CALDER

AIR

NATIONAL CENTER for ANALYSIS of LONGITUDINAL DATA in EDUCATION RESEARCH

TRACKING EVERY STUDENT'S LEARNING EVERY YEAR

A program of research by the American Institutes for Research with Duke University, Northwestern University, Stanford University, University of Missouri-Columbia, University of Texas at Dallas, and University of Washington

# Pension-Induced Rigidites in the Labor Market for School Leaders

Cory Koedel,
Jason A. Grissom,
Shawn Ni,
and Michael Podgursky

WORKING PAPER 67 • JANUARY 2012

# Pension-Induced Rigidities in the Labor Market for School Leaders

Cory Koedel
*University of Missouri*

Jason A. Grissom
*Vanderbilt University*

Shawn Ni
*University of Missouri*

Michael Podgursky
*University of Missouri*

# Contents

## Acknowledgements

The authors are faculty at the Peabody College of Education and Human Development, Vanderbilt University (Grissom); and the Department of Economics, University of Missouri (Koedel, Ni, Podgursky). Podgursky is also a Fellow of the George W. Bush Institute at Southern Methodist University. They thank Julie Cullen, Peter Dolton, and Peter Mueser for useful comments and suggestions.

This research was supported by the National Center for the Analysis of Longitudinal Data in Education Research (CALDER) funded through Grant R305A060018 to the American Institutes for Research from the Institute of Education Sciences, U.S. Department of Education. The authors also gratefully acknowledge support from the DC State Office of the Superintendent.



CALDER • American Institutes for Research
1000 Thomas Jefferson Street N.W., Washington, D.C. 20007
202-403-5796 • www.caldercenter.org

**Pension-Induced Rigidities in the Labor Market for School Leaders**

Cory Koedel, Jason A. Grissom, Shawn Ni, and Michael Podgursky

CALDER Working Paper No. 71
January 2012

## Abstract

Educators in public schools in the United States are typically enrolled in defined-benefit pension plans, which penalize across-plan mobility. We use administrative data from Missouri to examine how the mobility penalties affect the labor market for school leaders, and show that pension borders greatly reduce leadership flows across schools. Our most conservative estimates indicate that removing a pension border that divides two groups of schools will increase leadership flows between the groups by roughly 100 percent. We consider the implications of our findings for workforce quality in schools near pension borders in Missouri. Our results are of general interest given that thousands of public schools operate near pension boundaries nationwide.

## Introduction

A burgeoning research literature shows that school leaders play an important role in the education production process (Brewer, 1993; Cannon et al., 2010; Clark et al., 2009; Coelli and Green, 2010; Dhuey and Smith, 2010; Grissom and Loeb, 2011). However, we know relatively little about the labor market for school leaders.[1] One largely unexamined aspect of the labor market is that most school leaders, and prospective school leaders, are heavily invested in defined-benefit (DB) pension plans. A well-known feature of DB plans is that they restrict mobility by penalizing individuals who leave the system prior to retirement eligibility. For most school leaders the mobility penalties are severe. This is because the range over the career cycle when transitions to leadership typically occur coincides with the range when the mobility penalties in DB plans are greatest.

We use administrative panel data from Missouri to examine the effects of pension-system boundaries on interschool leadership mobility. Missouri is an interesting case to study because educators belong to three different pension systems. Roughly 90 percent participate in the state's Public Service Retirement System (PSRS), and the remaining 10 percent are enrolled in one of the district-specific systems in Kansas City (KC) and St. Louis (STL). A unique feature of the pension landscape in Missouri is that there is no reciprocity between the pension systems. Thus, the three systems are separate, and differ in the same way that pension systems differ across states.

We begin by illustrating the costs associated with crossing a pension border for prospective school leaders, focusing primarily on the transition from teaching to school leadership (teaching is the most common entry path into leadership). Relative to a within-system move from teaching to leadership, an across-system move would typically result in pension-wealth losses on the order of hundreds of thousands of dollars. Moves from one leadership position to another are similarly costly if a pension border is crossed.

[1] Notable recent studies include Cullen and Mazzeo (2008) and Loeb et al. (2010).

Next we examine data on leadership mobility within and across pension systems in Missouri. The two urban districts in the state, each operating a separate and independent pension system, are dramatically less likely to hire school leaders from an outside district relative to PSRS districts, which all share the same pension system. In fact, the average PSRS district is roughly five times more likely than KC or STL to hire a school leader from the outside. This simple difference suggests that the pension borders affect labor mobility, but inference is limited because the city districts differ from PSRS districts in many ways, most notably by size. To minimize the confounding influence of the differences in size and other characteristics between the city and non-city districts, we exploit disaggregated data on leadership flows between schools over the course of an 18-year data panel. We analyze the data using two different approaches, both of which lead to the same conclusion: pension borders significantly lower interschool leadership flows. For two groups of schools separated by a pension border, our most conservative estimates indicate that removing the border will roughly double leadership mobility between them.

We conclude by examining the consequences of the mobility barriers for leadership quality in schools. We use available quality measures to show that the pension borders separate the city districts from a more-qualified leadership pool on the outside. While the pension boundaries may not be the primary cause of the observed quality gaps between the city and non-city districts – our data do not permit a direct investigation of the cause of the gaps – they must be recognized as a significant impediment to their elimination. Given recent evidence on the importance of leadership quality for school success, this is a noteworthy result. It suggests that any reform efforts in the urban districts that aim to improve leadership quality will be significantly hampered by the pension structure in Missouri.[2]

[2] Of course, our findings do not rule out the success of internal efforts by the city districts to improve leadership-specific human capital within their pre-existing applicant pools – however, the descriptive statistics that we present in the Section VII imply that in the absence of the pension borders, efforts to improve leadership quality in the city schools would likely involve outside recruitment.

In a closing discussion we consider the implications of our findings in a larger national context. Across the United States, 22,000 miles of state boundaries, which are also pension boundaries, create substantial barriers to the efficient flow of leadership talent across schools. The mobility barriers have become stronger in recent decades due to benefit enhancements in most educator pension plans, and work against the growing reciprocity in educator licensing across states.[3] The rationale for licensing reciprocity is that it makes little economic sense to impede a fully-licensed teacher or school leader in moving across state lines. However, we show that the general structure of retirement benefits in the education sector is at odds with the objective of making the labor market more fluid.

## Motivation

Basic economic theory suggests that pension barriers in the leadership labor market have efficiency implications. Furthermore, they can reinforce or exacerbate inequitable distributions of leadership talent across schools. We briefly illustrate these points in Figures 1 and 2. Both figures show results from a simple, static model of the leadership labor market where two contiguous school districts are separated by an impermeable pension border. The figures illustrate two mechanisms by which pension borders will affect leadership quality in schools: (1) by shrinking the applicant pools from which administrators in border regions can draw, and (2) by reinforcing differences in applicant-pool quality across pension systems.

In the model, administrators fill leadership vacancies by choosing the highest-quality candidate from the pool of applicants, where quality is perfectly observed and measured by *Z*. The two districts are equal-sized, and grow together from containing 1 to 30 schools. Figure 1 shows average leadership quality in schools, with and without the pension border, when applicants are drawn from the same

[3] Specifically, the National Association of State Directors of Teacher Education and Certification (NASDTEC) Interstate Agreement is designed to reduce barriers to movement of certified teachers and administrators between states and jurisdictions.

underlying quality distribution in both districts ($Z \sim N(160,50)$).[4] When the pension border separates small districts, the reduction in leadership quality is substantial. This result is relevant for rural districts near state lines. For larger districts, the border effect on the size of the applicant pools alone is small.

In Figure 2 we adjust the underlying distribution of applicant quality in District 1 so it has a stronger pool ($Z_1 \sim N(165,50); Z_2 \sim N(160,50)$). Even as the districts become large, overall leadership quality is lower with the pension border. This is because some applicants in District 1 who are not hired into leadership positions have higher values of $Z$ than some leaders in District 2. The effect of the border on overall leadership quality is increasing with the gap in applicant-pool quality between districts (not shown in graph). An additional concern is that the border will reinforce and potentially exacerbate inequality if District 2 is more disadvantaged generally. Note that the quality gain for the disadvantaged district that results from removing the pension border exceeds the loss for the advantaged district; which, of course, is simply another way of saying the pension border is inefficient.

The model clearly oversimplifies the leadership labor market, but is useful to illustrate the ways in which pension borders can distort the distribution of leadership quality across schools.[5] Recent work by Jackson (2010) raises an additional concern related to the mobility-dampening effects of pension borders; namely, leader/school match quality is likely to be an important and independent determinant of leadership effectiveness. By preventing some leader/school matches from occurring, pension borders lower system-wide match quality. The result will be less effective leadership in schools near pension borders.[6]

---

[4] The distribution of $Z$ was selected to mimic the distribution of licensing exam scores, one of the observable measures of leadership quality that we consider in Section VII.

[5] The model treats differences in applicant-pool quality across districts as exogenous, but they may not be. For example, in Missouri, because the pension borders restrict mobility in the urban districts, they may affect teacher recruitment. Jacob (2007) discusses the general challenges that urban districts face in recruiting teachers – for KC and STL, the restricted mobility is another concern. The pension borders may also affect the returns to investing in leadership-specific human capital in the different systems. These issues are beyond the scope of the present study.

[6] Jackson's (2010) work examines the importance of teacher/school match quality, and shows that the quality of the match explains a large fraction of what we commonly describe as "teaching effectiveness." Of course, in addition to

## Background

Educators in public schools throughout the nation, including those in Missouri, are nearly universally enrolled in DB pension plans. In this section we demonstrate why educators who change plans typically incur large financial losses.

The general formula used to compute the annual annuity at retirement in a standard DB plan is:

$$Y_i = F_i * YOS_i * FAS_i \quad (1)$$

In (1), $Y_i$ is the annual pension payment for individual *i* at the time of retirement. The payment depends on a formula factor, $F_i$; years of system service, $YOS_i$; and the individual's final average salary, $FAS_i$, which is calculated based on the average of the highest few years of system earnings. Table 1 reports the parameters for each system in Missouri in the year 2000 (Appendix Table C.1 shows changes to the system rules over the course of our data panel). While some specific features differ across the systems, their general structures are similar.

A key difference that separates the city systems from PSRS is that educators in the cities participate in Social Security and educators in PSRS do not. This difference facilitates our analysis; without it there would likely be reciprocity between the pension systems. For example, in other city districts that operate their own pension plans (e.g., New York, Chicago), educators can move between the state and local systems without incurring large pension-wealth losses. But reciprocity is challenging in Missouri because benefits and contributions are very different across the systems that do and do not incorporate Social Security. The result is that the three pension systems within Missouri differ in the same way that systems differ across states.[7]

---

the likelihood that Jackson's findings are relevant for school leaders, the pension-induced rigidities also directly affect teachers.

[7] State and local workers were originally excluded from Social Security, but Congress passed legislation in the early 1950s that permitted states to include their employees. Teachers in most states are either in or out of Social Security. Missouri is unusual in this regard (Munnell, 2000). Although benefit levels and contributions differ across systems because of the Social Security difference, the structural components that cause the rigidities are unaffected (see below).

Two common features of the pension systems penalize mobility. First, like other educator pension systems nationwide (Costrell and Podgursky, 2009), all three systems offer generous retirement provisions that depend on within-system experience, *YOS*. These provisions include the rules of 80, 75 and 85 in PSRS, KC and STL, respectively. The rule amounts refer to individuals' combinations of age and system experience – once the sum of these two numbers reaches the level of the rule, the member can collect full retirement benefits. For example, although the official (and in earlier decades, typical) retirement age in KC is 60, an age-50 teacher with 25 years of system service is eligible to retire without penalty and begin collecting benefits immediately. PSRS also has an additional provision called "25-and-out," which stipulates that an individual can exit and begin collecting benefits immediately, regardless of age, so long as he or she has 25 years of system experience (with a penalty that is much less than is actuarially appropriate). In all three systems, and in most other educator pension systems, if an individual exits the pension system too soon she loses the option to exit under one of these lucrative provisions. As we illustrate below, this is very costly.

