it puts all teachers' assessment of effectiveness on relatively equal footing (because every teacher is grouped based on exactly 1 MGP score). Second, we need not limit the analysis to a subset of the school years, so we can use all available years of outcome data, as well as all 4,109 teachers who possess at least 1 MGP score. Approach C also the downside of being quite imprecise for everyone, since a single-year MGP score is typically not estimated with precision. Another downside is that a given teacher's effectiveness is now allowed to change from year-to-year, making it more difficult to conceptualize a connection between his or her overall skill and his or her decision to continue teaching. Nonetheless the benefits of Approach C eliminate some of the concerns with both Approach A and Approach B used in the body of the paper, and we therefore reanalyzed the data using Approach C to categorize teachers (see below):

Figure 17. Retention Rates Since Fall 2006, Separately for Top, Middle, and Bottom-Performing Teachers. ("Approach C" to Grouping Teachers)



In Figure 17 ("Approach C" to grouping), the results look quite similar to those reported in Figure 6 ("Approach A" to grouping): While all three groups experienced downward trends in the post-period, low-performing teachers were consistently the least likely to be retained. Because Figure 17 is based on the same sample of 4,109 teachers as Figure 6 (but different ways of grouping those teachers), it is not entirely surprising that the results are somewhat similar. These findings differ meaningfully from the findings presented in Figure 7, which used the most restrictive sample of 165 teachers and "Approach B" to group them into categories.

Appendix E: Overview of ProComp Incentives

The ProComp system includes 10 distinct financial incentives, divided into four broad categories: Knowledge and Skills, Student Growth, Market Incentives, and Comprehensive Professional Evaluations. The dollar amounts awarded for each of the 10 ProComp incentives are based on a pre-negotiated percentage of an overall index, which has ranged from $33,300 in the 2005-2006 school year to $38,118 in the 2013-2014 school year[15].

Payment for meeting the various ProComp incentives comes in the form of either a one-time bonus or a permanent, "base-building" salary increase, depending on the incentive and the teacher's years of experience. Of the ten incentives currently offered by DPS, six are offered only as a one-time bonus, two are offered only as base-building salary increases, and two are applied as either a bonus or a salary increase, depending on the circumstances. In addition, three of the ten bonuses are awarded at the school-level (High Growth, Top Performing, and Hard to Serve), while the rest are awarded based on teacher-level criteria.

All teachers and student service professionals (SSPs) who are covered by the DCTA collective bargaining agreement are eligible to join ProComp. ProComp-eligible employees include social workers, psychologists, school librarians, nurses, therapists, and intervention teachers, in addition to conventional classroom teachers. Charter school employees, however, are not eligible to join ProComp. For the purposes of this document, the term "teachers" refers to all ProComp-eligible educators.

The ProComp system officially began during the 2005-06 contract year, and all DPS teachers hired after January 1, 2006 are automatically enrolled in ProComp. Eligible professionals who

[15] Although information regarding the amount of each incentive is available via the ProComp website (http://denverprocomp.dpsk12.org/history) for the 2014-15 school year, in this report we only address those years for which we had data (i.e. 2005-06 through 2013-14).

were hired before 2006 were allowed to opt into ProComp during seven "windows" from 2005 through 2011. Table 4 shows that the number of ProComp participants has increased in recent years.

Table 4. ProComp Participation Rates, by School Year

	2005-06	2006-07	2007-08	2008-09	2009-10	2010-11	2011-12	2012-13	2013-14
Total Eligible Teachers	4,532	4,536	4,552	4,761	4,967	4,995	5,089	5,210	5,364
Total Participants	731	1,757	2,280	3,127	3,515	3,791	4,179	4,393	4,673
Participation Percentage	16.1%	38.7%	50.1%	65.7%	70.8%	75.9%	82.1%	84.3%	87.1%

Table 5 highlights the reason for this increase as new hires, for whom participation in ProComp was compulsory, replaced veteran teachers over time. Table 5 also shows the extent to which participation rates increased among those hired prior to 2006, both through opting-in and attrition of non-participants. The rows represent school years; the columns represent teacher cohorts, defined by Spring of the year in which the teacher first appeared in our data set. So, for example, the 2004 cohort (those teachers not present in the Spring of 2003 who were present in the Spring of 2004) had 12.0% of its teachers enrolled in ProComp in the inaugural 2005-06 school year. However, by the 2013-14 school year, the last year for which we have data, more than half (55.3%) of the remaining 2004 cohort members were enrolled in ProComp.

