Evidence of Learning Gains at Seattle’s South Shore School

November 2013

Prepared for:
The League of Education Voters

ECONorthwest
ECONOMICS • FINANCE • PLANNING
Summary

The South Shore School, a public PreK-8 school in southeast Seattle, is the result of an innovative public/private partnership between the Seattle Public Schools district and the New School Foundation (now merged with the League of Education Voters Foundation). South Shore's program is based on rigorous, evidence-based interventions, including high-quality and integrated pre-Kindergarten, small class sizes in early grades, and extra supports for students and faculty.

ECONorthwest conducted an evaluation of South Shore's impacts on student performance, comparing South Shore student performance to that of observably similar Seattle Public Schools students.

Examining student-level test data across multiple school years, ECONorthwest finds that attendance at South Shore is generally associated with improved scores on the WASL (Washington Assessment of Student Learning) and MSP (Measurements of Student Progress) assessments for grades 3-8 and an increased likelihood of meeting the state standards in math and reading. The South Shore effects are large and statistically meaningful, particularly in math.

The effects of attending South Shore are larger for students who enrolled during pre-Kindergarten or Kindergarten, and effects on achievement are consistent across racial and ethnic groups.
What is the South Shore School?

The South Shore School, a public PreK-8 school in southeast Seattle, is the result of an innovative public/private partnership between the Seattle Public Schools district and the New School Foundation that began in 1998. South Shore’s program is based on rigorous, evidence-based interventions, including high-quality and integrated pre-Kindergarten, small class sizes in early grades, and extra supports for students and faculty.¹

In 2011, the New School Foundation merged with the League of Education Voters (LEV) Foundation, which now administers an annual grant to South Shore and works statewide on policy, advocacy, and implementation related to PreK-3rd alignment projects.

In the 2011-12 school year, South Shore enrolled approximately 700 students in grades pre-K through 8. Students at South Shore are predominantly African-American and Asian and are drawn mostly from southeastern Seattle.

Purpose of this Report

LEV commissioned ECONorthwest to provide an updated evaluation of the effectiveness of South Shore’s program. ECONorthwest completed similar evaluations twice before, in 2009 and 2010.

In this report, as in our previous analyses, we present the results of an analysis that is more rigorous than simple cross-school comparisons, although it stops well short of the “gold standard” of randomized controlled trials. Specifically, we use student-level demographic and educational data provided by the Seattle Public Schools district (SPS) to compare South Shore students to similar students attending other SPS schools.

Our analysis was designed to answer the following questions:

- Do South Shore students perform better than observably similar SPS students?
- Are there meaningful differences in student performance between long-term students and new arrivals?

Methodology and Data Sources

Fundamentally, an evaluation of South Shore (or any school) consists of comparing the outcomes of South Shore students to the outcomes of students that did not attend South Shore. However, it is not particularly informative to simply compare the outcomes of all South Shore students to all other SPS students or the students attending other specific schools. Raw differences between South Shore outcomes and the outcomes at other schools reflect much more than differences in the educational environment. In addition to school effectiveness, student performance depends on individual, family, and other contextual characteristics. As such, it is possible (even likely) that two equally effective schools will have significantly different student outcomes due to large differences in the composition of their student populations.
The goal of this evaluation is to isolate the
effect of South Shore by accounting for
underlying differences that affect academic
performance, such as student demographic
characteristics, family situations, and other
contextual factors.

Researchers face two challenges when
undertaking such evaluations. First, school
performance data are noisy (i.e., they can
fluctuate fairly widely from year to year for
reasons unrelated to school quality).2 Such
fluctuations can reflect a variety of factors,
such as random differences among cohorts,
changes to the composition of the cohort as
students move in and out, or a dog barking
outside while students are completing an
assessment. These random fluctuations can
make it difficult to accurately measure
school performance. To minimize the
influence of random fluctuations, we
examine multiple years of data in each
analysis described in this report.

Second, no other school (or collection of
schools) provides a perfect comparison for a
given school and its programs. That is, there
is no school that is identical to South Shore
in every way except for its South Shore
programs. Student populations differ across
schools, and the available data do not
capture all relevant differences between
students. For example, students with more-
educated parents tend to perform better in
school. As such, if the data analyzed do not
include and account for parents’ education,
an analysis might falsely attribute certain
outcomes stemming from differences in
parents to differences in school quality.

