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

AIR

Investigating the Role of Human Resources in School Turnaround: A Decomposition of Improving Schools in Two States

MICHAEL HANSEN

WORKING PAPER 89 • FEBRUARY 2013

Investigating the Role of Human Resources in School Turnaround: A Decomposition of Improving Schools in Two States

Michael Hansen
American Institutes for Research

Contents

Acknowledgements

The research presented here was originally performed under contract with the Institute of Education Sciences (ED-04-CO-0025/0020). The study team was led by American Institutes for Research, with the Urban Institute, Decision Information Resources, and Policy Studies Associates. Additional analyses and the preparation of this manuscript were not supported by this contract. Tommy Gonzalez provided superior research assistance, and Mike Garet, Jane Hannaway, and Rebecca Herman provided critical guidance on the issues addressed here. I acknowledge the helpful comments of participants at presentations at the University of Virginia, the SREE Conference, and the INVALSI – Improving Education Conference. I also acknowledge the North Carolina Education Research Data Center and the Florida Education Data Warehouse Center for providing access to the data utilized in the study. This work does not necessarily represent the views of any affiliated institutions, and any and all errors are mine. Author contact: mhansen@air.org.

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

Investigating the Role of Human Resources in School Turnaround: A Decomposition of Improving Schools in Two States

Michael Hansen
February 2013

Abstract

Using longitudinal data on spanning the 2002-03 through 2007-08 school years in Florida and North Carolina, this paper decomposes the workforce dynamics among teachers and principals in low-performing schools that significantly improved their performance. In general, I find strong, consistent evidence of human capital development (i.e., improvements in the productivity of the teachers and principals already in the school) accounting for the increased performance in turnaround schools. These findings are robust to the inclusion of school random effects, alternative categorizations of both teachers and turnaround schools, and are observed across elementary and middle school samples in both states. There is also general evidence of productive incoming teachers helping to improve these turnaround schools, but little evidence to support negative attrition specific to these schools played a role. These findings are important as they document large improvements in the joint productivity of teachers in low-performing schools, a finding which is out of step with current federal efforts to improve schools that implicitly assume teacher productivity is essentially fixed over time.

I. Introduction

Human resources—both principals and teachers—are commonly presumed to play a key role in school turnaround. Turning around the nation's lowest performing schools has become a key priority to the U.S. Department of Education in recent years, and this presumption between human resources and school improvement is made in the department's official turnaround strategies to improve these chronically low-performing schools. Specifically, two of the four strategies (the turnaround and transformation models) explicitly require districts to replace the principal and/or teachers in low-performing schools to qualify for federal support under the recent Race to the Top (RTT) and School Improvement Grant (SIG) programs.1 Yet, the evidence documenting the relationship between principal or teacher quality and school turnaround is very weak overall (see Herman, et al., 2008).

This study uses longitudinal administrative data from Florida and North Carolina to investigate how changes in the human workforce correspond to school turnaround. Using data spanning a six-year period from the 2002-03 through the 2007-08 school years, I decompose the value-added productivity of teachers and principals in schools that are identified as chronically low-performing schools to determine which staffing patterns were associated with turnaround. In particular, the primary research question motivating this paper is whether the improvements in performance in these low-performing schools appear to be attributable more to workforce turnover (i.e., removing ineffective teachers and principals with effective ones) or human capital development (i.e., improving the productivity of the current stock of teachers or principals). In summary of the results, I find strong, consistent evidence of human capital development accounting for the increased performance in turnaround schools. These findings are robust to alternative specifications of teacher and principal mobility, the inclusion of school

[1] The other two turnaround strategies require the low-performing traditional school to be closed down (permanently, as in the *closure* model; or, to be reopened as a charter school, as in the *restart* model). Both of these strategies could be considered as implicitly removing the whole school staff.

random effects, and are observed across elementary and middle school samples in both states. There is also modest evidence of productive incoming teachers playing a role in these turnaround schools.

II. Background

Prior Research

Much is unknown about the process of changing school performance, in spite of research on the topic spanning nearly 20 years and large-scale intervention efforts dating from the mid-1980's (Kowel and Hassell, 2005). Prior work on school turnaround comes primarily from two distinct sources: qualitative case studies of successful turnaround schools and organizational turnaround principles developed in business settings applied to education (see reviews in Rhim, et al., 2007; Smarick, 2010). Neither source appears to provide a definitive picture of how to scale and sustain turnaround in failing schools—the case study literature lacks empirical data and comparison schools, while the organizational turnaround literature may not generalize well to public institutions such as schools. The few empirical studies on the topic primarily track schools flagged for low performance, and provide follow-up success rates (e.g., Brady, 2003; Meyers, et al., 2012; Stuit, 2010). Thus, given the generally low rigor of evidence in the area, the Institute of Education Science's turnaround practice guide concludes the research base on how to turn schools around is largely inadequate (Herman, et al., 2008).

The role of human resources in school turnaround is one of these many areas on which there is prior suggestive evidence of a relationship, but no definitive evidence. The turnaround field guide (Herman, et al., 2008) issues two recommendations related to human resources: "signal the need for dramatic change with strong leadership" (p. 10), and "build a committed staff" (p. 27), but clearly states that the level of evidence on these recommendations is low, based on only 10 case studies that were reviewed. No empirical studies investigating human resource practices in school turnaround were available at the time the practice guide was prepared for publication.

A recent study in the turnaround literature warrants particular attention. Dee's (2012) evaluation of turnaround efforts in California using the prescribed turnaround models from the SIG program presents quasi-experimental estimates of being targeted for intervention. Using a fuzzy regression discontinuity design, the author shows targeted schools significantly improved school performance by 0.32 school-level standard deviations on the state's Academic Performance Index, which the author approximates to 0.10 standard deviations of student achievement. This estimate is an overall effect for all schools targeted for intervention, but note that districts have a choice of four turnaround models to implement (one of which closes the school entirely). Interestingly, the schools showing the greatest jump in performance were those that adopted the *turnaround* model, which compels schools to replace at least 50 percent of the school's teaching staff in addition to providing school principals with the flexibility to fully implement a comprehensive approach to improve student outcomes, which may include the adoption of social-emotional student supports, increased learning time, or other school-level interventions. This study provides the best evidence that workforce turnover can dramatically improve school performance—yet, this does not fully isolate the causal effect of workforce turnover alone given the district's selection in determining which model to implement and the addition of complementary school-level interventions.

The focus on human resources in attempting to turnaround low-performing schools has face value, given the education production literature identifying teacher effectiveness as the most significant schooling input into student learning (e.g., Goldhaber, et al., 1999; Hanushek and Rivkin, 2010). An emerging literature on principal effectiveness also signals principals as having a large effect on student learning, second only to teachers (Clark, et al., 2009; Branch, et al., 2009).

One may also view efforts to turnaround low-performing schools by means of workforce turnover as a small-scale analog to the larger policy debate on improving the American public education system in general through workforce turnover. Through this lens, the turnaround policies compelling schools to turnover their workforce are comparable to proposals to "deselect" teachers from the workforce based on low performance (Hanushek, 2009). Whether such policies are efficient depend on the costs of replacing unproductive teachers relative to the costs of improving them. Investigating this issue, Staiger and Rockoff (2010) conclude the opportunity cost of waiting for teachers to improve drawfs the direct cost of replacing teachers, and suggest workforce turnover should be much more dramatic (replacing upwards of 80 percent of teachers based on their initial-year performance) to realize the largest potential gains to workforce productivity. Applying this same workforce turnover approach to turnaround low-performing schools appears straightforward.

In light of the evidence on teacher and principal quality in general and the suggestive evidence of human workforce turnover enabling low-performing schools to improve, the prescriptions for staff turnover in the Department of Education's official turnaround models seems warranted. Yet, this approach to improving schools implicitly rests on an assumption of static teacher and principal quality. In other words, the productivity of school staff is more or less fixed; therefore, a district should replace a low-performing school's unproductive staff with more productive staff to turn it around.

Recent evidence on teacher effectiveness, however, suggests the model of fixed teacher quality does not accurately describe teacher performance over time, which appears to be partially fixed and partially dynamic (Goldhaber and Hansen, forthcoming). In addition, this approach toward school improvement fails to recognize that some element of productivity may be context or peer-specific (e.g., Jackson and Bruegmann, 2009; Jackson, 2010). Thus, teachers' performance certainly changes over time; whether these within-teacher changes can be coordinated enough to improve a low-performing school's overall performance is unclear.

This paper makes a key contribution to this prior literature because it is the first large-scale, empirical study specifically investigating the dynamics of the teacher and principal workforce associated with school turnaround. By investigating the workforce in past turnaround schools, future policy decisions may be informed on reasonable expectations for how to improve school performance. It is also timely, given the current policy interest in scaling turnaround strategies to the lowest five percent of schools across the country. Though important, this study has some key limitations; namely, that it is a descriptive study of the workforce in schools identified as turnaround in retrospect. It is not an evaluation of any specific concerted efforts to improve the schools, aside from what was required under each state's accountability system. Thus, it is cannot be directly generalized to inform of the efficacy of current turnaround policies that rely primarily upon workforce turnover.

