

**An HLM Analysis of the Effects of a Teachers' Contract on More
Equitable Distribution of Experienced Faculty in Urban Schools**

Robert M. Offenberg, Ed. D.

**Independent Consultant
and
Adjunct Professor
St. Joseph's University
Philadelphia PA**

roffenb@aol.com

215-972-6792

*This research was a component of an
evaluation managed by Research for Action,
Philadelphia, PA.*

*Convention of the American Education Research Association
2008*

An HLM Analysis of the Effects of a Teachers' Contract on More Equitable Distribution of Experienced Faculty in Urban Schools

This paper presents a family of three hierarchical analyses of trends in the distribution of experienced teachers among schools. The analyses were used to assess whether implementation of key provisions of the 2004 contract between the Philadelphia School District and the Philadelphia Federation of Teachers (PFT) led to a more equitable teacher-experience distribution. HLM was applied to three years of pre-contract school data to create an Elementary/K-8, a Middle School and a High School function describing the relationship between characteristics of the students attending schools and the experience of the teachers in their faculties. The pre-contract equations were then applied to appropriate school data, yielding an estimate of what the experience of each school would have been post-contract if the contract had had no effect. Residuals, the differences between the HLM-based predictions and the actual experience of each schools faculty, were computed and then subjected to additional analyses. The results showed that the teachers' contract provisions were associated with an increase in school faculty experience in the school district. However the increase occurred mainly at more middle class schools, and not in the ones that were subject to contract provisions designed to attract experienced teachers.

This paper presents a family of three hierarchical analyses of trends in the distribution of experienced teachers among schools. The analyses were used to assess whether implementation of key provisions of the 2004 contract between the Philadelphia School District and the Philadelphia Federation of Teachers (PFT) were followed by a more equitable distribution of experienced faculty. They comprised a component of a larger study, *Closing the Teacher Quality Gap in Philadelphia: New Hope New Hurdles* (Useem, Offenber, and Farley, 2007) that explored many aspects of the school district's efforts to improve the quality of the faculties of its schools.

The School District of Philadelphia is a large, urban system. It had about 14,000 teachers in 250 schools from fall 2002 and fall 2005, when this study was conducted. Like nearly all large urban systems, school achievement, racial and income data on its schools are easily available to teachers from many sources including the Pennsylvania and local Philadelphia internet web pages, and are therefore common knowledge that teachers could use when considering school-to-school transfers. Under the seniority-based school-staffing provisions of contracts prior to the new 2004 agreement, teachers could use this information when involved in school transfers they initiated or ones required by district management. According to these contracts, employed teachers received their assignments first in order of school district seniority, and then new teachers, ranked by eligibility criteria, were assigned to schools where vacancies remained. Although the transferring and new teachers could choose among schools needing staff, there was little or no input from the principal or staff of the receiving school. Research showed that when these policies were in effect, veteran teachers tended

to transfer or to be ‘force-transferred’ to schools where students were higher achieving, had higher incomes and were less likely to be a minority than were students in the schools they left (Chester, Offenber and Xu, 2001). As a result, inexperienced teachers tended to be over-represented in faculties of schools serving lower achieving, lower income, heavily minority enrollments; schools where experienced teachers were most needed.

The 2004 contract between the school district and the Philadelphia Federation of Teachers contained three elements, all implemented beginning in fall 2005, that were supposed to weaken these patterns, and make the distribution of experienced faculty more equitable. All three were based on the assumption that school-based committees would be better able to staff their sites more equitably than the centralized procedures that preceded them had. First, all schools were allowed to participate in ‘Partial Site Selection’, that is the centrally-managed system could be bypassed for all new hires, and the seniority-based rules could be bypassed for half of the remaining teacher appointments. The new-hire and the seniority-bypass staffing decisions would be the responsibility of the school based committees. Second, a staffing approach called ‘Full Site Selection’ in 2005 was carried over from the preceding contract, and given greater emphasis than in the past. This approach required an annual, two-thirds vote of the faculty, but then allowed all new hires and all transfers to be the responsibility of the school-based committee. The third approach, ‘Incentive Schools’ in addition to giving school-based committees full responsibility for hiring new personnel, provided some incentives to teachers who came to or stayed at schools—principally tuition reimbursement for graduate courses and additional personal leave. Useem, Offenber and Farley (2007) discusses the implementation of the three approaches in detail.