The ideal way to accurately measure school
performance is through a randomized
controlled trial. In the case of South Shore,
using this methodology would involve
randomly assigning a large population of
students to South Shore (the treatment
group) or to other SPS schools (the control
group). Under this scenario, there would
ideally be no underlying, unobserved
differences between the treatment and
control groups (i.e., membership in each
group would be random).3 As such, any
differences between the two groups could
be clearly attributed to attendance at South
Shore.

In practice, students choose (or their parents
choose for them) to attend South Shore,
either explicitly by selecting South Shore in
SPS’s school choice program, or implicitly
by choosing to live within the South Shore’s
neighborhood-school boundaries. Because
of this “self-selection” into South Shore, we
cannot entirely rule out the possibility that,
compared with SPS students overall, South
Shore students have underlying differences
that affect their performance in school.
However, we can minimize the possibility
that our results reflect such underlying
differences by obtaining additional
information about South Shore students and
SPS students.

Specifically, we can use statistical
information about characteristics that are
typically correlated with performance on
achievement tests, such as race,
socioeconomic status, and attendance, to
statistically control for the effects of these
underlying differences using regression
analysis.4 This method allows us to compare
groups of students who have mostly similar
observable, relevant, and measurable
characteristics except for their exposure to
“treatment,” in this case, South Shore
attendance.
But even after accounting for observable differences, differences between South Shore students and non-South Shore students could remain. While our analytical methodology is subject to more error than a randomized controlled trial methodology, it nevertheless helps us understand the relationships between relevant characteristics that we can observe and measure. Appendix B includes additional detail about comparison populations and effect measurement.

**Compared to What?**

**Two Approaches to Examining Student Performance**

Even after accounting for differences among students, evaluating school performance requires some criteria to determine what constitutes good performance. That is, student and school outcomes need to be placed in a larger context that helps people understand what the results mean. Typically, there are two ways to judge performance:

- **Absolutely**—This approach compares student performance to a set of predetermined benchmarks
- **Relatively**—This approach compares student performance to the performance of other students in the same grade and year

Each of these approaches has important strengths and weaknesses. If the benchmarks are clearly understood, the absolute approach provides an easily interpretable metric for school performance—X% of students met the defined standards. This simple comparison allows performance to be easily compared across time and space (assuming testing or other measurement tools—and the benchmarks—remain constant).

The problem with the absolute approach is that it ignores what is happening away from the benchmark. Students can exceed or fall short of the benchmark by wide margins. The absolute approach ignores this information (and thus creates incentives for schools to divert resources from those students who clearly fall above or below the line to those who are likely to fall near the line). It is possible for two schools to have identical shares of students meeting the benchmark, but drastically different overall performance. For instance, one school could have most of its students exceeding the benchmark by a wide margin, while the other could have most of its students just barely meeting the benchmark.

The relative approach helps address this problem. In the relative approach, the full distribution of student performance is used to evaluate student and school performance. Schools are evaluated based on their students’ average performance relative to the average performance of students in the whole population. In this type of metric, every student’s performance matters (not just those near the benchmark). Furthermore, because students are typically compared relative to those in the same year and in the same grade, this type of metric does not require absolute consistency in the quality of the testing (and associated benchmarks).

The problem with the relative approach is that it is “zero-sum,” so it may be difficult to interpret. When students and schools are moving relative to each other, for one student or school to move up, another must move down. As such, the relative approach...
does not identify any widespread gains (or losses) in student performance. Thus, unless the distribution of student performance remains constant over time, it can be difficult to understand the information provided by this approach. A student with the same level of understanding and performance can receive very different relative scores from year to year.

Neither of these approaches is sufficient by itself to provide a full picture of student and school performance, so throughout this report we examine both how South Shore students perform relative to statewide benchmarks and how they perform relative to other SPS students in the same grade and year.

