MATHEMATICA Policy Research

REPORT

FINAL REPORT

Longitudinal Analysis of the Effectiveness of DCPS Teachers

July 8, 2014

Elias Walsh

Dallas Dotter

Submitted to:

DC Education Consortium for Research and Evaluation
School of Education and Human Development, George Washington University
2134 G Street, NW
Washington, DC 20052

Project Officer: Heather Harding
Contract Number: 13-SC08R

Submitted by:

Mathematica Policy Research
1100 1st Street, NE
12th Floor
Washington, DC 20002-4221
Telephone: (202) 484-9220
Facsimile: (202) 863-1763

Project Director: Elias Walsh
Reference Number: 40185.533

ACKNOWLEDGMENTS

This report would not have been possible without the support of the DC Education Consortium for Research and Evaluation and the National Research Council. We are especially grateful to Michael Feuer, Heather Harding, Alexandra Beatty, and Stuart Elliot for their leadership throughout this project and to the District of Columbia Public Schools for providing the data for the report. We are also grateful to Steven Glazerman and Duncan Chaplin for their helpful comments. Adele Costigan, John Hotchkiss, and Eli Kaplan provided excellent programming support. Betty Teller edited the report, and Jackie McGee provided word processing and production support.

CONTENTS

I INTRODUCTION 1

II THE DCPS IMPACT EVALUATION SYSTEM 4

III TRENDS IN RETENTION RATES 7

A. How dismissals of teachers vary over time 7

B. How retention rates of teachers vary over time 9

C. How retention rates of teachers vary across schools 11

IV TRENDS IN EFFECTIVENESS OF ENTERING AND EXITING TEACHER COHORTS 14

V TRENDS IN EFFECTIVENESS OF NOVICE TEACHERS RELATIVE TO VETERAN TEACHERS 18

VI CONCLUSION 22

REFERENCES 24

APPENDIX: SUPPLEMENTAL RESULTS 25

TABLES

II.1 Components and features of the DCPS IMPACT system, by school year 5

III.1 Dismissals of teachers in DCPS, by school year 8

IV.1 Trends in effectiveness of teachers who entered or left DCPS 16

V.1 Trends in effectiveness of novice DCPS teachers relative to veteran teachers 20

A.1 Retention rates of teachers in DCPS, by school year 26

A.2 Trends in effectiveness of novice DCPS teachers who remained in DCPS for at least two years, relative to veteran teachers 27

FIGURES

III.1 Percentage of teachers who left DCPS, by dismissal criteria eligibility and school year 8

III.2 Retention rates of teachers in DCPS, by effectiveness category and school year 10

III.3 Numbers of teachers retained and not retained in DCPS, by effectiveness category and school year 10

III.4 Number of schools by school-wide retention rate of all DCPS teachers and school year 12

III.5 Number of schools by school-wide retention rate of effective and highly effective DCPS teachers and school year 13

I. INTRODUCTION

The enactment of the Public Education Reform Amendment Act (PERAA), coupled with changes in the leadership of the DC Public Schools (DCPS)—principally the appointment of Michelle Rhee to the newly created post of chancellor—resulted in reforms in the governance structures, human capital policies, and resource management of DCPS and DC charter schools. PERAA also required periodic reports describing the impact of these reforms on DCPS business practices and strategies, human resources and human capital strategies, academic plans, and student achievement. The first two reports written for PERAA chronicled outcomes for the 2010–2011 and 2011–2012 school years (Education Consortium for Research and Evaluation 2013a and 2013b). Both enumerated outcomes and strategies for a single school year and did not address trends over time. In contrast, this report describes trends in outcomes related to human resources and human capital strategies between the 2008–2009 and 2012–2013 school years.

The major human resources/human capital change in DCPS after the enactment of PERAA was the creation of the IMPACT system for teacher performance evaluation. In this report, we focus on how teacher retention rates and teacher effectiveness—as measured by IMPACT scores—have evolved since DCPS implemented the IMPACT system for the 2009–2010 school year. DCPS has used IMPACT to evaluate teachers and make decisions about teacher retention and pay. Although the research questions we examine are not intended to identify the specific policies responsible for the outcomes or trends we analyze, our work may inform the following questions about policies aimed at improving teacher effectiveness:

- Have DCPS teacher retention and dismissal policies successfully retained the most effective teachers and removed the least effective teachers?
- Given the number and effectiveness of teachers who left DCPS each year and the need to hire replacements, have hiring practices in DCPS led to more effective teachers in the district?
- Have DCPS practices to support teachers early in their career helped teachers improve?

Because PERAA and IMPACT introduced multiple new policies to DCPS, and also because other changes and trends in DCPS may have affected the outcomes in our analysis, no specific outcomes or trends can be attributed to PERAA or IMPACT alone. The research questions examined here aim at helping education policymakers and practitioners understand how teacher retention and teacher effectiveness have changed over time in DCPS.

The measures of teacher effectiveness used in the analyses are based on DCPS teachers' IMPACT scores. A key challenge for our analysis of trends in the effectiveness of DCPS teachers is that year-to-year comparisons of IMPACT scores may not be meaningful because of changes in how the scores were calculated over time, such as changes to the weights given to the components used in the calculation of the scores and to the calculation of the component scores themselves (we describe the changes to IMPACT in Chapter II). Our main approach to address this issue is to examine changes in the gaps in average IMPACT scores between groups rather than to directly compare IMPACT scores across years. For example, to examine trends in the effectiveness of new teachers to DCPS, we compare the average IMPACT scores of the new teachers to the scores for a benchmark group of "core teachers" whose identity is consistent over

time. Whereas the average IMPACT scores of new teachers could change over time only because of changes to the calculation of IMPACT scores, changes to IMPACT are less likely to affect the gap in scores between new teachers and the benchmark group. This is because using a consistent comparison group removes the consequences of year-to-year changes to IMPACT if the changes affected scores for both groups in the same way. For example, after the first year of IMPACT, DCPS reduced the weight given to the classroom observation component in the calculation of IMPACT scores. Our approach accounts for this change if the IMPACT scores of new teachers changed similarly compared to those of core teachers.

Using administrative data from DCPS, our study addresses the following research questions:

1. **Have retention rates of effective teachers changed since the first year of IMPACT**? Because teachers who earn an ineffective rating in one year or a minimally effective rating for two consecutive years are dismissed under IMPACT, overall retention rates may fall even as effective teachers are retained at higher rates. We first examine how many teachers met IMPACT criteria for dismissal and were dismissed. We then compare the overall retention rate for teachers in the 2008–2009 school year, before IMPACT scores were calculated, to the overall retention rates in each of the first three years of IMPACT (the 2009–2010 to 2011–2012 school years). We also compare the rate at which DCPS retained its most effective teachers (measured by IMPACT) in the first year of IMPACT in 2009–2010 to the same rates in 2010–2011 and 2011–2012. Finally, we examine whether the rate at which individual schools retained teachers changed over time.
2. **How effective are teachers who leave DCPS and teachers who are hired to replace them**? Whether the average DCPS teacher was more effective four years after IMPACT began depends in part on whether DCPS retained its most effective teachers and hired new effective teachers to replace its least effective teachers. We examine the effectiveness of teachers who entered and exited DCPS since IMPACT was initiated compared to those "core teachers" who remained in DCPS from the 2008–2009 through the 2012–2013 school years. We calculate the gap between the average IMPACT scores of teachers who left DCPS in a given year and the core teachers in each year. We also calculate the gaps between the average IMPACT scores of teachers who were new to DCPS each year and the average IMPACT scores of core teachers.
3. **How does the effectiveness of novice teachers change with experience**? We measure changes in effectiveness for novice teachers as they become more experienced relative to the change in effectiveness for veteran teachers. Teachers typically become more effective as they gain experience, which is why novice teachers are usually less effective than veteran teachers. The success of IMPACT in improving teacher effectiveness may depend in part on whether and how much novice teachers improve over time compared to veteran teachers.

