



Canadian Labour Market and Skills Researcher Network

Working Paper No. 90

How Important are School Principals in the Production of Student Achievement?

Elizabeth Dhuey
University of Toronto

Justin Smith
Wilfrid Laurier University

December 2011

CLSRN is funded by the Social Sciences and Humanities Research Council of Canada (SSHRC) under its Strategic Knowledge Clusters Program. Research activities of CLSRN are carried out with support of Human Resources and Skills Development Canada (HRSDC). All opinions are those of the authors and do not reflect the views of HRSDC or the SSHRC.

How Important Are School Principals in the Production of Student Achievement?*

Elizabeth Dhuey
Centre for Industrial Relations and Human Resources
Department of Management
University of Toronto

Justin Smith
Wilfrid Laurier University
Department of Economics

Abstract

As school leaders, principals can influence student achievement in a number of ways, such as: hiring and firing teachers, monitoring instruction, and maintaining student discipline, among others. We measure the effect of individual principals on gains in student math and reading achievement between grades four and seven. We estimate that a one standard deviation improvement in principal quality can boost student performance by approximately 0.2 standard deviations in both math and reading. We also show that principal experience does not exert a significant influence on student performance. Our results imply that isolating the most effective principals and allocating them accordingly between schools can have a significant positive effect on reducing achievement gaps.

*We thank Kelly Bedard, Mati Dubrovinsky, David Johnson, Harry Krashinsky, Abigail Payne, Simon Woodcock, and various seminar participants for helpful comments.

JEL Code: I20, I21

Keywords: Economics of education, principals, education

I. Introduction

The summer of 2008 saw widespread media coverage of controversial Washington D.C. school chancellor Michelle Rhee and her move to fire roughly 40 school principals in her district. Many of the schools in question were failing to meet the academic proficiency standards of the No Child Left Behind Act, and this change in school leadership was meant to help these struggling schools improve their performance. Not only does this move highlight the value that school districts place on quality leadership within a school, but it also demonstrates an underlying belief that principals have an effect on student achievement within their schools. As the main administrator in a school, the principal is responsible, among other things, for maintaining and improving teacher quality, monitoring and enforcing student conduct, and ensuring the curriculum is properly implemented. However, despite the importance of these functions in contributing to the educational experience, much of the economics research to this point has focused on teacher quality rather than principal quality.

In this paper we quantify the importance of principals in the production of student test score gains between fourth and seventh grade using principal fixed effects estimated from longitudinal administrative data from the province of British Columbia (BC), Canada. Principal fixed effects measure the impact of all unchanging characteristics of a principal that are transferrable between schools, such as leadership ability, personality, and gender. We obtain these estimates for math and reading scores using a value-added model that controls for observable time-varying student, school, and neighborhood factors, along with fixed school effects. The frequent mobility of principals between schools in BC allows us to identify principal effects using within school variation, which is important because school fixed effects remove the influence of fixed school characteristics that might bias estimated principal effects if principals sort across schools based on those attributes. We also extend the model to allow for “match effects,” where part of the principal effect is portable across schools, and part depends on the school where they are employed. Finally, we estimate the effect of experience as a principal within a school (which we call tenure), and overall experience as a principal on student achievement.

We estimate that a one standard deviation move up the distribution of principal quality leads to an improvement of approximately 0.2 standard deviations in both math and reading score gains. When we extend the model to include match effects, we find that roughly half of that effect is due to the match between principal and school, and that the standard deviation of the match effects is 0.160. Given the importance of good principals and good matches between principals and schools, properly allocating principals to schools is an important policy consideration, especially in terms of reducing achievement gaps. In terms of principal experience, we find that more experienced principals and those with a longer tenure in a school have no significant impact on student performance. This implies that, at least in BC, what matters for student achievement is finding a principal with good fixed attributes and assigning that individual to the correct school.

II. Existing Literature on the Effect of Principals

Despite the extensive literature on teacher quality, relatively little attention has been paid by economists to the importance of principals in the production of student achievement.¹ The existing economics literature related to our research can be separated into three strands.² The first focuses on estimating the overall effect of the principal on achievement, the second on the effects of specific principal attributes on achievement and the third on the effects of specific principal training programs³

Most closely related to our research are two recent papers by Coelli and Green (forthcoming) and Branch, Hanushek, and Rivkin (2010) focus on the overall effect of the principal on achievement.⁴ Coelli and Green (forthcoming) use a sample of students entering grade 12 drawn from administrative data in

¹ Some recent studies regarding teaching quality are: Aaronson, Barrow, and Sander (2007), Buddin and Zamarro (2009), Harris and Sass (2006), Kane, Rockoff, and Staiger (2008), Nye, Konstantopoulos, and Hedges (2004), Rivkin, Hanushek, and Kain (2005), Rockoff (2004), and Sanders and Rivers (1996).

² Education scholars have studied principal effectiveness extensively. See Hallinger and Heck (1998) for a discussion of the more qualitative education research regarding principal effectiveness that has generally relied on teacher reports of principal attributes and leadership styles.

³ In addition, Cullen and Mazzeo (2008) find that for mobile principals, the information provided by the school's test scores allow principals whose schools do better to earn more by getting promoted to higher paying positions at other schools and at higher levels within the district.

⁴ Grissom and Loeb (2009) find that only the self-assessment of organizational management skills predicts growth in state-assigned school grades. Horng, Klasik, and Loeb (2009) similarly find that time spent dedicated to organizational management also is related to state-assigned grades.

BC to estimate the effect of principals on high school graduation and grade 12 provincial exam scores in English.⁵ In their main analysis, they estimate the lower bound of the variance of principal effects on the outcomes above using the method formulated by Rivkin, Hanushek, and Kain (2005), which was originally used to estimate teacher effects. They find that moving one standard deviation up in the distribution of principal effects can improve both outcomes, but has a particularly significant effect on grade 12 English scores. They also estimate a dynamic model of principal effects and find that that a principal's full impact is not immediate, but instead increases gradually over time. Even though we both draw our samples from BC data and estimate the dispersion of principal effects on outcomes, our research is best viewed as a complement to Coelli and Green (forthcoming) because of fundamental differences. The main difference is in the populations under analysis: we study elementary school students and they study high school students. Secondly, though our empirical strategies rely on similar identifying assumptions, they are different in that we use a value-added approach rather than a levels approach. We focus on value-added because our outcomes will depend on cumulative inputs from various sources. Finally, in addition to providing estimates of fixed principal effects, we also estimate match effects.

Branch, Hanushek, and Rivkin (2010) use data from Texas to estimate the importance of principals on student math and reading test score gains using data from the first three years of a principal's tenure at a school. In their main analysis, they estimate principal fixed effects in a value-added model, and present summary statistics based on those effects. They find a significant amount of dispersion in the principal effects overall, and that this dispersion is higher in low-income schools. They also find that principals in the middle of the quality distribution are generally more likely to stay at the school after three years, and that this pattern does not vary with the school's poverty level. They support their conclusions that what they estimate is a principal quality effect by showing that when high quality principals lead schools, the teachers who exit are of comparably worse quality. While similar to our research, the main specification in Branch, Hanushek, and Rivkin (2010) excludes school fixed effects.

⁵ Using the same data, Coelli, Green, and Warburton (2007) examine the effect of graduating on Income Assistance receipt in Canada. As one of their identification strategies, they use principal fixed effects as instrumental variables for graduation.

Though they mention that variance estimates are similar when school effects are included, it is difficult to imagine that principal mobility is uncorrelated with school quality, so excluding these controls will confound the effects of principal and school.

The second strand of the literature on principal effects focuses on the relationship between principal observable characteristics and school performance. Many of these papers focus on the role of principal education and experience on school performance, and have found mixed results. In terms of education, some researchers find that a more educated principal worsens the school's performance (see Ballou and Podgursky, 1993 and Eberts and Stone, 1988), whereas other researchers find no relationship (see Clark, Martorell, and Rockoff, 2009). In addition, Eberts and Stone (1988) and Clark, Martorell, and Rockoff (2009) find a positive correlation between teaching experience and school performance, but Brewer (1993) finds no correlation.

The third strand of the literature examines the effects of principal training programs. For example, Corcoran, Schwartz, and Weinstein (2009) show that principals trained in the New York City Aspiring Principals Program have positive effects on school performance. Clark, Martorell, and Rockoff (2009), however, find mixed evidence on the relationship between formal principal training/development programs and school achievement.

