Merit pay, case-by-case: Variables affecting student achievement, teacher retention, and the problem of standardized tests

Sung Joong Kim

Abstract: This study encompasses the history, definition, and implementation of merit pay to identify the variables that affect student achievement and teacher retention. By reviewing 13 studies on American education relevant to the policy, this study aims to determine how merit pay influenced student achievement and teacher retention rates. In the United States of America (USA), merit pay effectively improves student achievement and teacher retention rates under certain circumstances. However, this policy could also be ineffectual. An analysis of a review of 13 studies identified the variables that affected student achievement and teacher retention. These include year, location, policy duration, merit pay type, students’ grade level, and participating teachers’ ethnicities. The conclusions drawn from this study suggest implementing an institutionalized evaluation system that fosters the individual growth of students and teachers and an American educational administration that considers each school’s needs and situations to increase the merit pay policy’s effectiveness and fairness.

Subjects: Educational Research; Education Studies; Education Policy & Politics

Keywords: Merit pay policy; student achievement; teacher retention; standardized tests; American education

ABOUT THE AUTHOR

Sung Joong Kim is an assistant professor of educational administration at Presbyterian University and Theological Seminary in South Korea. He teaches Christian education, educational administration and management, and leadership. He is also in charge of global ministry at the Glocal Ministry Center at Presbyterian University and Theological Seminary. His research interests include educational administration and management, church administration and management, educational policy, and leadership. He has published fourteen books and articles in Cogent Business & Management, Mission and Theology, Journal of Christian Education in Korea, Korea Presbyterian Journal of Theology, The Journal of Asian American Theological Forum, and Journal of Youngsan Theology. He has an Ed.D. in education from the University of Florida, an S.T.M. in theology from Boston University, and an M.A. in Christian education and an M.Div. from Presbyterian University and Theological Seminary, and a B.A. in theology from Yonsei University.

PUBLIC INTEREST STATEMENT

Currently, the merit pay policy is a crucial policy being implemented in companies and schools. Two persistent issues present in American education are low student achievement and low teacher retention. Merit pay policy in education aims to raise student achievement and teacher retention. However, to increase the merit pay policy’s effectiveness and fairness, variables that affect student achievement and teacher retention should be determined. Therefore, this study aims to identify the variables that influence merit pay policy on student achievement and teacher retention. By determining these variables, this study suggests an institutionalized evaluation system that fosters students’ and teachers’ individual growth and presents an educational administration tailored to the needs and situations of educational fields with diverse cultures, environments, and components.
1. Introduction

Two persistent issues in present American education are low student achievement and low teacher retention. In the 2011 Program for International Student Assessment (PISA) rankings, the United States (US) placed 32nd in mathematics with a 32% student proficiency rate, where 22 countries significantly outperformed them based on the number of proficient students (Peterson et al., 2011). In 2015, American students placed 35th (Kennedy, 2016). For the reading component, American students ranked 17th with a 31% proficiency rating (Peterson et al., 2011) and were similarly ranked in 2011 (Kennedy, 2016). According to the latest data, American students ranked 37th in mathematics with an average score of 478 in 2018. Compared to the 2003 mathematics PISA score, the average mathematics score in 2018 (478) was lower than the 2003 average (483) and the 2018-OECD average mathematics score (489) (National Center for Educational Statistics, 2018).

For teacher retention, the national average annual turnover rate from 2000–2010 was around 15%, which saw 40-50% of teachers leaving within the first five years of teaching (Halstead, 2013). Although comparable to the turnover rates of the late 1980s to mid-1990s (12.4% to 13.5%; Johnson et al., 2005), the recent turnover rates are higher than the previous decade (Digest of Education Statistics, 2017). The Department of Education reported that in the 2017–2018 school year: “a majority of states lack teachers in mathematics (47 states and the District of Columbia), special education (46 states and the District of Columbia), science (43 states), and world languages (40 states and the District of Columbia)” (Strauss, 2017). Currently, the turnover rate of new teachers ranges from 19% to 30% during their first five years of teaching (Mulvahill, 2019).

One reform proposed by numerous schools that aimed to address these problems was merit pay; a policy that positively correlated with student achievement and teacher retention (G. Ritter et al., 2016). Also known as performance-based pay, the policy aimed to “adjust salaries to reward higher levels of performance,” according to the American Association of School Administration (van Looszen, 1983, 8). This principle is also evident in Kim’s research (2003), which revolved around merit pay at Korean junior colleges. In Kim’s study, merit pay was shown to attract and retain employees to accomplish an organization’s mission, motivate and reward job performance, and differentiate salary levels based on relative performance and contribution (Kim, 2003).