As in the previous reports for the 2010–2011 and 2011–2012 school years, this study uses administrative data from DCPS on teachers. To answer questions about trends in teacher effectiveness over time, we examine data provided by DCPS that include (1) IMPACT scores for all DCPS teachers in the 2009–2010 to 2012–2013 school years and (2) a list of teachers

teaching in the 2008–2009 school year.[1] Although there are limitations to using IMPACT scores to measure teacher effectiveness, the IMPACT score was specifically designed to evaluate DCPS teachers. Three main limitations could affect the interpretation of our results:

1. IMPACT scores are not guaranteed to be comparable year to year. For example, the actual effectiveness of teachers classified as highly effective could change. We address this first limitation by examining changes in the effectiveness of key groups of teachers using a consistent comparison group of teachers to remove year-to-year differences in how IMPACT scores are calculated.
2. The effectiveness of the teachers in our comparison group may change over time. If so, our findings about improvements in teacher effectiveness could be too large or too small. For example, we would find too little improvement in a cohort of novice DCPS teachers if performance pay or other features of IMPACT improve the effectiveness of veteran DCPS teachers—our comparison group for the novice teachers.
3. Even if IMPACT scores provide highly accurate measures of teacher effectiveness, no measure can provide a perfect evaluation. Misclassification errors in IMPACT could affect the interpretation of some of our results.

In Chapter II, we describe the details of the IMPACT evaluation system. In Chapter III, we present the results of our analysis of teacher retention, in Chapter IV, we examine the effectiveness of DCPS leavers and new hires, and in Chapter V, we examine how the effectiveness of novice teachers changes with experience. We present conclusions in Chapter VI.

[1] The IMPACT data provided by DCPS included some teachers with an incomplete IMPACT rating in one or more years. Teachers who received incomplete ratings were not subject to consequences under IMPACT. In preparing the data for analysis, we excluded all records for 281 teachers with incomplete ratings.

II. THE DCPS IMPACT EVALUATION SYSTEM

During the years of our study, IMPACT evaluation scores in DC were based on the following components (1) evaluations by school administrators and other trained observers using a classroom-observation rubric (Teaching and Learning Framework); (2) an individual value-added (IVA) measure of student achievement growth;[2] (3) an alternative measure of student achievement growth based on achievement targets determined by the teacher and principal (Teacher-Assessed Student Achievement Data); (4) a principal-assessed measure of the teacher's collaboration with colleagues and support of school initiatives and programs (Commitment to the School Community); (5) a principal-assessed measure of the teacher's attendance, adherence to school policies, and professionalism (Core Professionalism); and (6) a school value-added score (SVA).[3]

The composition of a teacher's IMPACT evaluation score is based on that teacher's IMPACT group. Group 1 consists of general education teachers of math and reading/English language arts (ELA) in grades 4 through 8—the grades and subjects for which test score data needed to calculate teacher value added are available. DCPS expanded group 1 to include reading/ELA teachers in grades 9 and 10 for the 2012–2013 school year. Group 2 includes all other general education teachers. Groups 3 through 7 (or 8 in the 2009–2010 school year) include other teachers, such as those with special education or English language learner (ELL) students. Over 75 percent of DCPS teachers are included in groups 1 and 2.

The calculation of IMPACT scores depended on the school year. This is because the weights for each component in the total score changed over time, as did the measurement of some components.[4]. We describe the weights by school year in Table II.1 for group 1 teachers (top panel) and group 2 teachers (middle panel). For example, prior to the 2012–2013 school year, the IVA score constituted 50 percent of the total evaluation score for group 1 teachers. The weight on the IVA score was reduced to 35 percent for the 2012–2013 school year. The components used for teachers in groups 3 through 8 vary too widely to include in the table. For example, in the 2009–2010 school year, the TLF score constituted 50 to 100 percent of the weight for these teachers; the remaining weight was given to a variety of components that included some of the same measures used for group 1 and 2 teachers and/or measures that are specific to the teacher's role, such as measures of the quality and timeliness of individualized education plans for special education teachers. For all groups, the total evaluation score ranged from 100 to 400 points. Based on this score, a teacher received one of four possible effectiveness ratings: highly effective (350 to 400 points), effective (250 to 349 points), minimally effective (175 to 249 points), or

[2] Value added is a measure of teacher effectiveness that seeks to isolate a teacher's contribution to student achievement from any confounding factors outside the teacher's control (Isenberg and Hock 2012).

[3] The SVA component was not used in teachers' evaluations for the 2012–2013 school year.

[4] For example, teacher and school value-added estimates were measured relative to the average DCPS teacher in the school year, and so would not reflect overall improvements in DCPS teacher effectiveness. Isenberg and Hock (2011, 2012), and Isenberg and Walsh (2014) provide technical details of the value-added models used in IMPACT for each school year.

ineffective (100 to 174 points). For the 2012–2013 school year, DCPS teachers could receive a new rating of developing (250 to 299 points).

Table II.1. Components and features of the DCPS IMPACT system, by school year

	Weight on component in school year (percentage)			
	2009–2010	2010–2011	2011–2012	2012–2013
Components for group 1				
Teaching and learning framework	40	35	35	40
Individual value-added score	50	50	50	35
Commitment to the school community	5	10	10	10
Teacher-assessed student achievement data	0	0	0	15
School value-added score	5	5	5	0
Core professionalism	Adjustment	Adjustment	Adjustment	Adjustment
Components for group 2				
Teaching and learning framework	80	75	75	75
Teacher-assessed student achievement data	10	10	10	15
Commitment to the school community	5	10	10	10
School value-added score	5	5	5	n.a.
Core professionalism	Adjustment	Adjustment	Adjustment	Adjustment
Feature				
Performance bonus pay	n.a.	Included	Included	Included
Performance base pay increase	n.a.	Included	Included	n.a.
Career ladder	n.a.	n.a.	n.a.	Included

Source: DCPS IMPACT Guidebooks for the 2009–2010 through 2012–2013 school years.

Notes: The components are scored on a continuous scale from 1.0 to 4.0, except for the core professionalism component, which is scored on a scale of 0 (best) to -40 (worst). The total IMPACT score is calculated by multiplying the teacher's score on each component (except core professionalism) by its weight and summing the results. A teacher's core professionalism score is applied as an adjustment after all other components have been combined.

The teacher-assessed student achievement component was called "non-value-added student achievement growth" in the 2009–2010 school year.

n.a. = not applicable

Under IMPACT, teachers who earn a highly effective rating receive performance pay; those who earn an ineffective rating one year or a minimally effective rating for two consecutive years are dismissed.[5] The amount of performance pay highly effective teachers receive is higher for teachers in schools where at least 60 percent of students are eligible for free or reduced-price lunch (FRL); bonuses in the years studied were as large as $25,000 for group 1 teachers and $20,000 for teachers in other groups. Additionally, teachers with consecutive highly effective ratings received an increase in base pay. For the 2012–2013 school year, DCPS introduced a more extensive career-ladder system: advances were based on receiving effective and highly

[5] Starting with ratings given in the 2012–2013 school year, teachers who earn a developing rating for three consecutive years will be dismissed.

effective ratings. The performance pay features of IMPACT by school year are shown in the bottom panel of Table II.1.

The DCPS IMPACT Guidebooks for these school years provide additional details about the IMPACT components, weights, and features.

III. TRENDS IN RETENTION RATES

New policies relating to dismissal and performance pay may have affected the retention of teachers in DCPS. As a result of dismissals under IMPACT, overall retention rates may have fallen and the additional exits may have been concentrated among less-effective teachers. Additionally, retention rates of more-effective teachers may also have fallen if the possibility of dismissal or other aspects of IMPACT that teachers view unfavorably led teachers to exit DCPS. On the other hand, performance pay could have provided incentives for highly effective teachers to remain in DCPS. Previous research has found evidence that performance pay incentives may have led to higher retention of the most effective teachers and that the threat of dismissal under IMPACT may have led more low performing teachers to leave DCPS even though they were eligible to remain (Dee and Wyckoff 2013).[6]

A. How dismissals of teachers vary over time

We can identify the teachers who were dismissed under IMPACT for low IMPACT ratings, although some of these teachers might have left DCPS even in the absence of IMPACT. Under IMPACT, teachers in the minimally effective category for two years or in the ineffective category for a single year are subject to dismissal.