Our research adds to the principal literature in several ways. Our main contribution is to measure the effectiveness of individual principals within a framework that controls for potential confounding school factors. We also relax the restriction that principal effects are fixed across schools by estimating the contribution of the principal-school match to student achievement. There are many reasons to think that principal effectiveness depends on the school environment where they work, so this is an important step forward in determining why principals matter for student achievement. Finally, we are also able to explore the effect of principal tenure at a school and overall principal experience, as these factors have the potential to matter if principals tend to improve over time.

III. Education in British Columbia

The public school system in BC is very similar to other jurisdictions across Canada and the United States. There are 60 school districts in BC, many of which correspond to city boundaries in the populated areas (Vancouver and its surrounding area and Victoria). In the more rural areas, districts correspond to a much wider area. Prior to 2003, students attending public schools were required to attend the school located in their catchment area. Since 2003, legislation has existed that allows parents to choose any public school. However, it is not clear that this legislation has been successful in creating widespread public school choice (DeCicca and Smith, 2011). Approximately, 30 percent of students attend a middle or junior high school that begins in either sixth or seventh grade, and the other 70 percent attend a school whose grades extend through seventh grade and higher (Dhuey, 2011). Therefore, the majority of the students in our sample will have the opportunity to stay at the same school between grade four and seven.⁶

Each year since 1999, students in fourth and seventh grade are tested in reading, writing, and math using the Foundation Skills Assessment (FSA) tests. All students are expected to participate in the tests, with the exception of some ESL students and students with special needs. These tests are low stakes in the sense that no funding is tied to the outcome of these exams, nor do they contribute to student course grades. Therefore, despite recent work regarding teacher and principal cheating, principal cheating on these exams is unlikely, mainly due to the fact that principal compensation and school resources are set according to guidelines that are not based on student outcomes (Coelli and Green, forthcoming).

Principals in British Columbia are appointed by the school board, and their duties are outlined in various parts of the School Act from the Revised Statutes of BC and are specifically listed in BC Regulation 265/89 and the many amendments thereof. BC principals are typical of most principals across jurisdictions. They are responsible for carrying out orders from the school district or from the Ministry of Education, and they are also responsible for making sure the instructional practices of their school

⁶ We include a robustness check that includes only students that attend the same school (see Table 6, Panel D). The results are not substantially different.

conform to the School Act. In addition, they are responsible for the smooth functioning of the province's various standardized testing programs. Unlike some other jurisdictions, principals do not directly hire and fire teachers; instead they provide information to the school district about teacher performance, and any disciplinary action is then decided at this upper level. Finally, they carry out various administrative tasks such as making teacher timetables, maintaining school records, and monitoring the conduct of students.

IV. Empirical Specification

To estimate the principal effects, we use the following value-added model of students' test scores.

$$y_{ispt}^7 = \beta_0 + y_{ispt}^4 \beta_1 + x'_{ispt} \beta_2 + z'_{spt} \beta_3 + \sum_{j=1}^J d_{pt}^{\text{exp}=j} + \delta_p + \phi_s + \eta_t + \varepsilon_{ispt} \quad (1)$$

where y_{ispt}^7 is the grade seven math or reading score for student i in school s with principal p ; y_{ispt}^4 is the student's corresponding grade four score; x'_{ispt} is a vector of student-level characteristics; z'_{spt} is a vector of school-level characteristics; $d_{pt}^{\text{exp}=j}$ is a set of dummy variables equal to 1 if principal p has j years of experience in year t ; δ_p , ϕ_s , and η_t are fixed effects at the principal, school, and year level, respectively; ε_{ispt} is an idiosyncratic error term.⁷

To describe the effect of principals on student achievement, we examine summary statistics based on the set of principal fixed effects, in particular their standard deviation. These principal fixed effects are interpreted as the effect of each individual principal on achievement that is invariant across time, students, and schools conditional on other time-varying and time-invariant student and school-level factors. This effect will measure the impact of any fixed attribute of the principal, such as leadership ability, gender, education, and training.

⁷ We model value-added over three years, whereas other models in the literature use one-year gains. This might lead our estimates to be higher than they would be if we used a comparable one-year value-added model. We explore this issue in various robustness checks in Section VII.B.

We estimate equation 1 using fixed effects methods, treating $\{\delta_p\}$ as parameters to be estimated. Normally, fixed effects are used to control for unobserved heterogeneity when researchers are trying to identify other coefficients in the model. In that scenario, the fixed effects are generally never explicitly estimated, but rather swept away by a within transformation. Our primary focus, however, is on the estimates of the principal fixed effects themselves, which adds several complications in terms of estimation. If we first consider the case where we do not include the school effects, we cannot separately identify all of the principal effects when the model contains a constant (the “dummy variable trap”); an identifying restriction must be imposed. The most obvious restriction is to simply set one principal effect equal to zero. This is unsatisfactory for our purposes because it forces the interpretation of the remaining fixed effects to be deviations from the dropped principal, which is sensitive to which principal is dropped. Instead, we restrict the fixed effects to sum to zero as in Mihaly et al. (2010). With this restriction, the principal effects are interpreted as deviations from the average of the principal effects.

The problem becomes more complicated when we include both principal and school effects. Principals and schools are implicitly separated into many disconnected groups because not every principal works at every school. Each group contains all principals that have ever worked at any school in the group, and all schools that have ever employed any principal in the group. Principals in one group are never observed working at a school in any other group, and schools in one group never employ a principal from any other group. When including school fixed effects, one principal effect in each group is not identified, and additional restrictions must be imposed.⁸ Following Mihaly et al. (2010), when the model contains principal and school effects, we restrict the principal effects to sum to zero within each group. The principal effects in this case are interpreted as deviations from the within-group mean.

One consequence of including school effects in the model is that principal effects must vary within schools to be identified. In our case, at least one school in each group must employ multiple

⁸ In the estimation, within each group the sum of the principal dummies equals the sum of the school dummies. Thus, one principal dummy in each group is dropped. Alternatively, one could drop one school dummy from each group. Because this model does not include a constant, we keep all school dummies in the regression.

principals over time to contribute to estimation. This arises mainly through principal mobility across schools, but can also occur when new principals enter the sample. In the extreme case where a group contains exactly one principal and one school (*i.e.* the principal does not move and the school does not employ any other principal), principal effects are not identified and therefore not estimated. The reported standard deviations based on the school fixed effects specification use only identified principal fixed effects. Thus, principal mobility is crucial for our estimation. Fortunately there is enough mobility to identify an effect for most principals in the school effects model (see Table 1). Despite the fact that we can only identify principal effects for the subset of identified principals, it is our preferred specification because controlling for fixed school factors is crucial. We examine the difference between “identified” and “unidentified” principals in Section V.C.

Our estimates of the standard deviation of the principal effects may be upwardly biased due to sampling error. As explained in Kane and Staiger (2002), and noted in Branch, Hanushek, and Rivkin (2010), Aaronson, Barrow, and Sander (2007), and Jacob and Lefgren (2005 and 2008), even in the absence of any real principal effect, we still might observe variations in the estimated effect due to random differences between samples of students. Such sampling variation is a problem particularly when a principal fixed effect is based on a small number of students. To correct for this, we adjust the standard deviation downward using an estimate of the variation of the sampling error.⁹ Following Aaronson, Barrow, and Sander (2007) and Jacob and Lefgren (2005 and 2008), we assume the estimated principal effect is the sum of the true effect and a mean-zero, independent, normally distributed error

$\hat{\delta}_p = \delta_p + \nu_p$. Due to independence, the variance of the estimated principal effect is the sum of the variances of the true effect and the error $\sigma_{\hat{\delta}}^2 = \sigma_{\delta}^2 + \sigma_{\nu}^2$. We obtain the true principal effect variance

⁹ While shrinking our estimates are important in this estimation, shrinking of estimates will generally be much more important in the case of teachers as they have more variation in class size and in the number of students tested each year which will generate more variation in the precision of teacher fixed effects versus principal fixed effects.

by subtracting the error variance from the variance of the estimated principal effect. Because σ_v^2 is unknown, we estimate it by taking the square of the average of the standard errors of $\hat{\delta}_p$.