The second common feature of the systems that penalizes mobility – and again, a feature common to teacher DB systems generally – is that *FAS* is frozen at the time of exit. To illustrate why this is important, consider two individuals who end up with the same wage profile over the course of their respective 30 year careers. The first individual stays in the same system and her final payment is equal to $30*F*FAS$, where *FAS* is calculated using her last few years of earnings. The second individual switches systems after 15 years. Her final payment comes from the two systems and is equal to $\{15*F*FAS_1 + 15*F*FAS_2\}$, where $FAS_1$ is her final average salary at the time of her exit from the first system, unadjusted for inflation or life-cycle pay increases.

We illustrate the mobility costs associated with crossing pension-system boundaries by comparing pension-wealth accrual under various career scenarios for a representative teacher in Missouri. The key comparisons are between within- and across-system moves from teaching to leadership. We do not explicitly document the costs associated with moves from one leadership position to another, but they are substantively very similar.[8]

Pension wealth is calculated as the present value of the stream of pension payments. Pension wealth at time *s*, with collection starting at time *j* where *j* ≥ *s*, can be written as:

$$\sum_{t=j}^{T} Y_t * P_{t|s} * d^{t-s} \tag{2}$$

In (2), $Y_t$ is the annual pension payment in period *t*, $P_{t|s}$ is the probability that the individual is alive in period *t* conditional on being alive in period *s*, *d* is a discount factor, and T is set to 101. In KC and STL, the pension payment is the sum of the system and Social Security payments; we allow the individual to begin collecting payments from each source at the optimal time. Details about our pension-wealth calculations are provided in Appendix A.

Figure 3 shows pension-wealth accrual over the career cycle for a representative teacher who is currently 37 years old and began teaching at age 25 in PSRS. Her pension-wealth profile is illustrated under several different career scenarios. First, the solid black line shows wealth accrual under the scenario where she remains in teaching within PSRS. The uneven accrual rate displayed in the figure has been well-documented (Costrell and Podgursky, 2009). The remaining pension-wealth profiles correspond to moves from teaching to school leadership at age-38.[9] The dashed red line shows pension-wealth accrual when the move occurs within PSRS. The within-system move corresponds to a large increase in pension wealth driven by higher leadership wages (see Appendix A for details). The other two lines show pension-wealth accrual in scenarios where the move to leadership involves crossing a pension boundary; the lines marked by circles and triangles are for moves to KC and STL, respectively. In both border-crossing scenarios the move corresponds to a dramatic reduction in pension wealth. For

[8] The pecuniary costs associated with crossing a pension-system border depend primarily on age and experience at the time of the move.

[9] Age-38 is the median age-at-entry into leadership in the Missouri data. The average entry age is slightly lower in rural districts, but the difference is not substantively important. Appendix Table C.2 shows calculations for moves at different points in the career cycle.

example, movement to a leadership position in KC from PSRS produces a 63 percent loss in pension wealth when compared to a similar leadership promotion within PSRS ($283,638 in 2009 dollars), holding the retirement age constant at the peak-value of pension wealth in PSRS.[10] If we amortize this over the remainder of her career it amounts to an annual loss of roughly $25,500.

The pension losses illustrated in Figure 3 are not unique to the move scenarios considered in the figure. In Appendix Table C.2 we show that moves across any pension border, and over most of the range of the career cycle when leadership work typically occurs, are associated with large losses in pension wealth.[11] The extent to which labor flows are actually restricted by the mobility penalties is an empirical question which has gone previously unaddressed, and to which we now turn our attention.

## Mobility Data

We use an 18-year administrative data panel from the entire state of Missouri. Personnel data are available from the 1992-1993 through 2009-2010 school years, and include information about each public school employee's position type, salary, plus information on the schools and districts in which the individual worked. The length and breadth of our data panel is a key feature of our study as it ensures that we observe a sufficient number of leadership moves.

We group principals and assistant principals together, and refer to them synonymously as "school leaders" for our analysis.[12] In total, we observe 11,034 leadership hires over the course of the data panel, where we define a leadership hire to have occurred whenever a leader starts in a new school. Nearly 50 percent of new hires (5,508) represent first-time leaders, and of those, 67 percent

[10] Most educators retire fairly close to the peak-value year (the year when pension-wealth is maximized), although there is some variability (Podgursky and Ehlert, 2007). See Coile and Gruber (2007) for a discussion of the peak-value construct and alternatives (also see Stock and Wise, 1990a,b).
[11] The exception is that more-senior school leaders can marginally benefit in some cases by switching pension systems to double-dip; however, in Appendix Table C.3 we show that double-dipping does not explain the limited cross-border mobility that we observe in the Missouri data.
[12] None of our findings are substantively sensitive to how we define "school leader."

come directly from teaching in Missouri schools. The remainder come mostly from other school- and district-level supervisory positions (e.g., special education administrators, program coordinators) and guidance counselor positions. Table 2 provides summary statistics for the data.

Table 3 shows within- and out-of-district hire rates for school leaders in KC, STL, PSRS, and three subsamples of PSRS districts (out-of-state hires are included with out-of-district hires in the table). The PSRS subsamples were selected to be comparable to the city districts along a particular dimension: column (4) shows hire rates for "neighbor" PSRS districts that are geographically close to the city districts, column (5) shows hire rates for districts that are similarly disadvantaged, and column (6) shows hire rates for the ten largest districts in PSRS.[13] Because KC and STL operate their own pension systems, all out-of-district leadership hires in the city districts necessarily cross a pension border. Districts within PSRS can share hires without the occurrence of a border crossing.

Consistent with the mobility penalties illustrated above, the raw data show that the city districts have the lowest out-of-district hire rates.[14] In fact, the out-of-district hire rate for the typical PSRS district is roughly five times that of either of the city districts. However, it is difficult to draw inference about the effects of the pension borders because the city districts differ from the PSRS districts in many ways, even for the PSRS subsamples (see the bottom panel of the table).[15] Ultimately, while the raw data are consistent with the pension borders reducing mobility, inference about the border effects is clouded by the lack of comparability between districts across the pension systems.

[13] The neighboring PSRS districts are well within a commutable distance to the city. They include districts that touch the city districts, or touch the districts that touch the city districts.
[14] Appendix Table C.3 also shows within- and out-of-district hire rates for several subgroups of school leaders – the patterns in Table 3 are replicated regardless of how we divide the data.
[15] Comparability remains an issue for the PSRS subsamples. The neighboring districts are too small and too advantaged, the disadvantaged districts too small, and the largest PSRS districts too small and too advantaged. In addition, some of the largest PSRS districts are isolated geographically, which likely reduces their out-of-district hire rates.

## Empirical Strategy

Before considering the details of our empirical strategy it is important to consider what non-pension factors would lead districts to make inside versus outside hires. A factor of obvious importance is district size: mechanically, one would expect larger districts to have more intra-district hires. To illustrate, imagine an exogenous network of hiring flows between schools. Ignoring out-of-state hires, if the entire state were a single district then all hires must be intra-district. As the single district is carved up into a grid of smaller and smaller districts, labor will flow across district lines more often.

The relationship between district size (log enrollment) and the intra-district hire rate is plotted in Figure 4 for the districts in the pension-border regions of Missouri (the city districts plus the "PSRS neighbors" from Table 3). The trend line is fitted using data from the PSRS districts only; the city-district data points are plotted *ex post* to ensure that the pension borders do not adjust the slope of the line. A clear positive relationship between district size and the intra-district hire rate is evident in the figure, and both city districts (at the far right) have more internal hires than would be predicted based on the data from their PSRS neighbors. But again, the city districts are major outliers in terms of size (as well as in terms of student poverty and percent minority, see Table 3), which makes it problematic to infer that the pension borders influence mobility from the plot in Figure 4 alone.

In the empirical work described below we attempt to improve inference about the pension border effects by disaggregating the data and focusing on leadership flows between individual schools rather than entire districts. While the *groups* of schools on different sides of the pension borders are clearly very different, there is much more overlap in the characteristics of *individual schools*. We undertake two separate, disaggregated approaches in our analysis. First, we perform a simulation exercise where we disaggregate the schools in each border region, then randomly re-aggregate them into equal-sized groupings of schools that differ in the extent to which they are divided by the real pension border. Within the school groupings we examine the relationship between pension-border segmentation and within-group shared hiring. Second, we use models analogous to gravity models from the international trade literature to examine labor flows between individual pairs of schools.

We focus our empirical analysis on the two border regions in Missouri. In each region our analytic sample includes data from all of the city schools and a fraction of the neighboring PSRS schools. The PSRS-neighbor subsample is selected so that in each region the share of city schools is 0.50 (the reason will become clear shortly). Figure 5 shows the locations of the schools that we use for this portion of our analysis. The darker flags in each region are city schools and the lighter flags are PSRS schools. In the bottom-middle panel of the figure we overlay the Labor Market Area in each region as defined by the Bureau of Labor Statistics (BLS, 2011). The BLS describes a Labor Market Area as "an economically integrated geographic area within which individuals can reside and find employment within a reasonable distance or can readily change employment without changing their place of residence." The samples of schools in each region are comfortably within their Labor Market Areas.

*Simulations*

There are two related issues that make it difficult to isolate the pension-border effects on mobility. Both are driven by the differences between the city and non-city schools, combined with the fact that the pension borders coincide with the city-district lines. The first issue is straightforward conceptually – the city districts are larger and more disadvantaged than their PSRS neighbors. Figure 4 shows that the city districts will be more likely to hire from within based on their size alone. Due to their disadvantage, the city districts are also likely to be less desirable places to work for educators who might otherwise move in. In fact, one could hypothesize that the low mobility into the city districts that we have documented thus far is driven entirely by these factors and has nothing to do with the pension borders.

The second issue is related to the first but more subtle – there could be border-crossing transition costs that are independent of baseline educator preferences. For example, it could be that teachers do not systematically prefer to work at advantaged or disadvantaged schools, but once they start in one school-type it is costly to switch to the other. This scenario would be akin to the idea of firm-specific human capital, but in our case the human capital would be specific to either city or non-city schools. The general concern is that for one reason or another, the labor pools inside and outside of the city simply do not mix well, and would not mix well even in the absence of the pension border.

Our simulation analysis addresses these two issues separately. The first issue is addressed in the construction of the simulated dataset, and the second issue by *ex post* adjustments when we analyze the data. We now briefly describe our methods for constructing and analyzing the simulated data. Then we return to the discussion of how the simulations deal with these threats to identification.

We perform the simulation exercise separately for each region. We begin by randomly re-grouping the schools in Figure 5 into ten equally-sized groups of schools irrespective of schools' actual district assignments.[16] We repeat the reassignment process 1,000 times to construct a simulated dataset that contain 1,000 batches of 10 randomly-drawn school groupings. The random assignment process generates natural variation in the characteristics of the school groups, including the degree to which the groups are segmented by the real pension border.