The proportion of teachers dismissed under IMPACT was higher in year 2 of IMPACT compared to year 1. According to Table III.1, DCPS dismissed 1.8 percent of teachers at the end of year 1 of IMPACT and 5.5 percent at the end of year 2 (row 1, columns 1 and 2). In year 1, DCPS only dismissed teachers with ineffective ratings. Teachers received their second rating in year 2 of IMPACT, so DCPS could base its retention decisions on two years of IMPACT data for the first time at the end of that school year. Thus, the higher rate of dismissals at the end of year 2 of IMPACT reflects additional separations among this first cohort of teachers who could have received IMPACT scores in the minimally effective category for two consecutive years. The 126 teachers with consecutive minimally effective ratings in year 2 represent over two-thirds of the dismissals in that year. Excluding those teachers would result in a rate similar to the rate from year 1 (row 2, columns 1 and 2).

The proportion of teachers dismissed under IMPACT was lower in years 3 and 4 of IMPACT compared to year 2. DCPS dismissed 2.5 percent of teachers at the end of year 3 for having two consecutive years of minimally effective ratings or an ineffective rating (row 1, column 3), just over half the rate of dismissals at the end of year 2. The proportion of teachers who were eligible for dismissal under IMPACT at the end of year 4 (row 1, column 4) was also smaller than in year 2. This is reflected in the lower percentages from years 3 and 4 of both teachers who were rated as minimally effective in two consecutive years (row 3) and those who were rated as ineffective (row 2), compared to the same percentages in year 2.

[6] Dee and Wyckoff (2013) focused only on teachers with IMPACT scores near the boundaries of effectiveness categories because doing so allowed them to infer that the changes in retention rates were caused by IMPACT. In contrast, our focus is on describing overall trends in the retention of DCPS teachers rather than on attributing changes in retention rates to IMPACT policy.

The proportion of DCPS teachers dismissed under IMPACT was smaller than the proportion of teachers leaving DCPS for other reasons. In addition to being dismissed on the basis of IMPACT ratings, teachers also left DCPS for other reasons. Teachers who left DCPS despite being eligible to remain represented 15.5 percent of DCPS teachers at the end of year 1 and increased to 16.9 percent by the end of year 3. As shown in Figure III.1, even at the end of year 2, when the dismissal rate peaked at 5.5 percent, there were nearly three times as many leavers who did not meet IMPACT criteria for dismissal as there were teachers who met the criteria.

Table III.1. Dismissals of teachers in DCPS, by school year

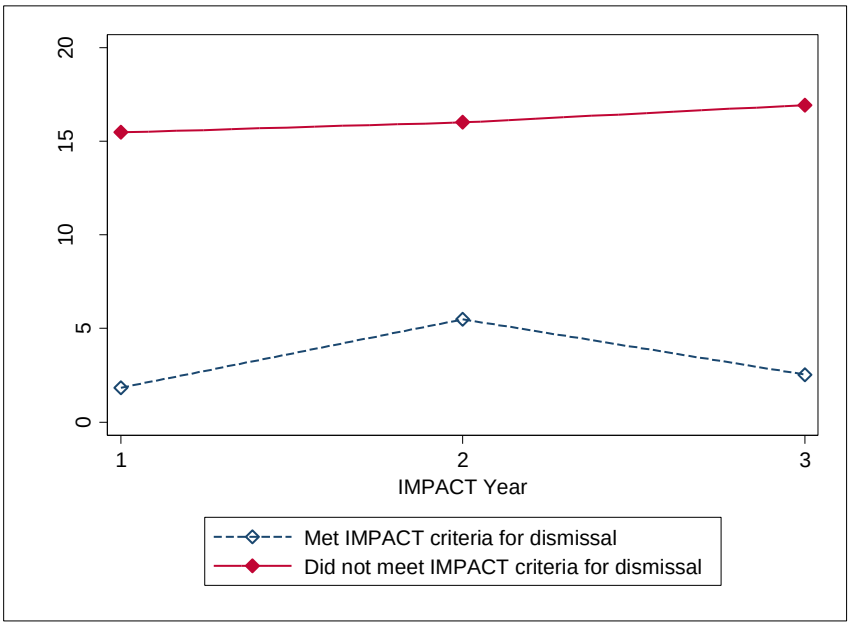
Category	Year 1	Year 2	Year 3	Year 4[a]
Percent dismissed	1.8	5.5	2.5	2.3
Percent ineffective	1.8	1.7	1.0	1.3
Percent minimally effective in consecutive years	0.0	3.8	1.5	1.0
Number of teachers dismissed	62	182	83	76
Total number of teachers	3,378	3,315	3,270	3,264

Source: Mathematica calculations based on administrative data from DCPS.

Notes: Under IMPACT, teachers who earn an ineffective rating in one year or a minimally effective rating for two consecutive years are dismissed. Teachers could be dismissed for consecutive minimally effective ratings beginning in year 2.

[a] Whereas the percentages and counts in the columns for years 1 through 3 reflect teachers who were actually dismissed by DCPS, those in year 4 reflect only teachers' eligibility for dismissal. In years 1 through 3, DCPS dismissed all eligible teachers.

Figure III.1. Percentage of teachers who left DCPS, by dismissal criteria eligibility and school year



Source: Mathematica calculations based on administrative data from DCPS.

Notes: Under IMPACT, teachers who earn an ineffective rating in one year or a minimally effective rating for two consecutive years are dismissed. Teachers could be dismissed for earning consecutive minimally effective ratings beginning in year 2.

B. How retention rates of teachers vary over time

The overall rate of retention for DCPS teachers was 71 percent in the year before IMPACT (the 2008–2009 school year).[7] The average rate of retention in the first three years of IMPACT (the 2009–2010 through 2011–2012 school years) was 80 percent, an increase of 12 percent (9 percentage points) pre- and post-IMPACT. However, this result should be interpreted with caution for two reasons. First, the best available information did not always allow DCPS to distinguish classroom personnel from other DCPS personnel in the 2008–2009 school year. Consequently, the list of teachers teaching in the 2008–2009 school year may include some personnel who were not classroom teachers and may exclude some who were. This could result in a pre-IMPACT retention rate that is too low or too high. Second, even if the difference reflects actual differences in retention rates of teachers between these years, it could be due to other policies or trends in DCPS and cannot be attributed solely to IMPACT.

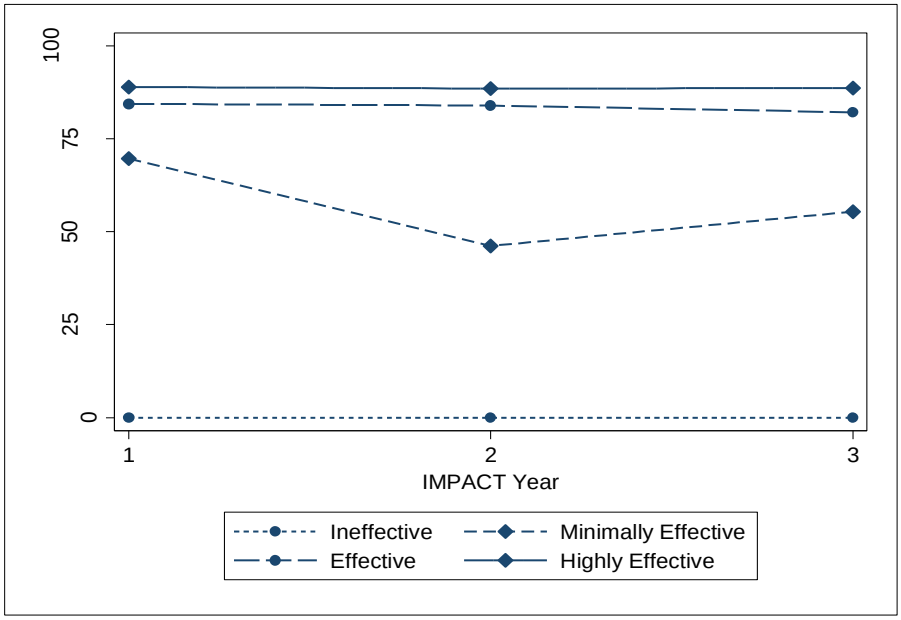
Although, we cannot examine retention by effectiveness prior to IMPACT, we can examine whether trends for more- or less-effective teachers are responsible for changes in retention rates since IMPACT. To do so, we compare the proportion of teachers who were retained in each IMPACT effectiveness category in 2009–2010—the first school year of IMPACT—to the same proportions in 2010–2011 and 2011–2012.