In addition to shrinking our original estimate of the standard deviation of principal effects estimated above, we also compute the standard deviation of an Empirical Bayes estimate of the principal effects. As in Jacob and Lefgren (2005), if $\delta_p \sim N(0, \sigma_\delta^2)$, then conditional on observing our “noisy” estimate of the principal effect, $\hat{\delta}_p = \delta_p + \nu_p$, a Bayesian estimate of the mean of the principal effect would shrink our existing estimates based on a signal to noise ratio:

$$\delta_p^* = E[\delta_p | \hat{\delta}_p] = \left(\frac{\sigma_\delta^2}{\sigma_\delta^2 + \sigma_\nu^2} \right) \hat{\delta}_p. \text{ If our estimate } \hat{\delta}_p \text{ contains mostly noise, then the estimates shrink}$$

toward zero.¹⁰ Using our estimates of σ_δ^2 and σ_ν^2 , we can construct the Empirical Bayes estimate δ_p by substituting those values into the equation above. Estimates will shrink more toward zero the higher the noise matters relative to the signal.¹¹

V. Data and Analysis Sample

A. Data Sources

Our main datasets are three linked files obtained from the Ministry of Education in BC. The first contains observations on all students writing the FSA test in the period 1999 through 2006. For these students, we know the percentage score on each test, the school in which the test is written, and whether

¹⁰ If we assume more generally that $\delta_p \sim N(\bar{\delta}_p, \sigma_\delta^2)$, then $E[\delta_p | \hat{\delta}_p] = \left(\frac{\sigma_\nu^2}{\sigma_\delta^2 + \sigma_\nu^2} \right) \bar{\delta}_p + \left(\frac{\sigma_\delta^2}{\sigma_\delta^2 + \sigma_\nu^2} \right) \hat{\delta}_p$

and a noisy estimate would be shrunk towards the prior mean. Our assumption that $\bar{\delta}_p = 0$ is akin to reinterpreting the principal effects as deviations from means, which is the correct interpretation of the estimates we produce.

¹¹ Sampling error makes the variance of the Empirical Bayes estimates lower than the variance of the actual principal quality (see Jacob and Lefgren, 2005). Therefore, we adjust this variance upward by adding in the estimate of the variance of the sampling error.

the student was “excused” from the test.¹² Student scores are linked over time via an encrypted student identifier.

Student test scores are linked via the student identifier to a file containing the administrative records of all public school students in BC from 1999–2006 in grades four to seven. These data include information on gender, aboriginal status, participation in Special Education or English as a Second Language (ESL) programs, and each student’s residential postal code. In urban areas, postal codes are very small area consisting sometimes of one side of a city block. In less-populated areas, they can coincide with part or all of a town. We use postal codes to link Dissemination/Enumeration Area (DA) level information from the Canada Census as proxies for students’ Socio-Economic Status (SES).¹³ DAs are relatively small areas that are designed to contain roughly 400–700 people, and as such act as a reasonable proxy for a student’s SES. We attach information on household income, average dwelling value, education levels, unemployment rates, ethnic and immigrant composition, and the age distribution of each DA.

The third file contains information on all public schools in British Columbia from 1995 through 2006. For each school we know information such as what grades are offered, the number of teachers, and each school’s exact address and postal code. Most importantly, we know the name of the principal, which we use to create principal fixed effects in the econometric specification.¹⁴ To this dataset we append principal names and school numbers dating back to 1989, transcribed from the *Public and Independent Schools Book*, a document produced by the Ministry of Education in BC. These additional data allow us to create a more precise measure of principal experience.

¹² Students with poor English abilities and some students with severe disabilities are excused from writing provincial exams.

¹³ This process is described in detail in a data appendix that is available from the authors upon request.

¹⁴ The data were checked extensively to ensure that typos, misspellings, and name changes (due, for example, to marriage) did not result in identifying one individual as two different individuals.

B. Descriptive Statistics on Principals

We begin by analyzing descriptive statistics for the principals in our data, with a focus on mobility, which is key for our estimation strategy. In this section, we restrict our attention to the years 1999–2006, since those are the years for which we have data on school and student demographics. We also restrict the sample to elementary, middle, and elementary-secondary schools to keep the focus on the student population of interest. Table 1 contains basic statistics about the number, mobility and experience of principals in British Columbia. Between 1255 and 1362 principals are employed each year, and between 117 and 178 of them are newly hired into the system. The main factors affecting the number of principals hired each year are retirements in the previous year (which depends on the age of those within the cohort of principals at that time) and school openings or closings in a particular year.

Between 69 and 77 percent of principals stay in the same school from one year to the next. Eleven to fifteen percent moved to a different school within British Columbia and 10–18 percent leave the sample. This latter group includes those principals who retired, those who moved out of province, and those who stop working as a principal but remain in the system in some other capacity.

Each year, roughly one quarter of the sampled principals are in their first year of tenure at a school, and another quarter remain at the same school for five or more years. The other half remain at the same school between two and four years. This shows considerable turnover of principals within schools, which helps our estimation strategy because we need multiple principals at each school to be able to identify the principal fixed effects. Finally, about one third of principals are in their first three years of employment as a principal, and another third have been a principal for 10 or more years.¹⁵

¹⁵ Given that we can observe principals back to 1989, the longest a principal employed in 1999 can be seen in our sample is 10 years, while in the 2006 sample, the longest a principal can be seen is 18 years.

C. Regression Sample

Because we estimate a value-added model in test scores, we restrict our focus to students who wrote both the fourth- and seventh-grade mathematics and reading exams.¹⁶ Since we observe test scores between 1999 and 2006, and since three years pass between each test, most such students are observed between 2002 and 2006. There are 229,337 students observed in the grade seven test data between 2002 and 2006. We drop 27,286 of these students (11.9 percent) who are not observed in all years between grades four and seven, and a further 14,692 students (6.5 percent) who do not have a valid test score in both years. We also drop 657 students (< 1 percent) who are missing school information in grades five or six. Finally, we drop 3,079 students (1.3 percent) who are in schools that have less than ten test takers. Our final analysis sample consists of 183,623 student observations.

The summary statistics in Table 2 regarding test scores, demographics, and neighborhood census characteristics are based on this sample of 183,623 students. The fourth- and seventh-grade math and reading test scores are standardized to have a mean of zero and standard deviation of one in the population. The means of these test scores are slightly positive in our analysis sample, which indicates that the students who were excluded from the samples scored marginally worse than the students who were included in the samples.

As we explained in Section IV, in the school fixed effects model, principal effects are only identified for a subset of principals. In Table 3 we compare sample means of observable characteristics for identified and unidentified principals from the regression sample. Column 1 pools all principals, and columns 2 and 3 restrict the sample to identified principals and unidentified principals, respectively. On average, the identified principals have slightly less experience and have shorter tenure in schools. This is not surprising because a principal needs to be mobile to be included in our sample. The other differences are that unidentified principals are more likely to be in rural areas (where it is more difficult to move),

¹⁶ We exclude the writing test due to the nature of the test and the distribution of the scores for that test. It is more difficult for graders to judge the “correctness” of answers to a writing test as compared to math and reading. In practice, the distribution shows that scores on the writing test tend to cluster around certain values, indicating that graders probably used a rubric to evaluate performance. Such clustering makes it difficult to differentiate between students and makes it difficult to justify using writing scores as an outcome variable.

which in turn means the percent visible minority, immigrant, and ESL are lower. These principals also tend to belong to schools with slightly lower test scores.

VI. Principal Mobility

In our preferred specification, principal effects within schools are estimable because of principal mobility. The identifying assumption is that mobility is exogenous conditional on all the variables that appear in our model. In particular, if principals prefer schools located in high-income areas, or prefer schools with certain fixed attributes, this will not affect the causal interpretation. Rothstein (2010) shows that such random assignment is tenuous in the context of identifying *teacher* effects, because, for example, students may be sorted into classrooms based on lagged test score gains.¹⁷ The identifying assumption, however, is much less tenuous in the context of *principal* effects. First, identification of a principal effect is not compromised when students are non-randomly assigned to teachers within a school; indeed, we posit that this is likely one of the mechanisms driving the principal effect. Second, the sorting of principals across schools is apt to be driven by school factors that are relatively fixed over the time period we consider in this analysis (five years), and since we control for fixed characteristics, this does not compromise identification. Finally, we control for many time-varying observed characteristics that may be related to principal mobility.

In spite of the above, one may still worry that mobility is endogenous. For example, districts may remove principals from schools who perform poorly on current and lagged test scores, and if those scores are mean-reverting, we may mistakenly attribute a subsequent improvement in scores after a move to the incoming principal. As such, in this section we attempt to provide more information on the principal mobility process by first discussing the administrative rules governing principal mobility, and then using the data to assess how principals move based on observables.

¹⁷ Rothstein (2010) tests the identifying assumptions for teacher effects by checking whether current teacher assignment is related to past test score gains, conditional on past teacher assignment. In theory, one could repeat that exercise with principals. Due to data limitations (having only 2 time observations for each student), we cannot replicate that test in our paper because we cannot construct a lagged gain score.