At each iteration, we capture several pieces of information about the groups of schools. First, we calculate the geographic proximity of the schools within each group, measured as the average distance between schools. Next, we measure each group's basic demographic characteristics including district enrollment, minority share, and free/reduced-lunch share.[17] We use the information about the minority and free/reduced-lunch shares to construct measures of the within-group similarity of schools. Consider the minority-share variable as an example. As the average minority share across schools within the group approaches 0.50, within-group diversity increases; as it approaches zero or one, diversity

[16] More precisely, we create groups with as close to an equal number of schools as possible. The first nine groups have the same number of schools and the tenth group is slightly larger or smaller.
[17] The districts are constructed to include the same number of schools, but there is still variation in total enrollment.

decreases. We construct the similarity measures as the absolute value of the difference between the across-school average of the given characteristic in group *j,* and 0.50. The larger the difference, the more similar are the schools within the group, and presumably the greater the rate of within-group hiring.[18]

Next we measure the amount of real pension-system sharing within groups. Given that schools in each region belong to one of two real pension systems, we construct the pension-sharing measure for each group as *PS* = max{*CS*, (1 – *CS*)}, where *CS* is the share of city schools. The variable *PS* indicates the degree to which each group of schools is segmented by the actual pension border. Larger values indicate more pension-system sharing among the schools; smaller values indicate less pension-system sharing.[19] Importantly, *PS* is a bidirectional measure of pension sharing – for example, groups of schools with city shares of 0.30 and 0.70 have equal values of *PS.*

The function that maps values of *CS* to values of *PS* is V-shaped, with the minimum value of *PS* occurring when *CS* = 0.50. This is illustrated in Figure 6. Therefore, the correlation between *PS* and *CS* can be broken in the simulated data by constructing the samples of schools so that E[*CS*] = 0.50 for each school group. This is an important benefit of the simulation-based approach. By constructing the initial samples of schools in each region so that exactly half are city schools, we can break the correlation in the data between pension sharing and the share of city schools.

Finally, for each group we record the number of shared leadership hires between schools over the course of our data panel, which is the key outcome variable. We hypothesize that groups of schools which are more segmented by the real pension border – where *PS* is closer to 0.5 – will have fewer shared leadership hires.

We analyze the simulated data using the following regression model:

$$WC_j = \delta_0 + X_j\delta_1 + D_j\delta_2 + CS_j\delta_3 + PS_j\delta_4 + \varepsilon_j \quad (3)$$

[18] This measure of similarity would not be useful in a world where all schools are similarly diverse (e.g., if every school had a 50-percent minority share), but this is not the case in our data.
[19] More pension sharing within the group opens up more unobstructed conduits for mobility between the schools.

In (3), $WC_j$ measures the number of within-group hires for group *j*, measured as a count. $X_j$ includes controls for total group enrollment and the similarity of the schools within the group, and $D_j$ measures the average distance between the schools. $CS_j$ indicates the share of the group comprised of city schools, for which the effect can be separately identified per the above discussion.[20] $PS_j$ measures pension sharing and is the variable of interest.[21]

Returning to the issue of identification, the key benefit of the simulation design is that pension sharing within school groups is uncorrelated with other key group characteristics by construction. Table 4 confirms this feature of the simulated data. It shows that the bidirectional variation in within-group pension sharing is essentially independent of the variation in groups' minority shares, free/reduced-lunch shares, enrollment levels, and their shares of city schools. Whereas some of the limited mobility into the city districts that we document in the Table 3 likely reflects educator preferences for advantaged schools, and differences in district size, the pension-border effects that we estimate using the simulated data will not be confounded by these factors.

The second threat to identification, which is not addressed in the construction of the simulated data, involves the potential for non-pension-related, bidirectional transition costs between the city and non-city schools. Indeed, Table 5 shows that pension sharing, which is synonymous with clustering on *either side* of the pension border, is correlated with the similarity of school groups. The correlations in Table 5 reflect the fact that school groupings with more schools on either side of the border are more similar to each other than school groupings that are more stratified (e.g., groups where *PS* = 0.7 will be more similar to each other than groups where *PS* = 0.5). The differences between schools on different

[20] We also estimate models where we replace $CS_j$ with a Herfindahl index for each group that indicates its stratification across actual district lines, and obtain substantively similar results. We cannot include the Herfindahl indices and city shares at the same time because they are highly correlated (in the KC region the correlation is roughly 0.94; in STL it is 0.98).

[21] We report our results using the simple linear model, but estimate nearly identical mobility effects if we use a count model instead. Obtaining the correct standard errors for the coefficients in (3) is not straightforward. Appendix B provides the technical details on how to compute the standard errors.

sides of the border in the real data are larger in the KC region, and correspondingly, so are the correlations between the similarity measures and pension sharing in Table 5.

Because the differences between schools on different sides of the real pension borders represent a fixed feature of the real data, the positive correlations in Table 5 cannot be mitigated by any empirical procedure. Instead, we deal with this issue by directly controlling for the within-group similarity of schools in equation (3). Recall that our similarity measures are based on schools' student compositions, and we also control for the average distance between schools in each group. If an unobserved factor related to cross-border transition costs is going to bias our estimates, its influence needs to be poorly proxied by the similarity controls. Although we cannot rule out bias from such a factor with certainty; there are no obvious, unobserved confounders not related to the pension system that remain unaccounted for.

*Pairwise Models*

As an alternative to the simulations we also perform a pairwise analysis using the universe of pairs of schools in each regional dataset. In the KC region our pairwise dataset contains 16,836 unique pairs of schools; in STL there are 35,245 pairs. For the pairwise analysis we compute the geographic distance between the schools in each pair, and combined enrollment. We also use information about each school's student body to construct similarity indices analogous to those in the simulations. The pairwise analysis is similar to the simulation-based approach in that we aim to improve inference via disaggregation. The pairwise models mirror gravity models in the international trade literature (see, for example, Linnemann, 1966; Tinbergen, 1962), but instead of modeling trade flows between countries we model labor flows between schools.

A limitation of the pairwise analysis is that most school pairs do not share any leadership hires over our panel period, so the outcome variable is mostly populated with zeros. The high prevalence of zeros makes linear regression unappealing. Furthermore, very few pairs of schools share more than one leadership hire, so the data are not well-suited for count models either.[22] Therefore, we convert our measure of leadership sharing to a simple binary indicator and estimate a logit model that predicts the occurrence of one or more shared hires between pairs of schools.

We estimate the following model with the pairwise data:

$$AS_j = \gamma_0 + X_j\gamma_1 + D_j\gamma_2 + DL_j\gamma_3 + PL_j\gamma_4 + u_j \quad (4)$$

In (4), $AS_j$ is an indicator equal to one if any shared hire is observed between the schools in pair $j$, $X_j$ includes the similarity measures and combined enrollment, and $D_j$ measures the distance between the schools. We include linear and quadratic distance terms in the pairwise regressions because both are generally significant. $DL_j$ and $PL_j$ are indicators for the schools being divided by a district line and a pension line, respectively. Pension-line crossings always involve crossing a district line, but the reverse is not true.

The model in (4) is similar to the model in (3). There are two key benefits of the pairwise approach. First, no re-sampling is required because the data include the exhaustive set of pairs of schools. Second, we directly and separately control for district-line crossings in the regression equation.[23] The weakness of the pairwise approach is that by reducing the outcome variable to be a binary measure we throw out some information about shared hires (i.e., school pairs that share more than one hire are treated as sharing just one hire). Throwing out this information is likely to lead to an understatement of the pension-border effects because schools that share more than one hire are more

[22] We experimented with using zero-inflated Poisson and negative binomial distributions to model the data (Greene, 1994; Mullahy, 1986), but they provide a poor fit. The problem arises from the fact that the count variation in our data is small. For example, in STL, of 35,245 school pairs, only 579 have one or more shared hire over our sample period. Of these, only 88 have more than one shared hire, and 27 have more than two shared hires. The distribution of counts is similarly sparse in KC.

[23] As noted in footnote 20, our simulation results are qualitatively similar if we replace the city-share control with a Herfindahl index for real-district diversity. The Herfindahl index control in (3) is analogous to the district-line control in (4) but is a less-direct measure.

likely to share the same pension system. For this reason, we view the estimates from the pairwise models as more conservative.

## Results

Table 6 presents results from the simulation regressions separately for the KC and STL regions. The models are increasingly detailed moving from left to right in the table. All of the control variables in the models are important predictors of labor-market connectivity, and their effects are in the expected directions.[24] But the key finding from Table 6 is that regardless of which other features of the groups of schools are included in the model, pension sharing remains a statistically significant and economically meaningful determinant of interschool leadership sharing in both pension-border regions.

How large are the pension border effects that are implied by our estimates? We answer this question by extrapolating from the simulated data to compare cases where $PS_j$ = 0 and $PS_j$ = 1. The first case is mechanically ruled out by the construction of our simulations, but can be conceptualized as a polar case where each school in the group has its own DB system. Thus, there is no pension sharing, and all interschool labor flows represent pension-border crossings.[25] At the other extreme is the case where $PS_j$ = 1, where there are no pension borders and all schools share the same DB system. The pension-border effect on mobility is the difference in leadership flows between these two cases.

In each region we use the regression coefficients from the full model to compare the predicted number of shared leadership hires when $PS_j$ = 0 to the predicted number when $PS_j$ = 1. We hold all non-

[24] In the KC results, it is seemingly counterintuitive that within group similarity in the share of students who are eligible for free/reduced lunch enters negatively into the equation. However, when all of the similarity controls are considered simultaneously (including the city share of schools), within-group similarity overall is strongly positively associated with leadership sharing.
[25] In the real world, in a region with one pension border, this case would be equivalent to measuring labor flows only between schools on different sides of the border (ignoring within-system moves).

pension variables fixed at their sample averages.[26] Our calculations indicate that going from complete segmentation to total integration increases leadership flows between schools by 97 (in STL) to 163 (in KC) percent.[27]

Table 7 reports estimates from the pairwise logit models (the estimates reported in the table are the raw logit estimates, not marginal effects). For brevity we report estimates from the fully-specified models only, and for ease of presentation the coefficients and standard errors are multiplied by 100. Our interest is in the pension line variable. Like with the other estimates in the table, the pension-line effects are consistent in sign and significance with what we find in the simulations.

We estimate the magnitudes of the border effects implied by the logit models similarly to the simulations. Specifically, we compute the marginal effect of changing the pension-line variable, fixing the district-line variable at one and all other variables at their sample averages. We estimate that removing the pension border between two schools increases the probability of a shared leadership hire by 85 percent in the KC region, and 96 percent in the STL region.

In the STL region the implied border effects are qualitatively and quantitatively similar using both empirical approaches. In the KC region the simulations imply larger border effects. We briefly investigate the extent to which the discrepancy in the KC region can be explained by the information loss in the pairwise models. Specifically, we re-run the simulations on a modified dataset with the same information loss (the modified dataset allows for, at most, one instance of a shared hire between any two schools). The implied border effect in the KC region is 26 percent smaller in the limited-information simulation (moving from 167 to 123 percent). If we place rough confidence intervals around the

[26] That is, for each region we use our estimates from the full model to compare $(\hat{\delta}_0 + \overline{X}_j\hat{\delta}_1 + \overline{D}_j\hat{\delta}_2 + \overline{CS}_j\hat{\delta}_3 + 0*\hat{\delta}_4)$ and $(\hat{\delta}_0 + \overline{X}_j\hat{\delta}_1 + \overline{D}_j\hat{\delta}_2 + \overline{CS}_j\hat{\delta}_3 + 1*\hat{\delta}_4)$.