**Acquiring and Organizing the Data**

ECONorthwest’s first task in conducting an evaluation of Seattle’s South Shore School was to acquire, assemble, and organize the necessary data. This section describes that process generally; Appendix A provides additional detail about how we created the databases for our study.

Fundamentally, our analysis consists of examining the outcomes (e.g., test scores) for South Shore students and comparing them to similar students who did not attend South Shore. To complete such an analysis, we needed to obtain de-identified, student-level longitudinal data for South Shore students as well as SPS students for the same time period. SPS provided these data in the following categories: demographics, historical enrollment, coursework, state assessments, attendance, and school choice data.

We requested and received a similar dataset from the Washington State Education Research and Data Center (ERDC). The ERDC data contains most of the same elements as the SPS data, with the addition of FRL status and the exception of school choice data.

**Assessment data**

In this evaluation, we focus on student performance on state assessments. The Washington Assessment of Student Learning (WASL) was Washington’s state test from 1997 until summer 2009. Students in grades 3-8 and 10 took the WASL each spring in reading and math. Students were also tested in writing in grades 4, 7, and 10, and science in grades 5, 8, and 10. Our previous reports considered only reading and math WASL scores because relatively few South Shore students had reached 4th and 5th grades.

In spring 2010, Washington State replaced the WASL with the Measurements of Student Progress (MSP) and the High School Proficiency Exams (HSPE). The MSP is for students in grades 3-8 for reading and math, and grades 4 and 7 for writing, and grades 5 and 8 for science. The HSPE measure the basic proficiency of high school students in reading and writing, and serve as the state’s exit exams in those subjects.\(^5\)

This analysis primarily uses WASL and MSP results, as only a small number of South Shore students (i.e., the 2002-03 Kindergarten cohort) were in 9th grade in 2011-12, the latest year for which we have data.

As discussed above, we examine performance both absolutely and relatively. To measure absolute performance, we
examine whether or not students met the state benchmark for reading or math in their grade. To measure relative performance, we calculate a standardized score—in standardized units (SUs)—for each student's test results. A standardized score reports how far, in terms of standard deviations on a normal curve, the student is from the average student in his or her grade and year. For the purposes of the charts and tables in this study, we analyze and display standardized WASL and MSP scores together.

Standardized scores (SU) are interpreted in a variety of ways based on contextual factors. For student-level interventions, What Works Clearinghouse considers effect sizes of at least 0.25 SU to be substantively important.

Demographics, attendance, and school choice

To account for potential differences between students attending South Shore and those attending other SPS schools, we used the following information:

- Sex
- Ethnicity (as coded in seven categories: Asian, Black, Hispanic, Multi-ethnic, Native American, Pacific Islander, and White)
- Grade
- An indicator equal to one if a student is identified as residing in a non-English speaking household
- An indicator equal to one if the student is classified as living with both parents
- An indicator equal to one if the student has an IEP classification
- An indicator equal to one if the student is classified as gifted

- The number of SPS elementary or middle schools attended (a proxy for student mobility)
- School year
- School choice ranking: Seattle students are allowed to rank their preferences for the school they attend. Actual school assignment is largely a function of geography and sibling attendance. In some of our analyses, we limit the sample of students to include only those attending South Shore or who ranked it as their first choice at some point.

Data Limitations

The data used in this analysis have potential limitations that should be taken into consideration in interpreting the results. First, if the WASL or MSP tests do not accurately measure student ability, the results of the analysis will be affected. Second, the small sample sizes inherent in the South Shore population increase the possibility of measurement and sampling error for key variables.

Two other important analytical issues must be considered. First, as discussed previously, this analysis is subject to selection bias because students are not randomly assigned to either South Shore or any other school. The factors that cause a student to apply to and attend South Shore may be a result of unobserved underlying differences in those students compared to those who do not apply. Therefore, the analytical results may reflect unobserved differences in student characteristics rather than, or in addition to, the observed differences that are accounted for in the regression analysis.
Second, the analysis does not measure the effects of specific South Shore programs. We simply calculate the total effects of South Shore on student achievement, accounting for the observable factors outlined above. These effects are what parents are most likely to be concerned with. That is, they reflect the expected change in student performance contingent on attending the school. We do not attempt to explain precisely which policies produce which effects, although this is what administrators are most likely to care about. We also cannot rule out the possibility that the results may reflect differences in context (or peer effects) and not policy.

