

# The Impact of School Choice Reforms on Student Achievement

Gary Miron, Stephanie Evergreen, and Jessica Urschel

The Evaluation Center, Western Michigan University

March 2008



## **EPRU** | EDUCATION POLICY RESEARCH UNIT

**Education Policy Research Unit**  
Division of Educational Leadership and  
Policy Studies  
College of Education, Arizona State  
University  
P.O. Box 872411, Tempe, AZ 85287-  
2411  
Telephone: (480) 965-1886  
Fax: (480) 965-0303  
E-mail: [epsl@asu.edu](mailto:epsl@asu.edu)  
<http://edpolicylab.org>

**Education and the Public Interest Center**  
School of Education,  
University of Colorado  
Boulder, CO 80309-0249  
Telephone: (303) 492-8370  
Fax: (303) 492-7090  
Email: [epic@colorado.edu](mailto:epic@colorado.edu)  
<http://education.colorado.edu/epic>

- This policy brief is available online at: <http://epsl.asu.edu/epru/documents/EPSSL-0803-262-EPRU.pdf>

One of a series of Policy Briefs made possible by funding from the Great Lakes Center for Education Research and Practice.

## **The Impact of School Choice Reforms on Student Achievement**

Gary Miron, Stephanie Evergreen, and Jessica Urschel

The Evaluation Center, Western Michigan University

### *Executive Summary*

This policy brief closely examines and summarizes the evidence regarding school choice and its impact on student achievement. After surveying studies across various choice forms, we selected 87 based on specific criteria. Each of these has been analyzed and assigned impact and quality ratings. Impact ratings indicate whether the choice schools' student performance was better or worse than comparison groups'; quality ratings reflect a study's score on a weighted scale that assesses six dimensions of research design. Results for both impact and quality ratings are summarized and mapped to facilitate reference and comparisons.

Key questions addressed in this policy brief were: What is the relative scope and quality of empirical research on school choice and student achievement? What are the overall conclusions that can be drawn from empirical research on school choice and student achievement?

Overall, the existing body of research on school choice reveals a mixed picture, with some studies suggesting positive impacts, and others indicating negative impacts. Large differences appear across school choice types in terms of the amount of research available, the overall quality of the research, and the conclusions the research supports. Voucher studies, generally of high quality, indicate a slightly positive impact, particularly for African American students. Studies of home schooling are few and fairly weak, with mixed or positive impact findings. There are also few studies of inter-, intradistrict choice and magnet programs, with mixed quality and impact scores. Charter school studies are most numerous, but their quality is also mixed; they indicate that charters generally perform similarly to traditional public schools, a finding that has not changed with time or with the addition of newer, higher-quality studies in recent years.

Aggregate findings across types of school choice hide considerable differences. Within all school choice models, there are certainly successful schools or cases as well as seriously flawed ones. At a macro level this policy brief can tell us what the body of research says and whether these models are worth replicating. Nevertheless, all forms of school choice could benefit from a better understanding of factors leading to success within particular schools or groups of schools.

## Recommendations

The mixed findings and quality of the studies analyzed leads to the following recommendations, which may help generate a better informed context for future research and policy design.

- **Improve research on school choice.** Attention to methods should include care in using lottery lists to stimulate random assignment, more matched student designs as state assessment systems improve and expand, and more longitudinal studies. More research is also needed on home schooling and on differences within and among forms of school choice. All researchers should be sure to articulate research design and limitations clearly, and they should offer appropriate cautions to readers about interpreting findings.
- **Improve the interpretation of research on school choice.** Policymakers and other research consumers should not evaluate school choice solely on the basis of outcomes from standardized tests. They should also be skeptical of sweeping conclusions and of press releases with no technical report to back them up. Instead, readers need to consider and reach their own conclusions about such methodological considerations as the population studied, sample size, and relevance of comparison groups. Studies weak in such areas, or that don't offer such detail, cannot be considered reliable. Research consumers should also consider whether the source of a study is an advocacy group—one that never sponsored a study with findings contrary to its position.
- **Reject any claims that research has produced definitive answers on school choice questions.** There are no definitive studies.

## **The Impact of School Choice Reforms on Student Achievement**

Gary Miron, Stephanie Evergreen, and Jessica Urschel

The Evaluation Center, Western Michigan University

### ***Introduction***

One of the most common—and most widely disputed—claims about school choice is that it will lead to improved student learning and performance on standardized tests. With growing interest in school choice and the expansion and improvement of state accountability systems, an increasing number of studies have taken up the question of whether student performance improves in the many school choice models relative to performance in comparable non-choice schools. This policy brief closely examines a wide range of evidence regarding school choice and its impact on student achievement.

Studying student achievement in school choice is complicated by a number of factors. First, there is limited evidence for many types of school choice. Generally, as is evident in other briefs in this collection, there are six choice models: vouchers/tuition tax credits, charter schools, cyber schools, home schooling, interdistrict choice, and intradistrict choice (including magnet schools and open enrollment plans). The scope of evidence on home schooling, cyber schools, and varied forms of inter- and intradistrict choice programs is very limited.

Another factor that complicates a synthesis of research evidence on school choice is that considerable weaknesses appear in available data. In fact, a majority of the studies available on school choice are limited by the researchers' access to student-level data and availability of relatively similar evidence that can be linked from year to year. In recent years, studies of school choice have been aided by the expansion of state assessment programs under the No Child Left Behind Act (NCLB), which now require testing in grades 3 to 8. A growing number of states are also moving to value-added accountability models that require student-level data sets. While excessive testing and preparation for testing is clearly taking away from time for instruction, a substantial evidence base that researchers and evaluators can draw upon is also accumulating.

A third factor that overshadows the body of evidence on school choice is the predominance of partisan researchers and activist organizations that carry out the research. Especially in the areas of home schooling, vouchers, and charter schools, the bulk of studies that find positive impacts in favor of school choice have been conducted by advocacy groups. That is not to say that research commissioned by advocacy groups and conducted by professional researchers will all result

## The Impact of School Choice Reforms on Student Achievement

in positive findings. What we can see, however, is that not one study released by groups advocating for school choice found that school choice had a negative impact on student achievement. Given the role of advocacy and opposition groups in pursuing research on student achievement to justify their agenda, it is not surprising to find that the two most polarizing and widely disputed forms of school choice (vouchers and charter schools) have been most studied.

In this paper we attempt to summarize what currently is known about the impact of various forms of school choice on student achievement. Key questions addressed are:

- What is the relative scope and quality of empirical research on school choice and student achievement?
- What are the overall conclusions that can be drawn from empirical research on school choice and student achievement?

The purpose of this brief is not to explore or explain the large differences in performance among diverse forms of school choice. Instead, we aim to provide answers to broad policy-related questions regarding whether the overall policies that promote school choice are likely to result in higher levels of student achievement. Our synthesis of findings follows the next section, which details our methodology and addresses such issues as how student achievement can reasonably be measured, which existing studies merit serious consideration, how the quality of studies can be assessed, and how findings can be reasonably combined into a “bottom line” statement of overall impact.

### *Methods*

The process of synthesizing existing research is dependent on several key methodological decisions. Most important are the selection criteria for studies to be included. That is, what characteristics make a study worth including, and how can the number of studies be limited in order to make a review a manageable—but still meaningful—undertaking? Most commonly, selection criteria deal with study design, quality issues, time limits (only studies between 2002 and 2004, for example), and/or geopolitical borders. The following discussion details selection criteria for this review as well as methods used to determine quality ratings and to synthesize findings. A discussion of limitations concludes this section.

#### **Selecting Studies**

In deciding which studies to include, we applied seven criteria.

1. Presence of a technical report offering a clear account of analytical procedures used.

## The Impact of School Choice Reforms on Student Achievement

2. Presence of aggregate analysis and conclusions. That is, we chose to exclude studies that would have required us to conduct our own analysis and draw our own conclusions based on others' data.
3. Use of standardized tests to measure student achievement. Standardized test results often provide the only way to compare achievement across a wide range of charter and noncharter schools.
4. Use of comparison groups. Any attempt to assess a given school's achievement impact requires some understanding of how choice students might have performed in the absence of choice schools. While randomized experiments with control groups are one of the most promising ways to determine impact, practical considerations have limited school choice researchers to observing "naturally" occurring comparison groups of non-choice schools.<sup>1</sup> In cases where studies included a variety of research designs, we considered only the methodologically strongest design.
5. Exclusion of duplicated studies. Only findings from the most recent study were included in cases where a particular author or group issued an update of earlier work using the same study design.
6. Exclusion of case studies or single school studies. These were excluded because it is unlikely findings can be reasonably generalized to the larger population of schools.
7. Exclusion of studies on school choice outside the United States of America. Although we recognize the importance of lessons that can be drawn from the experiences of other countries, we were concerned that we could not identify and consider a representative sample of international studies on school choice.

We considered and rejected two other selection criteria. First, we chose not to limit the time period because that would have resulted in few available studies for some choice models. And second, we chose not to exclude studies by advocacy or opposition groups, because doing so would have required making several difficult and subjective judgments. Instead, we have trusted that our quality rating methodology for weighting the evidence would—in part—reflect the inherent biases in research conducted by such groups.

### **Impact Ratings**

For the purpose of our analyses, the key finding for each study was its assessment of impact on student achievement. It is important to bear in mind that impact is not necessarily synonymous with absolute

## The Impact of School Choice Reforms on Student Achievement

achievement levels. For example, a magnet school with low test scores might still have significant positive impact if its students are gaining at a faster rate than similar students in other district schools. Conversely, a charter school with high test scores might have negative impact if its students are gaining more slowly than similar students in non-charter public schools. It is for this reason we considered comparison groups critical to assessing impact.

We assigned each study an impact rating according to the scale shown in Table 1. Positive values indicate that a study showed a particular school choice form to increase student achievement, and negative values indicate that it showed the model to decrease student achievement.<sup>2</sup>

**Table 1. Scale for Impact Ratings of Studies of Student Achievement in Diverse Forms of School Choice**

Scale Value	Description
2	Positive overall impact
1	Slightly positive overall impact
0	Mixed impact
-1	Slightly negative overall impact
-2	Negative overall impact

Due to the wide variety of measures and methods employed across the studies, it would be difficult, if not impossible, to derive an overall “effect size.”<sup>3</sup> Instead, we have systematically combined ratings of the studies’ findings with an assessment of their design quality.

### Assessing the Quality of the Studies

In a scheme similar to Scriven’s weight and sum methodology,<sup>4</sup> each study was rated on six weighted dimensions of overall quality: research design, duration of study, controls, measures used, scope of the study, and completeness of the technical report (see Table 2, following). Assigned weights ranged from 0 to 10 points depending on the importance of the dimension; scores on each dimension were added to produce a rating of overall study quality. All ratings are based solely on information in technical reports or publications.

Out of a possible 32 points, high quality studies generally scored 20 or more. The very weakest and least rigorous studies typically had quality scores ranging from 3 to 10.



**Table 2. Weighting Scheme for Quality Ratings of School Choice Studies on Student Achievement**

<i>Dimension</i>	<i>Values or Variations Within Each Dimension</i>	<i>Points</i>	<i>Total Possible Points</i>
<b>Research Design</b>	Randomized	10	<b>10</b>
	Matched students	8	
	Same cohorts	4	
	Consecutive cohorts	1	
	Cross sectional	0	
<b>Duration of Study</b>	More than 3 years of data	4	<b>4</b>
	2-3 years of data	3	
	1 year of data (cross-sectional)	0	
<b>Controls</b>	1 point for each of the following controls considered in the design: (i) family income, i.e., FRL; (ii) ethnicity; (iii) special education and/or LEP; (iv) starting performance level or use of gain score; (v) parents' education level; (vi) indicator of length of exposure	0-6	<b>6</b>
<b>Measure of Student Performance</b>	NPR, NCE, or Scaled Score	2	<b>2</b>
	Cut score (% meeting state standard)	1	
	General rating or grade	0	
<b>Scope of the Study</b>	Scope is based on 3 separate scores related to (i) relative size of the population studied, (ii) number of grade levels covered, (iii) number of subjects included.		<b>6</b>
	Relative size of population studied: 3 points for large comprehensive studies, 2 points for moderately comprehensive studies, 1 point for small studies, and 0 points for very small studies. <sup>5</sup>	0-3	
	Grade levels covered in the study (2 points for at least one grade at each of the three school levels; 1 point for at least one grade at two levels; 0 for at least one grade at one school level)	0-2	
	Subjects covered in study (1 for math & reading, 0 for one or neither)	0-1	
<b>Completeness of the Technical Report<sup>6</sup></b>	Technical report with clear and complete methods section	0-2	<b>4</b>
	Complete set of findings	0-1	
	Limitations of study included	0-1	
	<b>TOTAL POINTS</b>		<b>32</b>

Because the values or variations within each dimension of study design are specific and concrete, the process of assigning scores was an objective activity. However, the total scores assigned for each dimension

are subjective and reflect our judgments regarding the relative weight that each dimension should receive. For example, the overall study design can receive a maximum of 10 points, whereas the outcome measure is worth only 2 points. These judgments, though subjective, were informed by earlier syntheses of charter school research conducted by Miron and Nelson in 2001 and 2004.<sup>7</sup> The weighting system has evolved and become more elaborate to account for the characteristics of the broader field of school choice research.<sup>8</sup>

Given the wide variety of methodological rigor across studies, quality ratings are especially important in a synthesis of school choice research. Our approach includes studies with substantial limitations, but we give them less weight than other studies. We judged it important to include some of these less rigorous studies because they have attracted considerable attention and have been important in driving policy. By including them in our synthesis, we are able to present these weaker studies in a framework that allows readers to see the relationship between rigor and influence.