[27] We also estimated models separately for elementary and secondary schools (middle/high schools) because there is limited overlap in hiring across schooling levels. Those estimates, which we omit for brevity, show that the borders affect mobility at both schooling levels, but that the effects are larger at the secondary level. In the raw data, leadership hires at the secondary level are more likely to occur across schools and districts, which explains why the border effects are more pronounced there.

predictions from each approach using the limited data, we cannot reject that the implied border effects are equal in magnitude.[28]

*Summary of Empirical Findings*

After adjusting for differences in size and other factors between KC, STL and the surrounding districts, our estimates suggest that the effects of the pension borders on leadership mobility are statistically significant and large. Neither strategy is bullet proof, but even the more-conservative pairwise results indicate that the pension borders represent a substantial impediment to leadership mobility. Our estimates do not translate directly to the mobility gaps in the raw data as shown in Table 3, but to a rough approximation they suggest that the pension borders explain about a third of the gap in the within-district hire rate between the city and suburban districts in each border region. The remaining mobility gaps are likely driven by the other differences between the city and non-city districts including differences in size, student disadvantage, and urbanicity.[29]

## Policy Implications

We now turn to the policy implications of our findings for Missouri. In particular, we examine measured gaps in leadership quality between the urban and suburban districts. We begin by examining teachers' ACT scores.[30] The evidence relating ACT scores to value-added in the classroom is mixed; however, even without a direct link to classroom performance, observable differences in teachers' ACT scores likely to imply differences along other dimensions as well.[31] Table 8 shows average ACT scores for

[28] We also perform a similar simulation analysis with limited data for STL. The implied border effect declines slightly and remains comparable to the result from the pairwise model (the estimated effect on leadership flows falls from 97 to 88 percent).
[29] Jacob (2007) discusses the general recruitment problems faced by urban school districts, which are likely to be relevant for KC and STL.
[30] ACT scores are reported when available. Missing scores are fairly common. One reason is that many teacher preparation programs in Missouri do not require the ACT.
[31] Harris and Sass (forthcoming) find that entrance exams scores are not related to value-added performance measures for individual teachers, which suggests a limitation of using ACT scores. However, Ferguson and Ladd (1996) find a positive relationship using school-aggregated data.

incoming teachers between 2005 and 2009 for all PSRS districts, KC, STL, and the geographically close PSRS districts in each region.[32] City teachers clearly have the lowest ACT scores.

In Table 9 we turn to credentials for hired school leaders. First, we compare leaders using their scores on the School Leaders Licensure Assessment (SLLA), a standardized examination administered by the Educational Testing Service required of all applicants for school administration licenses in Missouri. The SLLA is designed to measure leadership capacity as defined by a set of national standards for practice (Reese and Tannenbaum, 1999). The standard deviation of SLLA scores for school leaders in Missouri is roughly 7.5 (mean ≈ 180).

We also use a supplementary data file from the Department of Elementary and Secondary Education to measure undergraduate-institution quality for school leaders. We consider Missouri graduates exclusively, who make up 75 percent of the leaders in the data, and all graduates. We identify four "high-quality" universities in Missouri: Truman State University in Kirksville, the University of Missouri in Columbia, the University of Missouri in Rolla, and Washington University in St. Louis.[33] We also identify the three public institutions in Missouri with the lowest entrance exam scores: Harris-Stowe State University in St. Louis, Lincoln University in Jefferson City, and Missouri Western State University in St. Joseph.[34] For the larger, nationwide sample of graduates we use the following (admittedly rough) guideline to categorize "high-quality" universities: if a university has an average entrance-exam score that is at least as high as at the University of Missouri Columbia, or if it is the

[32] The neighboring districts in Tables 8 and 9 are the same as in Table 3 (broken out by region).
[33] Within Missouri these are clearly the most prestigious universities for undergraduates. For example, Truman State University and the University of Missouri- Rolla have higher average ACT scores than the University of Missouri-Columbia, which is the state's flagship school.
[34] We do not have the data to categorize the bottom tail of the distribution of private universities. We selected these three public universities because there is a large difference in average ACT scores between them and the other public universities. For example, Missouri Western has the highest average ACT of the three, but it is still a half of a standard deviation below the average ACT for the next-lowest-ranked public university.

flagship university in its state, we code it as a “high quality” university. For the national sample we do not code low-quality universities.[35]

Table 9 shows that city leaders score lower on the SLLA exam and come from lower-quality universities relative to leaders on the outside. The differences across the pension systems are large, with a particularly notable difference being that city leaders are much more likely to have attended one of the least selective public universities in Missouri. The evidence is descriptive, and therefore merely suggestive, but it is consistent with the pension-induced rigidities adversely affecting leadership quality in the city districts. Moreover, even if the pension borders play little or no role in creating the initial quality differences, by imposing a large tariff on educators who cross boundaries they represent a major impediment to eliminating the gaps.

## Discussion and Conclusion

Leadership quality is increasingly recognized as a key contributor to student success in schools (Brewer, 1993; Cannon et al., 2010; Clark et al., 2009; Coelli and Green, 2010; Dhuey and Smith, 2010; Grissom, 2011; Grissom and Loeb, 2011). We examine an important and previously unaddressed feature of the labor market for school leaders – namely, school leaders, and prospective school leaders, are heavily invested in defined-benefit pension plans that restrict mobility. We show that pension borders introduce substantial rigidities into the leadership labor market, resulting in the misallocation of labor to school-leadership positions. When a pension border divides the labor market for a small school district, or when there is an underlying quality difference in the applicant pools on different sides of a pension border, leadership quality will be reduced. Furthermore, leader/school match quality will be lower near

[35] Like with teachers’ ACT scores, our measures of leadership quality are imperfect. In fact, Clark et al. (2009) find that the selectivity of a principal’s undergraduate institution, as measured by median SAT scores, is not related to principal quality in New York City. However, their analysis does not rule out the possibility that graduates from particularly unselective colleges perform worse. A large fraction of leaders in the border regions in Missouri, and in the city districts in particular, are graduates from such colleges (see Table 9).

pension borders. Applying recent evidence from Jackson (2010) to the leadership context suggests that the lower match quality will adversely impact students in K-12 schools.[36]

Missouri offers a unique opportunity to examine pension-border effects because educators belong to three different pension systems and there is not reciprocity between them. That is, the three systems within Missouri are separated in the same way that state pension systems across the country are separated. Our findings are of interest nationally given that 22,000 miles of state lines across the country serve as pension boundaries, bisecting school labor markets throughout.

The issues raised by our analysis are, in principle, straightforward to resolve. One solution would be for state governments to compensate school leaders for the costs of crossing pension system lines. Again noting the recurring finding that high-quality school leaders contribute greatly to student success, the benefits of such a policy may outweigh the costs. Furthermore, in Missouri, there would be an equity gain in addition to the efficiency gain because the urban districts educate disproportionately poor and minority students. The other option, which is conceptually straightforward but politically challenging to implement, is to remove the pension borders. This can be achieved by integrating pension systems nationwide, or by moving educators into retirement plans that do not discourage mobility (e.g., defined-contribution or cash-balance plans).

It is interesting to contrast the absence of inter-state reciprocity and collaboration in educator pension plans with the growing reciprocity in educator licensing across states, largely driven by the collaboration of certification staff in the various states.[37] Certification reciprocity arose from a recognition that it makes little economic sense to impede a fully-licensed teacher or school leader in moving from, say, a Kansas to a Missouri public school to practice her profession. It is ironic that while mobility restrictions due to educator licensing barriers have declined, the pension-driven tariffs on

[36] And, of course, the pension borders will also affect teacher/school match quality.

[37] See the National Association of State Directors of Teacher Education and Certification (NASDTEC) Interstate Agreement. There is little or no effort to create similar interstate reciprocity for educator pension plans. See Ruppert (2001).

interstate mobility of educators have grown over recent decades due to benefit enhancements in the state retirement plans. While this study represents what we hope is a useful beginning, much more research is needed on how the pension-induced rigidities in educator labor markets across the country affect workforce quality and school performance.

## References

Brewer, Dominic J. 1993. Principals and Student Outcomes: Evidence from U.S. High Schools. *Economics of Education Review* 12(4), 281-292.

Bureau of Labor Statistics. 2011. Labor Market Areas, 2011. Report available at: http://www.bls.gov/lau/lmadir.pdf

Cameron, A. C., J. B. Gelbach, and D. L. Miller. (2011). Robust inference with multi-way clustering. *Journal of Business and Economics Statistics* 29(2), 238-249.

Cannon, Sarah, David Figlio and Tim Sass. 2010. The Policy Choices of Effective Principals. Unpublished manuscript. Northwestern University.

Clark, Damon, Paco Martorell and Jonah Rockoff. 2009. School Principals and School Performance. CALDER Working Paper No. 38.

Clark, Robert and Craig, Lee A. (2011). "Determinants of the Generosity of Pension Plans for Public School Teachers, 1982-2006." *Journal of Pension Economics and Finance* 10(1), 99-118.

Coelli, Michael and David Green. 2010. Leadership Effects: School Principals and Student Outcomes. *Economics of Education Review* 31(1), 92-109.

Coile, Courtney and Jonathan Gruber. 2007. Future Social Security Entitlements and the Retirement Decision. *The Review of Economics and Statistics* 89(2), 234-246.

Costrell, Robert M. and Michael Podgursky. 2009. Peaks, Cliffs, and Valleys: The Peculiar Incentives in Teacher Retirement Systems and Their Consequences for School Staffing. *Education Finance and Policy* 4(2), 175-211.

Cullen, Julie Berry and Michael J. Mazzeo. 2008. Implicit performance awards: an empirical analysis of the labor market for public school principals. Unpublished manuscript. UCSD.

Dhuey, Elizabeth and Justin Smith. 2010. How Important Are School Principals in the Production of Student Achievement? Unpublished manuscript.

Ferguson, Ronald F. and Ladd, Helen F. 1996. "How and Why Money Matters: An Analysis of Alabama Schools," in *Holding Schools Accountable: Performance Based Reform in Education*, ed. Helen F. Ladd. Washington: Brookings Institution.

Greene, William H. 1994. Accounting for Excess Zeros and Sample Selection in Poisson and Negative Binomial Regression Models. New York University Department of Economics Working Paper EC-94-10.

Grissom, Jason A. 2011. Can Good Principals Keep Teachers in Disadvantaged Schools? Linking Principal Effectiveness to Teacher Satisfaction and Turnover in Hard-to-Staff Environments. *Teachers College Record* 113(11), 2552-2585.

Grissom, Jason A., and Susanna Loeb. 2011. Triangulating Principal Effectiveness: How Perspectives of Parents, Teachers, and Assistant Principals Identify the Central Importance of Managerial Skills. *American Educational Research Journal* 48(5), 1091-1123.

Harris, Douglas N. and Tim R. Sass (forthcoming). Teacher Training, Teacher Quality and Student Achievement. *Journal of Public Economics*.

Jackson, Kirabo. 2010. Match Quality, Worker Productivity, and Worker Mobility: Direct Evidence from Teachers. NBER Working Paper No. 15990. Cambridge, MA.

Jacob Brian. 2007. The Challenges of Staffing Urban Schools with Effective Teachers. *The Future of Children* 17(1), 129-153.

Linnemann, Hans. 1966. *An Econometric Study of International Trade Flows*. Amsterdam: North-Holland Publishing Company.

Loeb, Susanna, Demetra Kalogrides, and Eileen Horng. 2010. Principal Preferences and the Uneven Distribution of Principals across Schools. *Educational Evaluation and Policy Analysis* 32(2), 205-229.

Loeb, Susanna and Luke Miller. 2006. A Review of State Teacher Policies: What are they, what are their effects, and what are their implications for school finance? Policy Analysis for California Education. University of California – Berkeley.

Mitchell, Olivia, David McCarthy, Stanley Wisniewski, Paul Zorn. 2001. "Developments in State and Local Pension Plans." In O. Mitchell and E. Hustead. (eds) *Pensions and the Public Sector*. Philadelphia: University of Pennsylvania Press, 11-40.

Mullahy, John. 1986. Specification and Testing of Some Modified Count Data Models. *Journal of Econometrics* 33(3), 341-365.

Munnell, Alicia. 2000. The Impact of Mandatory Social Security Coverage of State and Local Workers: A Multi-State Review. AARP Public Policy Institute.

Podgursky, Michael and Mark Ehlert. 2007. Teacher Pensions and Retirement Behavior. CALDER Working Paper No. 5.