Project Context

This study was conducted as part of a larger project investigating potentially successful approaches to turning around chronically low-performing schools, a three-year effort sponsored by the Institute of Education Sciences. This larger project developed a method to retrospectively identify low-performing and turnaround schools in three states (Florida, North Carolina, and Texas).[2] After identifying these schools, three types of data (principal surveys, longitudinal administrative data, and qualitative data from site visits) were collected and analyzed to look for particular practices that may have been associated with turnaround in the past. This study constitutes the evidence from the longitudinal administrative data; findings from the project's other studies will also be discussed as relevant.

The project's identification method for low-performing and turnaround schools warrants some discussion, as this identification is a key explanatory variable in the analysis. Longitudinal data on student achievement from standardized tests were obtained from each state and separated into elementary (grades 3-5) and middle school samples (grades 6-8, with grade 5 as a pre-test score). The

[2] Note that this paper only uses data from Florida and North Carolina while the larger project additionally included data from Texas. Because Texas state data does not link students to teachers, it could not be used in the current investigation, which specifically investigates the productivity of teachers (as measured by value-added gains in their students).

administrative data spanned a six-year period (2002-03 to 2007-08 school years), which was separated into pre- and post-periods of two, three, or four years each. A three-level hierarchical linear model was estimated, providing estimates of a school's performance along both status and growth dimensions in the pre- and post-periods simultaneously. The hierarchical linear model used nested test observations over time within students within schools, including slope estimates for growth for each grade in each school. The status and growth estimates were adjusted to account for different sample size across schools. Full details on the identification method are presented in Hansen and Choi (2012).

The school performance estimates resulting from this model are used to identify low-performing and turnaround schools. Among all schools included in the original sample for the state, those with pre-period performance that fell below the 15th percentile in status and below the 40th percentile in growth were labeled chronically low performing in the given subject. Among the chronically low-performing schools, those with post-period performance that represented at least a five percentile point increase in status and showed growth exceeding the 65th percentile were labeled as turnaround schools (both increases were statistically significant using a one-tailed test with alpha level 0.05). Thus, turnaround schools were those that were low-status and low-growth during the pre-period and improved their status and had high growth in the post period.

III. Two Competing Human Resource Models of Improving School Performance

The improved performance in the turnaround schools identified in the data must come from somewhere, but it is unclear a priori to which teachers and principals they can be attributed. One may conceptualize the improvements as something that is either made (internally sourced with human capital development) or bought (externally sourced through workforce turnover). This section articulates these mechanisms for improvement over time.

Consider the mean performance ($\bar{A}_{st}$) of a particular school (s) over a given time period (t) as the average of the productivity of its individual teachers in the workforce (τ_i):

$$(1)\ \ \bar{A}_{st} = \frac{1}{J_t}\sum_{i=1}^{J_t} \tau_{ist}$$

The school's change in performance over time is therefore:

$$(2)\ \Delta\bar{A}_{st} = \bar{A}_{st+1} - \bar{A}_{st} = \left(\frac{1}{J_{t+1}}\sum_{k=1}^{J_{t+1}} \tau_{kst+1}\right) - \left(\frac{1}{J_t}\sum_{i=1}^{J_t} \tau_{ist}\right)$$

Note that the pool of teachers in each school need not be constant over time, though most teachers are generally retained each year. For clarity, one can separately identify teachers who leave the school (observed in time t but not t+1), from those who enter the school (observed in time t+1 but not t), and those who persist in the school for both periods:

$$(3)\ \Delta\bar{A}_{st} = \left(\frac{1}{IN_{t+1}}\sum \tau_{kst+1} + \frac{1}{STAY_{t+1}}\sum \tau_{kst+1}\right) - \left(\frac{1}{OUT_t}\sum \tau_{jst} + \frac{1}{STAY_t}\sum \tau_{jst}\right)$$

Performance in the second period (the first bracketed expression) is the sum of the productivity of incoming teachers and those who stayed in the school from the prior period; the first period's performance is the sum of the staying teachers' productivity and those that left the school. Separating groups of teachers this way enables me to decompose changes in performance associated with workforce turnover from that associated with improvements in teachers. Rearranging terms from above provides the following equation:

$$(4)\ \ \Delta\bar{A}_{st} = \left(\frac{1}{IN_{t+1}}\sum \tau_{kst+1} - \frac{1}{OUT_t}\sum \tau_{jst}\right) + \left(\frac{1}{STAY_{t+1}}\sum \tau_{ist+1} - \frac{1}{STAY_t}\sum \tau_{ist}\right)$$

The first bracketed expression represents the change in performance associated with workforce turnover (the difference in group productivity between those who left and those who replaced them) and the second expression represents the change observed among the group of teachers observed in both periods. This decomposition of change in school mean performance could also be used to account for productivity differences associated with the school principal, although the presentation above assumes a teacher workforce. Assuming the school's mean performance can represent principal productivity, improvements in the school's performance over time can either be attributed to turnover when the principal is replaced between periods, or to development when the principal is constant but performance improves.

Workforce Turnover

As discussed above, the prescribed turnaround strategies assume a model of fixed teacher productivity. That is, the second bracketed expression in Equation (4) is assumed to be zero. Therefore, improving school performance requires that the teacher workforce churn in such a way as to either remove the lowest performing teachers, fill any vacancies with highly productive teachers, or both. Graphically, this model may look like something akin to that presented in Panel A of Figure 1. This graph depicts the group value-added (on the y-axis) of different teacher groups over the time span of the data (x-axis) for a hypothetical school. The workforce turnover model prescribes selectively removing the worst teachers (hence, the outgoing series is lower than the stable teachers during the same years) and replacing them with high value-added teachers increasing the school's average performance. There is no prescribed change in the productivity of the stable group under this model of school improvement.

In practice, such a strategy may be difficult, as it is at odds with documented evidence on mobility from teacher labor markets. While ineffective teachers in general show a slightly higher likelihood of exiting the public school system as a whole (Goldhaber, et al., 2011), relatively productive teachers have been shown to leave disadvantaged school settings to teach in more affluent schools (Boyd, et al., 2011). In addition, schools' value-added performance is positively associated with its ability to retain productive teachers (Loeb, et al., 2011). Selecting high-productivity teachers to fill vacancies is also problematic, as principals generally have little information about teacher productivity (aside from prior experience) before observing them in the classroom (Staiger and Rockoff, 2010). Further, low-performing, disadvantaged schools are those least likely to have experienced teachers (that is, the most consequential observable characteristic that predicts teacher productivity) in their school's workforce

(Hanushek, et al., 2004). Similar dynamics have also been documented in the principal workforce—more productive or experienced principals are most likely to leave disadvantaged schools for more affluent schools (Branch, et al., 2009; Clark, et al., 2009).

Human Capital Development

Human capital development could feasibly be an alternative model to turning a low performing school around, though such a strategy has not been as prominent in current turnaround efforts. In fact, the Department of Education's transformation model, which has been identifiable primarily from its prescription to replace the principal in targeted schools, also contains some important elements of human capital development including comprehensive instructional reforms and intensive professional development. Under the human capital development strategies, the focus is on improving the entire stock of teachers in the school to make them more productive than they have been previously. Relating this to Equation (4) on school performance change above, the second bracketed expression would need to be positive in the absence of changes in the first expression to turn around a school. Graphically, this may look like Panel B of Figure 1, in which the stable teachers in the school improve considerably in a relatively short period (the graphic assumes this is a large one-time investment occurring at the turnaround point), while there is no prescribed change in the productivity of exiting teachers relative to incoming teachers.

Relying on this strategy to naturally improve schools is again, like the workforce turnover approach above, contrary to documented evidence from the teacher labor market. Teacher performance appears to be generally (though not perfectly) stable over a 10-year time span, and changes in performance within teachers off of the baseline (for better or worse) appear to be transitory and do not last longer than a few years (Goldhaber and Hansen, forthcoming). Prior studies generally find only small improvements in a teacher's performance beyond the first few years of teaching (e.g., Rockoff, 2004), and no gains in human capital associated with the attainment of additional credentials such as an advanced teaching degree or national board certification (Goldhaber and Anthony, 2007). Moreover, prior empirical studies linking professional development to teacher productivity on student test scores are generally mixed and not rigorous enough to determine whether professional development actually has a net positive impact on students (Wayne et al., 2008) and a recent evaluation of a professional development program in middle school mathematics using a randomized control design found no detectable effects associated with the program (Garet et al., 2011). Like teachers, principals also increase in productivity most rapidly in their first few years in the position (Branch, et al., 2009; Clark, et al., 2009), though I know of no prior research that investigates value-added productivity improvements in principals over time.

Which Method is Most Prominent in Past Turnaround Schools?

This study's primary research question is which of these two models accounts for improvements in the turnaround schools identified in Florida and North Carolina. As discussed, both strategies are inconsistent with documented evidence in the labor market. In short, in the absence of targeted efforts specifically countering the natural dynamics of the workforce, there is little reason to expect low-performing schools would be able to either selectively retain their best teachers (principals), select the best teacher (principal) candidates to fill available vacancies, or significantly improve and sustain the productivity of their teachers (principals).

Consequently, there is no predicted hypothesis of which strategy will be most prominent in past turnaround schools. Rather, this is an empirical question that I investigate here. Further, schools could feasibly engage in both strategies simultaneously to improve; these strategies are not necessarily mutually exclusive. Indeed, given districts' interests in raising performance in their lowest-performing schools in this age of school accountability, one may reasonably expect districts to intervene using both strategies.