This study reviewed 13 representative studies published from 1999–2015 to identify the variables that impact merit pay effectiveness for student achievement and teacher retention. The merit pay system’s effects were investigated by observing changes in student achievement and teacher retention rates. Specifically, 8 of 13 studies published from 2008–2012 were used because the Race to the Top program was implemented in 2009. The merit pay policy was nationally implemented through the Race to the Top program in 2009. Thus, studies on the merit pay program focused before and after 2009 were chosen for this study to verify the Race to the Top program’s effectiveness. The latest studies on specific merit pay programs were conducted from 2015, excluding meta-analysis studies, which comprehensively analyse previous studies on merit pay and explain why more recent studies on specific merit pay programs are not used. Each of the 13 studies was reviewed to determine the variables influencing the merit pay policy’s successes and failures through comparative analysis. The regions’ environment and conditions where the merit pay policy was implemented, the specific contents of the merit pay programs, and the merit pay program members’ characteristics were different. By individually examining the 13 studies that have investigated the effectiveness of each merit pay program, the features and conditions in which merit pay’s effectiveness appears in terms of student achievement and teacher retention were found. Next, the variables affecting student achievement and teacher retention were identified by comparing and analyzing each study’s results.

Based on the review and analyses, this study proposes improvements to the policy’s current implementation by identifying the problems with standardized tests. Namely, these tests cannot objectively verify and assess the teacher’s competency and cannot cover all subjects that students learn. Thus, this study suggests that an institutional evaluation system that fosters the individual
growth of students and teachers would be preferable. Finally, this study proposes an American educational administration that considers each school’s needs and situations to improve the merit pay policy’s effectiveness and fairness.

Consequently, this study aims to answer the following research questions. (a) What merit pay policy variables affect student achievement and teacher retention? (b) What are the problems of standardized tests relative to the merit pay policy? (c) What can be done to increase the effectiveness of the merit pay policy?

2. History and current state of merit pay in America
First officially introduced in 1908 for K-12 teachers in Newton, Massachusetts, merit pay in public education peaked in the early 1920s when half of the US school districts followed suit. By the 1930s, most school districts used uniform salary schedules. After Sputnik launched in 1957, concerns about the US educational system led to the revival of merit pay. By 1968, 11.3% of US school districts followed a merit pay system, though this declined during the 1970s. After the publication of “A Nation at Risk: The Imperative for Educational Reform,” in 1983, merit pay issues were reawakened by calls for educational reform for academic excellence (Kim, 2003). The report, published by the National Commission on Excellence in Education in 1983, showed that American schools were failing to improve student achievement and create a competitive workforce, and it aroused a desire for reforms at local, state, and federal levels (A Nation at Risk: The Imperative for Educational Reform, 1983). A significant number of public school districts in the United States began reconsidering merit-based pay as an alternative or supplement to the single salary schedule (Podgursky & Springer, 2007).

Similarly, the 2009 Race to the Top program, backed by the US Department of Education, implemented merit pay nationwide to overcome perceived issues (Williamson, 2010). It is a competitive grant program to incentivize and reward states that improve four areas in education: (1) standards and assessments; (2) data collection and use; (3) teacher effectiveness and teacher distribution equity; and (4) struggling schools (United States Department of Education, 2010). The US Department of Education stressed student growth as a relevant measure for evaluating teachers and principals. In turn, efficient educators are defined by student growth (Williamson, 2010). Here, student growth translated to positive changes in student achievement, and student growth data should be a substantial factor when evaluating the summative performance of teachers and principals for at least three evaluation criteria. The evaluation measures should include classroom-based, school-based, district-based, and state-based tools (Washington State Legislature, 2019). Teacher evaluations are thus linked to their income, as the underlying premise of merit pay is that financial incentives may encourage teachers to adopt practices that improve their students’ abilities.

In 2012, former President Obama supported merit pay programs with a US$5 billion budget to encourage state reforms for teaching and commensurate monetary compensation. Buttressed by governmental support, several of the nation’s largest school districts, including New York City, Washington, D.C., and Denver, implemented merit pay programs (Research for Action, 2012). During the 2011–2012 school year, 12.3% of public school teachers earned income from merit pay bonuses and state supplements (Digest of Education Statistics, 2017). By increasing the merit pay budget, teacher salaries increased by 13 percentage points between 2017 and 2018. In 2019, before the COVID-19 pandemic, support for increasing teacher compensation was higher than at any point since 2008 (Henderson et al., 2019).