Reese, Clyde M. and Richard J. Tannenbaum. 1999. Gathering Content-Related Validity Evidence for the School Leaders Licensure Assessment. *Journal of Personnel Evaluation in Education* 13(3), 263-282.

Ruppert, Sandra S. 2001. *Improving Pension Portability for K-12 Teachers.* Washington DC:

State Higher Education Executive Officers (SHEEO) (February).

Stock, James H. and David A. Wise. 1990a. Pensions, the Option Value of Work, and Retirement. *Econometrica* 58(5), 1151-1180.

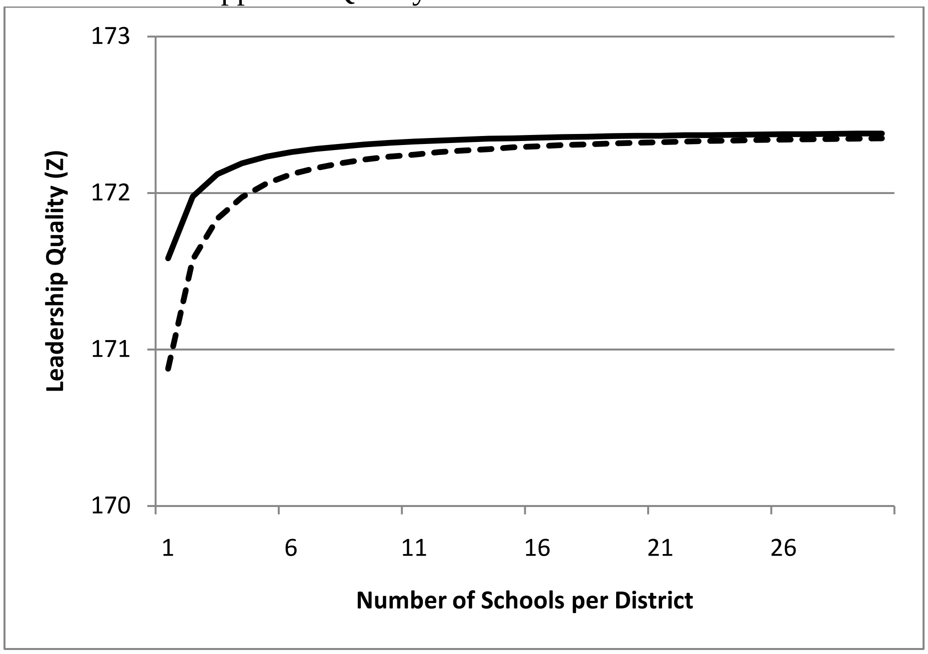
-- 1990b. The Pension Inducement to Retire: An Option Value Analysis. In David A. Wise, (Ed.) *Issues in the Economics of Aging* (Chicago: University of Chicago Press).

Thompson, Samuel B. (2010). Simple formulas for standard errors that cluster by both firm and time. *Journal of Financial Economics* 19(1), 1-10.

Tinbergen, Jan. 1962. *Shaping the World Economy: Suggestions for an International Economic Policy*. New York: The Twentieth Century Fund.

## Tables and Figures

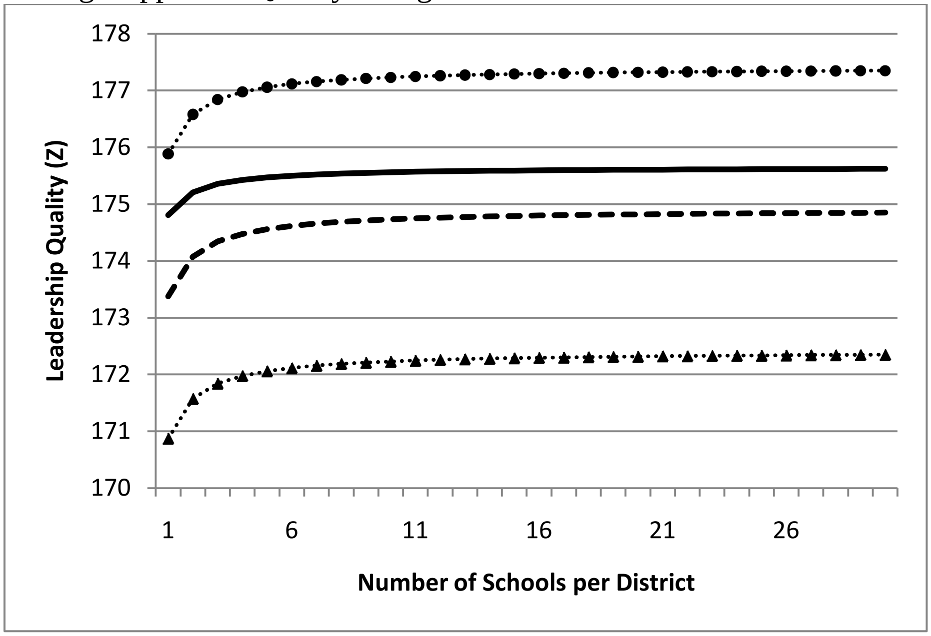
Figure 1. Average Leadership Quality in Schools from a Simple Two-District Model Where the Distribution of Applicant Quality is the Same on Both Sides of the Pension Border.




Notes: The dashed line indicates average leadership quality with the pension border; the solid line indicates average leadership quality without the pension border.

Brief model details: District 1 has $n_1$ schools, and therefore $n_1$ leadership positions. It has $N_1$ = ($n_1$*$s$) potential candidates, where $s$ indicates the number of candidates per school (i.e., teachers). Similarly, District 2 has $n_2$ schools and $N_2$ potential candidates. We set $n_1 = n_2 = n = 1$ and allow the districts to grow up to the point where $n = 30$. We hold $s$ fixed at 10 throughout. Candidate quality follows the same normal distribution in each district, $Z \sim N(160,50)$.

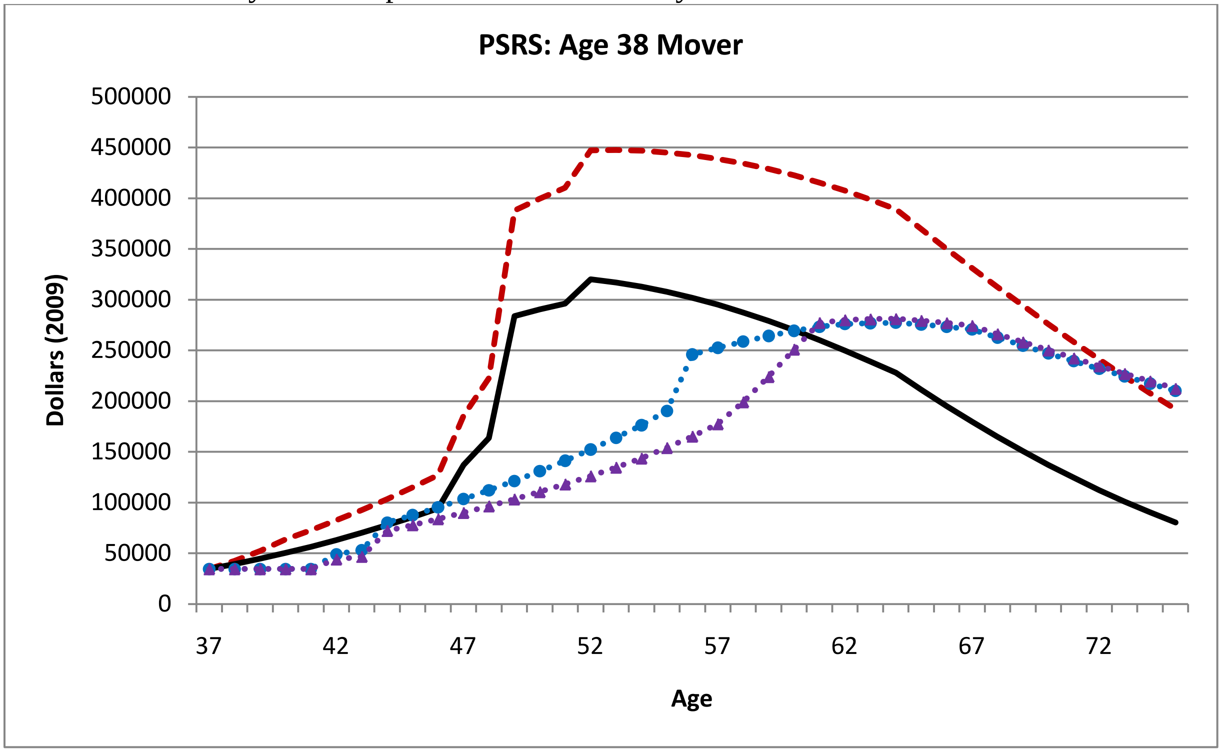
Figure 2. Average Leadership Quality in Schools from a Simple Two-District Model Where Average Applicant Quality is Higher in District 1.




Notes: The dashed line indicates average leadership quality with the pension border; the solid line indicates average leadership quality without the pension border; circles indicate average leadership quality at District 1 (stronger applicant pool); triangles indicate average leadership quality at District 2 (weaker applicant pool).

Brief model details: Same as in Figure 1 except that applicants in District 1 come from a new distribution: $Z_1 \sim N(165,50)$. Applicants from District 2 continue to come from the original distribution: $Z_2 \sim N(160,50)$.

Figure 3. Illustration of pension-wealth profiles for an age-38 teacher under various career scenarios. Based on year-2000 pension rules in each system.




Notes: Career scenarios: the teacher (1) remains in teaching in PSRS for her entire career (solid black line), (2) moves from teaching to leadership within PSRS at the age of 38 (red dashed line), (3) moves from teaching to leadership at the age of 38, but moves from PSRS to KC in the process (blue dotted line, circles), (4) moves from teaching to leadership at the age of 38, but moves from PSRS to STL in the process (purple dotted line, triangles).

Figure 4. District Size (Logged) Plotted Against the Within-District Hire Rate for Districts in the Pension Border Regions in Missouri. The Far-Right Circles Show the City Districts.

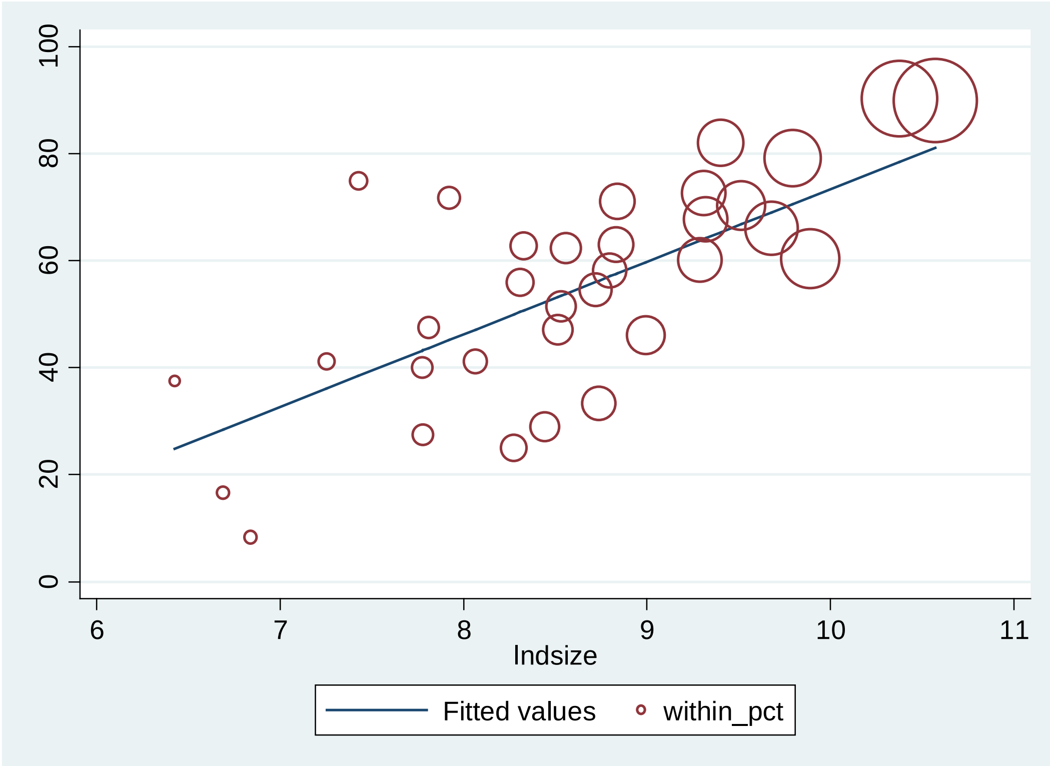


Figure 5. Illustration of Data Samples for the Simulations. Left: Schools Used in KC-Area Simulations. Center Top: Zoomed-Out View of KC- and STL-Area Schools. Center Bottom: KC- and STL-Area Schools Relative to the BLS Labor Market Areas. Right: Schools Used in STL-Area Simulations. The Darker Flags in Each Region Correspond to City Schools.