The administrative rules governing principal mobility in BC are discussed in detail by Coelli and Green (forthcoming). They find that mobility is determined at the school district level and is generally overseen by the Superintendent of Schools. Of the 60 districts in BC, 14 are found to mention principal rotation somewhere in their district policies; these 14 districts are larger than average in size. Their main example of a district policy towards rotation comes from the Vancouver School Board, whose written procedure is to transfer principals between schools roughly every five years. Other districts are sometimes less precise, stating in their district policies minimum and maximum times until principal transfer, leaving much discretion to the school district. Smaller districts that do not have written rotation procedures are found not to have included them in the district policy because there are too few schools to make such a strategy work.¹⁸

In Table 4 we empirically evaluate the determinants of principal mobility. We take a sample of all schools that offer the grade 7 FSA test between 1999 and 2006, and use a regression to relate principal mobility between those schools to observable characteristics of the principals and schools. We regress an indicator for a principal move between t and $t+1$ on school and neighborhood characteristics at time t (and in some specifications, $t-1$). In Columns 1-3 the dependent variable equals 1 only if the principal moves from one school to another school in the sample. In columns 4-6, the indicator equals 1 for any move, including switches to schools out of the sample or exits from the sample. Column 1 includes a full set of controls but excludes school fixed effects. Results show no evidence that principal turnover is related to any school and neighborhood characteristics, including math and reading test scores. Note in particular that the test score coefficients are very small and imprecise.¹⁹ The main significant predictors of mobility are principal experience indicator variables. When we add school effects in Column 2, the conclusions do not change. Column 3 adds lagged test scores, and there continue to be no significant relationships. The

¹⁸ The British Columbia School Trustees Association has gathered every school district's policies at the following website: <http://www.bcsta.org:8080/docushare>

¹⁹ The unit of measurement of each test score variable is student-level standard deviations. A one standard deviation increase in student test scores is roughly equivalent to a 0.5 standard deviation increase in test scores at the school level. Thus, these results should be interpreted as roughly half of their reported size.

same exercise is conducted in columns 4-6 for all switches. Once again, the only significant determinant of principal mobility is experience. Thus, while we cannot be certain that mobility is unrelated to unobservable factors that change during our time period, these results provide evidence that current and lagged school performance does not determine moves.

VII. Results

A. Variation in Principal Quality

Table 5 presents the baseline estimates from equation 1, and includes details on the distribution of δ_p including the standard deviation along with the 10th, 25th, 50th, 75th, and 90th percentiles. Both the standard deviation and the adjusted standard deviation are included. In addition, the gap between the 75th percentile and the median principal fixed effect is calculated as a representation of the difference between a "good" principal and an average principal. The number of principals and the number of schools used in the regression are listed along with the p-value for the F-test of the joint significance of δ_p (i.e., $\hat{\delta}_p = 0$ for all p).

Columns 1 and 6 present estimates of the distribution of the principal fixed effects from a specification that includes no demographic or census control variables, and no school fixed effects. This specification essentially computes the deviation of each principal's average score across students from the mean principal. The estimates in column 1 imply that moving one standard deviation up the distribution of principals increases test scores by 0.336 standardized units in math. Comparing the principal at the 75th percentile to the median leads to a more modest increase of 0.202 standardized student points. We find similar results in column 6, which shows the results from the reading test. The F-test of the joint significance of δ_p easily rejects the null of no principal fixed effects at a high level of significance in both cases. After adjusting our estimates for sampling error, the standard deviation in column 1 drops to 0.321 for math and 0.302 for reading, implying that roughly five percent of the math and reading estimate are

due to sampling error. Columns 2 and 7 add demographic and census control variables.²⁰ Estimates shrink slightly, but the story is largely unchanged. Interestingly, our estimates do not see the same kind of attenuation that was found in teacher effects when adding demographic control variables (see Aaronson, Barrow, and Sander 2007).

In Columns 3 and 8 we drop the principal effects that would be unidentified in a school fixed effects specification to assess the effect of limiting the sample of principals before explicitly including school fixed effects. This restriction decreases the number of principal effects estimated from 1323 to 1077. The standard deviation does not change dramatically compared to columns 2 and 7. This provides some evidence that our sample used to estimate both school and principal effects is not biased due to sample selection.

Columns 4 and 9 report estimates when school fixed effects are added to the model. Including school fixed effects is preferred because it holds constant all fixed school components like curriculum, school size, fixed family, and some community characteristics in computing principal effects. This is important because these factors may be highly correlated with principal effects, and because both fixed school components and principal effects impact student achievement. The estimated standard deviations are 0.231 and 0.224, and the adjusted standard deviations are 0.196 and 0.188 for math and reading, respectively. As before, the F-test rejects that the within-school principal fixed effects jointly equal zero at the one percent level.

Columns 5 and 10 show the Empirical Bayes estimates. These results indicate that an increase of one standard deviation increases math scores by 0.167 standardized units and reading scores by 0.158

²⁰ These variables include an indicator for whether the student is male, aboriginal, or in special education, or has repeated a grade. It also includes the number of students enrolled in the school and the number of teachers, along with the number of students excused from the math test and reading test. In addition, it includes the percentage of students at each school who are males, who are in special education, and who are ESL. Finally, the following DA level community characteristics are also included: average household income, percentage of individuals with a high school diploma, percentage of individuals with a university degree, percentage of individuals 25 years or older who are unemployed, percentage of the population who are immigrants, percentage of the population who are visible minorities, average value of dwellings, and percentage of the population who are older than 65 years.

standardized units. Comparing the principal at the 75th percentile to the median gives a much more modest 0.076 unit increase for math scores and a 0.070 unit increase for reading. Note that when calculating Empirical Bayes estimates, the sampling error makes the variance of the estimates lower than the variance of the actual principal quality (see Jacob and Lefgren, 2005). Therefore, we adjust this variance upward by adding in our estimate of the variance of the sampling error. The adjusted estimates for math is an increase of 0.207 standardized units and an increase of 0.199 standardized units in reading.

The results in Table 5 show that overall, individual principals can have a substantial impact on student achievement in reading and math. How large are these effects? One way to interpret their size is to compare them against other inputs to student cognitive ability. Using a similar research design, Rockoff (2004) estimates teacher fixed effects using one-year test score gains. He estimates that the unadjusted standard deviation of estimated teacher effects is roughly 0.21 in reading and 0.29 in math, and the adjusted effect is 0.10 for reading and math. Aaronson, Barrow, and Sander (2007) estimate a standard deviation of 0.13 for teacher effects on math scores in Chicago. Note that compared against these teacher effects, our magnitudes are likely larger in part because we use gains over three years rather than one year as in most other studies (see the results in Table 6 below).

What types of fixed principal characteristics might contribute to the effects we observe? Principals are not likely to have a direct impact on student performance through their teaching ability, as they generally do not interact directly with students in the classroom. Thus, principals affect students through indirect channels; the most likely such channels are through teachers and overall school policies such as those that include changes to the curriculum. Good principals will hire the best teachers and fire the worst (Jacob 2010), match teachers to students based on ability, set appropriate timetables, and evaluate/monitor teacher performance. Principals are also in charge of setting school-level policies on issues such as instruction and discipline. In addition, principals are responsible for acting as a liaison between the school and the district, a role which will involve passing information from school to district

and vice versa as well as making sure that district policies with respect to things like curriculum are implemented correctly.

All of these factors likely contribute in some way to what makes one principal better than another. Because we do not have data on the decisions and policies of these principals, we cannot uncover the independent influence of any of these factors, even though they would be of interest to policy makers. Nevertheless, our results are very useful in the sense that they can identify who are the good and bad principals in terms of increasing school-level achievement scores. Even without knowing the precise mechanism by which principals affect student performance, it is still possible for policy makers to use this information to improve student achievement in particular schools. This could be accomplished, for example, by identifying the best principals and allocating them strategically to underperforming schools in order to boost performance in those schools. Because the principal effects we estimate are fixed across time, this would not have any effect on aggregate achievement, but could help to reduce achievement gaps between good and bad schools.

B. Robustness Checks

To check the robustness of our results, Table 6 replicates the results of Table 5 using different specifications or different samples.²¹ We report only estimates that include school fixed effects. These estimates correspond to columns 4 and 5 for math and columns 9 and 10 for reading in Table 5. Therefore, the estimates in all columns include the demographic and census control variables in addition to school fixed effects.