3. Research data

3.1. Merit pay and student achievement
Studies that show positive correlations between merit pay and student achievement (Figlio & Kenny, Ritter et al., Springer, Lewis et al. (2010), and Sojourner, Mykerezi, & West) were published in 2006, 2008, 2010, and 2014 respectively. Meanwhile, studies with no correlation between merit pay and student achievement (Springer, Ballou et al., Goldhaber & Walch, and Glazerman &
Seifullah) were identified in 2010, 2012, and 2012 respectively. These studies were representative examples that identified the merit pay policy’s effectiveness in terms of student achievement.

In addition, the varied results may occur due to different merit pay programs. One outlier with a mixed correlation is the work of Springer et al. (2008), which utilized data obtained from around 1,200 students from school years 2002–2006 to estimate the effects of the Teacher Advancement Program (TAP) by comparing score gains in TAP and non-TAP schools. The data, retrieved from the Northwest Evaluation Association (NWEA) website, originated from the Measure of Academic Progress (MAP)—one of the most widely used standardized tests in the United States with reading, language, mathematics, and science as its subjects (Springer et al., 2008). The mixed correlation was attributed to the merit pay’s positive effect on elementary student achievement (grades 2–6) and the non-effect on secondary level students (6–10).

Figlio and Kenny’s (2006) study attempted to show that merit pay for teachers resulted in higher Florida Comprehensive Assessment Test (FCAT) scores using survey data obtained from 534 out of 1,319 Floridian public and private schools, which were sponsored by the US Department of Education since 1988. The teacher rankings based on the FCAT determined their bonuses from the Special Teachers Are Rewarded (STAR) plan (Buddin et al., 2007), where the 2006–2007 budget included US$147.5 million, appropriated within the Florida Education Finance Program (FEFP) (Podgursky & Springer, 2007). The data collected included the frequency and magnitude of the participant schools’ salary incentives, which were analysed concerning student achievement (University of Florida News, 2007). The results showed that students in schools with merit pay programs for teachers scored an average of 1–2 percentile points higher on standardized tests.

Similarly, G. W. Ritter et al. (2008) sought the same with the Achievement Challenge Pilot Project (ACPP) program of the Little Rock School District in Arkansas, which awarded substantial year-end bonuses to teachers based on students’ improvement in standardized exams. The comparison and analysis covered students’ test scores in the math, language, and reading components of the Iowa Test of Basic Skills (ITBS) from 2005–2007 and were reported in normal curve equivalent (NCE) units. The students showed improvements in all subjects. In math, students whose teachers were eligible for bonuses outperformed students in schools whose teachers were ineligible by 3.52 NCE points, and the rising effect of merit pay was nearly 7.00 percentile points. In language and reading, with eligibility criteria of 4.56 and 3.29 NCE points, respectively, the rise nearly reached 9.00 percentile points for language and 6.00 percentile points for reading (G. W. Ritter et al., 2008).

The study conducted by Springer, Lewis et al. (2010) focused on schools that were volunteered in the Texas District Awards for Teacher Excellence (DATE), a state-funded program that can provide approximately US$197 million worth of grants to districts for implementing a locally designed merit pay program. Data from the 2008–2009 (Year 1) and 2009–2010 (Year 2) showed that students in DATE schools obtained higher scores on the Texas Assessment of Knowledge and Skills (TAKS) than students in non-DATE schools. Specifically, participation in both years of DATE was associated with improved student achievement with .03 of a standard deviation in math and with .01 of a standard deviation in reading (Springer, Lewis et al., 2010).

Sojourner et al. (2014) focused on teacher Pay-for-Performance (P4P) that occurred after the introduction of the state’s Quality Compensation program (Q-Comp) in 2005. This state-funded merit pay program of US$86 million allowed teachers to earn raises based on merit, additional duties, and student achievement. Their study used existing data from the Minnesota Comprehensive Achievement test (MCA) and the NWEA Measures of Academic Progress achievement test in reading and math, obtained from third to eighth-grade students in Minnesota. Student achievement was measured in individual panels covering the 2003–2009 academic years, and different districts applied for and adopted Q-Comp in different years. The study also utilized information on the population of teachers linked to their district per year; Q-Comp program
data coded from archives of official documents; the American Schools and Staffing Survey (SASS); and data on district characteristics and finance obtained from the Minnesota Department of Education (MDE). The study determined that the Q-Comp program produced an average increase in reading achievement with a 3% standard deviation, and a similar effect on math achievement was also evident in the full sample.