### **Combining Impact Scores and Study Quality Scores**

The impact ratings can be combined to provide a single impact score. Rather than simply calculating a mean impact rating, however, we have calculated a weighted mean in which each study is weighted by quality.

Additionally, we have developed a method to map the impact and quality ratings for each study analyzed. In our maps, each study is charted on a horizontal axis according to its relative quality, and along a vertical axis according to its impact rating for a particular choice model. Impact ratings range from strongly negative to strongly positive, as Table 1 above indicates. The results maps illustrate three important facets: (1) breadth and scope of available research, (2) overall quality of research, and (3) overall concentration of findings in terms of impact ratings.

### **Limitations**

We are mindful of some important limitations in this synthesis of research on student achievement across diverse forms of school choice. First, any thorough evaluation of schools and school choice models should examine non-achievement outcomes, such as equity, student and family satisfaction and market accountability, curricular quality and relevance, and instructional effectiveness. Moreover, even when student achievement is the only concern, standardized test results are but one of many ways to assess it.

Like any review or meta-analysis, our portrayal of the existing literature is colored by the selection of studies for examination. We have made explicit the rules that guided our search for and selection of studies,

and where possible, we sought to test our findings' sensitivity to these assumptions. Generally, we have been somewhat surprised at the paucity of studies of student achievement and the difficulty we had in obtaining some of the studies. As readers will see in the next section, we attempted to set out clear evaluative criteria and to apply them fairly to all studies reviewed.

### ***Student Achievement for Vouchers Programs***

Since first being proposed in the 1950s by Milton Friedman,<sup>9</sup> vouchers have been discussed and debated widely. Publicly funded school voucher programs have been established in Milwaukee (as of 1991); Cleveland, Ohio (as of 1996); and Washington, D.C. (as of an act of the U.S. Congress in 2003). Some small, privately funded programs have also generated evidence regarding the effect of vouchers on student achievement in New York City; Dayton (Ohio); Washington, D.C.; and Charlotte (N.C.).<sup>10</sup>

In theory, we would have grouped tuition tax credit programs with vouchers because of their similar nature. However, because we could find no empirical studies examining academic achievement in tuition tax credit programs,<sup>11</sup> we focus here solely on voucher programs.

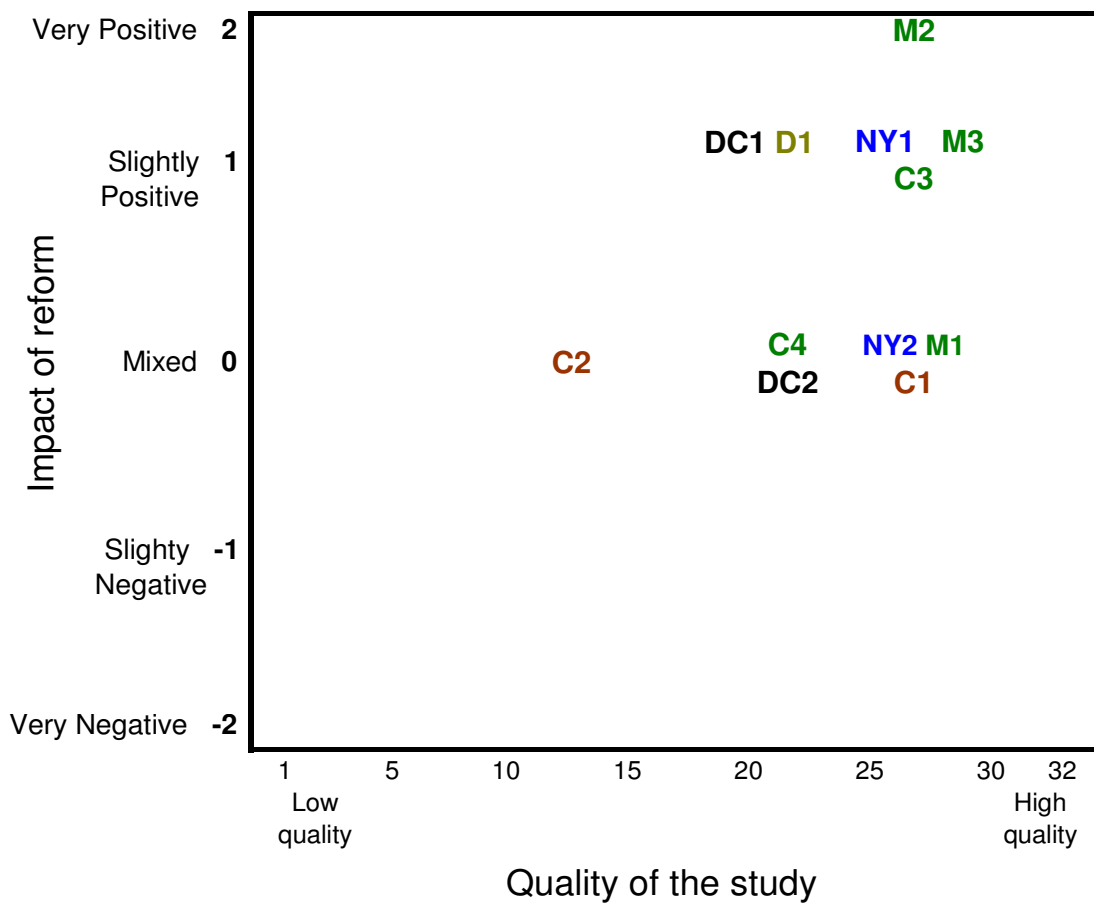
Figure 1 (following) illustrates our findings for 11 studies across the various voucher programs. We grouped the studies based on the particular program, and assigned letter codes for each program accordingly. Our discussion also groups studies by specific programs.

***Milwaukee.*** Milwaukee has the longest running voucher program in the nation. The program was started in 1991, and Witte and colleagues from the University of Wisconsin-Madison were contracted to evaluate the program.<sup>12</sup> Their evaluation used a host of demographic controls to match students from the Milwaukee Public School District. The final conclusion from their multiyear evaluation found that voucher students generally performed similarly to comparable students in math and reading.<sup>13</sup>

In 1996, Green, Peterson, and Du analyzed the Milwaukee data and came to a different conclusion than Witte. Because the program had more applicants than spaces available, a lottery was used to randomly select students to be admitted into the program. When comparing scores of those students who were lottery winners against those of lottery losers, Greene, Peterson and Du found a significant difference in favor of the admitted and enrolled choice students in both math and reading. A more polished update of their findings was published in 1999.<sup>14</sup> Witte questioned their randomized approach on a number of grounds, including that some students who were refused subsequently enrolled in private schools, thereby biasing the control group through attrition, and potentially leaving the remaining control group with lower-performing students.<sup>15</sup> Rouse was the third party to reanalyze the same data. She also

## The Impact of School Choice Reforms on Student Achievement

used students from the lottery lists as a control group and found that voucher students were making gains in math but not in reading.<sup>16</sup>



**Figure 1. Quality and Impact Ratings for 11 Studies of Student Achievement in Voucher Programs**

Note: This map illustrates estimated impact and quality ratings for 11 studies completed during the last decade.

M1 = Witte (1998). [Milwaukee]  
 M2 = Greene et al. (1999). [Milwaukee]  
 M3 = Rouse (1998). [Milwaukee]  
 C1 = Metcalf et al. (2004). [Cleveland]  
 C2 = Greene, et al. (1999). [Cleveland]  
 C3 = Plucker et al. (2006). [Cleveland]

C4 = Belfield (2006) [Cleveland]  
 NY1 = Mayer, et al. (2002). [New York City]  
 NY2 = Krueger & Zhu (2004). [NYC]  
 DC1 = Wolf et al. (2000) [District of Columbia]  
 DC2 = Wolf et al. (2007). [D.C.]  
 D1 = West et al. (2001). [Dayton]

There were a number of limitations in all of these studies. While Witte and Rouse carefully presented detailed methods, rationales for

decisions during the course of their analysis of data, and limitations in the interpretation of findings, this was not the case with the Green, Person, and Du study.

**Cleveland.** The Cleveland voucher program was approved in 1995 and started operating in 1996. This program had a much smaller voucher amount available to help cover tuition at a private school than did the Milwaukee program, but it made available a larger number of vouchers.

The Ohio Department of Education hired Metcalf and colleagues from Indiana University to evaluate this program; the most recent publication led by Metcalf was in 2004.<sup>17</sup> In 2006, another group of evaluators led by Plucker released an updated report on the Metcalf work. The Plucker group added one more year of data and also altered some of the analysis techniques for imputing missing data. Its evaluation found that voucher students had made noticeable gains relative to the comparison groups after entry into middle school.<sup>18</sup>

Similar to their work in Milwaukee, Greene and Peterson were quick to come up with their own analysis of the Cleveland results (see Greene, Howell & Peterson, 1997).<sup>19</sup> Their analysis of test data was limited to only two voucher schools, and they initially concluded that voucher students in those schools were making significant gains relative to a national norm. They updated the study in 1999 with additional years of data, at which point they found results to be mixed and in some cases negative. Nevertheless, the authors concluded that the program should be continued.<sup>20</sup> We have included only the second of their reports in our analysis, since both studies involved similar methods and authors, and the second reflected access to more data.

**Washington, D.C.** The D.C. Opportunity Scholarship Program was the first federally funded private school voucher program in the United States. The U.S. Congress created the program in 2003, providing scholarships of up to \$7,500 for low-income residents of the District of Columbia to send their children to local participating private schools. The U.S. Department of Education contracted a team of researchers led by Wolf to evaluate the program. The evaluation used a randomized controlled trial that compared students that received a place in a school via a lottery selection with students that did not. The third-year report concluded that there was no evidence of statistically significant differences in test scores between voucher recipients and students who applied but did not receive a voucher. This evaluation is ongoing and should yield more concrete results within the next few years.

In addition to this large, publicly funded voucher program, there is also a smaller, privately funded voucher program in Washington, D.C. An evaluation report after one year reported significant gains in math for African-American students who switched to private schools in grades 2 through 5.<sup>21</sup> Unfortunately, no subsequent reports have been released on the student achievement results from this program.

*New York City and Dayton, Ohio.* A study of a privately funded voucher program in New York City<sup>22</sup> concluded that the program was resulting in significantly higher test results for African-American voucher recipients, although no effects were seen for other ethnic subgroups. Krueger and Zhu<sup>23</sup> reanalyzed the data and found some serious shortcomings, including what they reported as exclusion of students and an inappropriate method for categorizing race. Their reanalysis indicated no effect favoring voucher students.

This program also was studied initially by some of the same persons involved in the studies of the New York and Washington, D.C., private voucher programs. Findings from Dayton reported by West, Peterson, and Campbell<sup>24</sup> concluded that there were no differences between voucher recipients and non-recipients. The one exception was for African-American students, who gained more than similar non recipients.

Figure 1 illustrates that a moderate number of empirical studies have been completed on student achievement in voucher programs. Given the few voucher programs in the nation, however, the number of studies is surprisingly large. The figure also shows that most studies were of higher quality (with a mean quality rating of just over 25 points on a 0-32 scale). In fact, the quality ratings for the voucher research are considerably higher than the research for other areas of school choice. On the whole, the voucher studies suggest a moderate effect in favor of private schools that participated in the voucher programs; the weighted mean for the impact ratings was +0.62. It is important to note that nearly half of the studies had mixed findings, and three of the five with slightly positive findings had positive results only for African-American students. None of the studies, however, indicated that vouchers were deterring learning for students who switched from public to participating private schools.<sup>25</sup>

Appendix A contains details on the voucher studies included and their ratings.