The focus here is on the hierarchical analysis used to determine whether the fall 2005 distribution of teachers was more equitable than it had been in the past, and therefore consistent with the goals of the contract. The approach was to:

- Use Hierarchical Linear Modeling or HLM (Raudenbush, Bryk and Congdon, 2004) to create a three-year, pre-contract time series relating the average experience of the faculty of each school each year to the prior year’s summary of the school’s student racial, income and reading test score data—the information that teachers transferring among schools or being newly placed could easily obtain.
- Apply the equations yielded by the pre-contract time series model to 2004 student data to predict what the experience of the faculty of each school would have been in 2005 if the new contract’s staffing policies had had no effect.
- Compare the observed experience of each school’s faculty with its predicted value, searching for patterns among the residuals that confirmed or disconfirmed the contract’s having made the distribution of experienced teachers more equitable in 2005.

A two-level HLM analysis was used to derive the pre-contract, time-series prediction functions. At Level 1, the dependent variable was the average years of experience of teachers at a school in a given year. The student-population predictors were: the most recent value of three school enrollment variables that teachers could know when making a school choice decision: the proportion of the school’s enrollment that was

White (i.e. non-minority), the proportion that was Low Income, and the proportion in tested grades who met Pennsylvania reading proficiency standards the previous school year. A counter indicating the study year was the fourth Level 1 variable. As the study was three years long, there were three Level 1 ‘cases’ per school. Level 2 was used to organize the Level 1 records of schools into a related series, but no new independent variables were added.

This analysis was replicated for ‘Elementary/K-8’ schools, Middle Schools and High Schools. The elementary and K-8 schools were grouped together because they shared many characteristics, and the school district occasionally added or deleted grades from these schools during the years of the study. The grade organizations of the Middle and High schools did not change.

As will be shown in the next section, the HLM analyses yielded prediction equations that confirmed broad, pre-contract beliefs about the staffing of Philadelphia schools and yielded predictions of teacher experience during this period that were consistent enough for us to believe that they could, across the school district, be used to predict what would have happened during the first post-contract year if the contract failed to have an effect on the level of, or distribution of teacher experience among schools. We applied the prediction equations obtained from the HLM models to 2004 student data, increasing the value of the year-variable to reflect the passage of time, and derived 2005 ‘no contract effect’ staffing predictions for each school. In a residual analysis we then found how much the actual experience of individual school faculties differed from what we would have expected, first in general and then for subgroups of schools, if the PFT contract provisions had no effect and temporal trends continued.

Pre-contract Predictive Analysis. Table 1 shows the HLM-derived equations relating the average experience of teachers at schools in fall 2002, 2003 and 2004 to the characteristics of the schools enrollments in each of the previous school years. The three HLM-derived models are all similar, and all confirm the existence of school staffing patterns that led to the PFT contract in all three grade-ranges of schools. They are based on all schools in the district that operated continuously from September 2001 through September 2005.¹ They all show that the presence of more white students, fewer poverty students and, except at high schools, more reading proficient students one year was associated with a more experienced faculty the next year. The negative, significant, ‘Annual Trend’ predictor shows that, during the pre-contract study phase, if there were a school where the student background and achievement values were constant, the average experience of its teachers would have tended to decrease annually, a trend that had been noticed before this study.

Scaling of the variables was done with care, and so trends in the values of the intercepts in the three models could be compared informally, even though their differences were not tested for significance. They suggest that middle schools had the least experienced faculties while high schools had the most experienced faculties, a

¹ Some of the schools were divided into components that operated independently during the last year of the study. The component parts were reassembled in order to make the student information of 2004-05, and the fall staffing of 2005 consistent with the schools’ history. The grade organizations of two schools were changed in fall 2005. While these schools were kept in the prediction equations, no 2005 predictions were made for them.

finding that is consistent with other school district data. Given that both the school district and teachers have a multiplicity of reasons for choosing, transferring or making extended commitments to schools, correlation of the predictions these models yield with the actual experience of teachers at schools are very high. They are .719 for the Elementary/K-8 schools, .778 for the Middle schools, and .820 for the High schools. Thus these models explain 52% to 67% of the before-the-contract variance in teacher experience.