While their former study analysed DATE, the concurrent work of Springer, Ballou et al. (2010) on the Project on Incentives in Teaching (POINT) yielded contrasting results. POINT was a three-year study conducted in Metro Nashville public schools from 2006–2009. Here, 296 middle school mathematics teachers volunteered for the study at the beginning of the 2006–2007 school year, but only 148 remained at the third year’s conclusion. The incentives did not affect the Tennessee Comprehensive Assessment Program (TCAP) test scores across all years and grade levels. Although researchers found a slightly positive effect among fifth-graders for merit pay eligibility, it was not evident in other grade levels (Springer, Ballou et al., 2010).

After completing the four components of Denver’s Professional Compensation System for Teachers (ProComp) (knowledge and skills, professional evaluation, market incentives, and student growth), professionally evaluated teachers eligible for compensation could receive a maximum of approximately US$2,000 (Podgursky & Springer, 2007). Goldhaber and Walch (2012) utilized the Colorado Student Assessment Program (CSAP) math and reading test scores to assess third- to tenth-grade students’ achievements. The result showed a lack of a consistent pattern across grade levels and subjects. In some cases, the test scores of students with ProComp teachers were higher, whereas in other cases, students with non-ProComp teachers scored higher.

Similarly, Glazerman and Seifullah (2012) conducted a study that resulted in an inconsistent influence of merit pay on student achievement in the Teacher Advancement Program (TAP) of Chicago public schools (CP5s), or Chicago TAP. There was no consistent growth in student achievement in the Illinois Standards Achievement Test (ISAT) scores during the four-year implementation period. Specifically, there was evidence of both positive and zero test score impacts in selected subjects, years, and school cohorts. However, there was no noticeable impact on overall test scores in math, reading, and science. Though the program positively impacted science scores, this result was not statistically significant unless the researchers used one particular matching method that excluded some Chicago TAP schools from the analysis (Glazerman & Seifullah, 2012). For an exact comparative analysis, Glazerman and Seifullah (2012) collected administrative data on over 300 Chicago public schools in which the Chicago TAP was not implemented.

3.2. Merit pay and teacher retention

Though more disparate than the case of merit pay and student achievement, the earlier works studied in this section—Kirby, Berends, & Naftel; Clotfelter, Giennie, Ladd, & Vigdor; and Coates-McBride & Kritsonis, published in 1999, 2006, and 2008—displayed significant correlations between merit pay and teacher retention. Likewise, no correlation was observable among the studies of Springer, Lewis et al. (2009), Springer, Lewis et al. (2010), and Glazerman and Seifullah (2012). Meanwhile, some correlation was evident in the study conducted by Springer et al. (2015).

The strong correlation between the studies of Kirby et al. (1999) and Clotfelter et al. (2006) in terms of teacher retention rate had much to do with the faculty consisting of minorities or with the faculty working in academically failing or impoverished schools. Correlations found by Springer et al. (2015) were also related to teachers working in academically failing schools. Therefore, merit pay may have the most significant impact on teacher retention with such circumstances. Thus, where and when the programs were implemented should be considered, especially in Texas and North Carolina.

As for the studies that displayed no correlation, merit pay had a short-term effect on teacher retention. The studies of Springer, Lewis et al. (2009, 2010) observed a decrease in turnover during the first year, though the rate reverted in the second and third years. Although similar to the
findings obtained by Glazerman and Seifullah (2012), the result was more variable for teachers in schools that initiated the program in later years. In addition, Clotfelter et al. (2006) revealed that teachers with less than 10 years of experience were less influenced by the merit pay program. Alternatively, according to Coates-McBride & Kritsonis (2008), teachers with 0–4 years of experience were strongly influenced. Although these last two results are conflicting, it is inconclusive whether career length alone can vary results.

Kirby et al. (1999) analysed longitudinal data from the Texas State Department of Education on Texas public school teachers from 1979–1996. The complete work history of these teachers allowed examinations of minority groups, their behaviours, and their characteristics over time. Furthermore, researchers were able to define high-risk districts and teachers. As a result, a US $1,000 salary increase was associated with 2.9% overall attrition reduction and 5-6% among Latin American and African American teachers in Texas. Therefore, merit pay played an important role in improving teacher retention in minority groups.