### *Student Achievement for Home Schooling*

Research on the student achievement of home schoolers has been the most difficult area of school choice to assess. Some of the obstacles are due to an inability to accurately measure the home-school population, a lament well noted in home-school research. A more important difficulty that we encountered was locating and identifying studies that met our minimum criteria for inclusion. While there are many studies on home schooling—as Ray’s 2008 Annotated Bibliography<sup>26</sup> attests—not all examine academic achievement. Within the group that does, only a small percentage use standardized tests as the outcome measure.<sup>27</sup> Often, studies cited in home-schooling magazines or journals that appeared to have a rigorous design could not be obtained or located, even though they were cited by other home-schooling researchers.<sup>28</sup> Home-schooling research studies generally tended to cite the same literature and to include many

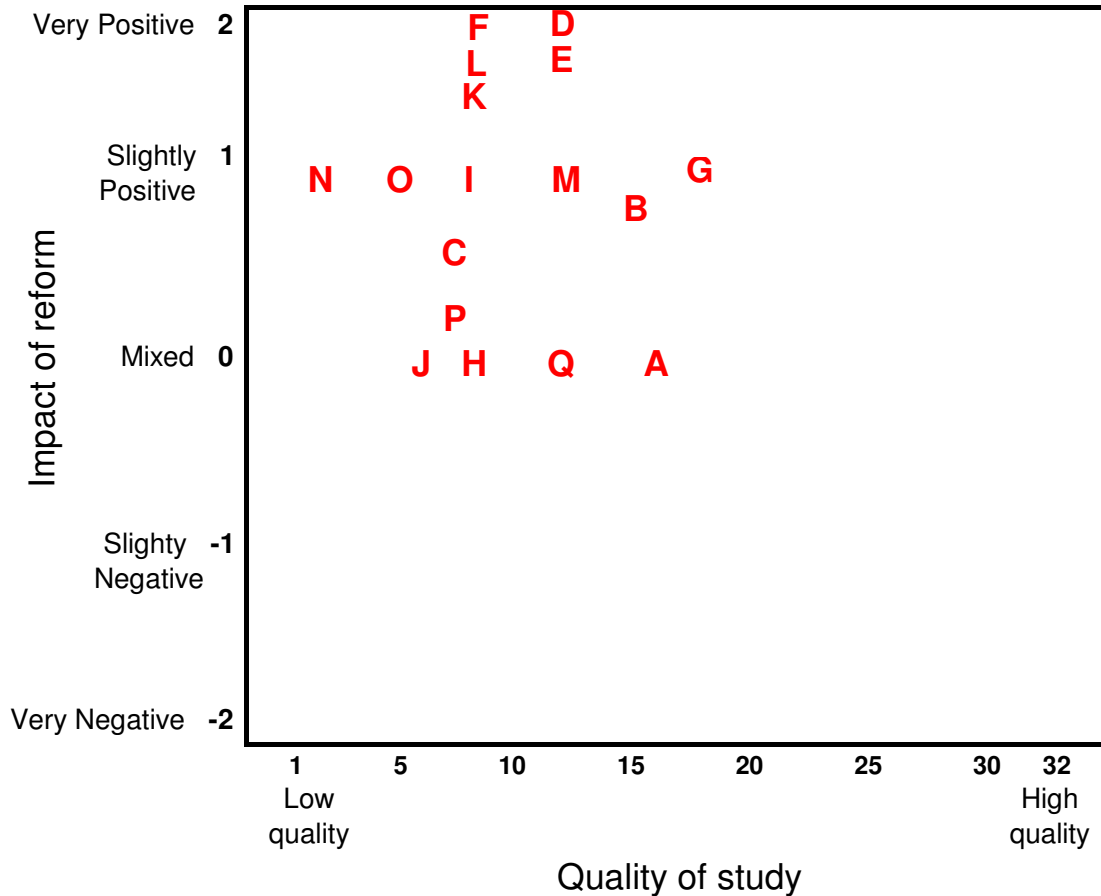
## The Impact of School Choice Reforms on Student Achievement

dated works (20 years old or older); many were also doctoral dissertations.<sup>29</sup> Other studies lacked such important items as a technical report, so that we were unable to discern quality. A few studies with strong designs were compromised by sample bias, researcher bias, or both. As noted earlier, however, we chose to include the studies with obvious bias, although they are down-weighted when these biases affect the design, scope of the study, or the completeness of the technical reports. Given the biases and errors built in to the existing body of home-schooling research, it comes as no surprise to learn that, on the whole, studies find high academic achievement among home schoolers. In fact, a home-schooling specialist we talked with said he couldn't think of any study on academic achievement among home schoolers that reported a negative finding.<sup>30</sup>

The scope of studies on achievement within home schooling is generally quite small. This is due, in part, to the difficulty faced in accurately defining populations. Because standardized tests are the comparison tool, population samples tend to be comprised of home-schooling families willing to have students tested. Often, researchers have obtained their samples through a testing center or a home-schooling advocacy group. In both situations, the fact that the sample lacks non-responders implies that it is biased, an issue more prominent in research on home schooling than in research on other forms of school choice. Generalizability, therefore, is very limited. An additional methodological issue is that home-schooling studies often use no demographic controls for comparisons.

We have included studies with sample biases because to exclude them would leave essentially no viable studies on home schooling for analysis. Nevertheless, it is important bear sample bias in mind when considering the overall report from the field that academic achievement among home schoolers is high. Routine standardized testing is not a part of the "set" curricula for home schoolers in the way that it is for, say, public school students, who all are tested at multiple grades. The first time that many home schoolers may take a standardized test is when they are preparing to enroll in college.

## The Impact of School Choice Reforms on Student Achievement



**Figure 2. Quality and Impact Ratings for Studies of Student Achievement in Home Schooling**

Note: This map provides an illustration of estimated impact and quality ratings for 17 studies completed during the past 22 years.

- |                      |                                       |
|----------------------|---------------------------------------|
| A = Boulter (1999)   | J = Holder (2001)                     |
| B = Collom (2005)    | K = Witt (2005)                       |
| C = Galloway (1995)  | L = Richman, Girtten, & Snyder (1990) |
| D = Ray (2000)       | M = Wartes (1990)                     |
| E = Rudner (1999)    | N = Rakestraw (1988)                  |
| F = Clemente (2006)  | O = Jones, & Gloeckner (2004)         |
| G = Qaqish (2007)    | P = Frost (1987)                      |
| H = Delahooke (1986) | Q = Belfield (2005)                   |
| I = Gray (1998)      |                                       |

Certainly, the characteristics of these children are different from those choosing not to pursue higher education, which leaves the field of home-school research with a large gap in its understanding of the students who are not tested or considered in these achievement studies.



## The Impact of School Choice Reforms on Student Achievement

Within the context of considerations detailed above, we found 17 studies on home schooling and academic achievement that met our minimum selection criteria. Figure 2 (preceding) charts them by their own report of impact and our rating of study quality.

In contrast to the graphs on the other forms of school choice, all the studies on home-schooling research are clustered in the upper left quadrant of the graph. Generally speaking, then, the body of home-schooling research on academic achievement is of low rigor and low overall quality.

As is apparent in Figure 2, we were unable to find any studies that found explicitly negative impacts of home schooling on academic achievement. However, the quality of the research designs that produced such positive findings is low. Within our 0-32 point rating scheme, the overall quality score for the home-schooling research studies we included was 9.88, indicting fairly low design rigor. The highest quality rating for an individual study was 18, a score still only slightly more than half the points available. The mean weighted impact rating for the studies was 1.0, indicating overall consensus among the researchers that home schooling as a method of school reform has had a positive impact.

One of the most widely cited studies in home-schooling literature that met our selection criteria was Ray, who found home-schooled children scored at or above the 80th percentile on standardized tests (Study D).<sup>31</sup> Even though he used a relatively large sample and his results may be true for the population he included, his sample left out home-schooled students who do not take tests. His technical report is also a bit misleading when it claims random selection of participants. A closer reading of the report shows that he gained access to the population through the mailing lists of home education organizations. He randomly selected from those mailing lists, not from the home-schooled population in general, leaving his work vulnerable to the same sample bias that runs through nearly all home-schooling research.

Ray has conducted much of the research in the field himself and is widely cited in nearly every study on home schooling. He is the founder and president of National Home Education Research Institute, and he edits and publishes a journal about home schooling, *The Home School Researcher*, in which many others have established their publishing record.<sup>32</sup> Though Ray's work is commonly considered the foundation of home-schooling research, only one of his studies qualified for our analysis, largely because most of his published work does not consist of original data.

Rudner's 1999 study of home schoolers<sup>33</sup> (Study E) is as frequently cited as Ray's work, though Rudner did not subsequently publish anything else on the topic. His original work was a large study that found home-schooled students scored in the 70th to 80th percentile on standardized tests. However, in addition to self-selection bias in his population, his sample was shaped by having been accessed through the

testing center at Bob Jones University, a southern Christian school with an overtly racist tradition. Welner and Welner<sup>34</sup> argued that the results of the study suffer from limited generalizability. The same critique can be applied to Galloway's popular 1995 study showing home schoolers' equal preparation for college, based on scores on the English subtest of the ACT (Study C). Her population sample came from an unnamed "large, private Christian University located in the Southeast,"<sup>35</sup> while her byline shows she was writing from Bob Jones University at the time. These two studies have served as foundational pieces in the field of home-school research, but their results reflect a largely white, Christian student population; reliance on them has skewed perceptions about home schoolers and their performance on standardized tests. An increasing number of families of color are home schooling, but they generally have been left out of nearly all empirical research on the topic.<sup>36</sup>

Thus, the most widely cited studies in home schooling are subject to researcher and sample bias, although we incorporated them in our analysis. Appendix B lists all of the studies we examined, including some that we originally intended to include but excluded when close examination revealed that they lacked critical components.<sup>37</sup> Appendix B also details points awarded for quality elements of each study along with the rationale for judgments made.

### ***Student Achievement for Interdistrict, Intradistrict, and Magnet School Programs***

This section examines diverse forms of school choice found within the traditional public school sector, including inter- and intradistrict choice programs. Magnet schools, which are part of a federally funded program, are also considered in this section since they are a form of intradistrict choice that is overseen by the local district school board.

Magnet schools<sup>38</sup> have received less attention, inspired less controversy, and generated less research than the other forms of school choice, although they remain the most common school choice option. While charter school accountability has been a prominent topic in school reform literature over the last 15 years, a review of American Education Research Association (AERA) conference programs reveals only one paper focusing on magnet schools from 1998 through 2006. Originally, magnet schools were devised as a strategy to decrease segregation in schools and as a response to violent protest against mandatory reassignment policies introduced in the mid-1970s.<sup>39</sup> Research literature on magnet schools is prolific in regard to nonacademic topics, including desegregation. However, very few isolate academic achievement and use standardized testing as a measure of success. From those few we have culled an even smaller number that met minimum selection criteria.

The majority of magnet school studies included in this analysis employ demographic controls, as any good study of school choice reform

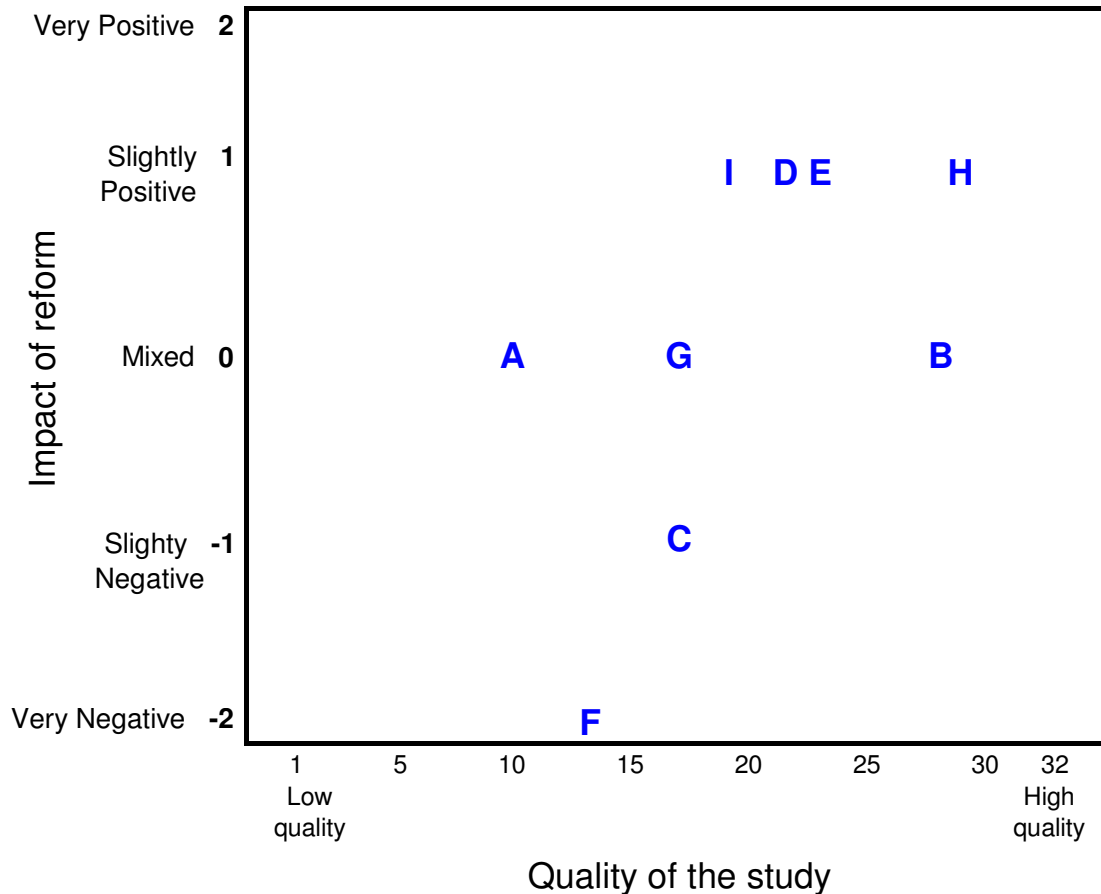
## The Impact of School Choice Reforms on Student Achievement

models should. Demographic controls are particularly important because such research is often subject to confounding variables. For example, parental involvement can be an interceding effect: parents who make the effort to research options and actively choose a school are likely to be more involved in a child's school life overall, contributing to higher academic achievement. Likewise, magnet schools tend to attract a greater percentage of students with high prior academic achievement, leaving non-magnet schools in the district with more at-risk students and rendering comparisons incompatible. Therefore, it is critical to strong research design that a study include a control for starting performance level or some other determination of a gain score.