Table 5. Participation Rates across School Years, by Teacher Cohort

	Teacher Cohort (Spring of Year of First Appearance)												
School Year	Before '03	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
2005-06	19.1%	11.5%	12.0%	10.1%	8.5%	NA	NA	NA	NA	NA	NA	NA	NA
2006-07	34.3%	22.2%	20.8%	17.5%	25.9%	98.5%	NA	NA	NA	NA	NA	NA	NA
2007-08	38.8%	23.9%	25.6%	24.9%	34.9%	99.5%	99.7%	NA	NA	NA	NA	NA	NA
2008-09	52.5%	35.5%	36.2%	38.8%	52.3%	96.7%	98.9%	98.3%	NA	NA	NA	NA	NA
2009-10	55.1%	39.7%	41.1%	40.6%	55.5%	96.8%	98.5%	98.6%	100.0%	NA	NA	NA	NA
2010-11	60.6%	47.3%	43.1%	48.3%	60.5%	96.1%	98.4%	98.5%	100.0%	99.4%	NA	NA	NA
2011-12	68.3%	52.2%	61.3%	58.0%	66.5%	95.8%	97.8%	98.5%	99.7%	99.6%	97.9%	NA	NA
2012-13	68.8%	54.4%	58.0%	58.6%	67.3%	96.0%	97.9%	98.5%	99.4%	99.7%	98.1%	99.0%	NA
2013-14	69.4%	57.0%	55.3%	61.3%	68.9%	96.7%	98.0%	98.6%	98.6%	99.7%	100.0%	99.6%	99.5%

As Table 6 below shows, Student Growth Objectives (SGOs) were the most commonly earned incentive in the last year for which we have data, followed by Professional Development Units (PDUs). Only a small minority of teachers earned the Exceeds Expectations (because this requires that a teacher be in a tested grade and subject, which is true of a minority of teachers) or Advanced Degree incentives, while most of the other ProComp incentives were achieved by 30-60% of teachers in each year for the period beginning with the 2008-2009 school year, when ProComp's current iteration took effect.

Table 6. Percentage of ProComp Participants Who Earn each Incentive, by School Year

		2005-06	2006-07	2007-08	2008-09	2009-10	2010-11	2011-12	2012-13	2013-14
Student Growth	Met 1 or 2 SGOs	0.0	0.0	13.8	59.8	74.3	80.9	72.1	74.6	68.2
	Exceeds Expectations	0.0	0.0	0.0	7.4	10.8	11.5	11.1	12.6	11.0
	High Growth	0.0	0.0	0.0	30.6	49.0	43.0	42.7	45.9	47.4
	Top Performing	0.0	16.8	17.5	30.2	42.7	33.5	39.4	41.4	44.6
Market Incentives	Hard-to-Serve	62.8	58.2	38.0	57.0	52.1	58.8	59.2	57.7	58.6
	Hard-to-Staff	32.9	42.7	42.1	37.6	38.4	37.3	36.9	32.1	31.3
Knowledge and Skills	Advanced Degree	30.1	12.3	8.2	9.3	8.0	7.9	7.1	8.7	9.3
	Tuition/School Loan Reimbursement	28.1	11.7	7.2	19.8	26.0	29.5	31.7	28.2	26.6
	PDUs	0.0	0.0	59.4	65.9	62.4	27.8	45.8	67.8	61.5
Evaluation	CPEs	0.0	0.0	2.4	53.8	62.7	52.7	47.8	38.3	43.3
Total ProComp Participants		731	1,757	2,280	3,127	3,515	3,791	4,179	4,393	4,673

Student Growth Incentives

The student growth component of ProComp consists of four separate incentives, two of which are awarded for school-wide achievements, and two of which are individual awards. At the school level, teachers are eligible to earn Top Performing and High Growth incentives; as individuals, teachers can earn Student Growth Objective incentives (SGOs) or an Exceeds Expectations bonus, both of which are based on measures of student achievement. Because the Exceeds Expectations bonus is tied to student performance on state tests, only those teachers in tested grades and subject areas are eligible to receive it. Therefore, a very small percentage of the overall participant population is eligible to receive Exceeds Expectations.