The analysis of merit pay’s effects in North Carolina by Clotfelter et al. (2006) compared turnover patterns before and after the three-year implementation, initiated in 2001. North Carolina awarded an annual bonus of US$1,800 to certified math, science, and special education teachers working in highly impoverished schools. The study results showed that this merit pay program could reduce the targeted teachers’ average turnover rates by 12%. As previously mentioned, teachers with 10–19 years of teaching showed a 37% reduction in turnover rates, suggesting incentivized behaviour for experienced teachers in high-poverty areas. On the other hand, teachers with less than 10 years of experience were influenced less.

Coates-McBride & Kritsonis (2008) analysed the effect of the merit pay program on the Houston Independent School District (HISD). In 2005, the HISD initiated a merit pay system for longstanding teachers, with thousands of dollars in bonuses for substantial academic improvement in children. According to HISD data, 1,554 out of 12,500 teachers left the district in 2005. After the program was implemented in the following year, the number of teachers leaving the district declined to 1,262, or by nearly 19%. The remaining teachers started in 2006, signifying a 25% decrease, with only 576 having left compared to the 773 who left in 2005, suggesting that merit pay incentivized new teachers in the HISD to continue teaching.

Regarding the Tennessee Governor’s Retention Bonus Program during the 2013–2014 school year, highly effective teachers received a US$5,000 retention bonus for working in low-performing priority schools. Springer et al. (2015) conducted their study with a sample of 587 teachers in 56 schools, which found that tested subject teachers who received bonuses were 20% more likely to remain in priority schools than teachers who did not receive bonuses. Nevertheless, there was no significant correlation between merit pay and teacher retention.

In comparison, Springer, Lewis et al. (2009) claimed that the Texas Governor’s Educator Excellence Grant (GEEG) during the 2005–2006 school year lowered the teachers’ predicted attrition rate by 3% in the first year of implementation. However, it was not sustained, as the rate returned to normal in the program’s succeeding years. The Texas GEEG program is the first of several multimillion-dollar, state-wide programs committed to developing performance incentives for high-performing educators. In the fall of 2006, the GEEG program allowed 99 schools to avail themselves of three-year, non-competitive grants ranging from US$60,000 to US$220,000 per year. The study on the Texas DATE program (Springer, Lewis et al., 2010) yielded similar causes and results: turnover decreased in the first school year of 2008–2009 among teachers who received awards, but the rate was not sustained and eventually returned to normal.

Teacher retention increase was evident in Glazerman and Seifullah (2012) study concerning the Chicago Teacher Advancement Program (Chicago TAP) during the 2009–2010 and 2010–2011 school years. Teachers working in Chicago TAP schools in 2007 returned in each of the following
three years at higher rates than teachers in non-TAP schools. Here, 67% (or 12 percentile points) of teachers in Chicago TAP schools who taught in the fall of 2007 returned to the same school in the fall of 2010 compared to about 56% of teachers in non-TAP schools. Chicago TAP teachers in the fall of 2007 were about 20% more likely to return than other teachers in those same schools three years later (Glazerman & Seifullah, 2012). However, this result was more variable for teachers in schools that initiated the program in later years, and the impacts of merit pay were not uniform or universal across years, cohorts, and subgroups of teachers.

4. Discussion and conclusion

Regarding the correlation between merit pay and teacher retention, the time—that is, the year implemented and span of the program—is an important variable for policy effectiveness. Moreover, demographics (whether the teachers belong to minority groups, practice in high-poverty areas, or work in academically failing institutions) were considered relative to the type of merit pay awarded and the students’ education level. Thus, the variables found in this study should be considered when reviewing the implementation of the merit pay policy.

Concerning the merit pay program’s effectiveness, the program played an important role in improving teacher retention in minority groups, impoverished, and academically failing schools. Furthermore, the merit pay program was effective in increasing the retention rate of new teachers or teachers with 10–19 years of experience. Finally, the merit pay program was effective in the early implementation stages.

However, it is crucial to note that these variables are subordinate to the covered studies’ essential commonality: assessing student achievement using standardized test scores. Together or separately, these variables cannot accurately determine the correlation between merit pay and student achievement because of the differences in standardized tests. Figlio and Kenny (2006) referred to FCAT scores. Meanwhile, G. W. Ritter et al. (2008) utilized scores from the ITBS. If the merit pay policy aims to improve student achievement by measuring it as objectively as possible, the means of measuring student achievement should be changed. Due to the fact that the evaluation of the teachers’ eligibility for merit pay was based on the outcome of standardized tests, one should consider if standardized test scores fully reflect student achievement and teacher competency.