Figure 3 maps the studies on magnet, intradistrict, and interdistrict choice schools that we analyzed. These three forms are grouped because too few studies met selection criteria in each category to allow for substantive individual analysis. For the combined forms, we identified nine studies of student achievement. Wide variety among the studies makes it important to consider them as a group. Some focused only on high school, while others focused on elementary or middle school levels. They also varied significantly in scope, from a single district to a national sample.

On average, these studies scored 19.56 on quality, although as Figure 3 indicates there was a significant spread in design quality as well as impact. Based on each study's perceived impact on student academic achievement, the overall impact rating for magnet schools is +0.26, reflecting the general view that magnet schools have had a slightly positive impact on student achievement, as measured by standardized tests. The highest quality score was for Ballou et al.<sup>40</sup> at 29 points (Study B). Appendix C offers a chart detailing quality ranking for studies in these categories.

## The Impact of School Choice Reforms on Student Achievement



**Figure 3. Quality and Impact Ratings for Studies of Student Achievement in Magnet or Interdistrict Choice Schools**

Note: This map provides an illustration of estimated impact and quality ratings for 9 studies completed during the last 13 years.

A = Beaudin (2003)

B = Ballou, Goldring, & Liu (2006)

C = Crain et al. (1992)

D = Gamoran (1996)

E = Heebner (1995)

F = Inst. for Assessment and Evaluation (2006)

G = Christenson et al. (2003)

H = Betts et al. (2006)

I = Eagle & Ridenour (1969)

Generally, studies discussed here tended to score high on design criteria. This is so in part because magnet school admission is typically decided by a lottery, in which many students submit a request to attend the school, and attendees are randomly selected from that pool (although sometimes preference is given to students in the school's surrounding neighborhood or to those with a sibling already admitted). Such an admission lottery facilitates random assignment for study design purposes. The target population is known (all students in the lottery pool), and the

experimental and control groups are clear—the latter being the students who were not selected in the lottery. However, the two comparison groups are not exactly random or similar. Students are awarded entry by the school, but there is still a self-selection bias that remains because students (and their families) can and do reject admission. Our weighting scheme cannot account for this slightly-less-than-random design, but such accommodations were made by Ballou et al., (Study B), Crain et al., (Study C), and Heenber (Study E).

Studies including a national sample tended to have only moderately high quality designs (Gamoran [Study D] and Christenson et al. [Study G]), largely because they did not take advantage of randomization. Those two studies will be discussed below. A statewide study from Connecticut (Study A by Beaudin) and four studies county-wide or smaller are also included.

Gamoran (Study D)<sup>41</sup> is characterized by a large sample size and the use of the same cohorts to track student achievement over time. The study also used demographic controls and considered students' starting performance to determine value added. However, only two years of test data are used to draw conclusions. The data are also quite dated, from test years of 1988 and 1990. Also, only two grades were tested, implying limited generalizability to K-7 education. The study was published in 1996, indicating the need for new, rigorous research with wide scope and longitudinal data.

The other national study in our analysis (Study G) was conducted in 2003 by Christenson and colleagues,<sup>42</sup> who were contracted by the U.S. Department of Education. The access to national data rendered a large sample size, but the data were limited to school-level information. Further, the analysis procedures highlight a difficulty in using national samples and standardized testing: individual states administer different standardized tests. This study addressed such incompatibility by converting multiple state tests to a common scale. As with the Gamoran study, this work has limited generalizability because it focused on only the elementary level.

Scoring details for the two national studies as well as the other five studies included appear in Appendix C. It is surprising that we do not have more studies meeting our minimum criteria, given the somewhat natural randomization of students in magnet schools and their more than three decades of existence. The studies that allowed for analysis, however, depict a rather neutral, though slightly positive, comparison with public schools in terms of student performance on standardized tests.

### ***Student Achievement for Charter Schools***

Today, charter schools have the largest number of studies examining student achievement. The mounting evidence is very welcome after so many years with few comprehensive evaluations or achievement studies. In 2001, Gill et al. found only three studies of charter schools that

met their criteria for a summary of evidence.<sup>43</sup> In the same year, Miron and Nelson<sup>44</sup> found 15 studies of charter school achievement; in a 2002 update (published in 2004<sup>45</sup>), they identified only 17 studies for analysis. Thus, the total 47 studies included here reflects significant growth in the field.<sup>46</sup>

As the number of studies on charter schools has increased over the last five years, so, too, has the overall quality of the studies. While there is only one study of two Chicago charter schools that uses randomized assignment based on oversubscribed waiting lists, there are now rather rigorous matched student designs for California, Delaware, Florida, and Texas. Older studies with weaker designs and few years of test data are being supplanted by studies with more rigorous designs and more years of data. Also, with the expansion of state testing systems to cover more grades, it is now easier for studies that rely on school level data to track cohorts or groups of students as they progress from grade to grade.

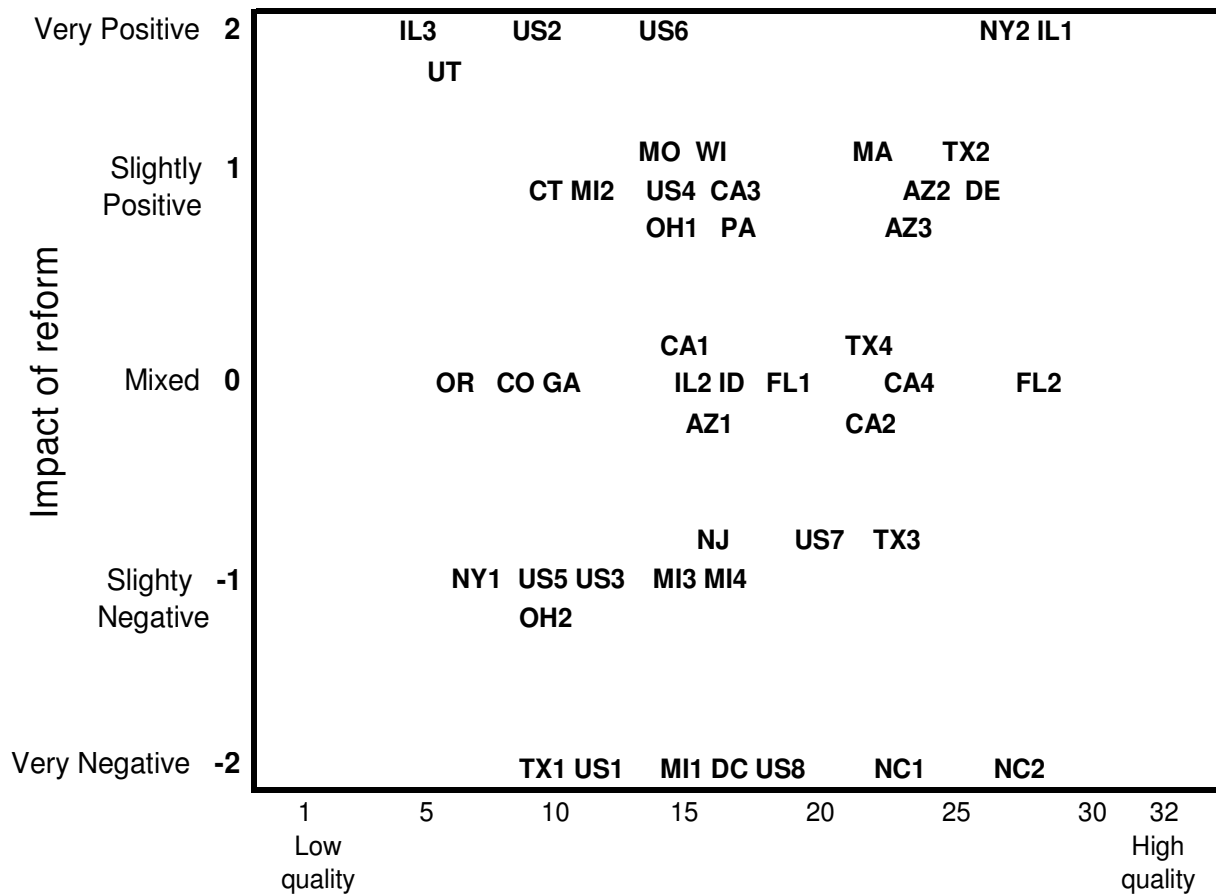
Close examination of Figure 4 reveals that studies vary widely in impact reported and design quality. Overall, 19 studies had positive findings, 12 studies had mixed findings, and 16 had negative findings. The mean impact rating for charters was +0.04. The weighted mean (adjusted for quality of studies) was +0.1. These findings indicate a mixed effect. Although not a strong or significant correlation, there is a very slight tendency for the studies with more rigorous designs to conclude that charter schools were outperforming their comparison groups. Appendix D details impact and quality scores assigned to each study.

Nearly all the charter studies are state studies. This is not surprising given that charter schools are a state-based reform model; 40 states and the District of Columbia have passed charter school laws, and more than 3,500 charter schools are now operating across the nation. Eight of the studies look at multiple states or use national data sets.

Figure 4 illustrates the number and geographic variation of the studies, which is impressive relative to the other forms of school choice. The impact ratings are more dispersed for charter schools than for the other forms of school choice we have examined. Similarly, quality ratings of the charter school studies vary widely. The mean quality rating is 17.5, much lower than that for voucher or district choice studies, but still much higher than that for home-schooling studies.

A few studies provide evidence of a substantial positive charter school effect. The Solmon and Goldschmidt (2004) analysis of Stanford Achievement Test (SAT9) scores in Arizona, for instance, found that charter schools had a significant positive impact on SAT9 scores in reading and a mixed to positive impact in math. At the other end of the spectrum, three of the four studies of achievement in Michigan charter schools are negative.

## The Impact of School Choice Reforms on Student Achievement



### Quality of the study

**Figure 4. Quality and Impact Ratings for Studies of Student Achievement in Charter Schools**

Note: This map provides an illustration of estimated impact and quality ratings for 47 studies completed during the last nine years.

AZ1 = Mulholland (1999)  
 AZ2 = Solmon & Goldschmidt (2004)  
 AZ3 = Garcia (2008)  
 CA1 = EdSource (2007)  
 CA2 = Rogosa (2003)  
 CA3 = Raymond (2003)  
 CA4 = Zimmer et al. (2003)  
 CO = Colorado Dept. of Education (2006)  
 CT = Miron (2005).  
 DC = Henig, et al. (2001)  
 DE = Miron, et al. (2007)  
 FL1 = Florida Dept. of Education (2006)

FL2 = Sass (2006)  
 GA = Plucker, et al. (2006)  
 ID = Ballou, Teasley, & Zeidner (2006)  
 IL1 = Hoxby & Rockoff (2004) [Chicago]  
 IL2 = Nelson & Miron (2002)  
 IL3 = Chicago Public Schools (2007)  
 MA = Massachusetts Dept. of Ed. (2006)  
 MI1 = Eberts & Hollenbeck (2002)  
 MI2 = Michigan Dept. of Ed. (2007)  
 MI3 = Bettinger (2005)

## The Impact of School Choice Reforms on Student Achievement

MI4 = Miron & Nelson (2002)	TX3 = Hanushek., Kain, & Rivkin (2002)
MO = Metis Associates. (2004)	TX4 = Booker et al. (2004)
NC1= Noblit & Corbett (2001)	UT = Was & Kristjansson (2006)
NC2 = Bifulco & Ladd (2006)	WI = Witte et al. (2004)
NJ = Barr (2007)	US1 = Finnigan et al. (2004)
NY1 = New York Board of Regents. (2003)	US2 = Hoxby (2004)
NY2 = Hoxby & Murarka. (2007)	US3 = US-DOE [NCES] (2004)
OH1 = Carr & Staley (2005)	US4 = Loveless (2003)
OH2 = Legis. Office of Ed.Oversight. (2003)	US5 = Nelson, Rosenberg, & Van Meter (2004)
OR = Bates & Guile (2006)	US6 = Greene, Forster, & Winters (2003)
PA = Miron, Nelson, & Risley (2002)	US7 = Miron, Coryn, & Mackety (2007)
TX1 = Maloney et al. (2007)	US8 = Braun, Jenkins, Grigg & Tirre (2006)
TX2 = Gronberg & Jansen, (2005)	

It is important to note that no studies have been completed on student achievement in cyber or virtual schools, which are typically charter schools catering to home-schooling families. In a 2003 study of California charter schools, Zimmer et al.<sup>47</sup> included some non-classroom-based charter schools and found that they had lower achievement scores than traditional public schools and other charters. In a 2002 evaluation of Pennsylvania charter schools, Miron et al.<sup>48</sup> similarly found that four virtual charter schools performed worse than or similar to comparison groups.