Teacher-Level Incentives: Student Growth Objectives (SGOs) and Exceeds Expectations.

At the beginning of each academic year, each DPS teacher confers with school leaders to establish two goals for their students' progress over the course of the year. These goals, called Student Growth Objectives (SGOs), may incorporate a wide variety of quantitative or non-quantitative measures, including nationally standardized tests, subject area exams created by DPS, or teacher-created tests, among other data sources.

Teachers who achieve both of their SGOs in a given year earn a salary increase equivalent to 1% of the ProComp index, which was $381.18 for the 2013-2014 school year. Teachers who achieve only one SGO earn a one-time bonus. In the period beginning with the 2008-2009 school year, the second year in which this incentive was awarded, more than half of participants received at least one SGO incentive each year, and in most years, it was much more common for participants to meet both SGO's.

Table 7. Percentage of Participants Earning SGO Incentives, by School Year

	2005-06	2006-07	2007-08	2008-09	2009-10	2010-11	2011-12	2012-13	2013-14
Did Not Meet SGOs	N/A	N/A	86.1	40.2	25.7	19.0	27.9	24.9	31.7
Met 1 SGO	N/A	N/A	10.7	8.3	10.8	12.1	72.1	11.5	10.5
Met 2 SGOs	N/A	N/A	3.2	51.5	63.5	68.9	0.00	63.6	57.8
N ProComp Participants	731	1,757	2,280	3,127	3,515	3,791	4,179	4,393	4,673

The Exceeds Expectations (ExEx) award is offered to teachers whose students achieved substantial growth on the applicable state assessment (e.g. Colorado State Assessment Program (CSAP) or Transitional Colorado Assessment Program (TCAP) exams). Because students take these exams only in certain grades and subjects, the ExEx incentive is available only to teachers in grades 4 through 10 who teach mathematics or language arts. The incentive is a one-time bonus of 6.4% of the index, which was $2,439.55 in the 2013-2013 school year. Table 8 shows the percentage of eligible (those with MGPs in at least one subject for a given year) teachers as well as total participants receiving ExEx for each year of the study.

Table 8. Percentage of Participants Earning ExEx, by School Year

	2005-06	2006-07	2007-08	2008-09	2009-10	2010-11	2011-12	2012-13	2013-14
Eligible Subjects/ Grades	0	0	0	25.9	34.2	36.2	29.7	34	31.3
All Participants	0	0	0	7.4	10.8	11.5	11.1	12.6	11
N ProComp Participants	731	1,757	2,280	3,127	3,515	3,791	4,179	4,393	4,673

School-wide Bonuses: High Growth and Top Performing.

Both the Top Performing and High Growth bonuses are based on the School Performance Framework (SPF), which rates schools in seven performance categories: Academic Growth, Academic Proficiency, College & Career Readiness, Improvement in College & Career Readiness Over Time, Student Engagement, Enrollment Rates, and Parent Satisfaction. These seven categories encompass dozens of variables, including CSAP and TCAP scores, parent satisfaction surveys, dropout rates, and attendance rates, among many others.

The Top Performing incentive is awarded to schools ranked in the top half of the annual SPF ratings. The High Growth incentive, however, is based exclusively on schools' Academic Growth score, which is determined by median student growth percentiles (MGP's) for each tested subject. Teachers earned a one-time bonus of 6.4% of the index ($2439.55 for the 2013-2014 school year) for each incentive.

As Table 9 illustrates, the number of teachers receiving the Top Performing incentive increased considerably in the 2009-2010 school year. Moreover, an overwhelming majority of teachers who earned the Top Performing bonus also earned the High Growth incentive. For example, during the 2013-2014 school year, 43.5% of all ProComp teachers received the Top Performing bonus, while 39.4% of teachers earned both the High Growth and Top Performing incentives; in other words, only 4% of teachers earned the Top Performing incentive but not the High Growth award.