DCPS retained over 80 percent of teachers classified as effective or highly effective by IMPACT, and these retention rates did not change over the first three years of IMPACT. Figure III.2 shows retention rates for all DCPS teachers by IMPACT rating.[8] DCPS retained 89 percent of teachers in the highly effective category in each of years 1, 2, and 3 of IMPACT (solid line). Of teachers in the effective category, DCPS retained between 82 and 84 percent in these years (dashed line with circles). Figure III.3 shows how many teachers were retained (white), left DCPS despite being eligible to remain (khaki), and left DCPS after meeting IMPACT dismissal criteria (blue) in each effectiveness category by year. Although the retention rates for the effective and highly effective groups of teachers were similar, the number of teachers classified as effective who left DCPS is substantially larger than the number classified as highly effective who left. This is because teachers in the effective category represented between 68 and 71 percent of all DCPS teachers during these years, whereas less than a quarter of teachers were classified as highly effective.

[7] Retention rates do not distinguish between teachers dismissed under IMPACT and teachers who exited voluntarily.

[8] Appendix Table A.1 shows retention rates by IMPACT group. Because there are relatively few group 1 teachers—for example, 459 in year 1 of IMPACT compared to 2,919 teachers in groups 2 through 8—some differences in retention rates between these groups that appear substantive are not statistically significant.

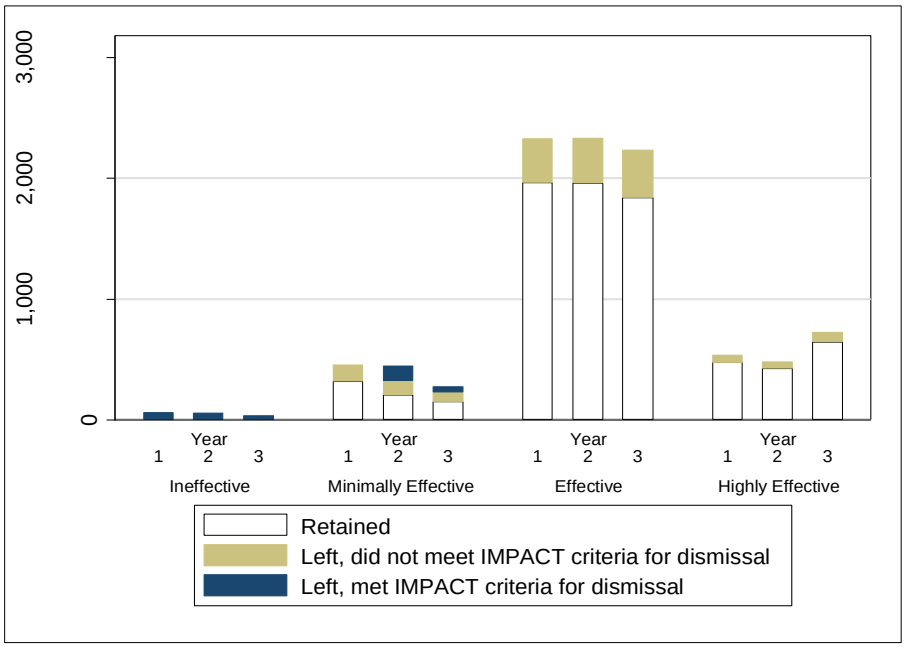
Figure III.2. Retention rates of teachers in DCPS, by effectiveness category and school year



Source: Mathematica calculations based on administrative data from DCPS.

Notes: IMPACT ratings are from the specified school year. Teachers are considered to have exited DCPS if they do not receive an IMPACT rating in the subsequent school year.

Figure III.3. Numbers of teachers retained and not retained in DCPS, by effectiveness category and school year



Source: Mathematica calculations based on administrative data from DCPS.

Notes: IMPACT ratings are from the specified school year. Teachers are considered to have exited DCPS if they do not receive an IMPACT rating in the subsequent school year.

Retention of teachers classified as minimally effective was lower in years 2 and 3 of IMPACT than in year 1. DCPS retained 70 percent of teachers in the minimally effective category in year 1 of IMPACT, but only 46 percent in year 2, a statistically significant drop.[9] The lower rate of retention for minimally effective teachers in year 2 of IMPACT reflects additional dismissals among this first cohort of teachers who could have received IMPACT scores in the minimally effective category for a second consecutive year. The retention rate for teachers rated minimally effective in year 3 of IMPACT was 55 percent, significantly lower than the rate in year 1, but not significantly different from the rate in year 2.

One possible limitation of the analysis of retention rates in this section is that IMPACT effectiveness categories are not guaranteed to be comparable year to year, so the actual effectiveness of teachers classified as highly effective could change. Changes could result from differences in the calculation of the total IMPACT score or in the scoring of the components, so rescaling IMPACT scores to be comparable across years is not straightforward. Although the effectiveness categories may not be consistent across years, our analysis provides information about how teachers are retained based on how they are classified into effectiveness categories by DCPS.

C. How retention rates of teachers vary across schools

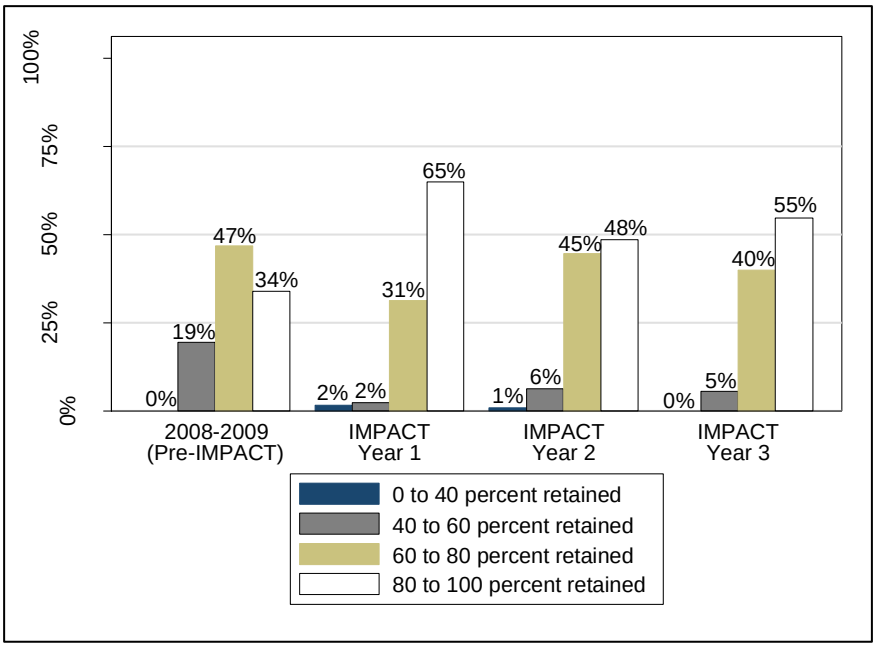
In addition to examining the retention of teachers in all DCPS schools, we also examined how retention rates differed between DCPS schools, and whether those differences changed over time. We measured the proportion of teachers in each school who remained in DCPS, so that these rates can be compared to the overall retention of teachers in DCPS from Figure III.2. Thus, a teacher who moved between two DCPS schools would count as retained. A smaller proportion of teachers may have remained in DCPS in some schools than others, and the number of schools with low retention rates may change over time. Additionally, schools may differ in how they retain the most effective teachers. Figure III.4 shows how retention rates of all teachers differed across schools and over time, and Figure III.5 shows how retention rates of teachers classified as effective and highly effective differed.