The unweighted average impact score across all studies was +0.04, suggesting that as a group the studies provide a mixed picture of the charter school effect. Two explanations are possible. One possibility is that large gains or losses in some charter schools are offset by losses or gains in other charter schools, yielding a mixed achievement impact. A second possibility is that impact is consistent across charter schools, but small. That the impact ratings are widely dispersed suggests an explanation that gains and losses are frequently offset.

### *Summary of Findings*

Overall, the existing research on school choice models and achievement provides a mixed picture, with some studies suggesting positive impacts and others indicating neutral or negative impacts. Except for the research on home schooling, the inclusion of relatively lower quality studies did little to change the overall findings.

There were large differences across school choice type, both in terms of the amount of research available as well as the conclusions that can be drawn from the research. The entire body of the literature leads to the following key findings.



## The Impact of School Choice Reforms on Student Achievement

### **Vouchers**

- Given that few voucher programs exist, a relatively large number of studies on them are available.
- The quality of the studies is reasonably high, with many relying on lottery lists to generate comparison groups of students.
- The results—on the whole—are slightly positive, particularly with regard to performance of African-American students.

### **Home Schooling**

- Relatively few studies exist.
- Most of the studies are especially weak in design quality.
- All findings are mixed or positive in favor of home schooling.

### **Inter-, Intradistrict Choice Programs and Magnet Schools**

- Relatively few studies exist.
- The quality of magnet school research is generally mixed, although the lottery lists from oversubscribed schools could facilitate more rigorous designs.
- Overall findings were mixed.

### **Charter Schools**

- The most studies are available on charter schools, with rapid growth in the literature appearing over the past six years.
- Design quality for research on charter schools varies considerably; for some half of the studies, relatively weak quality is due to the absence of—or inability to obtain—student-level data.
- Cumulative results from charter school research indicate that, on the whole, charters perform similarly to traditional public schools. Results from individual studies have remained mixed over time, even with the addition of newer and higher quality studies.

Table 3 and Figure 5 (following) facilitate comparison of findings across diverse choice models; together they summarize the total number of studies analyzed for each model as well as the impact and quality ratings in each category. Voucher studies had the highest overall quality ratings, and home schooling the lowest. Impact ratings include not only the mean but also the weighted mean, which takes into account study quality. On average, home school studies had the most positive impact ratings, and charter schools the least positive—although still mixed. Figure 5 charts the general position of these four broad forms of school choice in terms of relative quality and impact. On the whole, we could discern no correlation

between the studies’ quality and their findings relative to choice’s impact on student achievement.

**Some Observations on the Findings**

In considering import of the findings, we have been disappointed to note—especially relative to charter schools—that most of the media attention and public debate has focused on relatively weak cross-sectional studies. It appears that discussion is being shaped by research that does not merit the emphasis it is receiving.

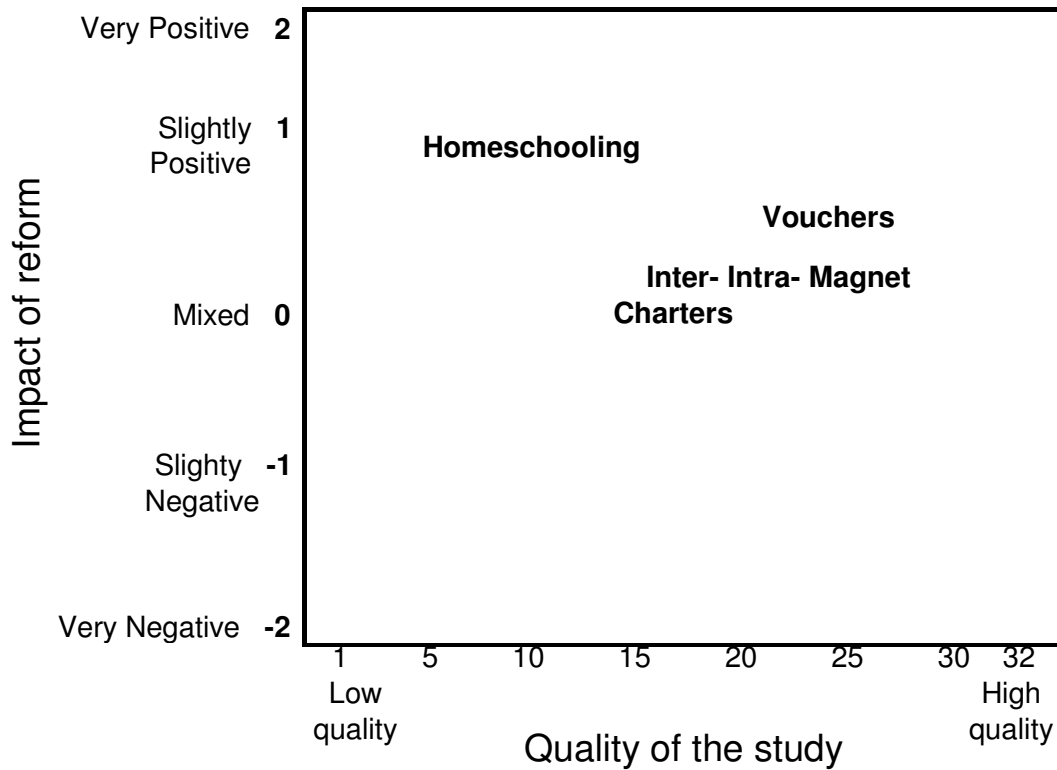
As this analysis demonstrates, the research and evaluation literature has not yet produced clear and unambiguous factual statements about achievement across any of the key types of school choice. Thus, stakeholders must weigh the strengths and weaknesses of the evidence. Since it is unlikely that there will ever be a single definitive study, the most reasonable approach for interpreting the evidence is to conduct a meta-analysis or assemble a picture of the findings across the broad body of research, as we have done here.

Still, it is important to simultaneously remain aware that aggregating findings across types of school choice hides considerable differences. Within all models, there are certainly successful choice programs and schools as well as seriously flawed ones. At a macro level this policy brief can tell us what the body of research says and whether these reform models are worthy of replication. Nevertheless, all forms of school choice could benefit from a better understanding of factors leading to success within particular schools or groups of schools.

**Table 3. Mean Quality and Impact Ratings Across Diverse Forms of School Choice Research**

Type of Choice	Studies (N)	Quality Rating (0 to 32 scale)		Impact Rating (+2 to -2 scale)		Weighted Mean Impact
		Mean	SD	Mean	SD	
<i>Voucher</i>	12	25.58	4.67	0.58	0.67	0.62
<i>Charter</i>	47	17.49	6.06	0.04	1.25	0.10
<i>Home school</i>	17	9.88	4.12	1.00	0.79	1.01
<i>Intra-/Inter-/Magnet</i>	9	19.56	6	0.11	1.05	0.26

## The Impact of School Choice Reforms on Student Achievement



**Figure 5. Illustration of the Overall Findings Across Diverse Forms of School Choice**

### **A Cautionary Word on the Role of Advocacy Research**

Given that school choice is a highly politicized issue, there are many attempts to influence policy with sensational claims about student achievement by advocacy or opposition groups. Typically, these claims allude to empirical research. Press releases with data charts and talking points appear in papers or on Web sites. Our preliminary review of the existing research had to weed through piles of these pseudo-studies/evaluations. Most were not included here because they lacked technical reports specifying the number of students and schools involved in a study or offering details on methodology. While technical reports may take many forms, the public should be aware that when one is not included, there is no way to determine whether conclusions are justified or findings can be verified and replicated.

Several advocacy studies of relatively high quality were included in this analysis. However, as with pharmaceutical companies doing rigorous research on their new drugs, findings that do not support the position of groups with a vested interest are often not released, so that caution in interpreting significance of results is advisable.

## ***Recommendations***

The analysis detailed here yields two core recommendations for researchers and policy makers. The first is to improve research on school choice, and the second is to improve interpretation of school choice research.

### **Improve Research on School Choice**

- Take care in creating comparison groups from lottery lists at choice schools. Selection bias may occur since some selected students may choose not to attend the choice school because of transportation or other barriers. The technique is promising, but researchers must still confirm and control for differences in the group of students who are accepted and those who are not.
- Consider using matched student designs, which are affordable and will be increasingly useful as state assessment systems improve and expand. Our analysis shows that studies using matched student designs often score high overall on quality ratings since they get more points for scope of study, demographic controls, completeness of technical reports than studies using other designs.
- Remember that impact can be adequately captured only with longitudinal designs. Thus, cross-sectional studies are most useful in assessing relative performance and describing the types of students enrolled in particular choice models.
- Promote more research on home schooling, especially as increasing numbers of home schoolers enroll in cyber schools.
- Promote research exploring differences across and within forms of school choice to help identify factors and conditions most likely to support successful school choice reforms.
- Clearly articulate research designs and methodologies, at the very least in an appendix or a Web document.
- Specify limitations and precautions that readers should consider when interpreting the findings.

### **Improve the Interpretation of Research on School Choice**

- Remember that performance on standardized tests is only one of several important outcome indicators. Standardized tests are the easiest but not necessarily the best way to evaluate student learning.
- Be skeptical of sweeping conclusions drawn from the body of existing research; the range of findings and relative weakness of many studies does not support such claims. Remember that there simply are no definitive studies.
- Be aware that many commonly discussed and debated studies have weak research designs, as evidenced by their failing to meet our minimal

## The Impact of School Choice Reforms on Student Achievement

selection criteria or, if included, by their quality ratings often being among the lowest.

- Consider the breadth of findings available regarding any single form of school choice, and when interpreting the research, remember the importance of study design, sample size, and the relevance of the comparison group.
- *Beware the press release.* Findings highlighted in press releases should be ignored if no technical report exists with details on the population studied or the study design used.
- When interpreting research, consider the source. Was the research funded or conducted by an advocacy group? Have the researchers ever released findings counter to their current results? Lead researchers of these studies typically have extensive experience, and the odds—for example, that someone would never have a finding that was in support of traditional public schools, or vice versa—speak loudly about the nature and purpose of their work.

### *Notes and References*

- <sup>1</sup> Inasmuch as many school choice programs have schools that are oversubscribed and regulations that require students to be selected at random from their waiting lists, randomized experiments ought to be possible, in principle. However, waiting lists often are not audited over time and are insufficient for the construction of a good randomized experiment since they are often out of date, contain an accumulation of names over a number of years, and often cannot be readily produced when requested.
- <sup>2</sup> Readers should bear in mind that our 5-point scale might understate the variation in impacts found across studies.
- <sup>3</sup> See, e.g., Light, R., & Pillemer, D. (1984). *Summing up: The science of reviewing research*. Cambridge, MA: Harvard University Press.
- Hedges, L., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.
- <sup>4</sup> Scriven, M. (1981). The “weight and sum” methodology. *American Journal of Evaluation*, (2); 85-90.
- <sup>5</sup> A large comprehensive study (3 points) requires more than 25 schools (school as unit of analysis) OR more than 1,000 students (if student is unit of analysis). A moderately comprehensive (2 points) requires between 11 and 25 schools, OR 121-1,000 students. A small study (1 point) is defined by 5-10 schools OR 50-120 students, and a very small study (0 points) considers 2-4 schools OR less than 50 students. Studies with only one school are not included.
- <sup>6</sup> Replication and verification are facilitated when there is a complete technical report. In interpreting findings, we generally perceive studies with comprehensive technical reports to be more credible. Some studies are most rigorous with randomized assignment, but they lose points due to the completeness of the technical report. It was surprising to us to see how many studies had weak or incomplete technical reports. Some of these had sparse details on methods and only reported on a restricted range of findings. Common in the weaker technical reports was that no mention was made of caveats or limitations that readers should taken into consideration.
- <sup>7</sup> Miron, G., & Nelson, C. (2004). Student achievement in charter schools. In K. Bulkley & P. Wohlstetter (Eds.), *Taking account of charter schools* (pp. 161-175). New York: Teachers College Press.
- Miron, G., & Nelson, C. (2001) *Student academic achievement in charter schools: What we know and why we know so little*. Occasional Paper No. 41. National Center for the Study of Privatization in Education, Teachers College, Columbia University.  
[[http://ncspe.org/keepout/papers/00041/590\\_OP41.pdf](http://ncspe.org/keepout/papers/00041/590_OP41.pdf)].
- <sup>8</sup> The scores assigned in the weighting scheme have also benefited from input from colleagues at the Western Michigan University Evaluation Center, where the weighting scheme has been vetted and opened up to public input in connection with two presentations (one in 2005 and one in early 2008).
- <sup>9</sup> Friedman, M. (1955). The role of government in education. In R.O. Solo, (Ed.), *Economics and the public interest*. New Brunswick, N.J.: Rutgers University Press.
- <sup>10</sup> We have not included research on the earliest voucher pilot program from Alum Rock, California, which was started in 1972 because of their complications in implementing the program and also because the published research on this program did not yield specific results regarding the academic achievement of students taking advantage of the voucher.
- <sup>11</sup> Zimmer & Bettinger, also confirm that they could not find any research on tuition or education tax credits and student achievement.