Table 9. Percentage of Participants Earning High Growth and Top Performing Incentives, by School Year

	2005-06	2006-07	2007-08	2008-09	2009-10	2010-11	2011-12	2012-13	2013-14
Neither	N/A	90.2	85	65.4	50.6	55.3	54.9	51.5	49.6
Top Performing	N/A	9.8	15	29.4	41.7	32.6	38.2	40.3	43.5
High Growth	N/A	N/A	N/A	29.7	47.8	41.8	41.4	44.7	46.2
Both Top Perf. & High Growth	N/A	N/A	N/A	24.5	40.1	29.7	34.6	36.5	39.4
Total Number of ProComp Participants	731	1,757	2,280	3,127	3,515	3,791	4,179	4,393	4,673

Teachers in elementary schools were much more likely to receive the Top Performing and High Growth awards than their counterparts in middle and high schools. As shown in Table 10, 56 percent of teachers in elementary schools received the Top Performing bonus during the 2013-14 school year, compared with only 33 and 38 percent of teachers in middle and high schools. Receipt of the High Growth Incentive is similarly varied across school levels. This suggests that teachers in elementary schools disproportionately benefit from certain ProComp incentives. While the High Growth incentive is based on MGPs, it is a school-wide award, so the relative percentage of teachers teaching in tested grades/subjects should not be the sole reason for the disparity.

Table 10. Percent of Participants Earning School Wide Awards for 2013-14, by School Level

	Top Performance	High Growth	Total Participants
Primary	56.1	58.6	2209
Middle	32.7	33.9	336
High	38.2	47.0	615
Alternative/Other	26.0	21.8	289

Market Incentives

Hard-to-Serve (School-Wide) and Hard-to-Staff (Teacher-Level).

The Hard-to-Serve bonus is designed to encourage DPS teachers to accept positions in high-needs schools. The Hard-to-Serve incentive is offered to teachers as a one-time bonus of 6.4% of the index ($2,439.55 for the 2013-14 school year) for working in schools with a high percentage of students living in poverty.

Like the Hard-to-Serve incentive, the Hard-to-Staff incentive was also paid as a one-time bonus. Each year, DPS designates positions as Hard-to-Staff if the supply of licensed professionals is low and the rate of turnover is high. In recent years, the list of eligible Hard-to-Staff positions has included special education, ELA-S, and mathematics assignments, among others[16].

During the last year of the study, more than half of all ProComp participants received the Hard-to-Serve bonus, while slightly less than one-third of participants served in Hard-to-Staff positions (see Table 11). Teachers who received the Hard-to-Staff incentive were somewhat more likely than other DPS teachers to receive the Hard-to-Serve bonus. In 2013-14, approximately two-thirds of teachers serving in Hard-to-Staff positions also received the Hard-to-Serve Bonus.

[16] For a complete list of the most recent Hard to Staff positions, please visit the ProComp website at http://denverprocomp.dpsk12.org/H2Staff2011-12.

Table 11. Percentage of Participants Earning Hard-to-Serve and Hard-to-Staff Incentives, by School Year

	2005-06	2006-07	2007-08	2008-09	2009-10	2010-11	2011-12	2012-13	2013-14
Hard-to-Serve	62.8	58.2	38.0	57.0	52.1	58.8	59.2	57.7	58.6
Hard-to-Staff	32.9	42.7	42.1	37.6	38.4	37.3	36.9	32.1	31.3
Earned Both	18.8	21.0	16.1	23.8	22.0	23.9	24.4	21.4	21.0
Earned Either	77.0	80.0	63.9	70.8	68.5	72.2	71.7	68.5	68.9
Total ProComp Participants	731	1,757	2,280	3,127	3,515	3,791	4,179	4,393	4,673

Knowledge and Skills

Higher Education Incentives: Advanced Degrees.