In years 1 through 3 of IMPACT, more DCPS schools retained at least 80 percent of teachers than was the case pre-IMPACT. In the year before IMPACT, 34 percent of DCPS schools retained at least 80 percent of teachers. These schools are shown in the white bar of Figure III.4 for the 2008–2009 school year. In contrast, more schools retained at least 80 percent of teachers in the three post-IMPACT years (the height of the white bar for the post-IMPACT school years is taller): 65 percent of schools retained at least 80 percent of teachers in year 1 of IMPACT, 48 percent did so in year 2, and 55 percent did so in year 3. However, we again caution that the pre-IMPACT retention rates may be too low or too high as a result of limitations in the quality of the pre-IMPACT data available from DCPS. Additionally, differences in retention rates pre- and post-IMPACT could be due to other policies or trends in DCPS, and cannot be attributed solely to IMPACT.

[9] We estimated standard errors assuming that retention decisions are uncorrelated across years.

In each of years 1 through 3 of IMPACT, over two-thirds of schools retained at least 80 percent of teachers classified as effective or highly effective. According to Figure III.5, school-wide retention rates of effective or highly effective teachers were similar in the first three years of IMPACT, although the number of schools that retained at least 80 percent of the most effective teachers gradually declined. In year 1 of IMPACT, 78 percent of schools retained at least 80 percent of teachers classified as effective or highly effective, 74 percent did so in year 2, and 68 percent did so in year 3. In contrast, a small number of schools retained fewer than 40 percent of the most effective teachers in the school in year 1 of IMPACT, whereas all schools retained at least 40 percent of these teachers in years 2 and 3.

Figure III.4. School-wide retention rates of all DCPS teachers by school year

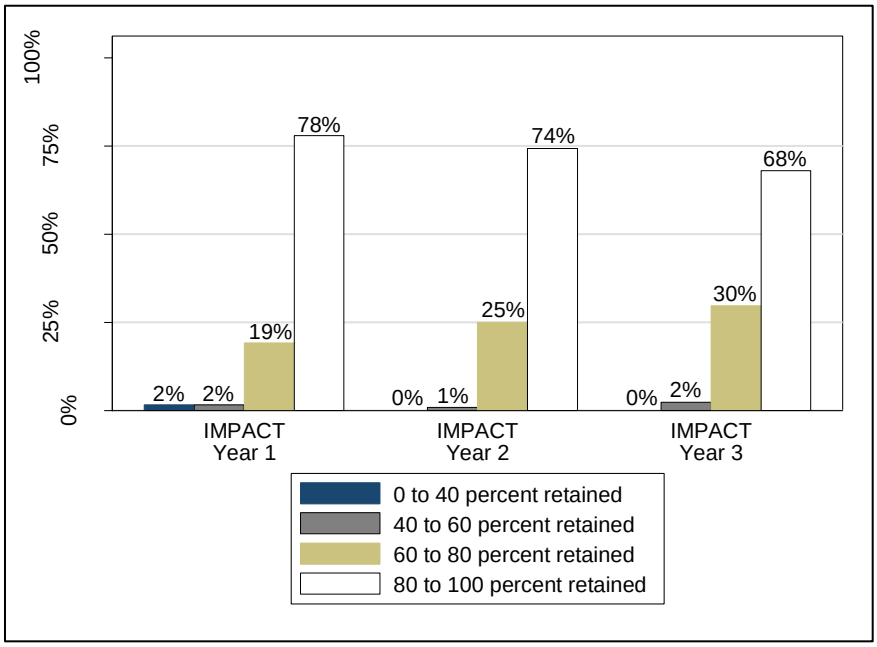


Source: Mathematica calculations based on administrative data from DCPS.

Notes: Retention rates are calculated as the percentage of teachers in the school who remained in DCPS the following year.

The figure includes 124 schools in the pre-IMPACT year, 131 in year 1 of IMPACT, 128 in year 2 of IMPACT, and 128 in year 3 of IMPACT. The figure excludes 7 schools in the pre-IMPACT year and one school in year 1 of IMPACT with fewer than 5 teachers in the school year.

Figure III.5. School-wide retention rates of effective and highly effective DCPS teachers by school year



Source: Mathematica calculations based on administrative data from DCPS.

Notes: Retention rates are calculated as the percentage of teachers in the school who remained in DCPS the following year.

The figure includes 131 schools in year 1, 128 in year 2, and 128 in year 3. The figure excludes one school in year 1 with fewer than five teachers in the school year.

IV. TRENDS IN EFFECTIVENESS OF ENTERING AND EXITING TEACHER COHORTS

This chapter investigates the extent to which the effectiveness of teachers in DCPS has changed since the introduction of dismissal and performance pay policies under IMPACT. To describe trends in the effectiveness of entry and exit cohorts since IMPACT went into effect, we examine the gap between the effectiveness of teachers who transitioned into or out of DCPS and the core group of 1,342 teachers who remained in DCPS from 2008–2009 through 2012–2013. The comparison group is necessary because year-to-year comparisons of IMPACT scores may not be meaningful. Using a consistent comparison group removes the consequences of year-to-year changes to IMPACT if the changes affected scores for both groups in the same way. We included the same teachers in the comparison group in each year so that changes in the gaps over time are more likely to reflect changes in the effectiveness of entry and exit cohorts rather than changes in the identity of teachers in the comparison group.[10]

In addition to examining teachers who left DCPS as a group, we distinguish between teachers who did and did not meet IMPACT criteria for dismissal. Teachers who met IMPACT criteria for dismissal will be among the least effective teachers based on IMPACT scores, because the criteria are based on these scores. However, the gap between teachers in this group and core teachers may change over time. For example, if DCPS dismissed the least effective teachers after the first year of IMPACT, the teachers dismissed in subsequent years may be relatively more effective, on average. We also compare teachers who left despite having IMPACT scores high enough to allow them to stay to core DCPS teachers. It might be difficult for DCPS to develop and maintain an effective teaching workforce if teachers who leave are relatively effective compared to core teachers.

As a group, new hires to DCPS were less effective than core teachers by a similar amount in all four years of IMPACT. According to Table IV.1, across the four years of IMPACT, new teachers to DCPS were between 26 and 33 IMPACT points less effective than core DCPS teachers (row 1).[11] Although these within-year gaps are statistically significant, none of the changes over time are significant, implying that new teachers in each of these years were similarly effective if there was no change in the effectiveness of core teachers.[12] The gaps remain similar despite substantial differences in the size of the new teacher cohorts over time. The first year of IMPACT (2009–2010) coincided with the entry of 1,135 new teachers, or 34 percent of all 2009–2010 teachers, whereas only 525 teachers, or 16 percent, were in their first year during year 4 of IMPACT.

[10] To address concerns that changes in the amount of dispersion in IMPACT scores over time could affect the results of our analysis using the raw IMPACT scores, we also conducted analysis based on IMPACT scores that were standardized to have the same mean and standard deviation in each year. The results from this alternative analysis did not substantively affect our main findings.

[11] The standard deviation of IMPACT scores ranges from 44 to 50 points across the four years in the analysis.

[12] Our standard error estimates in this chapter do not account for repeated observations of teachers across years.

Compared to core teachers, the teachers who left DCPS at the end of year 3 of IMPACT were relatively more effective than the teachers who left at the end of year 1, and this was true for leavers who both did and did not meet IMPACT criteria for dismissal. Teachers who left DCPS at the end of year 1 of IMPACT were less effective than the core teachers by 47 IMPACT points (Table IV.1, column 1, row 2). Teachers who left DCPS at the end of year 3 were less effective than the core teachers by 36 IMPACT points (column 3, row 2), implying that they were relatively more effective than the teachers who left at the end of year 1 by 11 IMPACT points if the effectiveness of core teachers did not change. Table IV.1 indicates that all three within-year gaps between leavers and core teachers are statistically significant. In addition, the decline in the gap from year 1 to year 3 is statistically significant.