Table 1
Average Years of Experience of Teachers at Schools as a Function of Student Characteristics and the Annual Trend, 2002-2004 School Years.

Predictor	School Type		
	Elementary and K-8 (174 Schools)	Middle (38 Schools)	High (38 Schools)
Intercept [in years]	14.671***	12.540***	19.588**
Portion of Enrollment that is:			
White	6.658***	8.040**	4.915*
Low Income	-5.544***	-4.944*	-6.781***
Reading Proficient	4.902***	5.866**	0.427
Annual Trend	-0.658***	-0.886***	-0.890***

***p<.001, **p<.01, *p<.05

Post-Contract (Fall 2005) Findings.

The three 2002 to 2004 functions that related the experience of schools' faculties in the fall to their previous years' student data were applied to the 2004-05 student data of each school to obtain an estimate of what the teaching experience level at each would have been in 2005 if the contract had had no effect and the pre-contract trends had merely continued. The residuals, the differences between predicted and actual teacher experience levels of schools, were then obtained. If the contract had had no general effect, the mean residual values would be about '0', so t-tests comparing the mean residuals of each type of school to '0' would, if significant, provide evidence that teacher-experience value of the average school in a class (Elementary and K-8, Middle, or High) had changed concomitantly with the new teacher contract rules.

A second aim of the contract was to reallocate the experienced teachers so that the distribution of teachers within school type would be more equitable; that is, the historical

income, race and achievement-related trends described by the previous analysis would be mitigated. Ideally, the experience level of teachers in the district would rise, and the increase would be due to increased experience in hard to staff schools. But even if the district-wide changes did not occur, assessing whether appropriate redistributions had occurred was important. This could be tested by examining the correlation, within each class of school, of the residual with an appropriate enrollment variable; and by examining the residuals of schools that were assigned staffing-advantages by the contract. The following are the key findings:

Mean Residuals. Table 2 shows the mean 2005 teacher-experience residual for each type of school. It shows that the mean residuals were always positive, and indicated that implementation of the staffing provisions of the contract was followed by increases in the experience of school faculties that ranged from an average of 1.14 years among Elementary and K-8 schools to 0.76 years of experience at high schools—amounts felt to be educationally meaningful. These average residuals were statistically significantly above ‘0’ among the Elementary and K-8 schools ($t_{172}=4.8$, $p<.001$) and among the Middle schools ($t_{38}=2.85$, $p<.007$); and nearly significantly above this no-effect value in the High schools ($t_{39}=1.61$, $p<.114$). This indicated that implementation of the staffing provisions of the contract by the district was followed by desirable changes in the school faculty experience trends that the historical models did not predict, first sign that the district-wide trends for faculties to become less experienced over time may have been coming to an end.

Table 2
Difference between Actual and Predicted Teacher Experience in Fall 2005.

School Type	Teacher Experience Difference (Years)		N of Schools	t	p<
	Mean	S.D			
Elementary & K-8	1.135	3.083	173	4.840	.001
Middle	0.926	2.364	37	2.38	.023
High*	0.755	2.951	40	1.617	.114

*Includes two high schools begun in 2004-2005, for which predictions were made despite their not being in the ‘Pre-Contract’ group.

Residual Distribution. One of the contract goals was to increase the experience of teachers in high poverty, minority and low-achieving schools. Relating residuals to poverty and reading proficiency showed that this did not happen. In Elementary/K-8 schools there were statistically significant trends—for the poverty levels of schools to be negatively correlated with residuals reflecting greater teacher experience ($r_{(172)} = -.26$, $p<.001$) and for the prevalence of proficient readers to be positively correlated with them

($r_{(172)} = .19$, $p < .02$). These relationships were, of course, in the opposite directions from what it was hoped the contract would achieve.