## The Impact of School Choice Reforms on Student Achievement

- Zimmer, R., & Bettinger, E. (2008). Beyond the rhetoric: Surveying the evidence on vouchers and tax credits. In H.F.Ladd & E.B.Fiske (Eds.), *Handbook of research in education finance and policy*. New York: Routledge.
- <sup>12</sup> A number of annual reports were issued to the Wisconsin State Legislature, but because these analyses were superseded by subsequent reports, we have only considered the most recent analysis, which was summarized in a refereed article published by John Witte in 1998 [see Witte, J. F. (1998)].
- <sup>13</sup> Witte, J.F. (1998). The Milwaukee voucher experiment. *Educational Evaluation and Policy Analysis*, 20(4), 229-251.
- <sup>14</sup> Greene, J.P., Peterson, P.E., & Du, J. (1999). Effectiveness of school choice: The Milwaukee experiment. *Education and Urban Society*, 31, 190-213.
- <sup>15</sup> Witte, J.F. (1999). The Milwaukee voucher experiment: The good, the bad and the ugly. *Phi Delta Kappan*, 81, 59-64.
- <sup>16</sup> Rouse, C. (1998). Private school vouchers and student achievement: An evaluation of the Milwaukee Parental Choice Program. *Quarterly journal of economics*, 113(2), 553-602.
- <sup>17</sup> Metcalf, K.K., Legan, N.A., Paul, K.M., & Boone, W.J. (2004, October). *Evaluation of the Cleveland scholarship and tutoring program: Technical report 1998-2003*. Bloomington: Indiana Center for Evaluation & Education, Indiana University.
- <sup>18</sup> Plucker, J., Muller, P., Hansen, J., Ravert, R., & Makel, M. (2006). *Evaluation of the Cleveland scholarship and tutoring program: Technical report 1998-2004*. Bloomington, IN: Center for Evaluation and Education Policy.
- <sup>19</sup> Greene, J.P., Howell, W.G., & Peterson, P.E. (1997). *An evaluation of the Cleveland scholarship program*. Cambridge, MA: Program in Education Policy and Governance, Harvard University
- <sup>20</sup> Greene, J.P., Howell, W.G., & Peterson, P.E. (1999). *An evaluation of the Cleveland voucher program after two years*. Cambridge, MA: Harvard University, Program on Education Policy and Governance.
- <sup>21</sup> Wolf, P.J., Howell, W.G., & Peterson, P.E. (2000). *School choice in Washington, DC: An evaluation after one year*. Cambridge, MA: Harvard University, Program on Education Policy and Governance.
- <sup>22</sup> Mayer, D.P., Peterson, P.E., Myers, D.E., Tuttle, C.C., & Howell, W.G. (2002). *School choice in New York City after three years: An evaluation of the school choice scholarships program* (No. 8404-045). New York: Mathematica Policy Research.
- <sup>23</sup> Krueger, A., & Zhu, P. (2004). Another look at the New York City school voucher experiment. *American Behavioral Scientist*, 47, 658-698.
- <sup>24</sup> Peterson, P.E., West, M.R., and Campbell, D.C. (2001). *School Choice in Dayton, Ohio After Two Years : An Evaluation of the Parents Advancing Choice in Education Scholarship Program*. Program on Education Policy and Governance Occasional Paper, Harvard University, PEPG 01-04.
- <sup>25</sup> Based on his extensive review of the research evidence on school vouchers, Carnoy concluded that “vouchers’ effects on student achievement are almost certainly smaller than claimed by pro-voucher researchers. Although programs in many cities were designed to be like randomized-trial medical experiments—with high validity and reliability—common problems in implementation may have compromised validity and produced misleading results. Moreover, the results are marked by broad inconsistencies across grades, academic subjects, and racial groups. See Carnoy, M. (2001). *School vouchers: Examining the evidence*. Washington, DC: Economic Policy Institute.
- <sup>26</sup> Ray, B.D. (2008). *Home centered learning annotated bibliography, seventeenth edition*. Salem, OR: National Home Education Research Institute.

## The Impact of School Choice Reforms on Student Achievement

<sup>27</sup> Several studies on home schooling have pointed to the inadequacy of standardized testing in measuring the success of home-schooled students, e.g.,

Meehan, N. & Stephenson, S. (1994). *Homeschooling in the United States; A review of recent literature*. (ERIC Document Reproduction Service No. ED 424 922);

Burns, J. (1999). *The correlational relationship between homeschooling demographics and high test scores*. (ERIC Document Reproduction Service No. ED 439 141);

Collom, E. (2005). The ins and outs of homeschooling: The determinants of parental motivations and student achievement. *Education and Urban Society*, 37(3), 307-335;

Ray, B.D. (2000). Home schooling: The ameliorator of negative influence on learning? *Peabody Journal of Education*, 75 (1&2), 71-106; and

Rudner, L.M. (1999). Scholastic achievement and demographic characteristics of home school students in 1998. *Education Policy Analysis Archives*, 7(8) (ERIC Document Reproduction Service No. ED 435 709).

Therefore, a portion of the field has focused on alternative measures of student success, such as first year college GPA (Jones, P., & Gloeckner, G. [2004]. First-year college performance: A study of home school graduates and traditional school graduates. *The Journal of College Admission*, 183, 17-20; Sutton, J.P. & Galloway, R.S. [2000]. College success of students from three high school settings. *Journal of Research and Development in Education*, 33(3), 137-146); perceptions of college admissions officers (Ray, B.D. [2001]. *Home education in Ohio: Family characteristics, academic achievement, social and civic activities, and college admissions officers' thoughts*. Salem, OR: National Home Education Research Institute); self-confidence (Sheffer, S. [1995]. *A sense of self: Listening to homeschooled adolescent girls*. Portsmouth, NH: Boynton/Cook Publishers); leadership skills (Montgomery, L.R. [1989]. The effect of home schooling on the leadership skills of home schooled students. *Home School Researcher*, 5(1) 1-10); and adulthood characteristics (Ray, B.D. [2004]. Homeschoolers on to college: What research shows us. *The Journal of College Admission*, 185, 5-11; Knowles, J.G. & Muchmore, J.A. [1995]. Yep! We're grown-up home-school kids – and we're doing just fine, thank you! *Journal of Research on Christian Education*, 4(1), 35-56).

<sup>28</sup> State department reports fell into this category. We saw Arizona Department of Education (1989). *Students taught at home: 1989 average grade equivalents*; North Carolina Division of Non-Public Education (1989). *North Carolina home school nationally standardized achievement test results 88-89 school term*; Arkansas Department of Education (1988); and Tennessee Department of Education (1987). Home school student test results: 1986 and 1987. These were reported often but we could not locate those reports anywhere. The major researcher citing these studies, Klicka, also did not have them any longer (Ridley, V.N., legal assistant to Christopher J. Klicka, Esq., personal communication, December 18, 2007).

<sup>29</sup> Some examples are Rakestraw, J.F. (1987). *An analysis of home schooling for elementary school-age children in Alabama*. Unpublished doctoral thesis, University of Alabama, Tuscaloosa;

Delahooke, M.M. (1986). *Home educated children's social/ emotional adjustment and academic achievement: A comparative study*. Doctoral dissertation, California School of Professional Psychology, Los Angeles;

Tipton, M. (1990). Untitled, unpublished MA thesis. Antioch University.

<sup>30</sup> Muchmore, J., personal communication, September 21, 2007.

<sup>31</sup> Ray, B.D. (2000). Home schooling: The ameliorator of negative influence on learning? *Peabody Journal of Education*, 75(1&2), 71-106.

<sup>32</sup> Further widely cited reports by Cisek, G. J. (1991). Alternative assessments: Promises and problems for home-based education policy. *Home School Researcher*, 7(4), 13-21; Lyn Boulter from Boulter, L. & Macaluso, K. (1994). Individualized assessment of home schooling education. *Home School*



## The Impact of School Choice Reforms on Student Achievement

- Researcher*, 10(2), 1-6; Rhonda Galloway from Galloway, R. A. & Sutton, J. P. (1995). Home schooled and conventionally schooled high school graduates: A comparison of aptitude for and achievement in college English. *Home School Researcher*, 11(1), 1-9; Wartes, J. (1987). Report from the 1986 home school testing and other descriptive information about Washington's home schoolers: A summary. *Home School Researcher*, 3(1), 1-4. Other well-known homeschool researchers not studying academic achievement also made their public reputation known through *Home School Researcher*, such as Marlee Mayberry and J. Gary Knowles. In this analysis, Qadish (2007) is sourced from *Home School Researcher*.
- <sup>33</sup> Rudner, L.M. (1999). Scholastic achievement and demographic characteristics of home school students in 1998. *Education Policy Analysis Archives*, 7(8). (ERIC Document Reproduction Service No. ED 435 709)
- <sup>34</sup> Welner, K.M. & Welner, K.G. (1999). Contextualizing homeschooling data: A response to Rudner. *Education Policy Analysis Archives*, 7(13). (ERIC Document Reproduction Service No. EJ 588 923)
- <sup>35</sup> Galloway, R.A.S. (1995). *Home schooled adults: Are they ready for college?* Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco. P. 8
- <sup>36</sup> Cooper, B.S. & Sureau, J. (2007). The politics of homeschooling. *Educational Policy*, 21(1), 110-131; Welner, K.M. & Welner, K.G. (1999). Contextualizing homeschooling data: A response to Rudner. *Education Policy Analysis Archives*, 7(13). (ERIC Document Reproduction Service No. EJ 588 923); McDowell, S.A., Sanchez, A.R. & Jones, S.S. (2000). Participation and perception: Looking at home schooling through a multicultural lens. *Peabody Journal of Education*, 75(1&2), 124-146.
- <sup>37</sup> Rothermel, P. (2004). Home-education: Comparison of home- and school-educated children on PIPS baseline assessments. *Journal of Early Childhood Research*, 2(3), 273-299; and Watkins, M.A.C. (1997). *Are there significant differences in academic achievement between homeschooled and conventionally schooled students?* Doctoral dissertation, University of Alberta. These initially looked promising, but were based on data outside of the United States.
- <sup>38</sup> In this section, we use the words “magnet schools” consistently, but we intend that term to also include interdistrict and intradistrict forms of school choice. We know the terms are not synonymous, but to use each term individually throughout the section would be laborious for the reader. If we were to put each of these forms of school choice into their own sections, the sections would be so small as to be not worth reporting. Finally, intradistrict choice also goes by the name of “open enrollment” in the research literature, and that is also included in this section.
- <sup>39</sup> See Fuller, B., Burr, E., Huerta, L., Puryear, S., & Wezler, E. (1999). *School choice: Abundant hopes, scarce evidence of results*. Berkeley, CA: Policy Analysis for California Education for a discussion of this aspect of magnet schooling.
- <sup>40</sup> Ballou, D., Goldring, E., & Liu, K. (2006, March). *Magnet schools and student achievement*. New York: National Center for the Study of Privatization in Education, Columbia University.
- <sup>41</sup> Gamoran, A. (1996). Student achievement in public magnet, public comprehensive, and private city high schools. *Education Evaluation and Policy Analysis*, 18(1), 1-18.
- <sup>42</sup> Christenson, B., Eaton, M., Garet, M. S., Miller, L. C., Hikawa, H., & DuBois, P. (2003). *Evaluation of the magnet schools assistance program, 1998 grantees*. Washington, DC: U. S. Department of Education, Office of the Under Secretary.
- <sup>43</sup> Gill, B., Timpane, P.M., Ross, K.E., & Brewer, D.J. (2001). *Rhetoric versus reality: What we know and what we need to know about vouchers and charter schools*. Santa Monica: RAND.
- <sup>44</sup> Miron, G., & Nelson, C. (2001) *Student academic achievement in charter schools: What we know and why we know so little*. Occasional Paper No. 41. National Center for the Study of Privatization in Education, Teachers College, Columbia University.