The third and fourth rows of Table IV.1 present gaps relative to the core group of teachers separately for the two types of leavers. Both groups of leavers were significantly less effective than core teachers during the first three years of IMPACT. Leavers who did not meet IMPACT criteria for dismissal were between 23 and 34 IMPACT points less effective than core teachers, whereas those who met the criteria were between 115 and 165 points less effective than core teachers. Between years 1 and 2, the gaps relative to core teachers for both groups declined; teachers in year 2 who met IMPACT criteria for dismissal were relatively more effective by 50 IMPACT points than those in year 1, and those who left DCPS but did not meet the criteria were relatively more effective by 6 points, although this second change is not statistically significant.[13] Between years 1 and 3, these same increases are 39 points for leavers who met dismissal criteria and 11 points for leavers who did not, and both increases are statistically significant. In part, the lower relative effectiveness of leavers who met dismissal criteria in year 1 is due to the fact that the year 1 dismissals included only teachers classified as ineffective by IMPACT, whereas the dismissals in years 2 and 3 included teachers classified as minimally effective.

New hires to DCPS in years 1 and 2 of IMPACT were more effective relative to teachers who left DCPS at the end of those years, but more similar to them in year 3. In all years, both new hires and teachers who left DCPS at the end of the year were less effective than core DCPS teachers. For example, teachers new to DCPS were less effective than core teachers by 30 IMPACT points in years 1 and 2 (row 1 of Table IV.1). Also, teachers who left DCPS at the end of year 1 of IMPACT were less effective than core teachers by 47 IMPACT points (column 1, row 2). Thus, new hires were 17 points more effective than leavers in year 1, although both groups were less effective than core teachers. In year 2, this difference was 20 points. By year 3, this difference was only 3 points and not statistically significant.[14]

[13] Although both groups of leavers were more effective relative to core teachers in year 2, compared to year 1, leavers as a group were relatively less effective in year 2 than in year 1. This is possible because relatively more leavers met IMPACT dismissal criteria in year 2 compared to leavers in year 1.

[14] The differences within a column reflect the performance of leavers in the year before they left to new teachers during the same year. An alternative comparison is between leavers in the year before they left and new teachers in the following year, although these differences could only be interpreted as gaps in effectiveness between these groups if there were no change in the effectiveness of core teachers. These alternative gaps are 18, 17, and 10 points in years 1, 2, and 3.

Table IV.1. Trends in effectiveness of teachers who entered or left DCPS

	Gap in effectiveness relative to core teachers (IMPACT points)			
Subgroup	Year 1	Year 2	Year 3	Year 4
Entering teachers (in their first year in DCPS)	-30*	-30*	-33*	-26*
Leavers (in their last year in DCPS)	-47*	-50*	-36*	n.a.
Met IMPACT criteria for dismissal	-165*	-115*	-126*	n.a.
Did not meet IMPACT criteria for dismissal	-34*	-28*	-23*	n.a.
Number of core teachers	1,342	1,342	1,342	1,342
Number of entering teachers	1,135	521	640	525
Number of leavers	585	713	637	n.a.

Source: Mathematica calculations based on administrative data from DCPS.

Notes: A positive gap indicates that teachers in the subgroup had higher IMPACT scores on average than core teachers.

Core teachers are teachers in DCPS for all five school years from 2008–2009 through 2012–2013.

Under IMPACT, teachers who earn an ineffective rating in one year or a minimally effective rating for two consecutive years are dismissed. Teachers could be dismissed for consecutive minimally effective ratings beginning in year 2.

*Statistically significant at the 5 percent level.

n.a. = not applicable

Although DCPS has not experienced a decline in the effectiveness of new hires relative to core teachers, these results suggest that the increase in the relative effectiveness of leavers over time is not offset by increases in the relative effectiveness of new hires. However, even in year 3 of IMPACT, the new hires were similarly effective compared to the leavers. Even if the new hires were less effective, this would not necessarily imply that the DCPS teacher workforce is becoming less effective over time, because the core teachers may be more effective in year 4 than in year 1, and new hires may become more effective over time. This first issue is one of two limitations of this analysis we discuss below. The second issue we explore in the next section.

Limitations. Although the gaps shown in Table III.1 indicate that the leavers became relatively more effective compared to both core and new DCPS teachers, the analysis has two limitations. First, the effectiveness on exits and entries is measured relative to the effectiveness of a core group of DCPS teachers. If these core teachers became more effective on average over the four years—as might be expected as they gain more experience or are positively affected by professional development through IMPACT[15]—the change in the gap for leavers from -47 to -36 IMPACT points from year 1 to year 4 could reflect an actual increase in the effectiveness of exiting teachers of more than 11 IMPACT points. We cannot directly measure changes in the

[15] Dee and Wyckoff (2013) find evidence that performance pay incentives under IMPACT led to improvements in the effectiveness of teachers who might receive them. However, because their results could be explained by changes in the identity of DCPS teachers over time (for example, the financial incentives could have led to relatively higher retention of highly effective teachers) rather than responses to the incentives for teachers who remain in DCPS, this has uncertain implications for our research design.

effectiveness of core teachers, because year-to-year comparisons of IMPACT scores may not be meaningful.[16]

Second, some differences between the average IMPACT scores of the groups we analyze could reflect changes to IMPACT components and weights between school years rather than actual gaps in effectiveness. For example, the changes may have led to higher scores on a component for more experienced teachers, but not for entering teachers. In this example, entering teachers would appear less effective than they actually are.

[16] The average IMPACT score of the core group of DCPS teachers ranges from 314 to 328 IMPACT points over the first four years of IMPACT.

V. TRENDS IN EFFECTIVENESS OF NOVICE TEACHERS RELATIVE TO VETERAN TEACHERS

Previous research has demonstrated that teacher effectiveness tends to improve during a teacher's first few years of teaching and then remain roughly constant afterwards (Rockoff 2004; Hanushek et al. 2005; Harris and Sass 2010). In DCPS, more-experienced teachers received higher IMPACT scores on average than less-experienced teachers in the 2010–2011 and 2011–2012 school years (EdCORE 2013a, 2013b). In the previous chapter, we found that the IMPACT scores of teachers in their first year in DCPS were about 30 points lower on average than a group of core DCPS teachers with more experience. Novice teachers may be less effective than veteran teachers, but they could become more effective over time. Nevertheless, a simple comparison of IMPACT scores for new teachers in one school year to more experienced teachers in another school year could be misleading, because there is no guarantee that the scores are comparable year to year.

This chapter investigates how the effectiveness of novice teachers changed with teaching experience since the introduction of IMPACT in DCPS. To identify the growth in effectiveness attributable to increased experience for novice teachers, we describe the change in effectiveness of novice teachers as they gain experience relative to the typical change in effectiveness of veteran DCPS teachers with experience. We calculate teacher experience in DCPS by linking teachers across consecutive years in the data. For example, first-year teachers are those who did not teach during the previous school year, and second-year teachers are those who taught during the previous school year but not during the school year two years previous.[17]

The group of veteran teachers used for comparison with each novice teacher cohort consists of more-experienced DCPS teachers in the same school year. In all school years, we include only teachers who remained in DCPS through 2012–2013 for the calculation of the gap, in order to produce samples of novice and veteran teachers that are consistent over time. For example, we compare novice teachers in their first year to veteran teachers in their second or higher year. Then we compare the same novice teachers when they were in their second year to the same veteran teachers when they were in their third or higher year, and so on.[18] Thus, the changes in the gaps reflect improvements in the effectiveness of the novice teachers with gained experience relative to the same group of veteran teachers, rather than differences in which teachers DCPS retains.

To describe changes in the relative effectiveness of novice teachers as they gain additional years of experience, we first calculate the gaps in IMPACT scores between the veteran and

[17] Our approach to calculating teacher experience reflects the best data available to us; we recognize, however, that any teachers on leave for one or more school years were misclassified as new or second-year teachers.