IV. Data

Longitudinal data on student test scores are utilized from Florida and North Carolina, spanning the 2002-03 through 2007-08 school years. Data from both states include information on student and teacher background characteristics. North Carolina additionally has data on its school principals, which was unavailable in the Florida data for this project, thus only North Carolina is used in the investigation of the principal workforce. In Florida, the test scores utilized are from the Florida Comprehensive Assessment Test – Sunshine State Standards (FCAT-SSS); North Carolina's test scores are from its End-of-Grade (EOG) test. Subject-specific tests in reading and math are available in both states, and raw scale scores are converted to normalized equivalents (z-scores).

The data samples from both states were constructed in a parallel fashion with similar restrictions to maintain comparability. Students in grades 4 through 8 are used (grade 3 scores are used as pretests), and each state is separated into two subsamples—one for elementary schools (grades 3-5) and middle schools (grades 6-8). The original data samples used to identify low-performing and turnaround schools are described in detail in Hansen and Choi (2012).

This study focuses specifically on those schools identified as chronically low performing, thus uses a subsample of the full dataset used in Hansen and Choi (2012). The sample used here is constituted of schools considered chronically low-performing in math, using a three-year pre-period and a three-year post-period. Only student observations that could be linked to their math teachers are retained for the analysis. Subsamples from reading outcomes and/or other cutpoints for the pre- and post-periods were also investigated using the workforce analysis presented here, but produced qualitatively similar results and are omitted for brevity.

Table 1 presents descriptive statistics of the four samples used in this analysis. Given that these samples represent those identified as the lowest performing schools, these variables are consistent with expectations. The schools contained in these datasets are generally disadvantaged—high levels of minority students and those eligible for free or reduced-price lunch. Mean test scores in math are low across all samples. The number of chronically low-performing (CLP) schools in each sample is noted, separately by turnaround (TA) status. Note that TA schools comprise between 10 to 20 percent of the total CLP schools included in the samples, and are relatively few in number. Yet, in spite of the small numbers of schools, note that there are 1,800 student-year observations or more associated with these TA schools over the six years of each sample.

Table 2 presents descriptive statistics of the schools' student bodies and the teacher workforce for each of the samples in the baseline pre-period (the 2002-03 through 2004-05 school years), by turnaround status. T-tests were performed to detect significant differences in means across these groups; those significant at 0.05 alpha level are indicated with asterisks. As shown, the student bodies in TA schools are not statistically distinguishable from those in non-TA CLP schools in all of the samples, with the exception of NC elementary schools where TA schools had significantly lower math test scores and lower re-enrollment rates than non-TA schools in the sample. Note that sample schools were identified as low performing based on performance in the baseline period, not on characteristics of the student bodies. Thus, these differences detected between TA and non-TA schools presumably reflect common baseline traits among schools that succeeded in eventually turning around and are not to be interpreted as causal. Interestingly, the teacher workforces in these schools showed no statistically significant differences in TA and non-TA schools during the baseline period (though the differences in experienced teachers in the FL middle school sample and the differences in teacher retention rates in the FL elementary sample are both marginally significant at the 0.10 alpha level). This evidence suggests the baseline teacher workforces in these schools are similar.

V. Methods

The primary approach used in this investigation decomposes the changes in school performance over time among the groups of outgoing, incoming, and stable teachers. Statistically, the decomposition approach I use here could be considered a modified difference-in-difference-in-difference (DDD) model.3 The three differences come from the following three comparisons in the study: 1) TA schools vs. non-TA CLP schools; 2) the pre- vs. the post-period; and 3) outgoing, incoming, and stable teachers in each school. Indicator variables are generated for each level of differences and fully interacted. The estimating equation takes the following form:

$$(5)\ A_{ijst} = \mathbf{A_{i,t-1}\beta_1} + \mathbf{X_{it}\beta_2} + TA_s\beta_3 + POST_t\beta_4 + TA_sPOST_t\beta_5 + OUT_j\beta_6 + OUT_jTA_s\beta_7 + IN_jPOST_t\beta_8 + IN_jTA_sPOST_t\beta_9 + \varepsilon_{ijst}$$

The dependent variable of current achievement in math (A_{ijst}) for student i taught by teacher j at time t in school s is modeled as a function of prior student achievement ($\mathbf{A_{i,t-1}}$), a vector of student-level explanatory variables ($\mathbf{X_{it}}$), and a set of interacted indicator variables that control for TA status (TA_s), observations in the post-turnaround period ($POST_t$), and outgoing (OUT_j) and incoming (IN_j) teachers.4 The reference category is stable teachers in the pre-period teaching in non-TA schools. This model is essentially a variant on a straightforward teacher value-added model where the teacher fixed effect is simply substituted with the decomposing indicator variables.

Note that the interacted variables in Equation (5) above include neither a standalone intercept for incoming teachers nor an interaction of outgoing teachers with the post period. This is intentional as

[3] Generally speaking, the DDD approach is a quasi-experimental methodology, intended to identify the effects of a particular intervention or change and validate difference-in-difference estimates. My use of a modified DDD model varies from this standard application, and I instead reference my approach as a "decomposition." In my analysis, I have no way to determine the causal relationship between these human resources workflows and turnaround, and I am only using this technique to decompose the workforce patterns most strongly association with turnaround. No claims of causal relationships are made or implied throughout any of the following analysis.

[4] Variables of prior student achievement in both subjects are included in the model (with those missing prior reading test scores being assigned a missing dummy variable). Student explanatory variables included in the model are race, gender, eligibility for free or reduced-price lunch, and designations for limited English proficiency or special education.

incoming teachers are observed in the post-period only and outgoing teachers are observed in the pre-period only. In the baseline specification presented in the following analysis (which I label Specification 1), any teacher observed in the pre-period exclusively (up to and including the 2004-05 school year) is categorized as an outgoing teacher; any teacher observed in the post-period exclusively (the 2005-06 school year or afterwards) is considered an incoming teacher; and teachers observed in both the pre- and post-periods for any number of years in considered a stable teacher. In the additional investigations presented in Section VI, I will also report the results of two alternative categorizations of teacher groupings.

The primary coefficients of interest for this analysis are the estimates for β_5, β_7, and β_9. These variables capture the differential in math student achievement in TA schools that is associated with stable teachers in the post-period, outgoing teachers in the pre-period, and incoming teachers in the post-period, respectively. Because TA schools are identified based on improvements in their post-period performance, relative to their pre-period performance, at least one of the following must be true: $\beta_5 > 0, \beta_7 < 0$, or $\beta_9 > 0$. The estimates of these decomposing coefficients provide evidence of which of the two competing models of school improvement have played a role in these schools. The workforce turnover model implies attrition of the least productive teachers in the pre-period ($\beta_7 < 0$) or selecting relatively productive teachers in the post-period ($\beta_9 > 0$), absent any increase in productivity among stable teachers in the post-period ($\beta_5 = 0$).[5] The human capital development model relies upon increasing the productivity of stable teachers in the post-period ($\beta_5 > 0$), in the absence of any differences in the attrition or selection of the teacher workforce ($\beta_7, \beta_9 = 0$). Lastly, a feasible outcome could be evidence of both workforce turnover and human capital development co-occurring simultaneously in these TA schools.

[5] Strictly speaking, the workforce turnover model needs either negative attrition from the school or positive selection into the school to produce a net improvement in the school's performance, so evidence of only one or the other could be interpreted as evidence of this model playing a role in past turnaround schools. However, turnaround interventions generally combine these efforts in a single strategy to improve the productivity of the workforce.

An important clarification to make is that the model's parameterization as presented here attributes all changes in a school's productivity to its teachers or principal. In other words, any turnaround efforts independent of the workforce will be soaked up by the workforce variables I am using here. If those efforts are school-wide (e.g., additional student supports outside of the classroom, expanding the school day), the productivity effect will presumably influence all teachers' performance. If any efforts are targeted to specific teachers or classrooms (e.g., pull-out instruction for struggling students), individual teachers' performance will be differentially affected according to how these intervention efforts are distributed across incoming, outgoing, and stable teachers' classrooms. Assuming most improvement efforts in these low-performing schools are applied broadly, I therefore consider improvements in stable teachers in TA schools over time to be equivalent to general improvements in overall school performance over time since these are not separable hypotheses in the data available to me.

VI. Results

The decomposition model was estimated on the teacher workforce in the samples of low-performing schools in both Florida and North Carolina, and was also applied to the principal workforce in North Carolina. The primary teacher results are presented first, followed by those from principals. After this, I conduct a series of additional investigations on the teacher data to investigate whether the results are robust and whether these patterns hold for other low-growth schools that turnaround their performance.

Primary Results

Table 3 presents the estimated coefficients on the decomposition parameters in the model. The results from Florida span columns 1-4; North Carolina results span columns 5-8. Each model was estimated with and without the inclusion of random school effects to control for the correlation of outcomes within a school (random effects models are in even-numbered columns).

The estimated coefficients in Table 3 suggest the most prominent model in these TA schools were improvements in the productivity of their stable teacher workforce—across all models, the estimated coefficients on TA schools in the post-period (i.e., the β_5 coefficient from above) are statistically significant and positive. The results are consistent across models both with and without school random effects. All show educationally significant magnitudes of improvement of 0.10 or more student standard deviations. The literature on teacher value-added generally equates one standard deviation of teacher productivity to between 0.11-0.36 student standard deviations in math (Hanushek and Rivkin, 2010). This implies an overall improvement in teacher productivity in a school of at least a third of a standard deviation of teacher productivity, and likely more.