The estimates presented in Table 5 do not take into the account that students may have different levels of exposure to their grade seven principals. Therefore, for our first robustness check, in Panel A we multiply each principal indicator variable by the number of years the student is exposed to their grade

²¹ Appendix Table 1 displays the descriptive statistics for each of these samples. Estimates of principal fixed effects without the inclusion of school fixed effects for this table are available from the authors upon request.

seven principal. The effects will be very similar to 1-year gains if the principal effects are linear. Columns 1 and 3 display the standard deviation and the adjusted standard deviation from equation 1 including the time with principal interactions. Columns 2 and 4 display both the Empirical Bayes standard deviation and adjusted standard deviation. The estimates in Panel A are 0.052-0.087 standard deviations lower than the corresponding estimates in Table 5. These estimates are the closest estimates to the single-year gains estimates from the teacher effect literature and have a similar magnitude to this literature.

One problem with the results from Table 5 and the above robustness check is that they include effects for only grade seven principals. Students may have a different principal in grades 5 and 6, and these principals may impact grade 7 scores. In Panel B we generalize the principal effects in Panel A to equal to the number of years each student has been with any given principal between grades five and seven.²² Including the additional principals increases the number of principal effects we can identify to 1660 from 1323.²³ The estimates in Panel B are slightly larger than the estimates calculated in Table 5.

In Panel C we estimate the set of principal fixed effects using only students who had the same principal in grades five through seven but had a different principal in grade four.²⁴ Using this sample, we can attribute most of the value added measure to the grade seven principal. These estimates are also slightly larger than the estimates found in Table 5.

Panel D limits the sample to students who stayed in the same school between grades four and seven. The purpose of this exercise is to see if the results are robust to student sorting across schools in search of a better principal. We have mostly been concerned with principals choosing to work in the best schools, but an equally plausible, although less likely, situation is that parents locate near a school

²² The sample size is smaller for in this regression because we needed to drop students who were in schools with very few students in grades 5 and 6.

²³ This increase in the number of principals arises for two reasons. First, there are a small number of principals who are not observed in grade seven, but are observed in grades five or six. Second, adding these principals creates additional connections between principals and schools, which allows us to identify principal effects among grade seven principals that were previously not identifiable.

²⁴ These estimates are similar to the estimates if we use the sample of students who had the same principal in grade five-seven but not necessarily a different principal in grade four.

because of the principal. The results from this exercise are similar to the main results. This suggests that student sorting is not causing substantial bias in our results.

Finally, our results might be driven by trends in test scores over time, which would not be absorbed by our fixed effects, and could persist even though we control for time-varying factors. In Panel E, we add school-specific linear time trends to the specification. Estimation is problematic in this case because a school is observed for a maximum of five years, and it is therefore difficult to statistically estimate these trends. Indeed, we were not able to identify a trend for schools that had a different principal in each year in the sample, or were observed for too few years (100 schools out of 912). The results show a higher standard deviation in principal effects in both math and reading. A significant amount of the increase in dispersion is driven by noise, as evidenced by the difference between the adjusted and unadjusted standard deviations. These results suggest that if there is any principal sorting due to trends, then our main results understate the value of principals.

C. Match Effects

Our analysis thus far provides estimates of principal effects that are fixed across schools (and time). It is entirely plausible, however, that a principal's effectiveness varies across schools. These "match effects" might arise if a principal's effect on student achievement depends on things like interactions with the existing teaching staff, demographic composition of the student body, location preference of the principal, or any other complementarity between the principal and school. If principals sort themselves into schools where they have good matches, our principal effect estimates will be biased upward since they partly measure the impact of fixed principal characteristics and partly the impact of the match. Match effects are also important in their own right because their existence implies that considering the school to which a principal is assigned is crucial in terms of improving student achievement in particular schools and reducing achievement gaps.

In this section we provide estimates of principal-school matches. Woodcock (2008) outlines two different methods for estimation. The first is the orthogonal fixed effects estimator, which identifies match effects by imposing that they are orthogonal to the principal and school effects.²⁵ This estimator has one major downside: it mechanically forces match effects to sum to zero for each principal and school. Thus, matches with one principal or school are assigned a zero match effect, which is not desirable. Therefore, while we provide orthogonal fixed effects estimates, we prefer the hybrid random effects estimator. This method avoids the issues with orthogonal fixed effects, but comes at the cost of imposing orthogonality between the principal, school, match, and residual.²⁶

Table 7 presents the match effects estimates. Columns 2 and 4 in Panel A report the standard deviations from the orthogonal fixed effects estimator, alongside the results from our main specification in columns 1 and 3 for comparison. For both math and reading scores, the principal and school effects are identical regardless of whether match effects are included. The match effects themselves are small relative to the principal and school effects: the standard deviation is 0.033 in math and 0.034 in reading.

Panel B of Table 7 presents the results of our preferred hybrid random effects estimator. For comparison, we also estimate the hybrid random effects model without match effects in columns 1 and 3. The standard deviation of the principal effects in columns 1 and 3 are much smaller than their corresponding fixed effect estimates from Panel A, but is of a similar magnitude to the Empirical Bayes estimates from Table 5. Columns 2 and 4 contain the hybrid random effects estimates with match effects. Here differences emerge in the size of the standard deviations. The standard deviation of the principal effect for math drops to 0.093, and the corresponding estimate for reading drops to 0.059. The school effect is roughly unchanged. Match effects are large with a standard deviation of about 0.160 for both

²⁵ The estimator proceeds by estimating the parameters in β and the experience effects by OLS using principal by school fixed effects (i.e., match fixed effects). It then take the residual plus fixed effects from that regression, computes within-match means, and regresses those on separate principal and school fixed effects. The residuals from this second regression are the match effects.

²⁶ This estimator has the same first step as the orthogonal fixed effects estimator, which allows arbitrary correlation between the fixed components and the regressors in the model. The variances of the principal, school, and match effects are then estimated via Restricted Maximum Likelihood (REML). The variances reported in Table 7 are the REML estimates.

math and for reading—roughly five times larger than the estimates derived from the fixed effects estimation.

Match effects have important implications for policy. That match effects appear to matter more than fixed principal effects (given random effects estimates) on their own implies that policy makers could not simply assign “good” principals to any school and expect them to have the same impact. A principal’s effect depends on the school where they are located, so if principals are rotated across schools in part to help improve student performance, the key will be to assign them to a school with the best match so they will have maximum impact. These estimates must be considered somewhat cautiously, however, because estimating them requires pushing the data very hard and/or imposing some restrictions. Fortunately, even when match effects are included, the fixed principal effects still exert a substantial effect on student outcomes, so assigning a principal to a school that is a good match will help bolster that effect.²⁷

D. Principal Experience

The effects of fixed principal characteristics on student achievement are the focus of this paper, but we also have data on principal experience, which is an important characteristic of principals that may impact student achievement. In our regressions, we separately analyze the effect on student performance of principal tenure in a school and overall principal experience. We model principal tenure with a set of dummy variables for each level of experience up to five years and a dummy for any amount of experience five years and over, with one year of experience as the excluded group. We model overall principal experience using dummies for groups of four to six, seven to ten, and ten or more years, with one to three years as the excluded group. With overall experience, these groupings reflect the fact that we are less interested in the yearly changes in experience, and more interested in comparing inexperienced to

²⁷ Unfortunately, the British Columbia data does not contain matches between students and teachers. Therefore, with these data we cannot explore interactions effects between “good” teachers and “good” principals.

somewhat experienced to very experienced principals. With principal experience at a school, we are more interested in how one-year changes in experience affect student scores.

Panel A of Table 8 reports the results from specifications that include our measures of principal tenure at a school. Columns 1 and 3 display estimates when school fixed effects are not included in the model. In these specifications, there is no distinct relationship between experience and student performance, except that principals who have been at a school for five years or longer boost test scores slightly. Including school fixed effects in columns 2 and 4 does little to change the pattern across years of experience, except that experienced principals no longer exert any influence on math scores.

These results contrast with those of Clark, Martorell, and Rockoff (2009), who find that principal experience still matters with math scores, but less for reading scores when principal and school fixed effects are held constant.²⁸ There are several factors that might explain differences in our results, most of which are speculative since we cannot formally test these differences. First, differences between the New York City and British Columbia school systems could be responsible. Their research suggests that New York City recently has placed a great amount of emphasis on their principals. This emphasis may not have been seen in British Columbia.²⁹ A second explanation is differences in the degree to which sorting causes bias. Both our study and theirs control for similar demographics and other characteristics in addition to fixed school and principal factors, so this is not likely to be the issue.

Panel B of Table 8 presents the results of using overall principal experience rather than school tenure. With no school fixed effects in the model, the most experienced principals increase student scores by 0.055 standard deviations in math and by 0.069 standard deviations in reading relative to the least experienced principals. Once we control for school effects, however, the pattern disappears, leaving no

²⁸ In their paper, Clark, Martorell, and Rockoff (2009) use principal by school effects rather than separate principal and school effects, as they are not interested in interpreting either fixed effect. In practice, both methods give similar results.