Standardized tests have many problems. First, standardized tests assess a student’s performance on a particular day without considering external factors, such as the student’s physical or mental condition on the exam day. Second, standardized tests only focus on evaluating the student’s individual performance instead of that student’s overall growth over the year. Third, these tests focus on reading, math, and science while neglecting other subjects, such as physical education, music, and art. One’s failure in another subject can be attributed to the strict focus on reading, math, and science. Likewise, standardized tests can neither verify nor assess the teacher’s competency. The irony is that teachers who focus primarily on raising standardized test scores may be incompetent because these tests do not necessarily require creative teaching, feedback skills, interpersonal skills, and common sense. Fairness is not guaranteed if teachers are evaluated for merit pay according to the student’s performance on standardized test scores. In addition, teachers who do not teach the subjects included in these tests can be easily marginalized within the merit pay system.

Thus, if policymakers want to justify implementing merit pay programs by objectively measuring student achievement, they should be interested in student performance consistently in all subjects learned. In the reviewed studies, the US Department of Education should develop standards and guidelines that evaluate student performance based on authentic assessments, which ask students to apply skills and abilities in real life. One example could be a portfolio that serves as a purposeful collection of a student’s work, demonstrating their efforts, progress, and achievements, as it includes their works in progress, revisions, and critical self-reflections on what they have learned. The US Department of Education should cooperate with state departments to
Lastly, verify the school's situation. In support of this study, each school's spending (Raikes & Darling-Hammond, 2019) can be compared to the average spending per student in New York (US$35,944) and Utah (US$6,968). The highest-spending district in New York is Kiryas Joel Village Union Free School District (US$213,130), while the highest-spending district in Utah is Franklin Discovery Academy School District (US$21,817) (Public School Review, 2021). In most states, children who live in low-income areas attend the most under-resourced schools (Raikes & Darling-Hammond, 2019). Because of this situation, it is difficult for state departments of education to take an individual interest in each district independently. Likewise, it is difficult for each district to make efforts to understand each school's needs. Therefore, by using the federal budget, each district should make an effort to understand the needs and situations of each school under the supervision and support of the states' departments of education, who should try to grasp the needs and situations of each district under the supervision and support of the US Department of Education. The American educational administration should aim to be more considerate of the needs and situations of each school.

Based on the analyses of the studies, the findings can be summarized as follows. First, the year a merit pay program's implementation, duration, demographics (minority groups, high-poverty areas, or academically failing institutions) affect the policy's effectiveness. Next, these variables cannot accurately determine the correlation between merit pay and student achievement because of different standardized tests. Furthermore, currently used standardized tests cannot objectively verify and assess the teacher's competency and cannot cover all subjects taught to students. Lastly, this study suggested an institutionalized evaluation system that focuses on student performance consistently in all subjects learned, based on authentic assessments, which ask students to apply their skills and knowledge in real life. This study also suggested the need for an educational administration system tailored to the needs and situations of various educational fields with diverse cultures, environments, and components.
Funding
This work was supported by the Presbyterian University and Theological Seminary [20201215-003].

Author details
Sung Joong Kim.¹
E-mail: newant99@gmail.com
ORCID ID: http://orcid.org/0000-0001-7615-7408
¹ Educational Administration, Presbyterian University and Theological Seminary, Seoul, Korea.

Citation information
Cite this article as: Merit pay, case-by-case: Variables affecting student achievement, teacher retention, and the problem of standardized tests, Sung Joong Kim, Cogent Education (2021), 8: 1920560.