Findings

Overall, we find that South Shore students perform at levels significantly higher than would be predicted based solely on their observable characteristics.

In our previous studies, we reported that, relative to similar students attending other schools, South Shore 3rd and 4th grade students were more likely to meet the WASL benchmarks in both reading and math, and they had higher standardized test scores. The sizes of the effects were larger for students who enrolled during pre-Kindergarten or Kindergarten compared to those who enrolled later. And South Shore’s effects were consistent across ethnic groups.

The current analysis finds similar results. Effects vary by grade, but South Shore students have higher test scores than demographically similar SPS students in math, reading, and writing. South Shore effects are particularly strong in math.

These effects are generally consistent across subgroups. We also find that students who attend South Shore in pre-Kindergarten and Kindergarten perform better than students who enroll in later grades and better than their demographically similar SPS counterparts, particularly in math.

We describe a simple set of results in the main body of the report. Appendix B contains additional analyses that use alternative estimation strategies and assumptions. The pattern of results described in Appendix B follows the same pattern as the results described here; however, the precise magnitude of the effects does vary somewhat.

Unadjusted Performance

Looking at overall performance on math and reading assessments without accounting for underlying differences between groups, the average South Shore student scores below the average SPS student in most grades (see Figure 1). In general, the average South Shore student performs slightly better than the average SPS student in 3rd grade math and 8th grade reading, and below the SPS average in every other grade.
Figure 1: Deviation of South Shore math and reading scores from Seattle Public Schools average

Source: ECONorthwest analysis of Seattle Public Schools data.

Observable Underlying Differences

When comparing South Shore students with students at other schools, it is important to understand the observable underlying characteristics of the groups, because these may explain some of the differences in test scores. It is likely that at least some of the observed differences in school performance reflect differences in sociodemographic characteristics of South Shore students.

Figure 2 shows comparisons of demographic characteristics for South Shore students compared to SPS students overall.

Compared to SPS overall, South Shore has a larger share of black students and Asian students and a smaller share of white students and Hispanic students. While the SPS population is about 40 percent white, the South Shore population is about 13 percent white.

South Shore also has a slightly larger proportion of students living in non-English-speaking homes, and a significantly larger share of students receiving free or reduced price lunch.
Comparison of Outcomes Accounting for Observable Differences

In this section we examine differences in performance between South Shore students and students at all SPS schools after considering the underlying observable differences discussed above. To account for these potential differences, we estimate the effect of attending South Shore by estimating regression equations that include a variety of controls for student characteristics. By doing so, we can assert with greater confidence that the differences between groups reported here are due to South Shore treatment effects and not to other observable differences.

3rd-8th grade assessment performance

Examining student-level test data across multiple school years, ECONorthwest finds that attendance at South Shore is associated with improved scores on WASL and MSP assessments in several grade and subject combinations, as well as an increased likelihood of meeting the state benchmarks in math and reading. These effects are large and statistically meaningful, particularly in math. Specific findings are discussed below.

Focusing on 3rd grade students, the first grade for which we have assessment scores and for which we have seven cohorts of students (about 375 students in total), we find that students enrolled at South Shore for 3rd grade score significantly higher in math (0.30 SU) and moderately higher in
reading (0.11 SU) than demographically comparable students attending other SPS schools (see Table 1 and Figures 3 and 4).

South Shore students are also significantly more likely to meet or exceed WASL and MSP benchmarks in both reading and math. Table 1 shows the share of South Shore 3rd graders we would expect to meet or exceed the benchmarks in math and reading based solely on student demographic characteristics, compared to the actual share of students meeting the benchmarks. The table indicates that South Shore students scored 12 percentage points above their expected scores in math and 5 percentage points above their expected scores in reading. Put another way, for every 100 South Shore 3rd graders, an additional 12 students met or exceeded the 3rd grade math benchmark and an additional 5 students met or exceeded the reading benchmark than would be expected had these students attended a different school.