²⁹ For example, several very unique and intensive programs designed to train principals to become effective leaders are available to individuals in the New York system, such as the Aspiring Principals Program (APP). It is not clear that British Columbia principals have access to similar types of intensive training programs.

relationship between experience and test scores. Overall, we find no evidence of any systematic relationship between experience and student achievement.

These results are interesting as they suggest that experience does not matter. This is important from a policy perspective since it suggests that if the goal is to boost performance of underperforming students, and reduce achievement gaps, a more effective approach would be to identify high-ability principals and then allocate them to schools, rather than relying on a principal to boost performance by gaining experience.

VIII. Conclusion

We estimate the impact of fixed principal characteristics on student performance. Results show that principals have a substantial impact on both math and reading scores. Our main results show that a one standard deviation shift up the principal quality distribution can increase achievement by approximately 0.2 standard deviations in math and reading. When we extend the model to allow the effectiveness of the principal to depend on the school where s/he is employed, our most robust model shows that a one standard deviation improvement in principal quality increases math scores by about 0.1 standardized units and reading scores by 0.06 standardized units. Additionally, a quality match improves student performance by approximately 0.15 standard deviations in math and in reading. We also estimate that a more experienced principal within a school or overall has no significant effect on student achievement.

These results have important implications for policy. The main implication is that shifting principals between schools has the potential to significantly reduce achievement gaps.³⁰ Some of this improvement depends on where the principal works, but a sizable portion is portable across schools. Policy makers could identify the most effective principals, and allocate them between schools to potentially reduce achievement gaps. Alternatively, we could identify the best principals and dig deeper to

³⁰ However, the optimal reassignment of existing principals, with respect to a gap-minimizing objective function, might affect the supply of effective principals.

learn more about the attributes that make them so effective. Information gleaned from that exercise could be used to train underperforming principals. What are the fixed characteristics that make one principal better than another? We cannot identify these in our data, but prior literature can shed some light on this issue. The results of Clark, Martorell, and Rockoff (2009) show that education level and pre-principal experience are not likely to be among the factors, while the evidence on training programs is mixed. Principals are most likely to affect student outcomes via their policies. Figlio and Sass (2010) show that when a new principal enters a school, that individual is most likely to change policies on teacher incentives, curriculum, and those policies that boost performance of low-achieving students. They further show that the most effective principals focus on policies that boost performance of low-achieving students. Finally, Jacob (2010) shows that principals are likely instrumental in the hiring and firing decisions about teachers, which may trickle down to affect student achievement.

IX. References

- Aaronson, Daniel, Lisa Barrow, and William Sander. 2007. "Teachers and Student Achievement in the Chicago Public High Schools." *Journal of Labor Economics* 25(1):95–135.
- Ballou, Dale, and Michael J. Podgursky. 1993. "What Makes a Good Principal? How Teachers Assess the Performance of Principals." *Economics of Education Review* 14(3):243–52.
- Branch, Gregory F., Eric A. Hanushek, and Steven G. Rivkin. 2010. "Estimating Principal Effectiveness." Unpublished manuscript.
- Brewer, Dominic J. 1993. "Principals and Student Outcomes: Evidence from U.S. High Schools." *Economics of Education Review* 12(4):281–93.
- Buddin, Richard, and Gema Zamarro. 2009. "Teacher Qualifications and Student Achievement in Urban Elementary Schools." *Journal of Urban Economics* 66(2):103–15.
- Clark, Damon, Paco Martorell, and Jonah Rockoff. 2009. "School Principals and School Performance." Working Paper 38. Washington, DC: The Urban Institute.
- Coelli, Michael, David A. Green, and William P. Warburton. 2007. "Breaking the Cycle? The Effect of Education on Welfare Receipt Among Children of Welfare Recipients." *Journal of Public Economics* 91 (7-8): 1369–1398.
- Coelli, Michael, and David Green. forthcoming. "Leadership Effects: School Principals and Student Outcomes." *Economics of Education Review*.
- Corcoran, Sean P., Amy E. Schwartz, and Meryle Weinstein. 2009. "The New York City Aspiring Principals Program: A School Level Examination." Working Paper. NYU Institute for Education and Social Policy.
- Cullen, Julie, and Michael J. Mazzeo. 2008. "Implicit Performance Awards: An Empirical Analysis of the Labor Market for Public School Administrators." Unpublished manuscript.
- DeCicca, Phil, and Justin Smith. 2011. "The Long Run Impacts of Early Childhood Education: Evidence from a Failed Policy Experiment." Unpublished manuscript.
- Dhuey, Elizabeth. 2011. "Middle School or Junior High? How Grade Level Configurations Affect Student Achievement." University of Toronto Working Paper.
- Eberts, Randall W., and Joe A. Stone. 1988. "Student Achievement in Public Schools: Do Principals Make a Difference?" *Economics of Education Review* 7(3):291–99.
- Figlio, David, and Tim Sass. 2010. "The Policy Choices of Effective Principals." Unpublished manuscript.
- Grissom, Jason A., and Susanna Loeb. 2009. "Triangulating Principal Effectiveness: How Perspectives of Parents, Teachers, and Assistant Principals Identify the Central Importance of Managerial Skills." CALDER Working Paper 35. Washington, DC: The Urban Institute.

- Hallinger, Philip, and Ronald H. Heck. 1998. "Exploring the Principal's Contribution to School Effectiveness: 1980–1995." *School Effectiveness and School Improvement* 9(2):157–91.
- Harris, Douglas N., and Tim R. Sass. 2006. "Value-Added Models and the Measurement of Teacher Quality." Working Paper. Florida State University.
- Hornig, Eileen Lai, Daniel Klasik, and Susanna Loeb. 2009. "Principal Time-Use and School Effectiveness." CALDER Working Paper 34. Washington, DC: The Urban Institute.
- Jacob, Brian A. 2010. "Do Principals Fire the Worst Teachers?" NBER Working Paper #15715.
- Jacob, Brian, and Lars Lefgren. 2005. "Principals as Agents: Subjective Performance Measurement in Education" *NBER Working Paper 11463*.
- Jacob, Brian, and Lars Lefgren. 2008. "Can Principals Identify Effective Teachers? Evidence on Subjective Performance Evaluation in Education" *Journal of Labor Economics* 26(1):101–36.
- Kane, Thomas J., and Douglas O. Staiger. 2002. "Volatility in School Test Scores: Implications for Test-based Accountability Systems." In *Brookings Papers on Education Policy 2002*, ed. Diane Ravitch, 235–69. Washington, DC: Brookings.
- Kane, Thomas J., Jonah E. Rockoff, and Douglas O. Staiger. 2008. "What Does Certification Tell Us About Teacher Effectiveness?" *Economics of Education Review* 27(6): 615–31.
- Mihaly, Kata, Daniel McCaffrey, J. R. Lockwood, and Tim R. Sass. 2010. "Centering and Reference Groups for Estimates of Fixed Effects: Modifications to felsdsvreg." *The Stata Journal* 10(1):82–103.
- Nye, Barbara, Spyros Konstantopoulos, and Larry V. Hedges. 2004. "How Large Are Teacher Effects?" *Educational Evaluation and Policy Analysis* 26(3):237–57.
- Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2005. "Teachers, Schools, and Academic Achievement." *Econometrica* 73(2):417–58.
- Rockoff, Jonah E. 2004. "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data." *American Economic Review* 94(2):247–52.
- Rothstein, Jesse. 2010. "Teacher Quality in Educational Production: Tracking, Decay, and Student Achievement." *The Quarterly Journal of Economics* 125(1): 175-214.
- Sanders, William L., and June C. Rivers. 1996. "Cumulative and Residual Effects of Teachers on Future Student Academic Achievement." Knoxville, TN: University of Tennessee Value-Added Research and Assessment Center.
- Woodcock, Simon. 2008. "Match Effects." Working Paper. Simon Fraser University

Table 1

Principal characteristics in British Columbia, 1999-2006

Year	1999	2000	2001	2002	2003	2004	2005	2006
Total Employed	1355	1362	1357	1325	1280	1274	1259	1255
New Hires	150	136	145	117	122	139	178	169
<i>Mobility</i>								
Same school	74%	77%	77%	72%	73%	72%	69%	
Different school	15%	12%	11%	15%	13%	11%	13%	
Out of sample	10%	12%	12%	13%	14%	17%	18%	
<i>Tenure at a school</i>								
1 yr	22%	25%	22%	21%	24%	26%	26%	29%
2 yrs	18%	19%	23%	21%	18%	21%	22%	22%
3 yrs	20%	14%	15%	19%	17%	15%	18%	16%
4 yrs	13%	15%	10%	12%	15%	12%	11%	11%
5+ yrs	27%	26%	29%	27%	25%	27%	23%	21%
<i>Experience as principal</i>								
1 to 3 yrs	30%	28%	29%	28%	28%	27%	32%	35%
4 to 6 yrs	26%	32%	23%	24%	22%	24%	21%	20%
7 to 9 yrs	16%	11%	19%	20%	23%	19%	18%	16%
10+ yrs	28%	29%	29%	29%	28%	30%	29%	29%

Notes: Numbers in this table are based on 10,467 principal-year observations, 2505 principals, and 1527 schools. The sample includes all elementary, middle, elementary-secondary, and junior secondary schools in British Columbia. Some principals are observed in multiple schools in the same year in certain rural locations. For these cases, we kept the principal's "main" school, which we define as the school in which the principal is observed most often. New hire is defined as a principal who is observed in the data for the first time as a principal.