There is mixed evidence of the workforce turnover model occurring in these TA schools. The first component of the turnover model is that of removing the lowest performing teachers from the school. In the Florida samples, negative attrition is observed among all outgoing teachers (this is significant in the middle school sample), but no statistically significant differential is associated with outgoing teachers in the TA schools (β_7). This suggests removing the lowest-performing teachers may not be a unique feature of TA schools only. The results from North Carolina, however, do not show any apparent negative attrition either among outgoing teachers generally or in TA schools specifically.

The second component of workforce turnover is an improvement associated with incoming teachers, and there is modest evidence of this positive selection playing a role. First, note that none of the coefficients on incoming teachers in TA schools in the post-period (β_9) are significantly positive (in fact, these coefficients are significantly negative in the NC elementary sample). Yet, this does not imply that there is no improvement realized through workforce turnover. First, note that the lack of statistical significance is partly a function of low power on these estimates (the outgoing teacher groups in TA schools are small). Second, notice these β_9 coefficients represent a two-way differential relative to the group of stable TA teachers in the post period (which is statistically positive and large) and relative to the group of incoming teachers in non-TA schools (consistently negative and smaller in magnitude). The combined effect implies that there was a net gain associated with the incoming teachers, though they were not as high performing as the stable teachers in the post-period. Directly comparing the incoming teachers in TA schools against those in non-TA schools during the same period (i.e., testing the significance of the sum $\beta_3 + \beta_5 + \beta_9$) shows that the incoming teachers to TA schools appear to have been generally more productive teachers than what other low-performing schools expected in the post-period. What is not clear from the data is whether these new teachers were more productive than other incoming teachers in non-TA schools at the point of hire, or whether they were no different at hire but improved because of professional development or other school-wide improvement effort; these hypotheses cannot be distinguished with the administrative data.

Figure 2 presents a graphical depiction of the changes in group value-added over time for the TA schools from the NC elementary (Panel A) and NC middle school samples (Panel B). These figures plot the same relationship as that depicted in Figure 1, using actual performance data for the teacher groups (comparable to the estimates in columns 5 and 7) rather than hypotheticals. These graphics show an overall improvement in the performance of the stable teachers in the post-period (2006-08 testing) compared to the pre-period (2003-05), which is larger in magnitude for the elementary schools than the middle schools. In both samples, there appears to be a net gain associated with the incoming teachers replacing the outgoing teachers, but the incoming teachers' average performance during the post-period is slightly lower than the stable teachers' during those years. Also, no clear attrition pattern among outgoing teachers is observed in Panel A, whereas Panel B shows some evidence of negative attrition among outgoing teachers.

Table 4 presents the coefficient estimates of the decomposing variables on the North Carolina samples using the school principals' mobility grouping as the determinant for the coding (recall principal data was not available in the Florida data for this analysis). As above, results with and without the inclusion of school random effects are reported. Note that these sample sizes have increased from those reported in Table 3, this is because Table 3 was limited to observed student-teacher matches while Table 4 does not require such a match.

The principal results are generally in line with what was seen in the teacher workforce—strong support for the human capital development model (i.e., prior-serving principals improving between the two periods) as seen in the statistically significant and positive estimates on the TA schools in the post-period. The estimates on outgoing principals in TA schools all fail to reject the null hypothesis, suggesting no negative attrition among exceptionally low-performing principals. Also, no evidence of highly productive principals coming to elementary schools to turn them around, but the group of incoming principals in middle TA schools appears to be relatively stronger (though this difference is marginally significant).

Taken together, both the teacher and principal results indicate the large improvements in performance in these turnaround schools appear to be primarily attributable to the human capital development model, or in other words, gains associated with the long-time staff in the school. This primary finding was consistent across the elementary and middle school samples in both states, as well as among the teacher workforce and principals. There is also some evidence of net gains associated with incoming teachers (principals) replacing outgoing teachers (principals), though the gains appear to be generally due to relatively strong incoming staff rather than removing the worst staff.

Additional Investigations

The primary results presented above show large, significantly positive estimates associated with the stable teachers and principals in schools. The magnitudes of these estimates are surprising and warrant further investigation to understand whether they are idiosyncratic or more generalizable. Specifically, I will first investigate whether the estimates are sensitive to the way teachers are grouped into outgoing, incoming, and stable categories. Second, I will investigate whether the improvements in teacher performance is accompanied by changes in the observable characteristics of the teacher workforce in these schools. Third, given the importance of the TA indicator variable as an explanatory variable in the model, which is itself estimated, I investigate whether the results are consistent for variations in the identification or grouping of turnaround schools. And finally, I investigate whether the patterns of improvement observed in these exceptionally low-performing schools are generalizable to other schools that show low growth. Note that due to small sample sizes among principals (and its availability in NC only), the following investigations were only conducted using the teacher personnel data.

First, the estimated coefficients showing large performance gains among stable teachers may be sensitive to how I identify 'stable' teachers, and I use two alternate categorizations to check for robustness. The first allows the definitions of incoming and outgoing teachers to cross over the turnaround cutpoint by one year to allow flexibility in the identification of the three teacher groups as workforce dynamics may have required more than one transition year in TA schools.6 The effect of this specification is that the stable group becomes slightly smaller, since more teachers are associated with the incoming and outgoing groups due to mobility. The second alternate categorization identifies teacher groups by their status specifically teaching a tested grade and subject within the school, rather than simply staying assigned to the school. This specification will help distinguish actual changes in the

[6] This specification additionally classifies teachers who leave in the year immediately following the turnaround cutpoint (i.e., exiting the school at the end of the 2005-06 school year) as outgoing teachers as long as they taught at least one year during the pre-period and classifies teachers entering the school in the year immediately before the turnaround cutpoint (i.e., entering the school at the beginning of the 2004-05 school year) and continuing to teach for at least two school years as incoming teachers.

performance of these different teacher groups from ‘staffing to the test’ that may occur strategically to help buoy test scores in low-performing schools (Cohen-Vogel, 2011; Fuller and Ladd, 2012).

Table 5 presents the results of the models using the alternate categorizations: the first categorization allowing for a flexible school transition period is used in columns 1-4, and the categorization specific to tested grades and subjects is presented in columns 5-8. All of the results presented here use school random effects. Aside from this alternate categorization of teachers in the decomposition indicator variables, the data sample remains the same as in Table 3.

The estimated coefficients in Table 5 are generally robust to these two alternate categorizations. In both cases, the coefficients again show strong support for a large, statistically significant improvement in stable teachers in TA schools in the post-period (relative to their pre-period performance), and are in the same range as the corresponding estimates on Table 3. Negative attrition is tendency among all low-performing schools in the samples in both categorizations (NC elementary schools excepted), though this difference is only statistically significant in the FL middle school sample. No statistically significant negative attrition is associated specifically with TA schools for either specification. As for TA schools attracting highly productive teachers in the post-period, the FL elementary school sample shows evidence of notably strong performance among incoming TA teachers when using the flexible transition specification; the remaining samples and the other specification show results that are similar to what was seen in Table 3 (net improvements associated with the incoming teachers, but performance that lagged behind the stable teachers in the post-period). Overall, these results are largely in line with those presented in Table 3. In other words, these gains in stable teachers’ performance do not appear to have been mistakenly attributed to teachers that were actually transitioning in or out of the school within a year of the turnaround point, nor do the gains appear to be the result of strategically shifting teachers in and out of tested classrooms within a school.

The second investigation asks whether there are any observable changes in the characteristics of the teacher workforce that may explain the performance changes that are observed in these TA schools. For this investigation, I include all teachers observed teaching in the schools in the sample and the level of observations is a teacher-year. I run the same decomposition estimating equation as above, but instead substitute binary teacher characteristics (whether a teacher has 4 or more years of experience, or holds a master's degree or higher) in lieu of the test scores as the dependent variable and estimate a logit model. These results are presented in Table 6. The estimated coefficients are reported as odds ratios, where values of 1 indicate no change in likelihood, and values less than 1 (greater than 1) indicate a negative (positive) association between the explanatory variable and the binary outcome.

Of the two observable characteristics included here, I am particularly interested in whether the TA schools differ in their attraction and retention of experienced teachers, as this is the one observable characteristic that most consistently predicts value-added productivity. Specifically, if TA schools tend to retain their teachers as they gain experience relative to non-TA schools, this may partially explain the gains in performance among the stable teachers in TA schools. The results on the observable experience levels in columns 1-4 do not provide any evidence that this is the case—the odds ratio from the FL middle schools sample is significantly less than 1, running counter to this hypothesis, and those from the other samples are not differentiable from 1. Nor is there evidence of compositional changes in experience associated with turnaround. Outgoing teachers in these CLP schools are generally less likely to be experienced across all samples, but no statistically significant association is noted with outgoing teachers in TA schools specifically. And the patterns of experience differentials associated with incoming teachers to TA schools are mixed (this difference was statistically positive and significant in NC elementary schools, but reversed in Florida schools though not statistically significant).