References
Buddin, R., McCoffrey, D. F., Kirby, S. N., & Xia, N. (2007). Merit pay for Florida teachers—design and implementation issues. Working paper WR-508-FEA, RAND Corporation, Santa Monica, CS. https://www.rand.org/pubs/working_papers/WR508.html
Choi, C., Glennie, E., Ladd, H., & Vigdor, J. (2006). Would higher salaries keep teachers in high-poverty schools? Evidence from a policy intervention in North Carolina. National Bureau of Economic Research. http://www.nber.org/papers/w12285.pdf.
Coates-mcbride, A., & Kritsonis, W. A. (2008). The M&M effect—assessing the impact of merit pay on teacher motivation: National implications. National Journal for Publishing and Mentoring Doctoral Student Research, 5(1), 1–6. https://filesERIC.ed.gov/fulltext/ED499772.pdf
Digest of Education Statistics. (2017). Table 211.10: Average salaries for full-time teachers in public and private elementary and secondary schools, by selected characteristics: 2011–12 and 2015–16. National Center for Education Statistics. https://nces.ed.gov/programs/digest/d17/tables/dt17_211.10.asp
Figlio, N. D., & Kenny, L. (2006). Individual Teacher Incentives and Student Performance. National Bureau of Economic Research. http://pypi.go.google/view/w12627.pdf.
Gius, M. (2012). Using a difference-in-differences approach to estimate the effects of teacher merit pay on student performance. Eastern Economic Journal, 39(1), 1–10. https://doi.org/10.1016/j.eeej.2012.02.
Glazerman, S., & Seifullah, A. (2012). An evaluation of the Chicago Teacher Advancement Program (Chicago TAP) after four years: Final report. Mathematica Policy Research. https://www.mathematica-mpr.com/our-publications-and-findings/publications/an-evaluation-of-the-chicago-teacher-advancement-program-chicago-tap-after-four-years.
Goldhaber, D., & Walch, J. (2012). Strategic pay reform: A student outcomes-based evaluation of Denver’s ProComp teacher pay initiative. Economics of Education Review, 31(6), 1067–1083. https://doi.org/10.1016/j.economy.2012.06.007
Halstead, E. O. (2013). Teacher satisfaction and turnover in WCPS. Data Trends: Data and Accountability Department. https://filesERIC.ed.gov/fulltext/ED559196.pdf.
Henderson, B. M., Houston, D., Peterson, E. P., & West, R. M. (2019). Public support grows for higher teacher pay and expanded school choice. Education Next. https://www.educationnext.org/school-choice-trump-era-results-2019-education-next-poll.
Johnson, S. M., Berg, J. H., & Donaldson, M. L. (2005). Who stays in teaching and why: A review of the literature on teacher retention. AARP.
Joongang Daily. (2015, September 3). [언론고시·교사봉사제, 교육부가 재개적 공고] Joongang Ilbo. https://newsJoins.com/article/18585825.
Kennedy, L. (2016, December). Decent PISA Numbers, but Can Mass. Students really do the math? WBUR Online. http://www.wbur.org/edify/2016/12/09/pisa-2015-mass-analysis.
Kim, Y. J. (2003). A study of the perceptions of administrators and faculty members toward merit pay for faculty at Junior Colleges in Korea [Unpublished doctoral dissertation], University of Texas.
Kirby, S. N., Berends, M., & Noffel, S. (1999). Supply and demand of minority teachers in Texas: Problems and prospects. Educational Evaluation and Policy Analysis, 21(1), 47–66. https://doi.org/10.3102/0162373720100104.
Mulvihill, E. (2019). Why teachers quit? Lack of respect, abominable working conditions, and more. We Are Teachers Online. https://www.weareteachers.com/why-teachers-quit.
National Center for Educational Statistics. (2018). PISA 2018 U.S. results. NCES online. https://nces.ed.gov/surveys/pisa2018/pdf/PISA2018_Complied.pdf.
The National Commission on Excellence in Education. (1983). A nation at risk: The imperative for educational reform. https://edreform.com/wp-content/uploads/2013/02/A_Nation_At_Risk_1983.pdf.
Petersen, P. E., Woessmann, L., Hanushek, E. A., & Lastra-Annadan, C. X. (2011). Globally challenged: Are US students ready to compete? The latest on each state’s international standing in math and reading. Harvard’s Program on Education Policy and Governance (PEG). https://sites.hks.harvard.edu/pepg/PDF/Papers/PEPG11-03_GloballyChallenged.pdf.
Podgursky, M. J., & Springer, M. G. (2007). Teacher performance pay: A review. Journal of Policy Analysis and Management, 26(4), 909–950. https://doi.org/10.1002/pam.20292
Public School Review. (2021). Average public school spending/Student. Public School Review Online. https://www.publicschoolreview.com/average-spending-student-stats/national-data.
Raikes, J., & Darling-Hammond, L. (2019). Why our education funding systems are derailing the American dream. Learning Policy Institute. https://learningpoliycinstitute.org/blog/why-our-education-funding-systems-are-derailing-american-dream.
Research for Action. (2012). Issue brief: Performance pay for teachers. Pennsylvania Clearinghouse for Education Research (PACER). https://www.researchoration.org/publications/issue-brief-performance-pay-for-teachers.
Ritter, G., Trivitt, J., Foreman, L., DeAngelis, C., & Denny, G. (2016). Merit pay programs for improving teacher retention, teacher satisfaction, and student achievement in primary and secondary education: A systematic review. The Campbell Collaboration. https://campbellcollaboration.org/media/22/attachment/Ritter_Merit_Pay_Protocol.pdf.
Ritter, G. W., Holley, M. J., Jensen, N. C., Riffel, B. E., Winters, M. A., Barnett, J. H., & Greene, J. P. (2008). Year two evaluation of the achievement challenge pilot project in the Little Rock Public School District. Department of Education Reform, University of Arkansas.
Sojourner, A. J., Miyazoe, E., & West, L. K. (2014). Teacher pay reform and productivity. Panel data evidence on adoptions of Q-Comp in Minnesota. Journal of Human Resources, 49(4), 945–981. https://doi.org/10.1353/jhr.2014.0029
Springer, M. G., Ballou, D., Hamilton, L., Le, V.-N., Lockwood, J. R., McCaffrey, D. F., Pepper, M., & Stecher, B. M. (2010). Teacher pay for performance: Experimental evidence from the project on incentives in teaching. National Center on Performance Incentives (NCPI). https://www.rand.org/content/dam/rand/pubs/reprints/2010/RAND_RP1416.pdf.