Table 2
Descriptive statistics for student data

	Mean	Std Dev.
<u>4th grade test scores</u>		
Math	0.071	0.922
Reading	0.080	0.891
<u>7th grade test scores</u>		
Math	0.020	0.934
Reading	0.067	0.900
<u>Demographics</u>		
Male	0.504	0.500
Aboriginal	0.103	0.303
ESL	0.209	0.406
In special education	0.059	0.236
Attends rural school	0.126	0.331
<u>Neighbourhood census characteristics</u>		
Percent over 65 yrs old	0.135	0.060
Percent visible minority	0.214	0.226
Percent immigrant	0.248	0.157
Percent unemployed	0.056	0.031
Percent no high school degree	0.220	0.072
Percent high school degree	0.279	0.036
Percent university degree	0.207	0.100
Average household income	60,823	16,815
Average dwelling value	306,313	139,547
Number of observations	183,623	

Notes: Test scores represent the z-score by year, grade, and subject among the population of students prior to sample exclusions. There are 1323 principals and 912 schools in the estimation sample.

Table 3

Average principal, school, and neighbourhood characteristics for identified and unidentified principals

	All principals (1)	Identified principals (2)	Unidentified principals (3)
<i><u>Principal characteristics</u></i>			
Principal tenure	3.16	2.87	4.40
Principal experience	6.65	6.28	8.23
Number of years in data	2.95	2.79	3.65
Number of principals	1323	1077	246
<i><u>School characteristics</u></i>			
7th grade math score	-0.040	-0.028	-0.095
7th grade reading score	-0.049	-0.041	-0.088
Excused from 7th grade math test	0.055	0.057	0.045
Excused from 7th grade reading test	0.053	0.056	0.043
Percent male	0.516	0.516	0.513
Percent aboriginal	0.138	0.135	0.147
Percent ESL	0.234	0.244	0.190
Percent special education	0.096	0.097	0.094
Rural school	0.215	0.206	0.255
<i><u>Neighbourhood census characteristics</u></i>			
Percent over 65 yrs old	0.133	0.131	0.143
Percent visible minority	0.203	0.213	0.160
Percent no high school degree	0.234	0.234	0.231
Percent high school degree	0.277	0.276	0.277
Percent university degree	0.198	0.198	0.196
Percent immigrant	0.237	0.243	0.210
Percent unemployed	0.062	0.062	0.065
Average household income	59071	59451	57407
Average dwelling value	293240	297450	274807

Notes: Data based on regression sample of 183,623 observations. Average school characteristics and average neighbourhood characteristics are computed by averaging all students observed in grade 7 in a school. The number of connected groups for the identified principals is 311.

Table 4
Regression of principal turnover on school and neighbourhood characteristics

	School to school switch			Any switch		
	(1)	(2)	(3)	(4)	(5)	(6)
7th grade math score	0.004 (0.015)	0.005 (0.020)	-0.007 (0.024)	0.002 (0.020)	-0.026 (0.026)	-0.017 (0.032)
7th grade reading score	0.026 (0.014)	0.005 (0.019)	0.022 (0.023)	0.000 (0.019)	0.028 (0.025)	0.031 (0.031)
Lagged 7th grade math score			-0.006 (0.024)			-0.014 (0.031)
Lagged 7th grade reading score			0.004 (0.022)			0.006 (0.029)
Percent excused from 7th grade math test	-0.002 (0.012)	0.003 (0.015)	0.003 (0.017)	-0.007 (0.016)	-0.011 (0.020)	-0.010 (0.022)
Percent excused from 7th grade reading test	0.003 (0.013)	0.000 (0.015)	-0.004 (0.017)	0.014 (0.017)	0.009 (0.020)	0.002 (0.023)
School enrolment	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000	0.000 0.000
Number of elementary teachers in school	0.001 (0.001)	0.001 (0.002)	0.001 (0.003)	-0.003 (0.001)	0.000 (0.003)	0.001 (0.003)
Percent ESL	-0.001 (0.003)	-0.002 (0.006)	-0.002 (0.007)	0.002 (0.004)	-0.001 (0.007)	-0.004 (0.009)
Percent male	-0.002 (0.003)	0.003 (0.004)	0.001 (0.005)	-0.005 (0.004)	-0.003 (0.005)	-0.003 (0.006)
Percent aboriginal	0.003 (0.003)	0.005 (0.006)	0.009 (0.007)	0.003 (0.004)	0.008 (0.007)	0.013 (0.009)
Percent special education	0.005 (0.004)	0.012 (0.006)	0.012 (0.008)	0.010 (0.005)	0.007 (0.008)	0.000 (0.010)
Average household income	-0.004 (0.004)	-0.005 (0.008)	-0.010 (0.009)	0.002 (0.005)	0.007 (0.011)	-0.004 (0.012)
Percent no high school degree	-0.020 (0.010)	-0.052 (0.021)	-0.039 (0.025)	-0.022 (0.013)	-0.045 (0.028)	-0.037 (0.033)
Percent high school degree	-0.015 (0.012)	-0.034 (0.022)	-0.028 (0.026)	-0.031 (0.016)	-0.046 (0.029)	-0.039 (0.034)
Percent university degree	-0.017 (0.010)	-0.002 (0.026)	0.006 (0.032)	-0.020 (0.013)	-0.006 (0.035)	0.000 (0.043)
Percent unemployed	0.010 (0.010)	0.027 (0.021)	0.018 (0.025)	0.010 (0.013)	-0.012 (0.028)	-0.027 (0.033)
Percent immigrant	-0.008 (0.011)	-0.022 (0.028)	0.012 (0.032)	0.006 (0.015)	-0.031 (0.036)	0.002 (0.042)
Percent visible minority	0.010 (0.008)	0.018 (0.021)	0.000 (0.025)	-0.002 (0.010)	0.015 (0.028)	0.007 (0.033)
Average value of dwelling	0.001 (0.001)	-0.001 (0.001)	-0.001 (0.002)	0.000 (0.001)	-0.004 (0.002)	-0.004 (0.002)
Percent 65+ years old	-0.044 (0.078)	0.022 (0.249)	-0.149 (0.296)	-0.056 (0.102)	0.222 (0.328)	0.164 (0.392)
4 to 6 yrs principal experience	0.052 (0.012)	0.116 (0.014)	0.113 (0.015)	0.064 (0.015)	0.178 (0.018)	0.183 (0.021)
7 to 9 yrs principal experience	0.031 (0.013)	0.139 (0.016)	0.135 (0.019)	0.062 (0.017)	0.263 (0.022)	0.285 (0.025)
10+ yrs principal experience	0.007 (0.011)	0.158 (0.017)	0.162 (0.020)	0.087 (0.015)	0.390 (0.022)	0.428 (0.026)
R squared	0.017	0.027	0.027	0.017	0.066	0.074
Number of observations	6458	6458	5254	6458	6458	5254
<u>Additional control variables:</u>						
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
School fixed effects	No	Yes	Yes	No	Yes	Yes

Notes: Data for this regression are at the school by year level, and include all schools that offer the grade 7 numeracy and reading tests from 1999-2005. A small number of principals who serve at 2 schools in the same year are excluded from the regression. School to school switch refers to all principal mobility from one school in the sample to another. Any switches refers to school switches plus switches to schools outside the sample or exits from the sample.