[18] Our definition of veteran teachers results in using a different group of veteran teachers for each cohort of novice teachers. To test the sensitivity of our results to our definition of veteran teachers, we repeated our analyses using an alternative definition that compared all cohorts of novice teachers to the same veteran teachers. We achieved this by using the same core group of 1,342 veteran teachers who remained in DCPS from the 2008–2009 through the 2012–2013 school years that we used in the analysis from the previous chapter. The results from this alternative analysis did not substantively affect our main findings.

novice teachers for each novice teacher cohort's first year in DCPS and for all subsequent school years through 2012–2013.[19] We then calculate the changes over time in this initial effectiveness gap between veteran and novice teachers for each novice teacher cohort.

Although novice teachers in DCPS were less effective relative to veteran teachers in their first year, this gap decreased significantly with an additional year of teaching experience. Compared to veteran teachers, the cohort of novice teachers who were in their first year during the 2009–2010 school year (cohort 1) improved by 8 IMPACT points between the end of their first and second years in DCPS (row 1, column 1 of Table V.1). Novice teachers entering DCPS in the 2010–2011 school year (cohort 2) improved by 15 IMPACT points between the end of their first and second years compared to veteran teachers; the analogous change for novice teachers entering in the 2011–2012 school year (cohort 3) was 16 points (rows 2 and 3, column 1).

Novice teachers continued to improve relative to veteran teachers during their third and fourth year in DCPS. Compared to veteran teachers, the first cohort of novice teachers improved by an additional 2 IMPACT points between the end of the novices' second year and the end of their third year, for a total of 10 points by the end of their third year (row 1, column 2 of Table V.1). The same cohort improved by an additional 6 points by the end of their fourth year in DCPS, for a total of 16 points over the three years in DCPS following their first year (row 1, column 3 of Table V.1). Similarly, the second novice cohort improved by a total of 20 points compared to veterans between the end of their first year in DCPS and the end of their third year (row 2, column 2). Although the annual improvements in novice effectiveness compared to veterans were smaller after the novices' second year in DCPS, the novices did continue to improve. These improvements reflect changes in the effectiveness of the same teachers, rather than changes in the teachers who remained in DCPS, because the identity of the novices and veterans in the comparison for a cohort do not change over time.

Teachers hired in the second and third years of IMPACT may have improved more quickly relative to veteran teachers than teachers hired in the first year of IMPACT. The results in the first column of Table V.1 suggest that the second and third cohorts of novice teachers improved almost twice as much as the first cohort between the end of their first and second year in DCPS. This difference persists through the end of the third year in DCPS for the first and second novice cohorts. We obtain this result despite evidence in the previous chapter that the new hires in their first year were similarly effective compared to core DCPS teachers. If the recent cohorts have improved more quickly, these results could suggest that more recent policies under IMPACT have aided the development of new teachers, or that DCPS hired teachers with more potential for improvement beginning in the second year of IMPACT, although other explanations could be responsible for this finding.

[19] As in the previous chapter, we also conducted analysis based on IMPACT scores that were standardized to have the same mean and standard deviation in each year, to address concerns that changes in the amount of dispersion in IMPACT scores over time could affect the results of our analysis using the raw IMPACT scores. The results from this alternative analysis did not substantively affect our main findings.

Table V.1. Trends in effectiveness of novice DCPS teachers relative to veteran teachers

Sample	Improvement in novice effectiveness compared to veteran teachers, by novice experience (IMPACT points)			Number of novice teachers	Number of veteran teachers
	From first year to second year	From first year to third year	From first year to fourth year		
Cohort 1: 2009–2010 teachers retained through 2012–2013	8*	10*	16*	444	1,342
Cohort 2: 2010–2011 teachers retained through 2012–2013	15*	20*	NA	290	1,786
Cohort 3: 2011–2012 teachers retained through 2012–2013	16*	NA	NA	502	2,076

Source: Mathematica calculations based on administrative data from DCPS.

Notes: Novice teachers are defined as first-year DCPS teachers in the base year. Veteran teachers are defined as returning teachers in the base year. The table includes only veteran and novice teachers retained through the 2012–2013 school year. The numbers of novice teachers included in the table represent between 39 and 78 percent of all first-year teachers in a given base year. The numbers of veteran teachers included in the table represent between 60 and 79 percent of all returning teachers in a given base year.

Compared to the sample of all first-year DCPS teachers, the novice teachers retained through the 2012–2013 school year are more effective on average in their first year.

The standard errors used to determine statistical significance account for repeated observations of teachers over time.

* Statistically significant at the 5 percent level.

NA = not available

Limitations. Four limitations of our analysis affect the interpretations of our results. The first two limitations are similar to those from the previous chapter. First, we cannot definitively distinguish improvements in novice effectiveness from declines in veteran effectiveness. This is because year-to-year changes in average IMPACT scores for veteran teachers could be caused by changes to IMPACT rather than by changes in veteran effectiveness. Improvements with experience in one cohort could be caused by a decline in the effectiveness of veteran teachers in the cohort, whereas the same pattern in another cohort could be caused by an increase in the effectiveness of novice teachers. However, average IMPACT scores for teachers in our sample of veterans do not decline over time, suggesting that the relative improvements are caused by gains from experience for novice teachers. The second limitation is that trends in IMPACT score gaps between novice and veteran teachers could be caused in part by the changes DCPS made to the calculation of IMPACT scores, if these changes provided a relative benefit to one group.

The third limitation arises as a consequence of our choice to compare the same teachers within a cohort over time. This limitation affects the interpretation of differences in the amount of improvement across cohorts. Because we restricted novice cohorts to those teachers who

remained in DCPS through the 2012–2013 school year, earlier cohorts have had more time for teachers to leave DCPS, resulting in fewer of the novices remaining in the analysis. For example, if less-effective novice teachers leave DCPS before the end of their fourth year in DCPS, the second novice cohort could include less-effective teachers on average compared to the first novice cohort. Thus, the effectiveness one cohort of novice teachers may not be comparable to the effectiveness of the novices in the other two cohorts. This selection of teachers out of the analysis sample based on their tenure in DCPS could also lead to differences in the amount of improvement for novices relative to veterans.

To address this concern, we examined novice cohorts defined a different way, by requiring that novices in all three cohorts remain in DCPS for a minimum of one year after their first year.[20] Defining novice cohorts in this way means that teachers we include in all three cohorts had the same amount of time to leave DCPS. Appendix Table A.2 reports both the gaps in effectiveness between novices and veterans and improvements over the first two years for these groups of teachers. The improvements for these novice teachers support the finding that the first cohort may have improved more slowly than the second and third cohorts. The first cohort defined this way improved by 5 IMPACT points compared to veterans between the end of their first and second years in DCPS (row 1, column 3 of Appendix Table A.2). The same improvements for the second and third novice cohorts defined this way were more than three times larger (rows 2 and 3, column 3).

Finally, our findings are also limited by the available data. We have IMPACT scores over four years for the first cohort of novice teachers, but over only two years for the third cohort. As additional years of data become available in the future, revisiting the changes in the novice-veteran gaps will provide important information about whether these improvements continue for the first three novice cohorts since IMPACT was implemented and whether improvements for subsequent cohorts are similar.

[20] It is not possible to require that teachers in all three cohorts remain in DCPS for additional years after the base year because we only have two years of IMPACT scores for teachers in cohort 3.

VI. CONCLUSION

Our findings describe important trends over time in the retention and effectiveness of DCPS teachers that may be relevant for education policymakers and practitioners. However, the questions we address in this report are not intended to provide conclusive evidence on whether IMPACT has been successful in meeting its goals, or to isolate its effects on students or educators. Such an analysis may not be possible, given that many policies coincided with the implementation of IMPACT in DCPS.

First, we investigated trends in retention rates for DCPS teachers. We found that the number of teachers dismissed under IMPACT for having an ineffective rating or consecutive minimally effective ratings was highest in year 2, and was under 3 percent of all teachers in years 3 and 4. The higher rate of dismissals in year 2 was driven by teachers with two consecutive minimally effective ratings; teachers in year 1 were not dismissed for this reason. The retention rates of the most effective teachers in DCPS were above 80 percent and did not change across the three years of IMPACT we studied.