Springer, M. G., Ballou, D., & Peng, A., (XJ). (2008). Impact of the teacher advancement program on student test score gains: Findings from an independent appraisal (Working Paper 2008–19). National Center on Performance Incentive (NCPI). https://my.vanderbilt.edu/performanceincentives/files/2012/10/200819_Springer_ImpactAdvancedProg1.pdf.

Springer, M. G., Lewis, J. L., Ehler, M. W., Podgursky, M. J., Crader, C. D., Taylor, L. L., Gronberg, T. J., & Stuit, D. A. (2010). District awards for teacher excellence (d.a.t.e.) program: Final evaluation report. National Center on Performance Incentives (NCPI). https://my.vanderbilt.edu/performanceincentives/files/2012/10/FINAL_DATE_REPORT_FOR_NCPI_SITE4.pdf.

Springer, M. G., Lewis, J. L., Podgursky, M. J., Ehler, M. W., Taylor, L. L., Lopez, O. S., & Peng, A., (XJ). (2009). Governor’s Educator Excellence Grant (GEEG) program: Year two evaluation report. National Center on Performance Incentives (NCPI). https://s3.amazonaws.com/vu-my/wp-content/uploads/sites/584/2012/10/04183715/GEEG_Year_2_Report_ES.pdf.

Springer, M. G., Swain, W. A., & Rodriguez, L. A. (2015). Effective teacher retention bonuses evidence from Tennessee. Educational Evaluation and Policy Analysis, 38(2), 199–221. https://doi.org/10.3102/0162373715609687

Strauss, V. (2017, September 18). Where have all the teachers gone? The Washington Post. https://www.washingtonpost.com/news/answer-sheet/wp/2017/09/18/where-have-all-the-teachers-gone?noredirect=on&utm_term=.7c0f691fd7a.

UF study: teacher merit pay boosts student standardized test scores. (2007, January 4). University of Florida News. http://news.ufl.edu/archive/2007/01/uf-study-teacher-merit-pay-boosts-student-standardized-test-scores.php.

United States Department of Education. (2010). Race to the top program guidance and frequently asked questions. U.S. Department of Education. http://www2.ed.gov/programs/racetothetop/faq.pdf.

van Loosen, L. (1983). Some points to consider when you discuss merit pay. (Ed 245 336). American Association of School Administrators.

Washington State Legislature. (2019). Minimum criteria for the evaluation of certificated employees—Four-level rating evaluation system—Procedures—Steering committee—Implementation—Reports—Comprehensive performance evaluation (RCW 28A.405.100). https://apps.leg.wa.gov/rcw/default.aspx?cite=28A.405.100.

Williamson, R. G. (2010, April 23). Merit pay for teachers. Ronald Williamson, Ed. D. Practicial Leadership Solutions, http://ronwilliamson.com/uploads/MeritPayApr2010.pdf.