Master's degrees are not generally associated with value-added productivity, though may indicate a more credentialed workforce with higher levels of investment into the profession among teachers. The evidence here does not support a systematic relationship between the teacher workforce's observed level of education and turnaround—both positive and negative associations are observed on all of the variables associated with TA schools. Overall, I find little systemic evidence of changes in observable teacher attributes associated with turnaround.7

The third investigation queries whether the results are sensitive to the subset of schools considered TA in the analysis. As described in Section II, the TA indicator is applied to CLP schools that significantly improved both their estimated status and growth measures in the post-period relative to the pre-period. But since this indicator is an estimate of performance changes only, some of the schools may only marginally be considered TA or miss the criteria to be considered TA, and commonalities among these similar schools around the threshold to be identified as TA may mask important changes in TA schools.

Table 7 addresses this issue in two ways. First, columns 1-4 drop schools that were considered moderately improving (MI) in the original Hansen and Choi (2012) identification; schools in this category improved in either status or growth, but not both. The remaining schools are those considered not improving (NI) and showed no significant improvement overall. By removing the schools with modest improvements, the coefficients are expected to show a sharper contrast in the important workforce dynamics associated with turnaround. Second, columns 5-8 alternately classify schools that met the growth turnaround criterion only as TA schools (rather than requiring improvements in status and growth). This more liberal categorization of turnaround will increase the number of TA schools. In both cases, the results are qualitatively consistent with those observed in Table 3.

Finally, I wish to investigate whether the patterns of improvement observed in the CLP school sample generalize to other schools as well, or whether these dynamics are particular to these

[7] In NC, I have access to data on other teacher characteristics, which I also tested for associations with turnaround, including national board certification, licensure status, and teacher absences. None of these variables showed any compelling evidence of a systematic association with school turnaround among various teacher groups.

exceptionally low-performing, high-need schools (pre-period status measures at or below the 15th percentile). Table 8 presents the key coefficients of interest for three different samples: Panel A limits the sample to low-growth schools (less than the 40th percentile in growth) falling in the 16th to 30th percentiles during the pre-period; Panel B limits to low-growth schools falling in the 31st to 50th percentiles; and Panel C limits to low-growth schools falling in the 51st through 70th percentiles. As in the main results, I still identify schools as TA if they significantly improve their status and growth in the post-period, applying the same identification rules.

Interestingly, the results of Table 8 present some substantive departures from the patterns observed in prior tables. First, the large gains in the performance of stable teachers in TA schools during the post-period are uniformly significant in elementary schools across all these samples, but are markedly smaller and/or not statistically significant in the case of middle schools. Thus, middle schools may be more context-dependent in determining which school improvement models are most likely to succeed. And second, there are consistent gains associated with the incoming teachers in the TA schools in Panels B and C, suggesting the introduction of effective teachers may become a relatively more important strategy for these schools nearer the average. Finally, note there is still no consistent evidence of negative attrition associated specifically with TA schools (estimates here are both positive and negative in different samples).

VII. Conclusion

This paper investigates the workforce dynamics among teachers and principals in low-performing schools that significantly improved their performance to understand how the workforce improved so dramatically over a short period of time. In general, I find strong, consistent evidence of human capital development (i.e., improvements in the productivity of the teachers and principals already in the school) accounting for the increased performance in TA schools. These findings are robust to the inclusion of school random effects, alternative categorizations of both teachers and turnaround schools, and are observed across elementary and middle school samples in both states. The human capital development model is also well supported using the principal workforce data in North Carolina.

The workforce turnover model also shows some modest evidence of playing a role in these turnaround schools. Specifically, the incoming teachers to TA schools in the post-period were statistically significantly more productive than those they replaced and other teachers new to non-TA schools in the same period; however, they in general did not outperform the stable teachers in TA schools during the post-period. Also, the results showed some evidence of a general trend of negative attrition among all of these low-performing schools, but there was little systematic evidence of TA schools being particularly effective at removing their lowest performers.

These findings are important as they document large improvements in the joint productivity of teachers in low-performing schools, a finding which is out of step with current federal efforts to improve schools that implicitly assume teacher productivity is essentially fixed over time. This finding invites follow-up questions of how all of these teachers jointly improved and whether these improvements were sustained long term (beyond the three-year post-period imposed by the model). It also invites further investigation into the cost effectiveness of improving existing teachers' productivity vis-à-vis utilizing workforce turnover to improve overall quality—even though this strategy may be most common in the experience of past TA schools it is not necessarily the strategy most likely to be successful or most cost effective long term.

I wish to note some limitations on these findings. First, because this is a retrospective investigation, I make no claim as to the causal nature of human capital development or workforce turnover as strategies in promoting school turnaround. A second and related caveat is that this analysis does not evaluate a specific turnaround strategy and as such cannot directly predict the efficacy of current Department of Education strategies in turning around these schools. And third, as noted in the

Methods section, I interpret the human capital development strategy here to be observationally equivalent to school-wide reform efforts that are independent of the workforce. Thus, this paper's primary contribution is not in advocating for greater professional development but instead in documenting that workforce turnover strategies played a considerably smaller role (particularly negative attrition) in the experience of these TA schools.

The larger project's related studies have produced some data that is helpful in distinguishing these various sources of improvements. Herman and Huberman (2012) report the results of a principal survey administered to 750 of the schools identified as chronically low performing across the three states included in the project (FL, NC, and TX). On the workforce turnover model, the few items related specifically to human resource management are generally in line with what I find in this analysis. Namely, they find no statistically significant difference between TA and non-TA CLP schools reporting whether "staff members uncommitted to change [left] the school," similar to the no-difference finding of negative attrition between outgoing teachers in TA and non-TA CLP schools here. Also, they report TA elementary schools were significantly more likely to report having "experienced, highly qualified teachers join the school," which is qualitatively similar to the significant difference I find in the productivity of incoming teachers in TA schools during the post-period relative to incoming teachers in non-TA schools during the same period and relative to the outgoing teachers they replaced. Contrary to survey respondents' claims, however, Table 6 simply finds mixed evidence that TA schools were more likely to attract experienced teachers.

Additionally, Arcaira and Turnbull (2012) analyze qualitative data gathered from 281 interviews with teachers and administrators in 36 of the CLP schools (24 TA schools, 12 non-improving schools) identified in the project. They find 58 percent of TA schools reported strategic staffing initiatives (which fall under the workforce turnover model described in this paper) compared to 33 percent of non-TA schools. Interestingly, they report both higher levels of implementation and larger differences between

TA and non-TA schools in their “strong leadership,” “intensive professional development,” “orderly academic climate,” and “collective teacher commitment” domains—all of which would fall under the more broad human capital development strategy presented here. Their evidence is consistent with these findings of the administrative data that suggest the human capital development strategy appears to have played a relatively larger role in these schools that have successfully turned around in the past.

In conclusion, I find evidence consistent with the model of human capital development—investing in current teachers to help them improve—playing the primary role in these past TA schools, with complementary evidence of relatively high-performing staff entering the school. There is little evidence here to support the hypothesis that the removal of low-performing staff was a systemic effort to turnaround these schools. These findings invite further research into policies and practices that can help promote productivity gains in low-performing schools, and particularly invite research into which policies would be more cost effective in turnaround. Ultimately, the optimal strategy would be one that could scale turnaround in low-performing schools and sustain improvements in the long term. According to Dee (2012), workforce strategies like those articulated in the RTT turnaround model promise to be considerably more cost effective than class-size reduction strategies at improving student performance. Yet, precious little is known about the expected success and cost effectiveness of whole-school improvement efforts in these low-performing, high-need schools.

References

Arcaira, E. & Turnbull, B. (2012). *Implementation of Turnaround Strategies in Chronically Low-Performing Schools.* Paper Presented at the 2012 Fall Research Conference of the Society for Research on Educational Effectiveness.

Boyd, D., Grossman, P., Lankford, H., Loeb, S., & Wyckoff, J. (2009). *Who leaves? Teacher attrition and student achievement.* CALDER Working Paper #23. Retrieved from http://www.caldercenter.org

Brady, R. C. (2003). *Can Failing Schools Be Fixed?* Washington, DC: Thomas B. Fordham Foundation. Retrieved from http://www.edexcellence.net

Branch, G., Hanushek, E., & Rivkin, S. (2009). *Estimating Principal Effectiveness*. CALDER Working Paper #32. Retrieved from http://www.caldercenter.org

Clark, D., Martorell, P., and Rockoff, J. E. (2009). *School principals and school performance.* CALDER Working Paper #38. Retrieved from http://www.caldercenter.org

Cohen-Vogel, L. (2011). Staffing to the Test Are Today's School Personnel Practices Evidence Based?. *Educational Evaluation and Policy Analysis*, 33(4), 483-505.

Dee, T. (2012). *School Turnarounds: Evidence from the 2009 Stimulus*. NBER Working Paper 17990.

Fuller, Sarah C., and Helen F. Ladd. (2012). *School Based Accountability and the Distribution of Teacher Quality Among Grades in Elementary Schools*. CALDER Working Paper #75. Retrieved from http://www.caldercenter.org

Garet, Michael S., Andrew J. Wayne, Fran Stancavage, James Taylor, Marian Eaton, Kirk Walters, Mengli Song, Seth Brown, Steven Hurlburt, Pei Zhu, Susan Sepanik, and Fred Doolittle. 2011. *Middle School Mathematics Professional Development Impact Study: Findings After the Second Year of Implementation.* NCEE 2011-4024. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Goldhaber, D., & Anthony, E. (2007). Can teacher quality be effectively assessed? National board certification as a signal of effective teaching. *The Review of Economics and Statistics*, *89*(1), 134-150.