Second, we investigated the effectiveness of teachers who left DCPS and of the teachers hired to replace them. Compared to core DCPS teachers, teachers who left DCPS at the end of year 3 of IMPACT were relatively more effective than the teachers who left after year 1. This was the case for teachers who met IMPACT dismissal criteria as well as for those who did not meet the criteria. In part, this change in the effectiveness of leavers can be explained by changes in the dismissal criteria between years 1 and 2 of IMPACT.

We also found that new hires were less effective than core DCPS teachers but more effective than the teachers who left DCPS in years 1 and 2 of IMPACT. In year 3 of IMPACT, the leavers and new hires were similarly effective; this was due primarily to an increase in the effectiveness of the leavers relative to core teachers. If the teachers who leave DCPS continue to become relatively more effective compared to new hires, it might be difficult for DCPS to develop and maintain an effective teaching workforce. However, even if leavers are more effective than new hires, this would not necessarily imply that DCPS teachers are becoming less effective. This is because both new teachers and veteran teachers who remain in DCPS may become more effective with additional experience.

In our final investigation, we found that novice teachers improved with experience in DPCS relative to how much veteran teachers improved with experience. On average, novice teachers improved relative to veteran teachers after one additional year of experience in DCPS for both groups and continued their relatively high improvement with a third or fourth year in DCPS. The largest improvement for novice teachers was between the end of their first and second years in DCPS. We also found evidence that suggests the two most recent cohorts of novice teachers we studied—those hired in years 2 and 3 of IMPACT—may be improving more quickly relative to veteran teachers than the cohort hired in year 1. This could suggest that IMPACT policies have helped new teachers develop, or that DCPS has hired teachers with more potential for improvement, although other explanations could be responsible for this finding. Additionally, we can only measure teacher effectiveness over three years for new teachers hired in year 2, and over just two years for those hired in year 3, so it remains to be seen whether the higher levels of effectiveness can be sustained and if the novice cohorts continue to improve.

Based on our analysis, we cannot say whether teachers in DCPS are more or less effective on average four years after IMPACT. This is primarily because we cannot measure changes in the effectiveness of core DCPS teachers—those who were teaching in DCPS prior to IMPACT and continued to teach through the 2012–2013 school year. We do find evidence that DCPS has retained many of its most effective teachers and has hired new teachers that are at least similarly effective to the teachers they replaced. These new teachers have improved over time with additional experience more than veteran teachers improved with experience. However, teachers who leave DCPS have become more effective over time relative to the comparison group used in this analysis. Perhaps offsetting the increased effectiveness of leavers, the most recent cohorts of teachers hired to replace teachers who leave appear to have improved more quickly with experience. Given this evidence of the evolving trends in the effectiveness of DCPS teachers since IMPACT, it will be important to continue to monitor these trends as DCPS continues to implement IMPACT.

REFERENCES

Dee, Thomas and James Wyckoff. "Incentives, Selection, and Teacher Performance: Evidence from IMPACT." Working paper # 19529. Cambridge, MA: National Bureau of Economic Research, 2013.

Education Consortium for Research and Evaluation. "Evaluation of the DC Public Education Reform Amendment Act: Report No. 1, School Year 2010–2011." July 2013a.

Education Consortium for Research and Evaluation. "Evaluation of the DC Public Education Reform Amendment Act: Report No. 2, School Year 2011–2012." September 2013b.

Hanushek, Eric A., John F. Kain, Daniel M. O'Brien, and Steven G. Rivkin. "The Market for Teacher Quality." Working paper # 11154. Cambridge, MA: National Bureau of Economic Research, 2005.

Isenberg, Eric, and Heinrich Hock. "Design of Value-Added Models for IMPACT and TEAM in DC Public Schools, 2010–2011 School Year." Washington, DC: Mathematica Policy Research, May 2011.

Isenberg, Eric, and Heinrich Hock. "Measuring School and Teacher Value Added in DC, 2011–2012 School Year." Washington, DC: Mathematica Policy Research, August 2012.

Isenberg, Eric, and Elias Walsh. "Measuring Teacher Value Added in DC, 2012–2013 School Year." Washington, DC: Mathematica Policy Research, January 2014.

Harris, Doug N., and Tim R. Sass. "What Makes for a Good Teacher and Who Can Tell?" Unpublished manuscript, 2010.

Rockoff, Jonah E. "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data." *American Economic Review*, vol. 94, no. 2, 2004, pp. 247–252.

APPENDIX: SUPPLEMENTAL RESULTS

Table A.1. Retention rates of teachers in DCPS, by school year

IMPACT Rating	2008–2009 (pre-IMPACT)		2009–2010 (IMPACT year 1)		2010–2011 (IMPACT year 2)		2011–2012 (IMPACT year 3)	
	Percent retained	Number of teachers	Percent retained	Number of teachers	Percent retained	Number of teachers	Percent retained	Number of teachers
Groups 1 Through 8 (all teachers)								
Highly effective (350 to 400 points)			89	536	89	480	89	725
Effective (250 to 349 points)			84	2,327	84	2,331	82	2,235
Minimally effective (175 to 249 points)			70	453	46	448	55	276
Ineffective (100 to 174 points)			0	62	0	56	0	34
All IMPACT ratings	71	3,142	82	3,378	78	3,315	81	3,270
Group 1 (teachers with IVA scores)								
Highly effective (350 to 400 points)			97	37	76	17	91	82
Effective (250 to 349 points)			87	289	83	301	79	251
Minimally effective (175 to 249 points)			76	123	50	127	69	96
Ineffective (100 to 174 points)			0	10	0	12	0	15
All IMPACT ratings	n.a.	n.a.	83	459	72	457	76	444
Groups 2 Through 8 (teachers without IVA scores)								
Highly effective (350 to 400 points)			88	499	89	463	88	643
Effective (250 to 349 points)			84	2,038	84	2,030	83	1,984
Minimally effective (175 to 249 points)			68	330	45	321	48	180
Ineffective (100 to 174 points)			0	52	0	44	0	19
All IMPACT ratings	n.a	n.a	81	2,919	79	2,858	81	2,836

Source: Mathematica calculations based on administrative data from DCPS.

Notes: The table excludes all records for 281 teachers with incomplete IMPACT ratings in one or more years. Overall retention rates that instead include these teachers are similar to those reported in the table. Otherwise, the table includes all group 1 through 8 teachers in each specified school year. IMPACT ratings are from the specified school year.

Teachers are considered to have exited DCPS if they do not have an IMPACT rating in the subsequent school year.

n.a. = not applicable

Table A.2. Trends in effectiveness of novice DCPS teachers who remained in DCPS for at least two years, relative to veteran teachers

Sample	Gap in effectiveness between novice and veteran teachers by novice experience (IMPACT points)		Improvement from first year to second year	Number of novice teachers	Number of veteran teachers
	First year in DCPS	Second year in DCPS			
Cohort 1: 2009–2010 teachers retained through 2010–2011	-16*	-11*	5*	898	1,850
Cohort 2: 2010–2011 teachers retained through 2011–2012	-19*	-2	17*	404	2,156
Cohort 3: 2011–2012 teachers retained through 2012–2013	-27*	-11*	16*	502	2,076

Source: Mathematica calculations based on administrative data from DCPS.

Notes: Novice teachers are defined as first-year DCPS teachers in the base year. Veteran teachers are defined as returning teachers in the base year. The table includes only veteran and novice teachers retained for one year after the base year. The numbers of novice teachers included in the table represent between 78 and 79 percent of all first-year teachers in a given base year. The numbers of veteran teachers included in the table represent between 77 and 82 percent of all returning teachers in a given base year.

A negative gap indicates that veteran teachers have higher IMPACT scores, on average, than novice teachers.

The standard errors used to determine statistical significance account for repeated observations of teachers over time.

* Statistically significant at the 5 percent level.

NA = not available

www.mathematica-mpr.com

Improving public well-being by conducting high quality, objective research and data collection

PRINCETON, NJ ■ ANN ARBOR, MI ■ CAMBRIDGE, MA ■ CHICAGO, IL ■ OAKLAND, CA ■ WASHINGTON, DC

MATHEMATICA
Policy Research