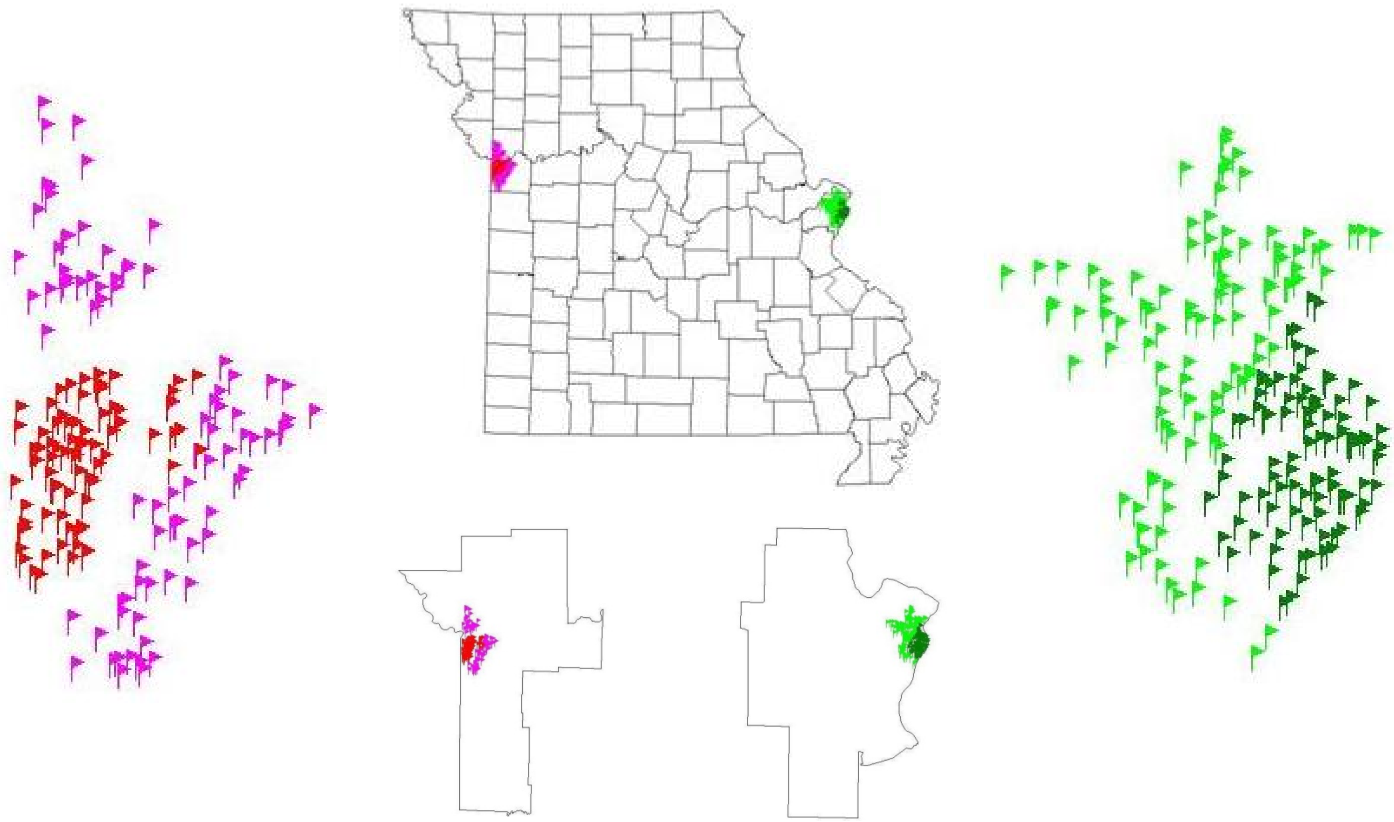


Figure 6. Function Mapping City Share (*CS* ) to Pension Sharing (*PS*) Within School Groupings from the Simulations.

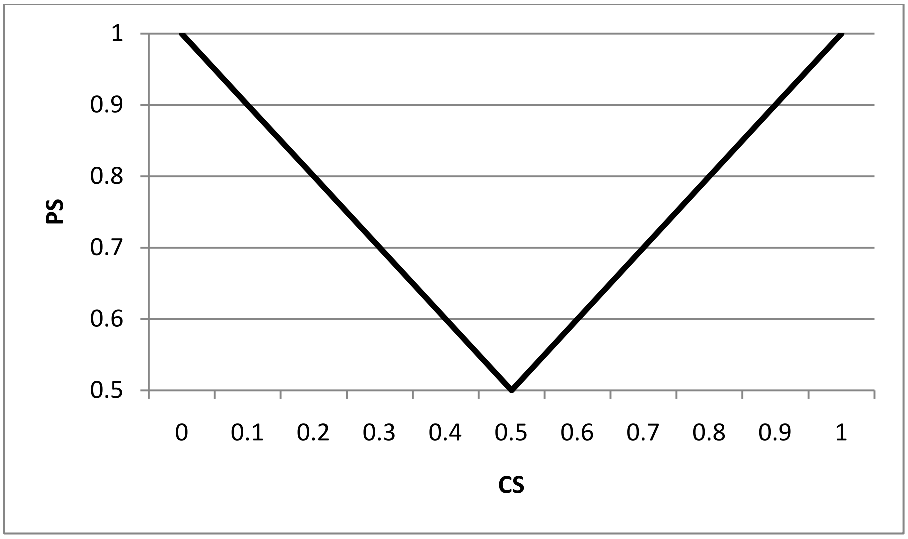


Table 1. Pension Plan Parameters in 2000.

| | PSRS | KC | STL |
|---|---|---|---|
| Vesting (Years) | 5 | 5 | 5 |
| Social Security | No | Yes | Yes |
| Retirement Eligibility (Normal and Early) | Full Retirement: Age 60 or 30 YOS; Early Retirement: Age 55, Rule of 80, 25-and-out | Full Retirement: Age 60; Rule of 75<br>Early Retirement: Any Age with YOS ≥30 | Full Retirement: Age 65; Rule of 85<br>Early Retirement: Age 60 |
| Formula Factor at Full Retirement | 0.025 | 0.020, *F*YOS* capped at 0.60 | 0.020, *F*YOS* capped at 0.60 |
| FAS | Highest three years | Highest four years | Highest three years (in last 10 years) |
| COLA | Yes (80% Cap) | No | No |

Table 2. Descriptive Statistics.

| | N | Percent of Total Hires |
|---|---|---|
| Total School Leadership Hires | 11034 | |
| Principal Hires | 6313 | 57.2% |
| Assistant Principal Hires | 4721 | 42.8 |
| | | |
| First-Time School Leadership Hires (In-State) | 5508 | 49.9 |
| From Teaching Positions | 3706 | 33.6 |
| From Supervisor/Other Administrative Positions | 1004 | 9.1 |
| From Guidance Counselor | 198 | 1.8 |
| From Other | 137 | 1.2 |
| Prior Position Not Recorded | 463 | 4.2 |
| Moves from Other Leadership Positions (In-State) | 5268 | 47.7 |
| From Principal Positions | 2255 | 20.4 |
| From Assistant Principal Positions | 2513 | 22.8 |
| From Other | 500 | 4.5 |
| Moves from Out-of-State | 258 | 2.3 |

| *Characteristics of School Leader Hires* | Mean | St. Dev |
|---|---|---|
| Female | 0.47 | 0.50 |
| Nonwhite | 0.15 | 0.35 |
| Salary (in $2009) | 68,307 | 15,485 |

Table 3. Within- and Out-of-District Leadership Hires for PSRS, KC, STL, and Subsamples of PSRS Districts.

| | PSRS All | KC | STL | PSRS Neighbors | PSRS Disadvantaged | PSRS Large Districts |
|---|---|---|---|---|---|---|
| <u>All Leadership Hires</u> | | | | | | |
| Within-District Hires | 54.5 | 90.4 | 90.0 | 61.0 | 58.7 | 72.8 |
| Out-of-District Hires | 43.7 | 8.5 | 9.3 | 37.8 | 40.1 | 26.2 |
| Unknown (in-state) | 1.8 | 1.1 | 0.8 | 1.2 | 1.2 | 1.0 |
| N | 9668 | 707 | 659 | 2214 | 414 | 1818 |
| | | | | | | |
| Avg. % Free/Reduced Lunch | 31.7 | 59.9 | 70.8 | 27.6 | 60.3 | 13.5 |
| Avg. % Disadv. Minority | 7.1 | 70.9 | 83.3 | 38.9 | 79.3 | 18.2 |
| Average Enrollment | 1495 | 32839 | 39948 | 6708 | 3121 | 17948 |
| Number of Districts | 539 | 1 | 1 | 32 | 9 | 10 |

Notes: PSRS Neighbors include districts that are geographically adjacent to either KC or STL, or are adjacent to the adjacent districts. Out-of-state hires are included with out-of-district hires. Unknown in-state hires are individuals with previous Missouri experience but for whom we cannot identify their place of prior employment.

Table 4. The Range of *PS* (Pension Sharing) in the Simulated Data, and Correlations Between School-Group Characteristics and *PS*.

| | KC | STL |
|---|---|---|
| Range of *PS* (pension sharing) | 0.50 – 0.94 | 0.52 – 0.81 |
| Correlation between *PS* and: | | |
| Share eligible for free/reduced lunch | 0.012 | 0.008 |
| Share nonwhite | -0.007 | -0.003 |
| Enrollment (log) | -0.023* | -0.013 |
| Share of city schools | 0.006 | 0.005 |

* Indicates statistical significance at the 5 percent level.

Notes: The larger the pension-sharing variable, the less segmented is the group of schools.