Table 5

Student level estimates of principal fixed effects, 4th to 7th grade gains

	Math					Reading				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Standard deviation	0.336	0.309	0.288	0.231	0.167	0.318	0.309	0.276	0.224	0.158
Adjusted standard deviation	0.321	0.293	0.270	0.196	0.207	0.302	0.292	0.258	0.188	0.199
10th percentile	-0.349	-0.298	-0.286	-0.265	-0.191	-0.259	-0.231	-0.227	-0.219	-0.154
25th percentile	-0.175	-0.150	-0.149	-0.120	-0.087	-0.097	-0.080	-0.094	-0.097	-0.068
50th percentile	-0.006	0.002	-0.008	0.010	0.007	0.039	0.039	0.027	0.002	0.001
75th percentile	0.196	0.162	0.157	0.116	0.084	0.167	0.148	0.136	0.102	0.072
90th percentile	0.395	0.335	0.319	0.265	0.191	0.273	0.250	0.244	0.225	0.159
(75th percentile - median)	0.202	0.159	0.165	0.106	0.076	0.128	0.109	0.108	0.100	0.070
number of principals	1323	1323	1077	1077	1077	1323	1323	1077	1077	1077
number of schools	912	912	912	912	912	912	912	912	912	912
P-value on F-test	0.000	0.000	0.000	0.000		0.000	0.000	0.000	0.000	
<u>Control variables:</u>										
Demographics	no	yes	yes	yes	yes	no	yes	yes	yes	yes
Census	no	yes	yes	yes	yes	no	yes	yes	yes	yes
School fixed effects	no	no	no	yes	yes	no	no	no	yes	yes

Notes: Figures in this table represent the distribution among all 1323 or 1077 principal effects estimated via equation 1. See Section VII.A. for a description of the demographic and census variables included in the regressions. Coefficients are measured in standard deviations. The p-value on the F-test is the test for the joint significance of all principal fixed effects in the model. Estimates in columns 5 and 10 are Empirical Bayes shrinkage estimates.

Table 6

Estimates of principal fixed effects using different samples, 4th to 7th grade gains

	Math		Reading	
	(1)	(2)	(3)	(4)
<u>Panel A: 7th grade principals*</u>				
Standard deviation	0.145	0.115	0.137	0.105
Adjusted standard deviation	0.129	0.133	0.120	0.124
<u>Panel B: All principals 5th-7th grades*</u>				
Standard deviation	0.314	0.196	0.263	0.124
Adjusted standard deviation	0.248	0.274	0.181	0.228
<u>Panel C: Students with same principal: 5th-7th grades</u>				
Standard deviation	0.274	0.204	0.241	0.165
Adjusted standard deviation	0.236	0.246	0.199	0.214
<u>Panel D: Students in same school</u>				
Standard deviation	0.284	0.204	0.276	0.196
Adjusted standard deviation	0.241	0.254	0.233	0.246
<u>Panel E: School-Specific Linear Trends</u>				
Standard deviation	0.410	0.220	0.382	0.179
Adjusted standard deviation	0.300	0.355	0.262	0.331
<u>Control variables:</u>				
Demographics	yes	yes	yes	yes
Census	yes	yes	yes	yes
School fixed effects	yes	yes	yes	yes

Notes: See Section VII.A. for a description of the demographic and census variables included in the regressions. Estimates in columns 2 and 4 are Empirical Bayes shrinkage estimates. Panel A includes sample of 1077 principals and 912 schools. Panel B includes sample of 1660 principals and 1191 schools. Panel C includes sample of 337 principals and 763 schools. Panel D includes sample of 947 principals and 798 schools. In Panel E, time trends are identified for 812 schools. *Estimation includes principal fixed effects multiplied by years with principal.

Table 7

Estimates of match effects

	Math		Reading	
	Principal/School	Match effects	Principal/School	Match effects
	(1)	(2)	(3)	(4)
<i>Panel A: Orthogonal fixed effects estimates</i>				
Standard deviation for principal	0.231	0.231	0.224	0.224
Standard deviation for school	0.327	0.327	0.319	0.319
Standard deviation for match		0.033		0.034
<i>Panel B: Hybrid random effect estimates</i>				
Standard deviation for principal	0.181	0.093	0.166	0.059
Standard deviation for school	0.220	0.211	0.208	0.202
Standard deviation for match		0.160		0.160

Notes: All regression include all demographic and census control variables listed in Section VII.A. Fixed Effects estimates are based on 1077 principals, 662 schools, and 1449 matches between 1999 and 2006. Random Effects estimates based on 1323 principals, 912 schools, and 1699 matches. Restricting the random effects estimates to include only principals, schools, and matches in the fixed effects regressions does not significantly change results. Principal/School refers to the pure principal and school effects model, and match effects refers to the model with principal school interactions. Standard deviations are based on the unadjusted fixed effects estimates. In the fixed effects models, principal, school, and match effects are centered within connected groups. Standard deviations in random effects models are estimated by REML.

Table 8

Regression coefficients for experience measures

	Math		Reading	
	(1)	(2)	(3)	(4)
<i>Panel A: Years of experience at school</i>				
4th grade score	0.522 (0.002)	0.520 (0.002)	0.483 (0.002)	0.480 (0.002)
2 yrs school experience	0.003 (0.006)	0.002 (0.006)	-0.002 (0.005)	-0.005 (0.006)
3 yrs school experience	-0.005 (0.006)	-0.018 (0.009)	0.007 (0.006)	-0.004 (0.009)
4 yrs school experience	0.012 (0.007)	-0.023 (0.012)	0.007 (0.007)	-0.020 (0.012)
5+ yrs school experience	0.051 (0.006)	-0.009 (0.016)	0.058 (0.006)	0.002 (0.016)
<i>Panel B: Years of experience as a principal</i>				
4th grade score	0.522 (0.002)	0.520 (0.002)	0.483 (0.002)	0.480 (0.002)
4 to 6 yrs principal experience	0.024 (0.007)	-0.016 (0.010)	0.024 (0.007)	-0.002 (0.010)
7 to 9 yrs principal experience	0.033 (0.009)	-0.022 (0.016)	0.041 (0.009)	-0.022 (0.016)
10+ yrs principal experience	0.055 (0.009)	-0.005 (0.022)	0.069 (0.008)	-0.004 (0.022)
<u>Control variables:</u>				
Principal fixed effects	yes	yes	yes	yes
Demographics	yes	yes	yes	yes
Census	yes	yes	yes	yes
School fixed effects	no	yes	no	yes

Notes: Coefficients in this table correspond to the models in Table 5. Bolded numbers represent significance at the 5% level. Coefficients are measures in standard deviations. See Section VII.A. for a full list of demographic and census measures.

Appendix Table 1
Descriptive statistics for the student data in Table 6

	Estimation Sample							
	7th grade principals		5th-7th grade principals		Same principal		Same school	
	Mean	Std Dev.	Mean	Std Dev.	Mean	Std Dev.	Mean	Std Dev.
<i><u>4th grade test scores</u></i>								
Math	0.073	0.920	0.079	0.915	0.087	0.915	0.114	0.904
Reading	0.081	0.889	0.089	0.881	0.099	0.874	0.122	0.857
<i><u>7th grade test scores</u></i>								
Math	0.021	0.933	0.029	0.929	0.106	0.933	0.121	0.931
Reading	0.069	0.898	0.076	0.890	0.118	0.879	0.129	0.871
<i><u>Demographics</u></i>								
ESL	0.209	0.406	0.210	0.408	0.246	0.431	0.242	0.428
Male	0.504	0.500	0.504	0.500	0.502	0.500	0.505	0.500
Aboriginal	0.102	0.303	0.099	0.299	0.087	0.282	0.082	0.274
In special education	0.059	0.235	0.059	0.235	0.054	0.226	0.053	0.224
Attends rural school	0.125	0.331	0.121	0.326	0.144	0.351	0.135	0.342
<i><u>Neighbourhood census characteristics</u></i>								
Percent over 65 yrs old	0.135	0.060	0.135	0.060	0.129	0.064	0.130	0.061
Percent visible minority	0.214	0.226	0.217	0.227	0.262	0.250	0.261	0.249
Percent no high school degree	0.220	0.072	0.219	0.072	0.223	0.080	0.220	0.079
Percent high school degree	0.279	0.036	0.279	0.037	0.277	0.040	0.277	0.041
Percent university degree	0.207	0.100	0.208	0.100	0.218	0.108	0.220	0.110
Percent immigrant	0.248	0.157	0.250	0.157	0.279	0.171	0.280	0.171
Percent unemployed	0.056	0.031	0.056	0.030	0.057	0.035	0.056	0.035
Average household income	60,837	16,816	60,993	16,837	62,618	19,482	62,968	19,330
Average dwelling value	306,454	139,531	308,123	139,449	323,423	158,408	327,997	158,527
Number of observations	183,066		179,062		57,360		92,156	

Notes: Test scores represent the z-score by year, grade, and subject among the population of schools prior to sample exclusions.