Goldhaber, D. D., Brewer, D. J., and Anderson, D. J. (1999). A Three-way Error Components Analysis of Educational Productivity. *Education Economics* 7(3).

Goldhaber, D., Gross, B., & Player, D. (2011). Teacher career paths, teacher quality, and persistence in the classroom: Are public schools keeping their best?. *Journal of Policy Analysis and Management*, *30*(1), 57-87.

Goldhaber, D., & Hansen, M. (forthcoming). Is it Just a Bad Class? Assessing the Long-term Stability of Estimated Teacher Performance. Forthcoming in *Economica.*

Hansen, M., & Choi, K. (2012*). Chronically Low-Performing Schools and Turnaround: Findings in Three States.* CALDER Working Paper #60. Retrieved from http://www.caldercenter.org

Hanushek, Eric A. (2009). "Teacher Deselection." In *Creating a New Teaching Profession*, edited by Dan Goldhaber and Jane Hannaway (165-80). Washington, D.C.: The Urban Institute Press.

Hanushek, E.A., Kain, J.F., & Rivkin, S.G. (2004). Why public schools lose teachers. *Journal of Human Resources*, *39*(2), 326-354.

Hanushek, E. A., Rivkin, S. G., Figlio, D., & Jacob, B. (2010). Generalizations about using value-added measures of teacher quality. *American Economic Review*, 100(2), 267-271.

Herman, R., Dawson, P., Dee, T., Greene, J., Maynard, R., Redding, S., & Darwin, M. (2008). *Turning Around Chronically Low-Performing Schools: A practice guide* (NCEE #2008-4020). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/publications/practiceguides

Herman, R. & Huberman, M. (2012). *Differences in the Policies, Programs, and Practices (PPPs) and Combination of PPPs across Turnaround, Moderately Improving, and Not Improving Schools.* Paper Presented at the 2012 Fall Research Conference of the Society for Research on Educational Effectiveness.

Jackson, C. Kirabo. (2010). Do Students Benefit from Attending Better Schools? Evidence from Rule-based Student Assignments in Trinidad and Tobago. *The Economic Journal*, *120*(549), 1399-1429.

Jackson, C. Kirabo, and Elias Bruegmann. (2009). Teaching Students and Teaching Each Other: The Importance of Peer Learning for Teachers. *American Economic Journal: Applied Economics*, 1:85-108.

Kowal, J.M., and Hassel, E.A. (2005).*Turnarounds with new leaders and staff.* Washington, DC: The Center for Comprehensive School Reform and Improvement.

Loeb, S., Kalogrides, D., and Beteille, T. (2011). *Effective schools: Teacher hiring, assignment, development, and retention.* NBER Working Paper 17177.

Meyers, C., Lindsay, J., Condon, C., & Wan, Y. (2012). A Statistical Approach to Identifying Schools Demonstrating Substantial Improvement in Student Learning. *Journal of Education for Students Placed at Risk (JESPAR)*, *17*(1-2), 70-91.

Rockoff, Jonah E. (2004). The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data. American Economic Review, 94(2):247-252.

Rhim, L.M., Kowal, J.M., Hassel, B.C., and Hassel, E.A. (2007). School turnarounds: a review of the cross-sector evidence on dramatic organizational improvement. Chapel Hill, NC: Public Impact, and Lincoln, IL: Academic Development Institute.

Smarick, A. (2010). The Turnaround Fallacy. *Education Next, 10 (1)*, 21-26. Retrieved from http://educationnext.org

Staiger, D.O and Rockoff, J.E. (2010). Searching for effective teachers with imperfect information. *Journal of Economic Perspectives, 24(3)*, 97-118.

Stuit, D. (2010). *Are Bad Schools Immortal? The Scarcity of Turnarounds and Shutdowns in Both Charter and District Sectors*. Washington, DC: Thomas B. Fordham Foundation. Retrieved from http://www.edexcellence.net

Wayne, A. J., Yoon, K. S., Zhu, P., Cronen, S., & Garet, M. S. (2008). Experimenting with teacher professional development: Motives and methods. *Educational Researcher*, *37*(8), 469-479.

Tables and Figures

Table 1. Descriptive Sample Statistics				
State	Florida		North Carolina	
School sample	Elementary	Middle	Elementary	Middle
Proportion of female students	51.1%	52.8%	49.1%	50.0%
Proportion of African American students	52.6%	40.4%	55.5%	62.3%
Proportion of Hispanic students	21.6%	31.6%	12.9%	9.2%
Proportion of students with limited English proficiency	6.6%	4.2%	7.7%	4.5%
Proportion of students ever eligible for free or reduced-price lunch program	89.8%	83.2%	74.4%	69.5%
Mean Student Achievement in Math	-0.37	-0.11	-0.46	-0.33
Unique CLP, Non-TA Schools	87	22	66	37
Total student-year observations in CLP, Non-TA schools	34,514	13,592	34,003	21,667
Unique CLP, TA Schools	17	3	8	5
Total student-year observations in CLP, TA schools	9,039	1,806	3,368	2,838
Total student-year observations	43,553	15,398	37,371	24,505

Note: Samples limited to student-teacher linked observations in math in chronically low-performing schools identified in Hansen and Choi (2012) using the 2005 turnaround point in math.

Table 2. Pre-period Comparison of TA and Non-TA Schools in Samples

State	Florida				North Carolina			
School sample	Elementary		Middle		Elementary		Middle	
School type	Non-TA	TA	Non-TA	TA	Non-TA	TA	Non-TA	TA
Student Body Characteristics								
Proportion of female students	50.5%	51.5%	55.3%	49.6%	49.3%	50.2%	49.8%	52.2%
Proportion of African American students	56.0%	54.2%	44.9%	39.4%	59.9%	62.1%	62.0%	69.2%
Proportion of Hispanic students	15.6%	22.5%	22.9%	21.4%	8.6%	10.0%	6.4%	5.4%
Proportion of students with limited English proficiency	5.8%	6.9%	4.8%	1.7%	4.4%	4.4%	2.5%	2.8%
Proportion of students ever eligible for free or reduced-price lunch program	88.8%	89.3%	81.1%	79.7%	77.4%	79.6%	70.0%	76.0%
Mean student achievement in math (standardized)	-0.418	-0.391	-0.175	-0.106	-0.437**	-0.577**	-0.314	-0.458
Mean re-enrollment rate (2003 and 2004 only)	66.3%	63.8%	78.5%	83.1%	80.0%**	70.4%**	80.7%	79.5%
Mean school enrollment (includes unlinked)	97	102	197	199	88	68	108	108
Teacher Characteristics								
Proportion of females	84.1%	81.4%	62.3%	49.4%	88.6%	94.4%	74.9%	81.5%
Proportion of black or Hispanic teachers	40.0%	47.2%	46.6%	53.9%	35.8%	35.0%	42.9%	46.7%
Proportion of teachers with 4 or more years of experience	59.5%	61.5%	57.1%	86.7%	77.7%	80.2%	76.4%	64.5%
Proportions of teachers fully licensed					90.3%	89.4%	82.0%	74.0%
Mean proportion of teachers returning in following year (2003 and 2004 only)	68.7%	77.0%	67.0%	71.1%	79.3%	73.5%	77.9%	78.2%
Total Schools	87	17	22	3	66	8	37	5

Note: *, ** represents differences in group means significant with p<0.05, p<0.01. Teacher licensure data was not available in FL data used for the study. Samples limited to student-teacher linked observations in math in chronically low-performing schools identified in Hansen and Choi (2012) using the 2005 turnaround point in math.

Table 3. Decomposition Estimates of School Improvement among Teachers

State	Florida				North Carolina			
School sample	Elementary		Middle		Elementary		Middle	
School Random Effects	No	Yes	No	Yes	No	Yes	No	Yes
TA (β_3)	-0.009	0.012	-0.069**	-0.078	-0.086**	-0.047	-0.014	0.008
	(0.014)	(0.034)	(0.019)	(0.046)	(0.020)	(0.030)	(0.020)	(0.034)
Post (β_4)	0.106**	0.103**	0.038**	0.033**	0.005	0.014	-0.024**	-0.020*
	(0.009)	(0.009)	(0.010)	(0.011)	(0.007)	(0.008)	(0.009)	(0.009)
TA*Post (β_5)	0.137**	0.139**	0.130**	0.153**	0.215**	0.187**	0.098**	0.092**
	(0.018)	(0.019)	(0.044)	(0.045)	(0.025)	(0.026)	(0.028)	(0.028)
Outgoing (β_6)	-0.009	-0.017	-0.036**	-0.048**	0.012	0.010	-0.005	0.003
	(0.010)	(0.010)	(0.011)	(0.012)	(0.009)	(0.009)	(0.010)	(0.010)
Outgoing*TA (β_7)	0.011	0.027	0.005	0.031	0.051	0.019	-0.016	-0.026
	(0.024)	(0.025)	(0.031)	(0.034)	(0.031)	(0.032)	(0.030)	(0.031)
Incoming*Post (β_8)	-0.039**	-0.049**	-0.052**	-0.053**	-0.025**	-0.043**	-0.049**	-0.056**
	(0.008)	(0.009)	(0.013)	(0.013)	(0.008)	(0.009)	(0.010)	(0.010)
Incoming*TA*Post (β_9)	0.004	0.022	-0.016	-0.010	-0.058*	-0.063*	0.007	0.014
	(0.018)	(0.018)	(0.047)	(0.050)	(0.027)	(0.028)	(0.028)	(0.029)
Observations	43,553	43,553	15,398	15,398	37,371	37,371	24,505	24,505
R-squared	0.575	0.575	0.623	0.623	0.644	0.644	0.681	0.681

Note: *, ** represents p<0.05, p<0.01. All regressions are estimated on student-year-level observations on math standardized test scores, and include the following covariates: prior math and reading scores, an indicator variable for missing reading scores, as well as indicators for race, gender, eligibility for free or reduced-price lunch, and designations for limited English proficiency or special education. Samples limited to student-teacher linked observations in math in chronically low-performing schools identified in Hansen and Choi (2012) using the 2005 turnaround point in math.