The effects of South Shore vary as students progress through the school’s grades, but South Shore students generally perform better than expected. In math, South Shore students score better than expected in all grades except 6th. In reading, South Shore students score near expectation in all grades except 3rd and 8th.

To date, South Shore has had too few 9th grade students to allow us to confidently analyze the effects of South Shore attendance beyond 8th grade. Also, it is generally more difficult to interpret an analysis for later grades because a significant number of older students transferred into South Shore in later grades and thus did not participate in many of the programs offered by South Shore. As full cohorts continue to age through the South Shore program, their performance data will illuminate the longer-term impacts of South Shore attendance.

### Table 1: South Shore effects on math and reading scores, and predicted and actual shares of South Shore students meeting state standards, by grade

<table>
<thead>
<tr>
<th>Math at South Shore School</th>
<th>3rd</th>
<th>4th</th>
<th>5th</th>
<th>6th</th>
<th>7th</th>
<th>8th</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standardized score difference</td>
<td>0.30</td>
<td>0.19</td>
<td>0.15</td>
<td>-0.10</td>
<td>0.05</td>
<td>0.31</td>
</tr>
<tr>
<td>Predicted meet rate</td>
<td>61%</td>
<td>51%</td>
<td>56%</td>
<td>60%</td>
<td>58%</td>
<td>53%</td>
</tr>
<tr>
<td>Actual meet rate</td>
<td>73%</td>
<td>62%</td>
<td>56%</td>
<td>46%</td>
<td>62%</td>
<td>58%</td>
</tr>
<tr>
<td>Difference (percentage points)</td>
<td>12</td>
<td>11</td>
<td>0</td>
<td>-14</td>
<td>4</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reading at South Shore School</th>
<th>3rd</th>
<th>4th</th>
<th>5th</th>
<th>6th</th>
<th>7th</th>
<th>8th</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standardized score difference</td>
<td>0.11</td>
<td>-0.03</td>
<td>-0.03</td>
<td>0.02</td>
<td>0.04</td>
<td>0.37</td>
</tr>
<tr>
<td>Predicted meet rate</td>
<td>67%</td>
<td>65%</td>
<td>65%</td>
<td>70%</td>
<td>59%</td>
<td>71%</td>
</tr>
<tr>
<td>Actual meet rate</td>
<td>72%</td>
<td>66%</td>
<td>61%</td>
<td>68%</td>
<td>60%</td>
<td>79%</td>
</tr>
<tr>
<td>Difference (percentage points)</td>
<td>5</td>
<td>1</td>
<td>-4</td>
<td>-2</td>
<td>1</td>
<td>8</td>
</tr>
</tbody>
</table>

Note: Standardized score differences (South Shore effects) are the deviation from predicted performance, accounting for individual characteristics. Meet rate is the share of students meeting the state proficiency standard. Source: ECONorthwest analysis of Seattle Public Schools data.
Figure 3: South Shore effects on math scores, by grade (deviation from expected performance, accounting for individual characteristics)

Source: ECONorthwest analysis of Seattle Public Schools data.

Figure 4: South Shore effects on reading scores, by grade (deviation from expected performance, accounting for individual characteristics)

Source: ECONorthwest analysis of Seattle Public Schools data.
The effects of starting at South Shore in Pre-K or Kindergarten

A key component of South Shore is high-quality pre-Kindergarten and Kindergarten. In this section, we examine the effects of these programs. Unfortunately, we do not currently have a reliable measure of student performance prior to third grade. Instead, we look at the effects of the pre-Kindergarten and Kindergarten programs by examining WASL and MSP scores for 3rd graders (the first grade level with test score data) by number of years attending South Shore.

Figure 5 summarizes the results of our analyses. Students who enrolled at South Shore in pre-Kindergarten or Kindergarten significantly outperformed demographically similar students attending other SPS schools in math and moderately outperformed similar students at other schools in reading. The strongest effect is for students who enrolled in the South Shore pre-Kindergarten program: math scores are 0.39 SU higher than scores for all SPS 3rd graders with similar consistent enrollment in a single school. The strongest South Shore reading effect is for students who enrolled during Kindergarten: these reading scores are 0.18 SU higher than scores for all SPS 3rd graders with similar enrollment characteristics.