ProComp participants receive a base-building salary increase of 9% of the index for earning each degree or approved professional license beyond a bachelor's degree. During the 2013-14 school year, earning an advanced degree incentive added $3,430.62 to a teacher's base salary. Teachers are eligible to earn this incentive only once in any three-year period. Since the 2007-08 school year, less than 10% of teachers have received the incentive in any given year (see Table 6). For the 2013-14 school year, 426 teachers, or 9.3% of participants, received this incentive.

Higher Education Incentives: Tuition Reimbursement.

In addition to the salary increase for completing an advanced degree, DPS also offers teachers a Tuition Reimbursement incentive, which can be used to pay for preexisting student loans, fees for conferences and professional development workshops, or tuition for an advanced degree program. Teachers can earn up to $1,000 per year in Tuition Reimbursement, with a lifetime cap of $4,000. During the 2013-14 school year, 946 teachers, 20.8% of participants, received reimbursement for student loans; 271 teachers, 5.9% of participants, received reimbursement for current tuition expenses (see Table 6).

Professional Development Units.

In an effort to encourage ongoing training, ProComp participants receive an incentive of 2% of the index ($762.36 for the 2013-14 school year) for completing approved Professional Development Units (PDUs). Teachers who have 14 or fewer years of credited tenure with DPS earn a base-building salary increase for successfully earning the PDU incentive; teachers with more than 14 years of service earn a one-time bonus.

Teachers may complete multiple PDUs, but they may only earn a single award in any given school year. However, PDU credits are "bankable", meaning that a teacher who completes two PDUs in a given school year can earn the award for the second completed PDU in a subsequent school year. This has been a relatively highly attained incentive in most years, with 61% of teachers receiving the award in the 2013-14 school year.

Evaluation

Comprehensive Professional Evaluations (CPEs).

DPS began to transition from the CPE framework, wherein teacher evaluations were conducted by school administrators and were the sole determinant of receipt of the CPE incentive, to Leading Effective Academic Practice (LEAP) evaluations in 2011-12. Teachers could achieve the CPE incentive by either earning a satisfactory rating on their CPE evaluation or by participating in the LEAP pilot program during that year. The exact payout for teachers who successfully earned the

CPE inventive varies, depending on the teacher's employment status. Probationary teachers—who are either early in their tenure at DPS or have recently earned an unsatisfactory score on a professional evaluation—are evaluated annually, and are therefore eligible for a base-building salary increase each year of 1% of the index ($381.18 in 2013-14). Teachers in their first 14 years who work in Innovation Schools also receive annual evaluations, and are therefore eligible for the CPE incentive each year, regardless of their probationary status. Most non-probationary teachers who do not work in Innovation Schools undergo a professional evaluation once every three years, and are therefore eligible for a base-building salary increase of 3% of the index ($1,144 for 2013-14) for achieving a satisfactory rating on their evaluation. Teachers with more than 14 years of credited tenure with DPS are never eligible for the CPE incentive, regardless of their performance on professional evaluations.

References

Braun, H., Chudowsky, N., & Koenig, J. (2010). Getting value out of value-added: Report of a workshop. In p. e. Committee on value-added methodology for instructional improvement, and accountability (Ed.): Center for Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press

Briggs, D. C. (2012). Making value-added inferences from large-scale assessments. In K. E. M. Simon, & M. Rousseau (Ed.), *Improving Large-Scale Assessment in Education: Theory, Issues and Practice*. London: Routledge.

Donaldson, M. L., & Papay, J. P. (2014). Teacher evaluation for accountability and development. *Handbook of Research in Education Finance and Policy*, 174.

Kane, T. J., McCaffrey, D. F., Miller, T., & Staiger, D. O. (2013). Have We Identified Effective Teachers? Validating Measures of Effective Teaching Using Random Assignment. Research Paper. MET Project. *Bill & Melinda Gates Foundation*.

McCaffrey, D. F. (2012). Do value-added methods level the playing field for teachers. *Carnegie Knowledge Network*.

McCaffrey, D. F., Sass, T. R., Lockwood, J., & Mihaly, K. (2009). The intertemporal variability of teacher effect estimates. *Education Finance and Policy*, *4*(4), 572-606.