## The Impact of School Choice Reforms on Student Achievement

- <sup>45</sup> Miron, G., & Nelson, C. (2004). Student achievement in charter schools. In K. Bulkley & P. Wohlstetter (Eds.), *Taking account of charter schools* (pp. 161-175). New York: Teachers College Press.
- <sup>46</sup> In more recent years, a number of reviews and Web-based systems have facilitated the tracking of charter school research. Most noteworthy is an online searchable database of studies maintained by the National Charter School Research Project at the University of Washington. This useful database now contains 70 studies that consider charter school student achievement, although many lacked the technical reports or comparison groups required for this synthesis. (See National Charter School Research Center. University of Washington. Retrieved March 7, 2008, from [http://www.ncsrp.org/cs/csr/print/csr\\_docs/pubs/achieve\\_wp.htm](http://www.ncsrp.org/cs/csr/print/csr_docs/pubs/achieve_wp.htm) )
- Hassel et al. (2007) also identified 70 charter school studies, although his particular methodology allowed for some studies to be counted 2 or 3 times. Diverse studies and publications on charter schools are also available from the searchable database on the US Charter Schools Web site—maintained by varied charter supporters. See Hassel, B., Terrell, M.G., Kain, A., Ziebarth, T. (2007). *Charter school achievement: What we know*. Washington, DC: National Alliance for Public Charter Schools. Retrieved March 7, 2008, from <http://www.publiccharters.org/content/publication/detail/2974/>
- <sup>47</sup> Zimmer, R., Buddin, R, Chau, D., Gill, B., Guarino, C., Hamilton, L., Krop, C., McCaffrey, D., Sandler, M., & Brewer, D. (2003). *Charter school operation and performance: Evidence from California*. Santa Monica: RAND.
- <sup>48</sup> Miron, G., Nelson, C., & Risley, J. (2002). *Strengthening Pennsylvania's charter school reform: Findings from the statewide evaluation and discussion of relevant policy issues*. Harrisburg, PA: Pennsylvania Department of Education.

## APPENDIX A VOUCHER STUDIES

Code Used in the Chart	Title of Study/Evaluation	Description of the Study	Key Findings	Study Design 0-10	Duration of Study 0-4	Controls Used 0-6	Measure of Performance 0-2	Scope of the Study			Completeness of the Technical Report			Quality Rating 0-32	Impact Rating -2 to +2
								Scope of study 0-3	Grades levels covered 0-2	Subjects covered 0-1	Clear and complete methods section 0-2	Complete set of findings 0-1	Limitations of study included 0-1		
<b>M1</b>	Witte, J.F. (1998). The Milwaukee voucher experiment. <i>Educational Evaluation and Policy Analysis</i> , 20(4), 229-251.	Student level data for more than 1,300 students (slightly different in math and reading); Iowa Test of Basic Skills scores in reading and mathematics collected for 1990-1994; because of problems with lottery losers as unreliable comparison group, Witte compares choice participants and MPS low-income students <i>Limitations:</i> Limited generalizability, because of matched design, controls used may not control for differences adequately	Mixed: No substantial difference over the life of the program between choice and MPS families, especially MPS low-income students	8	4	5	2	3	2	1	2	1	1	29	0
<b>M2</b>	Greene, J.P., Peterson, P.E., & Du, J. (1999). Effectiveness of school choice: The Milwaukee experiment. <i>Education and Urban Society</i> , 31, 190-213.	Individual level scores on math and reading Iowa Test of Basic Skills, lottery winners compared with lottery losers in a randomized design <i>Limitations:</i> Questions raised about legitimacy of comparison group; did not report findings based on low-income MPS students; nonrandomness of attrition.	Strongly positive: Statistically significant changes for winners in their third and fourth year in the program when demographic controls are used	10	4	4	2	2	2	1	2	0	0	27	2
<b>M3</b>	Rouse, C. (1998). Private school vouchers and student achievement: An evaluation of the Milwaukee Parental Choice Program. <i>Quarterly Journal of Economics</i> , 113(2), 553-602.	This analysis sample consists of African-American and Hispanic students who applied to the choice program between 1990 and 1993 for grades K-8; compares the test scores of students selected to attend a participating private school with those of unsuccessful applicants and other students from the Milwaukee Public Schools <i>Limitations:</i> Questions about the nonrandomness of attrition	Slightly positive: Students selected for the choice program scored approximately 1.5 to 2.3 percentile points higher per year in math compared with unsuccessful applicants and the sample MPS students. Math learning gains are higher for choice students and statistically significant; however, reading	10	4	6	2	2	1	1	2	1	1	30	1
<b>C1</b>	Metcalfe, K.K., Legan, N.A., Paul, K.M., & Boone, W.J. (2004, October). <i>Evaluation of the Cleveland scholarship and tutoring program: Technical report 1998-2003</i> . Bloomington: Indiana University, School of Education.	The study followed 780 first-grade scholarship students attending private schools, 541 first-grade public school applicant nonrecipients, and 1,233 first-grade nonapplicants; achievement data collected from same cohort each spring; reports findings from autumn, 1998 (early first grade) through spring 2003 (late fifth grade). <i>Limitations:</i> Questions about the nonrandomness of attrition	Mixed: Program does not show any substantial gains for voucher users relative to other comparison groups. The CSTP is not differentially effective for African-American students.	10	4	4	2	3	0	1	2	1	1	28	0
<b>C2</b>	Greene, J.P., Howell, W.G., & Peterson, P.E. (1999). <i>An evaluation of the Cleveland voucher program after two years</i> . Harvard University, Program on Education Policy and Governance.	California Achievement Test in fall 1996 and spring 1997 and spring 1998 scores for two academies were collected and group learning gains determined; 2 academies used were created in response to the Cleveland Scholarship Program; average student gains from these schools compared with national average <i>Limitations:</i> Only 2 academies from the program used; school level data; compared with national averages, not a specific comparison group with	Mixed: During first year, NPRs in both math and reading rose significantly but did not continue to rise during the second year; some actually declined, one score significantly declined. However, authors recommend that program is continued	4	3	0	2	1	0	1	1	1	1	14	0
<b>C3</b>	Plucker, J., Muller, P., Hansen, J., Ravert, R., & Makel, M. (2006). <i>Evaluation of the Cleveland Scholarship and Tutoring Program: Technical report 1998-2004</i> . Bloomington, IN: Center for Evaluation and Education Policy.	Student level data used in mixed model, longitudinal approach on Terra Nova standardized test scores; controls for prior achievement, student mobility, and poverty status included. Compares lottery winners and nonwinners over time. <i>Limitations:</i> Missing data for some students had to be mathematically estimated; controls are	Positive: In first and second grades, CSTP outperformed public school students; but with more exposure, differences disappeared (except for language arts, in which CSTP maintained higher scores)	10	4	4	2	3	1	1	2	1	1	29	1
<b>NY1</b>	Mayer, D.P., Peterson, P.E., Myers, D.E., Tuttle, C.C., & Howell, W.G. (2002). <i>School choice in New York City after three years: An evaluation of the school choice scholarships program (No. 8404-045)</i> . Princeton, NJ: Mathematica Policy Research	Compares Iowa Test of Basic Skills scores of lottery winners and nonwinners for baseline and for several subsequent years; uses several demographic controls <i>Limitations:</i> Response bias, attrition	Slightly positive: After 3 years, no significant difference; some positive results for African Americans	10	4	5	2	3	0	1	2	1	1	29	1

Code Used in the Chart	Title of Study/Evaluation	Description of the Study	Key Findings	Study Design 0-10	Duration of Study 0-4	Controls Used 0-6	Measure of Performance 0-2	Scope of the Study			Completeness of the Technical Reporting			Quality Rating 0-32	Impact Rating -2 to +2
								Scope of study 0-3	Grades levels covered 0-2	Subjects covered 0-1	Clear and complete methods section 0-2	Complete set of findings 0-1	Limitations of study included 0-1		
NY2	Krueger, A.B., & Zhu, P. (2004). Another look at the New York City voucher experiment. <i>American Behavioral Scientist</i> , 47(5), 658-698.	Data were collected from low income students in grades k-4 and their parents at baseline and in the spring of each of the next 3 years. Base weights constructed so sample was representative of the pool of eligible applicants. Students were given the Iowa Test of Basic Skills (ITBS) at baseline and in the spring of each of the 3 follow-up years. Study compares gains over time for lottery winners and losers. <i>Limitations:</i> Lack of generalizability to other grades and voucher programs	Mixed: When students with missing baseline scores are taken into account, results are insignificant	10	4	4	2	3	0	1	2	1	1	28	0
DC1	Wolf, P.J., Howell, W.G., & Peterson, P.E. (2000). <i>School choice in Washington, DC: An evaluation after one year</i> . Cambridge, MA: Program on Education Policy and Governance, Harvard University.	Involved 1,584 students in grade 2-8 who applied to scholarship and had not previously attended a private school; students tested at baseline and follow up sessions, scholarship winners and nonwinners were compared in terms of Iowa Test of Basic Skills gains in math and reading <i>Limitations:</i> Did not look at high school effects; legitimacy of comparison group questioned; attrition patterns may be threat to internal validity; no significance for any racial group except African Americans but conclusions reported as extremely positive	Slightly positive: African-Americans switching to private schools in grades 2 through 5 outperformed public school students by 3 percent in reading (not statistically significant), 7 percent in math (statistically significant); African American students attending private schools in grades six through eight scored 2 national percentile points higher in math (not statistically significant) but trailed their public school peers in reading by 8 points (statistically significant).	10	1	3	2	3	1	1	1	1	0	23	1
DC2	Wolf, P., Gutmann, B., Puma, M., Rizzo, L., & Eissa, N. (2007). <i>Evaluation of the DC Opportunity Scholarship Program: Impacts After One Year</i> . Washington: Institute of Education Sciences, U.S. Department of Education.	Randomized controlled trial used to assess the first-year impacts of the Program on those who applied for and were given the option. OSP impact sample group includes the randomly assigned members of the treatment and control groups and comprises 57 percent of all eligible applicants in the first 2 years of Program operation <i>Limitations:</i> only one year of data, not generalizable to other programs	Mixed: No statistically significant impacts, positive or negative, on student reading or math achievement for the entire impact sample in year 1, or on subgroups	10	1	3	2	2	2	1	2	0	0	23	0
D1	West, M.R., Peterson, P.E., & Campbell, D.E. (2001, August). <i>School choice in Dayton, Ohio after two years: An evaluation of the Parents Advancing Choice in Education scholarship program</i> . Cambridge, MA: Program on Education Policy and Governance, Harvard University	Included 458 of 803 included in Howell, & Peterson (2000). Statistical model estimated to take nonrandomness of the placement of students in public and private schools. Each student's status as a member of the treatment or control group was used as an instrumental variable in a two stage least squares regression in which the dependent variable in the first-stage regression was whether or not the student attended a private school <i>Limitations:</i> Positive for one subgroup in some areas, but expressed as positive rather than mixed; attrition may be important	Slightly positive: After two years African American students who attended private schools scored higher in reading and on combined reading and math score. Their score also increased in math, although not statistically significant. Non-African American students did not differ significantly	10	3	3	2	2	1	1	1	1	0	24	1
C4	Belfield, C. (2006). <i>The evidence on education vouchers: An application to the Cleveland Scholarship and Tutoring Program</i> . Occasional Paper 112. New York: National Center for the Study of Privatization in Education.	Compares TerraNova scores for scholarship users, non-users, rejected applicants and a public school comparison group <i>Limitations:</i> Lack of generalizability to other programs, only used second and fourth graders	Mixed: No academic advantages for voucher users in second or fourth grade; results do not vary according to: adjustments for prior ability, intention-to-treat versus treatment effects, and dosage differences; not differentially effective for African American students	8	3	4	2	3	0	1	1	1	0	23	0