Table 5. Correlations in the Simulated Data Between the School-Group Similarity and Distance Measures, and *PS* (Pension Sharing).

| | KC | STL |
|---|---|---|
| Correlation between *PS* and: | | |
| Free/reduced-lunch similarity | 0.350** | 0.073** |
| Racial similarity | 0.300** | 0.003 |
| Average distance between schools | -0.084** | -0.114** |

** Indicates statistical significance at the 1 percent level.

Table 6. Regression Results for Simulations. Dependent Variable: Shared Leadership Hires.

| | Model 1 | Model 2 | Model 3 | Model 4 | Model 5 |
|---|---|---|---|---|---|
| Kansas City | | | | | |
| **Pension sharing (PS)** | **6.08** | **5.97** | **4.08** | **6.10** | **6.21** |
| | **(1.21)**** | **(1.22)**** | **(1.24)**** | **(1.24)**** | **(1.24)**** |
| KC schools share | | 12.56 | | 12.51 | 12.87 |
| | | (0.71)** | | (1.02)** | (1.12)** |
| Free/reduced-lunch similarity | | | -13.10 | -3.34 | -3.53 |
| | | | (2.13)** | (2.16)** | (2.18)** |
| Racial similarity (nonwhite) | | | 20.40 | 4.17 | 4.20 |
| | | | (1.51)** | (1.89)* | (1.94)** |
| Enrollment (log) | | | 9.71 | 9.88 | 9.87 |
| | | | (0.48)** | (0.47)** | (0.47)** |
| Average distance (miles) | | | | | 0.07 |
| | | | | | (0.10) |
| St. Louis | | | | | |
| **Pension sharing (PS)** | **4.41** | **4.37** | **4.25** | **4.45** | **4.17** |
| | **(1.10)**** | **(1.09)**** | **(1.04)**** | **(1.02)**** | **(1.03)**** |
| STL schools share | | 4.77 | | 5.56 | 5.14 |
| | | (0.64)** | | (0.69)** | (0.70)** |
| Free/reduced-lunch similarity | | | 7.61 | 3.77 | 4.08 |
| | | | (1.68)** | (1.69)** | (1.69)** |
| Racial similarity (nonwhite) | | | 8.52 | 6.67 | 5.30 |
| | | | (1.47)** | (1.48)** | (1.65)** |
| Enrollment (log) | | | 9.91 | 10.38 | 10.43 |
| | | | (0.38)** | (0.39)** | (0.39)** |
| Average distance (miles) | | | | | -0.21 |
| | | | | | (0.11)* |

** indicates statistical significance at the 1 percent level
* indicates statistical significance at the 5 percent levels
Notes: Robust standard errors are in parentheses, clustered within simulations and adjusted for the effective number of simulations as described in Appendix B.

Table 7. Pairwise Logit Results. Marginal Effects from Full Models. Dependent Variable: Any Shared Leadership Hire.

| | Full Model |
|---|---|
| Kansas City (16,836 pairs) | |
| **Pension-Line Crossing** | **-0.568** |
| | **(0.280)*** |
| District-Line Crossing | -4.591 |
| | (0.980)** |
| Free/reduced-lunch similarity | 0.008 |
| | (0.004)* |
| Racial similarity (nonwhite) | 0.004 |
| | (0.005) |
| Enrollment (log) | 1.059 |
| | (0.144)** |
| Distance (miles) | 0.063 |
| | (0.038)† |
| $Distance^2$ (miles) | -0.008 |
| | (0.003)* |
| St. Louis (35,245 pairs) | |
| **Pension-Line Crossing** | **-0.360** |
| | **(0.108)**** |
| District-Line Crossing | -2.037 |
| | (0.302)** |
| Free/reduced-lunch similarity | 0.005 |
| | (0.002)* |
| Racial similarity (nonwhite) | 0.002 |
| | (0.002) |
| Enrollment (log) | 0.828 |
| | (0.077)** |
| Distance (miles) | -0.099 |
| | (0.024)** |
| $Distance^2$ (miles) | 0.004 |
| | (0.002)* |

** indicates statistical significance at the 1 percent level
* indicates statistical significance at the 5 percent level
† indicates statistical significance at the 10 percent level
Notes: The estimates reported in the table are not marginal effects. We use the standard logit estimates as reported to estimate the marginal effect of removing the pension border between two schools as discussed in the text. Robust standard errors are in parentheses. All coefficients and standard errors are multiplied by 100 for ease of presentation.

Table 8. Average ACT Scores for Incoming Teachers between 2005 and 2009: PSRS, KC, STL and PSRS Neighboring Districts.

| | | Kansas City Region | | St. Louis Region | |
|---|---|---|---|---|---|
| | PSRS All | KC Schools | Neighbor (PSRS) | STL Schools | Neighbor (PSRS) |
| Average ACT Score | 22.83 | 21.26 | 23.00 | 21.96 | 23.40 |
| | (3.87)[a,b] | (4.69) | (3.74)[a] | (4.72) | (4.11)[b] |
| N (teachers) | 11,501 | 177 | 1190 | 336 | 1304 |
| | | | | | |
| Avg. % Free/Reduced Lunch | 31.7 | 59.9 | 23.9 | 70.8 | 29.6 |
| Avg. % Disadv. Minority | 7.1 | 70.9 | 27.1 | 83.3 | 45.0 |
| Districts | 539 | 1 | 11 | 1 | 21 |

Notes: The neighboring districts are within a commutable distance to the city in each region. ACT scores are reported for all teachers for whom they are observed, which is roughly half of all of the new teachers who are identified in the statewide data panel. Standard deviations are in parentheses. Superscripts indicate statistically significant differences ($p < 0.01$) from Kansas City (a) and St. Louis (b), respectively (test results are reported within region and for each city district relative to PSRS).

Table 9. Licensure Exam Scores and College Quality for School Leaders: PSRS, KC, STL and PSRS Neighboring Districts.

| | | Kansas City Region | | St. Louis Region | |
|---|---|---|---|---|---|
| | PSRS All | KC Schools | Neighbor (PSRS) | STL Schools | Neighbor (PSRS) |
| Licensure Exam Scores | | | | | |
| Average Score | 178.2 | 172.8 | 179.8 | 175.1 | 178.8 |
| | (7.4)[a,b] | (7.2) | (7.0)[a] | (7.1) | (7.3)[b] |
| N (leaders) | 4,099 | 163 | 261 | 222 | 322 |
| | | | | | |
| College Quality: All | | | | | |
| High Quality | 0.174 | 0.112 | 0.184 | 0.077 | 0.233 |
| | (0.379)[a,b] | (0.316) | (0.388)[a] | (0.267) | (0.423)[b] |
| N (leaders) | 8,873 | 339 | 803 | 377 | 1101 |
| | | | | | |
| College Quality: MO Specific | | | | | |
| High Quality | 0.187 | 0.083 | 0.177 | 0.078 | 0.262 |
| | (0.390)[a,b] | (0.276) | (0.382)[a] | (0.269) | (0.440)[b] |
| Low Quality (Public) | 0.064 | 0.223 | 0.056 | 0.566 | 0.162 |
| | (0.245)[a,b] | (0.417) | (0.229)[a] | (0.496) | (0.369)[b] |
| N (leaders) | 7,089 | 193 | 575 | 295 | 809 |
| | | | | | |
| Avg. % Free/Reduced Lunch | 31.7 | 59.9 | 23.9 | 70.8 | 29.6 |
| Avg. % Disadv. Minority | 7.1 | 70.9 | 27.1 | 83.3 | 45.0 |
| Districts | 539 | 1 | 11 | 1 | 21 |

Notes: The neighboring districts are within a commutable distance to the city in each region. Licensure exam scores are available for school leaders from 2000-2009. College quality is available throughout the data panel and is coded based on the institutions where leaders obtained their initial bachelor's degrees. Standard deviations are in parentheses. Superscripts indicate statistically significant differences ($p < 0.01$) from Kansas City (a) and St. Louis (b), respectively (test results are reported within region and for each city district relative to PSRS).

## Appendix A
## Details for the Pension-Wealth Calculations

The pension-wealth calculations are based on the year-2000 rules for each system (see Appendix Table C.1 for rules in all years – the mobility penalties are large regardless of which set of rules we use). We calculate mobility costs in Appendix Table C.2 for moves in all possible directions in Missouri at three different ages for the representative teacher – age-38, age-45, and age-50. These ages correspond to the 50$^{th}$, 75$^{th}$ and 90$^{th}$ percentiles of the distribution of age-at-entry into leadership in Missouri. Our calculations use a real interest rate of four percent.[38]

We project the representative teacher's survival probabilities and wages, and some parameters of the Social Security system. The survival probabilities come from the Social Security Administration (Cohort Life Tables). We use survival probabilities for women; expected mobility costs for men are slightly lower. We project wages under different career scenarios using data for all teachers and leaders in Missouri over the course of our data panel. The remainder of this appendix describes the wage projections.

To project wages we begin with the sample-average wage among teachers in Missouri who are the same age as the mover in the year prior to the move. This wage is projected forward and backward to create a full wage profile over the career cycle under the different career scenarios. Projecting wages forward would suffice for the pension systems alone; however, Social Security wealth depends on a long history of earnings (up to 35 years) so we must also project wages backward.

For teaching years we project wages using the parameters from a wage growth function that is a cubic of experience. Moves to leadership correspond to a 30-percent wage increase, on average in the data. The prediction function for the immediate wage increase is again a cubic of teaching experience. Because we combine assistant principals and principals as "school leaders", the wage-increase function averages all moves from teaching to either position. However, the results are qualitatively similar if we focus only on full principals. A large share of the salary increase is associated with moving from a 9-month to 11-month contract.

[38] This discount rate falls somewhere in the middle of what others in the literature have used. For example, Coile and Gruber (2004) use a real discount rate of 6 percent, and Costrell and Podgursky (2009) use a real discount rate of 2.5 percent.

Once a switch occurs, the wage-growth function for school leaders differs from that of teachers. First, it depends on leadership experience, not teaching experience.[39] Second, wages for leaders do not grow smoothly so we estimate a step function that measures the average gain in salary going from one to two years of leadership experience, two to three years, etc. The real return to experience is highest early in leaders' careers, and flattens out at around two percent per year by year nine.[40] Third, because school leaders who remain in leadership longer are likely to be non-randomly selected (much more so than teachers), we include individual fixed effects in the leadership wage-growth function.[41]

[39] That is, teaching experience does not predict principal wage growth conditional on principal experience. This is consistent with evidence from Clark et al. (2009), who show that principal effectiveness is related to prior leadership experience, but not to prior teaching experience.
[40] We allow the final step to repeat indefinitely for principal spells lasting more than 9 years.
[41] For consistency we also estimated teacher-wage-growth models with individual fixed effects, but as expected, the individual fixed effects in the teacher projection function are less important. We have considered the sensitivity of our findings to using a model without fixed effects and this does not qualitatively affect our results.

## Appendix B
## Procedure for Determining the Effective Number of Simulations

We resample from the same underlying pool of schools to construct the groups of schools in each region. Therefore, in addition to the data being potentially correlated across groups within each simulation, the data are also correlated across simulations. Two-way clustering approaches (Thompson, 2010; Cameron et al., 2011) cannot be used because the groups within each simulation are unordered, and there are not cluster labels. This makes it inadvisable, and undesirable, to impose assumptions on the variance-covariance structure across simulations (although, as noted in the text, we do allow for within-simulation error covariance). To deal with the across-simulation correlation in the data, we develop a procedure to estimate the effective simulation counts given the actual number of simulations that we run in each model (1,000). The effective simulation calculations are based on an *ex post* analysis of the degree of similarity of error terms from the regression in (3) across simulations.

We compare the 1,000 batches of residuals from the regression (normalized to unit variance), where each batch contains the 10 residuals corresponding to the 10 groups, to analogous batches of residuals from a Monte Carlo. First, we sort the 10 regression residuals in each batch in ascending order and compute the variance of the *k*th largest term, $s_k^2$ (k=1,..,10), across the 1,000 batches. Then we conduct a Monte Carlo where we draw 1,000 batches of 10 errors from $q$ independent batches, where the Monte Carlo batches come from a comparable distribution to the distribution of errors in the simulated data.[42] We control $q$ and draw 1,000 batches at each level of $q$ from 1 to 1,000. As with the residuals from regression (3), we sort the 10 errors from each Monte Carlo batch in ascending order and compute the variance of the *k*th (k=1,..,10) largest term across the 1,000 batches. We repeat the procedure 10,000 times and compute the average variance $\sigma k$ 2 (k=1,..,10). $\sigma_k^2$ is increasing in the number of independent batches, $q$. In the extreme case with $q$ =1, all 1,000 batches are identical and $\sigma_k^2 = 0$ (for k=1,..,10). If the errors in each batch were iid, then as $q$ increases $\sigma_k^2$ approaches the variance of the k-order statistics.[43]

[42] The 10 errors in each batch in the Monte Carlo are drawn from a joint normal distribution with unit variance and a given within-batch correlation, which reflects the estimated within-batch correlation in the errors in the simulated data (e.g. -0.03 for schools bordering Kansas City).