Table 4. Decomposition Estimates of School Improvement among Prinicipals

School sample	NC Elementary		NC Middle	
School Fixed Effects	No	Yes	No	Yes
TA (β_3)	-0.049**	-0.029	-0.016	0.012
	(0.017)	(0.026)	(0.015)	(0.042)
Post (β_4)	-0.008	-0.007	-0.047**	-0.037**
	(0.008)	(0.008)	(0.008)	(0.008)
TA*Post (β_5)	0.158**	0.154**	0.074**	0.065**
	(0.023)	(0.023)	(0.024)	(0.023)
Outgoing (β_6)	0.017*	0.015	-0.033**	-0.018*
	(0.008)	(0.009)	(0.007)	(0.008)
Outgoing*TA (β_7)	0.028	0.013	0.016	-0.000
	(0.027)	(0.028)	(0.022)	(0.022)
Incoming*Post (β_8)	0.013	0.008	-0.017	-0.016
	(0.008)	(0.008)	(0.009)	(0.009)
Incoming*TA*Post (β_9)	-0.003	-0.005	0.039	0.034
	(0.025)	(0.025)	(0.026)	(0.026)
Observations	39,394	39,394	37,353	37,353
R-squared	0.640	0.640	0.682	0.682

Note: *, ** represents $p<0.05$, $p<0.01$. All regressions are estimated on student-year-level observations on math standardized test scores, and include the following covariates: prior math and reading scores, an indicator variable for missing reading scores, as well as indicators for race, gender, eligibility for free or reduced-price lunch, and designations for limited English proficiency or special education. Samples limited to all valid student observations in math in chronically low-performing schools identified in Hansen and Choi (2012) using the 2005 turnaround point in math.

Table 5. Decomposition Estimates among Teachers Using Alternate Teacher Categorizations

Specification	Flexibility in school transition period				Specific to tested grades and subjects			
State	Florida		North Carolina		Florida		North Carolina	
School sample	Elementary	Middle	Elementary	Middle	Elementary	Middle	Elementary	Middle
School Random Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
TA (β_3)	0.002	-0.065	-0.061*	0.005	0.014	-0.066	-0.051	0.010
	(0.034)	(0.037)	(0.030)	(0.037)	(0.033)	(0.044)	(0.031)	(0.035)
Post (β_4)	0.116**	0.039**	0.021**	-0.024*	0.102**	0.026*	0.019*	-0.024*
	(0.009)	(0.012)	(0.008)	(0.010)	(0.009)	(0.011)	(0.008)	(0.010)
TA*Post (β_5)	0.106**	0.135**	0.203**	0.101**	0.135**	0.149**	0.191**	0.100**
	(0.020)	(0.045)	(0.028)	(0.030)	(0.019)	(0.045)	(0.027)	(0.029)
Outgoing (β_6)	-0.008	-0.033**	0.014	-0.007	-0.019	-0.049**	0.015	-0.002
	(0.009)	(0.011)	(0.008)	(0.010)	(0.010)	(0.012)	(0.009)	(0.010)
Outgoing*TA (β_7)	0.041*	-0.011	0.044	-0.008	0.018	0.007	0.025	-0.030
	(0.021)	(0.033)	(0.029)	(0.028)	(0.024)	(0.033)	(0.032)	(0.031)
Incoming*Post (β_8)	-0.053**	-0.038**	-0.043**	-0.039**	-0.048**	-0.032*	-0.046**	-0.050**
	(0.009)	(0.013)	(0.008)	(0.010)	(0.009)	(0.013)	(0.008)	(0.010)
Incoming*TA*Post (β_9)	0.083**	-0.031	-0.024	0.001	0.025	-0.036	-0.070*	-0.011
	(0.019)	(0.050)	(0.027)	(0.030)	(0.018)	(0.050)	(0.028)	(0.029)
Observations	43,553	15,398	37,371	24,505	43,553	15,398	37,371	24,505
R-squared	0.575	0.622	0.644	0.681	0.575	0.623	0.644	0.681

Note: *, ** represents $p<0.05$, $p<0.01$. All regressions are estimated on student-year-level observations on math standardized test scores, and include the following covariates: prior math and reading scores, an indicator variable for missing reading scores, as well as indicators for race, gender, eligibility for free or reduced-price lunch, and designations for limited English proficiency or special education. Samples limited to student-teacher linked observations in math in chronically low-performing schools identified in Hansen and Choi (2012) using the 2005 turnaround point in math. See text for description of teacher coding under these alternate categorizations.

Table 6. Decomposition Estimates among Teachers Using Alternate Identification of TA Schools

Dependent Variable	4+ years of experience				Teacher holds master's degree			
State	Florida		North Carolina		Florida		North Carolina	
School sample	Elementary	Middle	Elementary	Middle	Elementary	Middle	Elementary	Middle
School Random Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
TA (β_3)	1.127	1.494	0.736	1.020	1.316	0.495	1.048	0.728
	(0.173)	(0.730)	(0.215)	(0.370)	(0.245)	(0.261)	(0.265)	(0.130)
Post (β_4)	1.380**	1.906**	1.822**	1.787**	1.002	0.973	1.265**	1.194**
	(0.071)	(0.220)	(0.109)	(0.142)	(0.031)	(0.063)	(0.053)	(0.058)
TA*Post (β_5)	1.184	0.660*	0.738	1.060	0.968	0.965	0.669**	1.196
	(0.136)	(0.114)	(0.216)	(0.339)	(0.053)	(0.095)	(0.094)	(0.157)
Outgoing (β_6)	0.716**	0.684	0.628**	0.558**	1.116	0.875	1.053	1.013
	(0.075)	(0.195)	(0.069)	(0.059)	(0.128)	(0.205)	(0.090)	(0.098)
Outgoing*TA (β_7)	0.858	0.692	1.393	1.110	0.716*	1.318	1.041	0.830
	(0.129)	(0.203)	(0.288)	(0.213)	(0.105)	(0.331)	(0.162)	(0.278)
Incoming*Post (β_8)	0.245**	0.144**	0.199**	0.228**	0.913	1.141	0.759**	0.727**
	(0.033)	(0.036)	(0.021)	(0.018)	(0.173)	(0.234)	(0.062)	(0.064)
Incoming*TA*Post (β_9)	0.894	0.577	1.644*	1.262	0.779	0.522*	0.922	1.739**
	(0.186)	(0.263)	(0.371)	(0.210)	(0.190)	(0.161)	(0.234)	(0.275)
Observations	10,744	3,393	13,217	10,485	11,318	3,604	13,479	10,831

Note: *, ** represents p<0.05, p<0.01. All regressions are estimated with a logit model on teacher-year-level observations on the stated binary dependent variables, with no covariates. Samples limited to all teacher observations in chronically low-performing schools identified in Hansen and Choi (2012) using the 2005 turnaround point in math.

Table 7. Decomposition Estimates among Teachers Using Alternate Identification of TA Schools

Changes in sample	Dropping Moderately Improving schools				TA conditional on improving growth only			
State	Florida		North Carolina		Florida		North Carolina	
School sample	Elementary	Middle	Elementary	Middle	Elementary	Middle	Elementary	Middle
School Random Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
TA (β_3)	-0.007	-0.077**	-0.078*	0.010	0.011	-0.068*	-0.008	-0.001
	(0.041)	(0.021)	(0.031)	(0.026)	(0.028)	(0.034)	(0.021)	(0.030)
Post (β_4)	0.031*	0.073**	-0.090**	-0.047**	0.088**	0.030*	0.009	-0.020*
	(0.015)	(0.017)	(0.012)	(0.014)	(0.010)	(0.012)	(0.008)	(0.009)
TA*Post (β_5)	0.211**	0.108*	0.292**	0.125**	0.117**	0.064*	0.117**	0.095**
	(0.022)	(0.045)	(0.027)	(0.030)	(0.016)	(0.025)	(0.019)	(0.024)
Outgoing (β_6)	-0.028	-0.010	-0.003	-0.009	-0.021	-0.046**	0.010	0.000
	(0.017)	(0.018)	(0.014)	(0.018)	(0.012)	(0.013)	(0.010)	(0.011)
Outgoing*TA (β_7)	0.037	-0.016	0.032	-0.012	0.022	0.007	0.003	-0.002
	(0.028)	(0.033)	(0.034)	(0.034)	(0.020)	(0.025)	(0.023)	(0.026)
Incoming*Post (β_8)	-0.068**	-0.111**	-0.006	-0.156**	-0.056**	-0.057**	-0.041**	-0.064**
	(0.014)	(0.021)	(0.014)	(0.018)	(0.010)	(0.015)	(0.009)	(0.011)
Incoming*TA*Post (β_9)	0.042	0.044	-0.099**	0.113**	0.035*	0.026	-0.052*	0.035
	(0.021)	(0.048)	(0.030)	(0.032)	(0.015)	(0.029)	(0.021)	(0.025)
Observations	21,316	6,827	16,462	10,574	43,553	15,398	37,371	24,190
R-squared	0.578	0.608	0.637	0.690	0.575	0.623	0.644	0.682

Note: *, ** represents $p<0.05$, $p<0.01$. All regressions are estimated on student-year-level observations on math standardized test scores, and include the following covariates: prior math and reading scores, an indicator variable for missing reading scores, as well as indicators for race, gender, eligibility for free or reduced-price lunch, and designations for limited English proficiency or special education. Samples limited to student-teacher linked observations in math in chronically low-performing schools identified in Hansen and Choi (2012) using the 2005 turnaround point in math. See text for description of alternate sampling or TA coding under these alternate resgressions.