At middle and high schools, the residual distribution correlations were not significant. Among the middle schools they were $r_{(36)} = -.10$, $p = .54$ for the poverty levels, and $r_{(36)} = -.04$, $p = .81$ for the prevalence of proficient readers. Among the high schools they were $r_{(38)} = .11$, $p = .51$ for poverty and $r_{(36)} = .12$, $p = .46$ for reading. These values clearly show that for the middle schools and high schools the historical trends that the new teachers' contract components sought to mitigate had, instead, continued into Fall 2005.

The last contract goal was to use the 'full site selection' and 'incentive' procedures to redistribute experienced teachers. The residual analysis summarized by Table 3 suggested that, except perhaps at middle schools where the results were at best inconsistent, schools involved in these procedures typically ended up with less experienced faculties than the HLM analysis suggested given the student population being served—preliminary evidence that the new contract components was keeping away experienced teachers.

Table 3

The Extent that the Pre-Contract Relationship between Teacher Experience and the Student Population of Schools was Changed in 2005, by Site Selection and Incentive Status.

	No. of Schls.	Percent of Schools Where the Difference Between Expected Actual and Predicted Experience was:			Total
		One or More SD's Above Expected	Within One SD of Expected	One or More SD's Below Expected	
Elementary & K-8					
'Regular'	131	32.8%	64.9%	2.3%	100
Incentive*	15	0.0	66.7	33.3	100
Full Site Selection**	26	7.7	88.5	3.8	100
Both I & FSS	1	0.0	100.0	0.0	100
Middle					
'Regular'	20	40.0	60.0	0.0	100
Incentive*	2	50.0	50.0	0.0	100
Full Site Selection**	11	45.5	36.4	18.2	100
Both I & FSS	6	0.0	83.3	16.7	100
High					
'Regular'	35	31.4	57.1	11.4	100
Full Site Selection	5	0.0	60.0	40.0	100

* Incentive schools that were not also Full Site Selection. ** Full Site Selection schools that were not also Incentive.

Implications: This study has two clear types of implications, one for those interested in using HLM as a policy evaluation strategy, the other for Philadelphia and similar school district managements and teachers' unions. For the HLM audience, this study shows that, fairly simple hierarchical models can yield useful and readily understood findings, especially when the goal is to use historical data to create a standard against which new policies can be evaluated. Using an approach that recognizes that data of schools will never meet traditional design assumptions because they are inevitably organized in hierarchies need not be avoided when dealing with a policy-making audience.

This audience knows that teachers of urban schools are aware of year-to-year changes in the communities that are being served, and are sensitive to publicity about extraordinarily good or poor school outcomes. It also knows that there are general historical trends. The HLM-analyses, which found general relationships among these factors, and then used them to predict that which would have occurred if the teachers' contract had no effect created a level playing field by applying the most recent information possible to recent trends..

From the point of view of Philadelphia and other school district managements, the findings of this study show that strategies that sound good on paper need to be carefully assessed. In this case, the results were mixed. The goal of increasing the experience of faculty was attained, but it did not lead to the redistribution of experienced teachers to low income schools and schools participating in programs for schools where experience was most needed. The policy of identifying schools as 'Full Site Selection' and 'Incentive' may have served, instead of attracting experienced faculty, to help them identify schools that should be avoided.

In conclusion, although this study was more complex than those found in the typical school policy evaluation, HLM yielded findings that were more valid than other approaches because it allowed us to use historical trends to temper current findings, while remaining reasonably transparent to the policy makers.

References:

Chester, M. D., Offenber, R., and Xu, M. D. (2001). *Urban Teacher Transfer: A Four Year Cohort Study of the School District of Philadelphia*. Paper presented at the AERA, Seattle, Washington.

Raudenbush, S., Bryk, A., and Congdon, R. *HLM for Windows, Version 6*. Lincolnwood, IL, Scientific Software International.

Useem, E. Offenberg, R. and Farley, E. (2007). Closing the Teacher Quality Gap in Philadelphia: New Hope and Old Hurdles. Philadelphia, PA, Research for Action.