[43] The errors within each batch are not iid because of the within-batch correlation. We also conduct simulations based on draws that are iid within batches. The estimated effective simulation counts do not change substantially.

By controlling $q$, it is straightforward to track the amount of underlying independent data in the Monte Carlo. We estimate the effective number of simulations in each model, $q$, by equating the cross-simulation variances of the ordered residuals, $s_k^2$, and the Monte Carlo variances, $\sigma_k^2$, for (k=1,..,10). We take the implied $q$ averaged over $k$. The resultant $q$*'s* for the sample range from 350 to 450. They are larger in the STL region where the underlying pool of schools is larger.[44]

[44] Sample programs for the Monte Carlo are available from the authors upon request.

# Appendix C
# Supplementary Tables

Appendix Table C.1. Changes to Key Parameters of Pension Systems, 1995 – 2009 (there were no changes after 2002). Initial Parameters as of 1995 are Reported in Row 1.

| | PSRS | KC | STL |
|---|---|---|---|
| 1995 | Formula factor 0.023, early retirement by 55-25 rule, COLA cap 65 percent | Formula factor 0.0175, Rule of 75, no COLA | Formula factor 0.0125, Rule of 85, no COLA |
| 1996 | Implement unrestricted “25 and out” | | |
| 1997 | COLA cap increased from 65 to 75 percent | | |
| 1998 | | | |
| 1999 | Formula factor raised to 0.025 for full retirement (with corresponding upward adjustments for early retirement) | | |
| 2000 | Implement Rule of 80<br><br>FAS changed to highest three years of salary | Formula factor increased to 0.020 | Formula factor increased to 0.020, *F*YOS* capped at 0.60 (previously uncapped) |
| 2001 | COLA cap increased to 80 percent | | Implement DROP provision |
| 2002 | Formula factor increased to 0.0255 if YOS $\geq$ 31 (new factor applies to *all service years* for eligible individuals) | | |

* The DROP provision in St. Louis allows teachers and principals who qualify for full retirement to delay retirement for up to four years without losing their pension payments for that time. Pension wealth is frozen, and those payments are “dropped” into an account that is paid out at the time of work stoppage.

Appendix Table C.2. Pension-Wealth Costs of Across System Moves into Leadership Positions Based on Peak Value. Year-2000 Pension Rules. Year-2009 Dollars. Women.

| | Constrained Retirement Date in Second System | | Work Until Age 65 in Second System | |
|---|---|---|---|---|
| | Comparison Between Across-System and Within-System Moves to Leadership | Comparison of Within-System Career-Teacher Scenario and Across-System Move to Leadership | Comparison Between Across-System and Within-System Moves to Leadership | Retirement Age Without/With Move (Determined by Peak value) |
| PSRS to KC | | | | |
| Move at: 38 | -283,638 (-63) | -156,093 (-49) | -171,978 (-38) | 53 / 65 |
| 45 | -339,682 (-65) | -204,658 (-52) | -216,116 (-41) | 53 / 65 |
| 50 | -85,868 (-14) | -52,486 (-11)* | 3,987 (+1) | 56 / 65 |
| PSRS to STL | | | | |
| Move at: 38 | -312,871 (-70) | -185,326 (-58) | -168,081 (-38) | 53 / 65 |
| 45 | -358,815 (-68) | -223,791 (-57) | -212,680 (-40) | 53 / 65 |
| 50 | -132,436 (-22) | -52,486 (-11)* | 6,738 (+1) | 56 / 65 |
| KC to PSRS | | | | |
| Move at: 38 | -151,768 (-44) | -63,522 (-25) | -39,766 (-12) | 54 / 65 |
| 45 | -164,237 (-40) | -73,922 (-23) | -37,570 (-9) | 54 / 65 |
| 50 | -53,453 (-12) | -30,033 (-8)* | 89,136 (+19) | 54 / 65 |
| KC to STL | | | | |
| Move at: 38 | -181,190 (-53) | -92,944 (-36) | -72,766 (-21) | 54 / 65 |
| 45 | -179,182 (-44) | -88,867 (-28) | -68,847 (-17) | 54 / 65 |
| 50 | -66,138 (-14) | -30,033 (-8)* | 45,598 (+10) | 54 / 65 |
| STL to PSRS | | | | |
| Move at: 38 | -141,445 (-41) | -51,310 (-20) | -46,003 (-13) | 55 / 65 |
| 45 | -198,204 (-48) | -103,736 (-33) | -87,382 (-21) | 55 / 65 |
| 50 | -183,799 (-39) | -96,915 (-25) | -56,970 (-12) | 55 / 65 |
| STL to KC | | | | |
| Move at: 38 | -142,957 (-42) | -52,822 (-21) | -82,901 (-24) | 55 / 65 |
| 45 | -193,231 (-47) | -98,763 (-32) | -122,095 (-30) | 55 / 65 |
| 50 | -187,890 (-40) | -101,006 (-27) | -105,065 (-22) | 55 / 65 |

Notes: Percentages of baseline pension wealth under the relevant no-move scenario are reported in parenthesis. An '*' indicates that under the constrained retirement scenario the individual will not become vested in the second pension system, which occurs in several cases when we use the career-teacher scenario as the baseline.

Appendix Table C.3. Within- and Out-of-District Leadership Hires for PSRS, KC, STL, and Subsamples of PSRS Districts. Various Subsamples of School Leaders.

| | PSRS All | KC | STL | PSRS Neighbors | PSRS Disadvantaged | PSRS Large Districts |
|---|---|---|---|---|---|---|
| First-Time Leadership Hires | | | | | | |
| Within-District Hires | 60.3 | 89.9 | 91.4 | 67.7 | 60.1 | 76.9 |
| Out-of-District Hires | 36.2 | 6.3 | 6.8 | 29.7 | 37.4 | 20.7 |
| Unknown (in-state) | 3.5 | 3.9 | 1.8 | 2.6 | 2.5 | 2.4 |
| N | 5023 | 207 | 278 | 1028 | 203 | 801 |
| | | | | | | |
| First-Time Leadership Hires Direct from Teaching | | | | | | |
| Within-District Hires | 60.6 | 92.8 | 91.2 | 67.8 | 54.2 | 78.7 |
| Out-of-District Hires | 39.4 | 7.2 | 8.8 | 32.2 | 45.8 | 21.3 |
| Unknown (in-state) | - | - | - | - | - | - |
| N | 3421 | 138 | 147 | 627 | 118 | 507 |
| | | | | | | |
| Full Principals Only, All Hires | | | | | | |
| Within-District Hires | 53.4 | 86.5 | 87.2 | 64.4 | 64.6 | 77.5 |
| Out-of-District Hires | 44.5 | 12.8 | 12.6 | 34.4 | 34.1 | 21.8 |
| Unknown (in-state) | 2.0 | 0.7 | 0.3 | 1.2 | 1.2 | 0.7 |
| N | 6342 | 282 | 366 | 1086 | 246 | 843 |
| | | | | | | |
| Age 45 or Younger (All Hires) | | | | | | |
| Within-District Hires | 54.4 | 88.1 | 86.7 | 59.8 | 56.7 | 70.5 |
| Out-of-District Hires | 44.6 | 10.1 | 12.2 | 39.7 | 43.3 | 29.0 |
| Unknown (in-state) | 1.0 | 1.8 | 1.1 | 0.6 | 0.0 | 0.5 |
| N | 6327 | 218 | 180 | 1382 | 217 | 1205 |
| | | | | | | |
| Avg. % Free/Reduced Lunch | 31.7 | 59.9 | 70.8 | 27.6 | 60.3 | 13.5 |
| Avg. % Disadv. Minority | 7.1 | 70.9 | 83.3 | 38.9 | 79.3 | 18.2 |
| Average Enrollment | 1495 | 32839 | 39948 | 6708 | 3121 | 17948 |
| Number of Districts | 539 | 1 | 1 | 32 | 9 | 10 |

Notes: PSRS Neighbors include districts that are geographically adjacent to either KC or STL, or are adjacent to the adjacent districts. The neighboring districts are within a commutable distance to the city in each region.

*Appendix Table C.1*

Appendix Table C.1 shows changes to the pension-system rules over the course of the data panel. Benefits became more generous during the late 1990s and early 2000s in each system, mirroring a larger national trend. Mobility costs have increased due to the benefit enhancements. The calculations corresponding to Figure 3 in the text are based on the system rules as of the year-2000.

*Appendix Table C.2*

Appendix Table C.2 shows the costs of pension-border crossings in Missouri for the representative teacher moving into leadership at three ages: 38, 45 and 50. These ages correspond to the 50$^{th}$, 75$^{th}$ and 90$^{th}$ percentiles of the unconditional distribution of age-at-entry for new school leaders in Missouri who come directly from teaching (for whom the simulations are relevant based on the wage projections). The pension-accrual profiles in Figure 3 correspond to the age-38 move scenarios for the PSRS teacher.

The table compares pension-wealth accrual under the scenario where an across-system move into leadership occurs to scenarios where (1) the move into leadership occurs without crossing a pension border, and (2) no leadership promotion occurs and the individual remains in teaching within the same pension system. There is some variability in the cost of the move depending on its direction, although the costs in most cases are high. One source or variability in the mobility costs is that it is more costly to move from the more-generous PSRS system, which further disadvantages the city districts. Also, differences in the early-retirement provisions across systems are important – it is more costly to leave the plans that have more-generous early retirement provisions for younger workers, and vice-versa for older workers.[45] On a related note, late-career leaders in PSRS and KC stand to lose less if they move, and in some cases can marginally gain pension wealth because they can double-dip. That is, they can begin collecting their first pension while working in the new system. However, Appendix Table C.3 shows that double dipping cannot explain what little cross-border mobility we see in the data.

[45] For example, consider the age-50 mover, for whom the constrained-retirement costs of exiting KC and PSRS are much smaller than the costs of exiting STL. This is because in both PSRS and KC the mover can take advantage of an early-retirement provision and begin collecting benefits immediately (25-and-out and the rule-of-75, respectively). In STL, the move is more costly because the age-50 mover is still 10 years away from the rule-of-85 at the time of the move. Many studies in the general retirement literature focus on replacement rates as a measure of pension generosity (e.g., Mitchell, et.al , 2011; Loeb and Miller, 2006; Clark and Lee, 2011). Appendix Table C.2 shows that the age when individuals can begin collecting is also important.

*Appendix Table C.3*

Appendix Table C.3 shows within- and out-of-district hire rates for subsamples of school leaders in Missouri. The mobility rates within each subsample are generally similar to what is reported for all leaders in Table 3 in the text.

On the issue of double-dipping, note that the out-of-district hire rate is not lower in the city districts for movers who are aged 45 or younger. Based on our calculations in Appendix Table C.2, one might expect lower mobility rates for this group if across-system mobility is primarily driven by older movers, who can draw on their first pension while working in the new system. But the data are not consistent with double-dipping playing an important role in the labor market. We offer two explanations for this result. First, it could be that late-career educators – who would be eligible to double-dip – are more likely to turnover after shorter leadership spells, which would discourage demand. Second, on the supply side, despite the potential for pecuniary gains associated with late-career switches across borders, individuals' marginal returns to labor and leisure during the late stage in their careers may be such that the move is undesirable. This explanation seems more plausible given that even without double dipping, educators who are eligible for full retirement in Missouri already collect generous pensions.