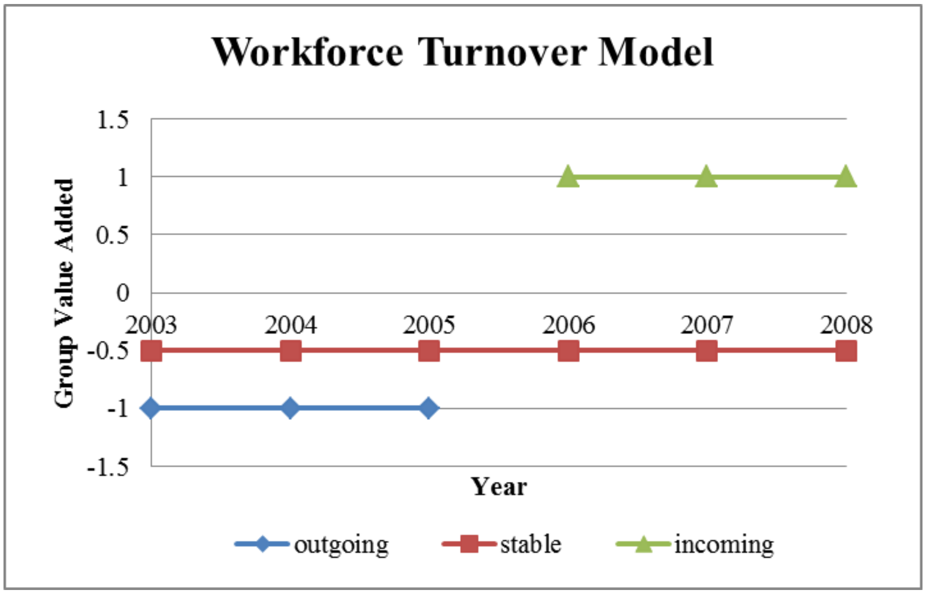
Table 8. Decomposing School Improvements in Higher Status Schools

State	Florida		North Carolina	
School sample	Elementary	Middle	Elementary	Middle
School Random Effects	Yes	Yes	Yes	Yes
Panel A. Limiting sample to schools with status measures in 16 - 30 percentiles				
TA*Post (β_5)	0.070**	0.108	0.179**	-0.010
	(0.019)	(0.132)	(0.018)	(0.030)
Outgoing*TA (β_7)	-0.011	-0.034	0.000	-0.006
	(0.025)	(0.042)	(0.023)	(0.031)
Incoming*TA*Post (β_9)	0.024	-0.236	-0.096**	0.158**
	(0.020)	(0.174)	(0.020)	(0.055)
Observations	38,064	26,411	43,829	14,819
R-squared	0.587	0.669	0.674	0.730
Panel B. Limiting sample to schools with status measures in 31 - 50 percentiles				
TA*Post (β_5)	0.070**	0.003	0.120**	0.026
	(0.015)	(0.017)	(0.013)	(0.040)
Outgoing*TA (β_7)	0.048*	0.027	-0.030	0.096*
	(0.021)	(0.021)	(0.019)	(0.047)
Incoming*TA*Post (β_9)	0.075**	0.053*	0.051**	0.244**
	(0.017)	(0.025)	(0.015)	(0.051)
Observations	49,617	33,487	67,216	13,623
R-squared	0.612	0.654	0.673	0.727
Panel C. Limiting sample to schools with status measures in 51 - 70 percentiles				
TA*Post (β_5)	0.097**	0.028*	0.141**	0.057**
	(0.013)	(0.012)	(0.014)	(0.020)
Outgoing*TA (β_7)	-0.071**	0.073**	-0.012	-0.024
	(0.020)	(0.022)	(0.018)	(0.026)
Incoming*TA*Post (β_9)	0.028	0.057**	0.019	0.091**
	(0.015)	(0.017)	(0.016)	(0.024)
Observations	62,339	47,571	59,885	28,166
R-squared	0.630	0.658	0.689	0.744

Note: *, ** represents $p<0.05$, $p<0.01$. All regressions are estimated on student-year-level observations on math standardized test scores, and include the following covariates: prior math and reading scores, an indicator variable for missing reading scores, as well as indicators for race, gender, eligibility for free or reduced-price lunch, and designations for limited English proficiency or special education. The remaining decomposition variables were also included but are not reported. Samples limited to all valid student observations in math in low-growth schools with status parameters in the range stated in each panel heading, as estimated in Hansen and Choi (2012) using the 2005 turnaround point in math.

Figure 1. Hypothetical Performance of Teacher Groups by School Improvement Model

Panel A



Panel B

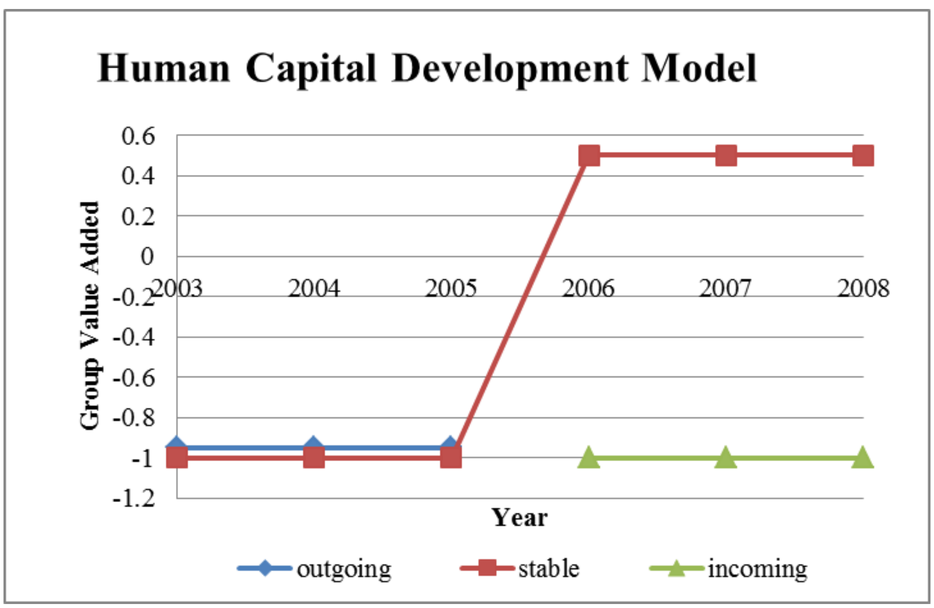
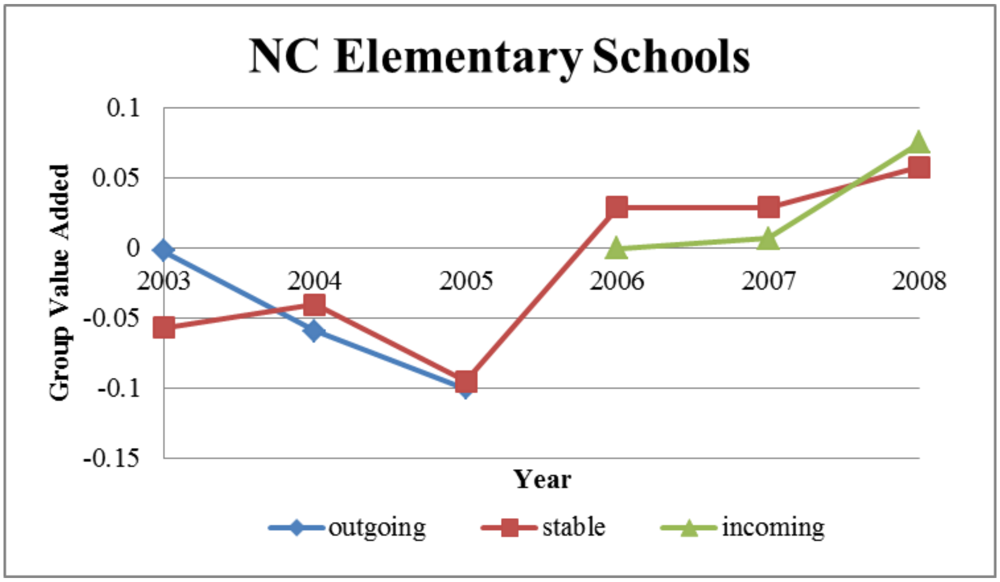


Figure 2. Observed Performance among Teachers in NC Turnaround Schools

Panel A



Panel B