These results indicate that the effects of attending South Shore are larger for students who enrolled during pre-Kindergarten or Kindergarten. While this pattern of results does not conclusively prove that South Shore’s pre-Kindergarten and Kindergarten programs are responsible for these differences (the differences could reflect other unobserved differences between early and late enrollees), our results are certainly consistent with the belief that attending South Shore’s pre-Kindergarten and Kindergarten programs significantly improves subsequent student performance.

The effect of South Shore on subpopulations

As discussed previously, South Shore attendance increases students’ average scores and the likelihood of meeting the state assessment benchmarks. An important question is whether that effect is the same for all groups of students, or if it has greater impacts on some groups compared to others.

Throughout SPS, African-American students score below non-African-American students in both reading and math. The pattern generally holds at South Shore as well (although the exact magnitudes of the differences are not identical). However, relative to demographically similar African-American students attending other schools, African-American students attending South Shore have higher average scores and are more likely to meet the state benchmarks in math and reading. For instance, in 3rd grade, South Shore African-Americans score 0.26 SU higher than demographically similar African-Americans at other SPS schools in math and 0.14 SU higher in reading. These effects are roughly in line with overall differences between South Shore 3rd graders and non-South Shore 3rd graders described in the previous section.

Figure 6 shows estimates for South Shore effects on 3rd grade math scores for various subgroups. Effect sizes vary, but they are all positive and not substantially different than one another (especially given the small
sample sizes for some groups). Among the subpopulations depicted here, white 3rd graders at South Shore have the largest estimated South Shore effect, scoring 0.43 SU higher than demographically similar white students at other SPS schools.

Figure 5: South Shore effect on math and reading scores, by students' earliest grade of enrollment at South Shore

![Bar chart showing South Shore effect on math and reading scores by grade level.]

Note: South Shore effect is the deviation from expected performance, accounting for individual student characteristics. Source: ECONorthwest analysis of Seattle Public Schools data.

Figure 6: South Shore effects on 3rd grade math scores for various subpopulations

![Bar chart showing South Shore effect on 3rd grade math scores for different subpopulations.]

Source: ECONorthwest analysis of Seattle Public Schools data.
South Shore compared to other SPS elementary schools

To provide context for the South Shore effect sizes, ECONorthwest conducted similar analyses for each SPS elementary school and estimated a school effect size for each school. As described in greater detail in Appendix B, one can compute such effects in a variety of ways, but regardless of approach, South Shore ranks highly.

For 3rd grade math, the South Shore effect ranks between 1st and 6th (out of 76 schools). For 3rd grade reading, the South Shore effect ranks between 1st and 26th. As an example, Figure 7 depicts one set of rankings for 3rd grade math, where South Shore is ranked 4th. As described in Appendix B, this range in rankings reflects differences in what types of school-level characteristics are included in the calculation. Each school’s estimated effects and rankings are sensitive to the outcome examined and control variables included in the analysis.

Finally, the differences in the estimated effect size among most of the other “top” schools are not statistically significant; that is, we cannot say with confidence that differences between closely ranked schools do not reflect randomness in the data. Thus, while we can reasonably assert that schools with large positive effects are improving student outcomes, we are less comfortable asserting that there is a meaningful difference between the effects of South Shore in particular and other schools with similar estimated effect sizes.

Figure 7: Distribution of SPS school effects for 3rd grade math scores (deviation from expected performance, given observable student characteristics)

Note: School effects for this particular figure are random effects estimates for a regression similar to the one described in the text, but including a school random effect for each school. See Appendix B for additional detail.
Source: ECONorthwest analysis of Seattle Public Schools data.
Conclusion

Examining student-level test data across multiple school years, ECONorthwest finds that attendance at South Shore is generally associated with improved scores on the WASL and MSP assessments for grades 3-8 and an increased likelihood of meeting the state standards in math and reading. The South Shore effects are large and statistically meaningful, particularly in math.

These effects are generally consistent across subgroups. In addition, students who attend South Shore in pre-Kindergarten and Kindergarten perform better than students who enroll at South Shore in later grades and better than their demographically similar SPS counterparts, particularly in math.