Appendix

In Table A-1 below, we describe the data collected by location. These data are described in greater detail below.

Table A-1

<table>
<thead>
<tr>
<th>Location</th>
<th>School Year Charter Schools Began Operating</th>
<th>Years of K-12 Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chicago</td>
<td>1997-98</td>
<td>1997-98 through 2006-07</td>
</tr>
<tr>
<td>Denver</td>
<td>1995-96</td>
<td>2001-02 through 2005-06</td>
</tr>
<tr>
<td>Milwaukee</td>
<td>1996-97</td>
<td>2000-01 through 2006-07</td>
</tr>
<tr>
<td>Ohio</td>
<td>1998-99</td>
<td>2004-05 through 2007-08</td>
</tr>
<tr>
<td>Texas</td>
<td>1996-97</td>
<td>1994-95 through 2003-04</td>
</tr>
</tbody>
</table>

**Chicago**

Chicago Public Schools provided the project team with student-level race/ethnicity information, test score data, as well as the school of attendance and grade enrolled for each year 1997-98 through 2006-07 school years. The provided grade range of test scores varied by year. For the 1997-98 through 2000-01 school years, math and reading test scores were provided for students in grades 1-8. For 2001-02 through 2006-07, math and reading test scores were provided for students in grades 3-8.

**Denver**

Denver Public School District provided the project team with student-level race/ethnicity information, test score data, as well as the school of attendance and grade enrolled for each year from 2000–01 through 2005–06 school years. In addition, the district provided a list of school
identifiers of each charter school and the year in which the school was established. Using this list combined with the school identifiers for each student, we were able to indicate whether a student attended a charter school for each year. However, it should be noted that because a substantial portion of students did not have school identifiers included in the 2000-01 school year, we deleted this year from our analysis and examined data from the 2001-02 through 2006-07 school years only.

The grade range of test scores provided varied across reading and math. For reading, test scores were provided for grades 3 through 10 for each of these years. For math, in the 2001–02 through 2003–04 school years, test scores were provided in grades 5 though 10. In the 2005–06 school year, test scores were provided in grades 3 through 10. We were therefore able to track the student achievement progress more completely in reading than in math.

**Milwaukee**

The Milwaukee public school district provided student-level test and demographic data from the 1997-98 through 2006-07 school years. During the course of the panel, state tests switched from the Terra Nova to the Wisconsin Knowledge and Concepts Examination (WKCE) assessment, which over the last few years has incorporated ever more state-developed items. Therefore, to enable us to measure student gain scores, we created z-scores using scale scores by year, grade level, and test subject.

We could not use all of the data that the district provided for a number of reasons, however. First, the data from 1997-1999 were too sparse to estimate gain scores, so we limited the analyses to the 2000-01 through 2006-07 school years. Second, although the district provided test data for students in grades 2-10 in math and reading, there were too many missing scores in grade 2 to estimate the models. Therefore, we focused the analyses on math and reading achievement in grades 3-10, which provided us with gain scores for grades 4-10. Third, we obtained data for only those charter schools that the Milwaukee Public School system chartered (which comprise the majority of charter schools in Milwaukee) and therefore were unable to incorporate into this study schools chartered by other entities (i.e., "independent" charter schools). Finally, it is important to note that the data we analyzed are on tested students only.

**Ohio**

State of Ohio provided the project team with student-level race/ethnicity information, test score data, as well as school of attendance and grade enrolled from year 2003-04 through 2007-08. However, these variables were not provided for all students for each year, but the subset of students tested in the particular year. In addition, which grades were tested varied by year and in some years students were tested in the fall. Because a small portion of students was actually tested in the fall and because it would have meant that we would not have consistent intervals of testing periods for all students, we eliminated the fall test. Furthermore, in the 2003-04 school year, only 3rd graders were tested in reading and no students were tested in math. Also, only a portion of 3rd graders were tested in reading in the 2003-04 school year. Therefore, we also eliminated all 2003-04 school year data from our analysis.

We also performed several checks on the data and dropped records outside of our population as well as records that had conflicting information. In the first cut of the data, we removed all records for subjects other than math and reading. We dropped all records for a student within math and reading if the student had more than two records for a single subject. In total, we dropped less than two percent of students in each year by selecting out our population.
and discarding student records with conflicting data. As noted above, in some cases, the research team was given both spring and fall tests for an individual student. Of the students with two records in a subject in a year, we kept a single student record for each subject when information on the student was consistent and dropped both student records when the data were incongruent. We dropped both records if: (1) gender, race, or school were discrepant; we had two different spring scores or two different fall scores for a student; or (3) neither record had school data. We dropped inconsistent student records across math and reading within a year—when gender, race, or school did not match, and we dropped student records with inconsistent gender data across years.

In the 2004-05 school year, the state provided math test scores in grades 3, 7, and 8 and reading test scores in grades 3, 4, 5, and 8. In the 2005-06 through 2007-08 school years, the state provided reading test scores for grades 3 through 8. It should also be noted that for many students, the state provided both raw scores and scaled scores. However, we only had either the raw or scaled scores (not both) for a small subset of students. To maximize the scores, we normalized the raw and scaled scores by year and grade. For students missing scaled normalized scores, we used the raw normalized values. It should be noted that the raw and scaled scores had a high correlation of 0.87 across students that had both scores.

Philadelphia

Philadelphia Public School District provided the project team with student-level race/ethnicity information, test score data, as well as the school of attendance and grade enrolled for each year from 2000–01 through 2006–07 school years. In addition, the district provided a list of school identifiers of each charter school and the year in which the school was established. Using this list combined with the school identifiers for each student, we were not only able to identify whether a student attended a charter school for each year. In the period under examination (2000-01 through 2006-07), students in Philadelphia took three kinds of annual achievement tests in reading and math, varying with the school year and grade as listed below:

- Pennsylvania System of School Assessment (PSSA) tests for math and reading for grades 5, 8, and 11 annually beginning in spring 2001 and grades 3 through 8 and 11 in spring 2006 and 2007;
- Stanford 9 tests in math and reading in grades 3, 4, 7, and 10 in spring 2001 and spring 2002.1
- Terra Nova tests in math in grades 2 through 10 annually in the springs of 2003 through 2005 and in grades 2, 9, and 10 in spring 2006.
- Terra Nova tests in reading in grades 1 through 10 annually in the springs of 2003 through 2005 and in grades 1, 2, 9, and 10 in spring 2006.

Although we were able to get the Stanford 9 test results for non-charter students in 2001 and 2002, we were unable to get these data for charter students. However, it is still valuable to have these data in these years because it can help track the performance of students who later enrolled in charter schools.

---

1 In the spring of 2002, the Stanford 9 4th grade test was only administered to K-4 schools and not to K-5 or K-8 schools (email correspondence with Philadelphia School District Director of Accountability Mike Schlesinger, February 16, 2008).
Because there is no consistent scale across the various tests on which to gauge absolute changes in student achievement over time, we convert all scaled test-score results into rank-based z-scores, by year and grade, with a mean of zero and a standard deviation of one. Specifically, we sort all student scores by rank and then convert them to z-scores that are normed across the entire district-wide population of tested students in that subject and grade. This conversion does not require that students have the same rank on one test as on another, but it assumes that differences in the distribution of students on different tests are not correlated with charter status. Random differences in student ranks across different tests would introduce noise, but not bias, to the analysis. The conversion of scaled scores to rank-based z-scores means that we cannot make claims about the absolute amount of learning in one school or another (lacking a psychometrically valid developmental scale), but it permits an examination of changes in rank with fewer assumptions than would be needed under other kinds of scaling. In cases in which students took both the Terra Nova (TN) and PSSA, we used the PSSA because it is the state accountability measure and, in recent years, has been administered in more grades.

San Diego
San Diego School District provided the project team with student-level race/ethnicity information, test score data, an indicator of whether the student attends a charter schools, as well as the school of attendance and grade enrolled for each year from 1997-98 through 2006–07 school years. For each school year and for both reading and math, test scores of students were provided in grades 2 through 11.

Texas
The State of Texas provided the project team with student-level race/ethnicity information, test score data, and school of attendance and grade enrolled from year 1995-96 through 2003-04. These data include math and reading test scores for students in grades 3-8 in all years for all public school students, including students in Texas charter schools. Using these data along with school identifiers, we were able to follow each student over time as they transitioned between schools. For 2001-02 and prior school years the test scores were the math and reading scaled scores on the Texas Assessment of Academic Skills (TAAS), and for 2002-03 and 2003-04 the test scores were the math and reading scaled scores on the Texas Assessment of Knowledge and Skills (TAKS). Because these tests are on different scales, we converted all scores to rank-based z-scores, by year, test, and grade.

---

2 For further discussion of the use of rank-based z-scores, see Gill et al. (2005).