## APPENDIX B HOMESCHOOL STUDIES

Code Used in the Chart	Title of Study/Evaluation	Description of the Study	Key Findings	Study Design 0-10	Duration of Study 0-4	Controls Used 0-6	Measure of Performance 0-2	Scope of the Study			Completeness of the Technical Report			Quality Rating 0-32	Impact Rating -2 to +2
								Scope of study 0-3	Grades levels covered 0-2	Subjects covered 0-1	Clear and complete methods section 0-2	Complete set of findings 0-1	Limitations of study included 0-1		
A	Boulter, L.T. (1999). <i>Academic achievement in home school education</i> . Salisbury, NC: Catawba College.	Compares scores of homeschool students on Woodcock Johnson revised test with national average <i>Limitations:</i> Incomplete methods section (lack of information on sampling procedure); sample was all white, middle or upper-middle class, and demographic controls used; sample includes fewer than 50 homeschoolers	Mixed: Homeschoolers were at or above 50th percentile on all subsets of test, but percentile scores for all four clusters were negatively correlated with years in home schooling; significant decline in broad written language and broad knowledge	4	4	1	2	0	2	1	1	1	0	16	0
B	Collom, E. (2005). The ins and outs of homeschooling: The determinants of parental motivations and student achievement. <i>Education and Urban Society</i> , 3(3), 307-335.	Compares scores of 175 homeschooled students on SAT9 to the national average <i>Limitations:</i> Limited design in one school, hinged on option to complete parental motivation survey that was merged with test data, school factor clouds results, cross sectional	Slightly positive: Homeschoolers scored in the 54th percentile on reading, language, and math	0	0	4	2	2	2	1	2	1	1	15	1
C	Galloway, R.A.S. (1995, April). <i>Home schooled adults: Are they ready for college?</i> Paper presented at the annual meeting of the American Educational Research Association, San Francisco.	Compares homeschooled graduates with both private and public graduates who all attend the same Christian university on ACT scores <i>Limitations:</i> Sample was taken from one Christian university, no demographic controls used, cross sectional	Slightly positive: Only significant difference was for English subset ACT scores—significantly higher for home school students over private school graduates ONLY; no other statistically significant differences were found between the groups	0	0	0	2	2	0	0	2	1	0	7	1
D	Ray, B.D. (2000). Home schooling: The ameliorator of negative influences on learning? <i>Peabody Journal of Education</i> , 75(1-2), 71-106.	Compares self-reported homeschoolers' scores on various tests obtained through home education organizations' mailing lists to national averages <i>Limitations:</i> Cross-sectional, uses self-report measures, sample obtained through home education organizations' mailing list so representativeness of all homeschoolers is in question, scores on various tests reported	Strongly positive: Homeschoolers scored at 87th percentile in reading, math 82nd, complete battery 87th	0	0	3	2	3	0	1	1	1	1	12	2
E	Rudner, L.M. (1999). Scholastic achievement and demographic characteristics of home school students in 1998. <i>Education Policy Analysis Archives</i> , 7(8).	Obtains sample from those homeschoolers using a particular testing center; compares scores of homeschooled children with national averages for "grade level" <i>Limitations:</i> Testing site at Bob Jones University so representativeness of all homeschoolers is questionable, cross-sectional, no demographic controls used	Strongly positive: Median scores for homeschoolers at 75th percentile	0	0	0	2	3	2	1	2	1	1	12	2
F	Clemente, D.F. (2006). <i>Academic achievement and college aptitude in homeschooled high school students compared to their private-schooled and public-schooled counterparts</i> . (UMI No. 3218862). Unpublished doctoral dissertation, Regent University, Virginia Beach.	Compares SAT scores of college freshmen who previously had been homeschooled with those who graduated from public and private high schools; sample obtained from 7 Christian colleges and universities <i>Limitations:</i> Limited generalizability due to sample used, questionable appropriateness of using a directional analysis of variance analysis, cross-sectional	Strongly positive: SAT scores for homeschoolers significantly higher using both data analyses; difference between public and private schooled freshmen's SAT scores not significant	0	0	0	2	3	0	0	2	1	1	9	2
I	Gray, D.W. (1998). <i>A study of the academic achievements of home-schooled students who have matriculated into post-secondary institutions</i> . (Doctoral dissertation, University of Florida, Sarasota, 1998). Dissertation Abstracts International, 59(021).	Compares SAT scores of random sample of public and private school graduates with population of previously homeschooled college freshmen at three Georgia universities <i>Limitations:</i> Homeschooled could not be separated from those with GED, limited generalizability due to sample used	Slightly positive: Slightly higher scores for homeschooled though not statistically significant	0	0	0	2	1	0	1	2	1	1	8	1
J	Holder, M.A. (2001). <i>Academic achievement and socialization of college students who were homeschooled</i> . Unpublished doctoral dissertation, The University of Memphis. (UMI No. 3829894).	Compares ACT scores for random sample of public school graduates and population of homeschooled from one university <i>Limitations:</i> Small sample size (N=34), limited generalizability due to sample being taken from one university, cross-sectional, no demographic controls used	Mixed: No statistically significant differences in ACT scores among homeschooled and public schooled students	0	0	0	2	0	0	1	2	1	0	6	0

Code Used in the Chart	Title of Study/Evaluation	Description of the Study	Key Findings	Study Design 0-10	Duration of Study 0-4	Controls Used 0-6	Measure of Performance 0-2	Scope of the Study			Completeness of the Technical Report			Quality Rating 0-32	Impact Rating -2 to +2
								Scope of study 0-3	Grades levels covered 0-2	Subjects covered 0-1	Clear and complete methods section 0-2	Complete set of findings 0-1	Limitations of study included 0-1		
K	Witt, V.L. (2005). A comparison and descriptive analysis of homeschool reading and vocabulary scores to the national average. <i>Dissertation Abstracts International</i> , 65(01), 1696. (UMI No. 3174333).	Compares homeschooled students' percentiles on reading and vocabulary subtests of California TerraNova with national averages. Data came from existing database, but participants were selected by parents who returned questionnaire <i>Limitations:</i> Small sample size (N=103), cross-sectional, representativeness of all homeschool students questionable	Strongly positive: Homeschooled math scores at 79th percentile, vocabulary at 78.5 percentile	0	0	0	2	1	2	0	2	1	1	9	2
H	Delahooke, M.M. (1986). Home educated children's social/emotional adjustment and academic achievement: A comparative study. <i>Dissertation Abstracts International</i> , 47 (2), 475A. (UMI No. 8608759).	Compares homeschooled students' scores to private school students' scores on parts of Wide Range Achievement Test-Revised <i>Limitations:</i> Small sample size (N=60), no random selection, participants chose to participate in study from private and homeschool settings, no demographic controls used, cross-sectional	Mixed: Study found no differences in test results on parts of Wide Range Achievement Test-Revised	0	0	0	2	1	1	1	2	1	1	9	0
G	Qaqish, B. (2007). An analysis of homeschooled and non-homeschooled students' performance on an ACT mathematics achievement test. <i>Home School Researcher</i> , 17(2), 1-12.	Compares homeschoolers' ACT mathematics scores to non-homeschoolers' ACT mathematics scores using matched student design <i>Limitations:</i> Cross-sectional, math only	Slightly positive: On average, non-homeschoolers performed better than homeschoolers, by about 2 items out of 60 items, on the ACT mathematics test that was analyzed	8	0	2	2	3	0	0	2	1	0	18	1
N	Rakestraw, J. (1988, December). Home schooling in Alabama. <i>Home School Researcher</i> , 4(4).	Compares homeschooled students' scores on SAT with "grade level" <i>Limitations:</i> limited generalizability because homeschooled participants were solicited through home education organizations/church ministries; small sample size; technical report is unclear about comparison groups, sample and sampling procedures; no limitations discussed and complete findings are not presented; cross-sectional	Slightly positive: The academic achievement of the homeschooled children in Alabama was at grade level or above in almost all subject areas, except mathematics for Grades 1 and 4 and in reading comprehension and vocabulary for Grade 5, in which homeschoolers were below grade level	0	0	0	0	1	0	1	0	0	0	2	1
L	Richman, H.B., Girtten, W., & Snyder, J. (1990). Academic achievement and its relationship to selected variables... <i>Home School Researcher</i> , 6(4), 9-16.	Compares homeschoolers' standardized test scores with national averages <i>Limitations:</i> nonrandom sampling (parents had to pay to take test), small sample size, cross-sectional	Strongly positive: Math score for homeschoolers corresponded to 73rd national percentile; reading score correlated with 86th national percentile rank for achievement test	0	0	0	2	2	2	1	0	1	1	9	2
M	Wartes, J. (1990). Recent results from the Washington homeschool research project. <i>Home School Researcher</i> , 6(4), 1-7.	Compares homeschoolers' scores on Stanford Achievement Test to national norms; multiple years of data gathered but no longitudinal analysis <i>Limitations:</i> Complete findings are not presented, cross sectional	Slightly positive: Homeschoolers scored comparably to public composite scores for 1986 69th percentile, 1987 65th or 66th, 1988 65th percentile, 1989 65th percentile	0	0	3	2	3	2	1	0	0	1	12	1
O	Jones, P., & Gloeckner, G. (2004). First year college performance: A study of home school graduates and traditional school graduates. <i>The Journal of College Admission</i> , 17-20.	Compares homeschooled and nonhomeschooled college freshmen ACT scores <i>Limitations:</i> Small sample size (N=108), insufficient demographics reported on sample, limited generalizability due to sample tested, cross-sectional, no control for demographics used	Slightly positive: Homeschoolers scored higher on ACT but not significantly so. More variance in homeschoolers' scores	0	0	0	2	1	0	1	0	1	0	5	1
P	Frost, E.A. (1987). A descriptive study of the academic achievement of selected elementary school-aged children educated at home in five Illinois counties. (Doctoral dissertation, Northern Illinois University, 1987). <i>Dissertation Abstracts International</i> , 48(7), 1589A.	Sample of 74 students from personal contacts with homeschool educators; uses group level characteristics to select comparison groups <i>Limitations:</i> Nonrandom sampling, limited generalizability, cross-sectional	Mixed: Homeschoolers were above grade level in reading, but below grade level in math. Findings ultimately presented as composite, masking inferior math test scores by combining them with test data on unusual subject areas like "work study skills"	0	0	1	1	1	0	2	2	1	1	7	0
Q	Belfield, C.R. (2005). Home-schoolers: How well do they perform on the SAT for college admissions? In B.S. Cooper (Ed.), <i>Home schooling in full view: A reader</i> . Charlotte, NC: Information Age Publishing.	Compares one year of national SAT scores with large national sample of homeschoolers *large sample of reported homeschoolers (N=6033) <i>Limitations:</i> Cross-sectional, description of access to population absent	Mixed: Homeschooled students scored high on reading but lower than comparison on math. When demographic controls introduced, there were no noticeable differences between groups	0	0	4	2	3	0	2	1	0	0	12	0

## APPENDIX C INTER-, INTRADISTRICT CHOICE AND MAGNET SCHOOL STUDIES

Code Used in the Chart	Title of Study/Evaluation	Description of the Study	Key Findings	Study Design 0-10	Duration of Study 0-4	Controls Used 0-6	Measure of Performance 0-2	Scope of the Study			Completeness of the Technical Report			Quality Rating 0-32	Impact Rating -2 to +2
								Scope of study 0-3	Grades levels covered 0-2	Subjects covered 0-1	Clear and complete methods section 0-2	Complete set of findings 0-1	Limitations of study included 0-1		
<b>A</b>	Beaudin, B. (2003). <i>Interdistrict magnet schools and magnet programs in Connecticut: An evaluation report</i> . Bureau of Evaluation and Educator Standards, Division of Evaluation and Research.	Compares cut scores of interdistrict magnet schools with statewide averages over two years of test data <i>Limitations:</i> No demographic controls used, no understanding of value added by reform, school level data	Mixed: Positive results for interdistrict magnet schools on one standardized test, negative results on the other standardized test	0	3	0	1	3	1	1	1	0	0	10	0
<b>B</b>	Ballou, D., Goldring, E., & Liu, K. (2006, March). <i>Magnet schools and student achievement</i> . New York: National Center for the Study of Privatization in Education, Columbia University.	Compares lottery winners with losers, adding controls for 7 potential confounding variables <i>Limitations:</i> One district studied, no data on magnet high schools	Mixed: Positive impact of magnet schools on mathematics scores until prior achievement and student demographics are taken into account, suggesting attrition patterns are causing differences in scores	10	4	4	2	3	1	1	2	1	1	27	0
<b>C</b>	Crain, R.L., Allen, A., Thaler, R., Sullivan, D., Zellman, G., Little, J.W., & Quigley, D.D. (1992). <i>The effects of academic career magnet education on high schools and their graduates</i> . Berkeley, CA: NCRVE	Aggregates student level data to program level and compares randomly accepted students' scores with randomly rejected students' scores <i>Limitations:</i> Sample of programs not defined, not generalizable to all magnet programs, cross-sectional	Slightly negative: Students in academic career magnet schools do not have higher or lower reading scores, but do have slightly lower math scores	10	0	0	0	3	1	1	2	1	0	17	-1
<b>D</b>	Gamoran, A. (1996). Student achievement in public magnet, public comprehensive, and private city high schools. <i>Education Evaluation and Policy Analysis</i> , 18(1), 1-18.	Using NELS test data, compares gains from eighth to tenth grade for magnet schools, public comprehensive schools, and Catholic schools <i>Limitations:</i> Old data, school level data	Slightly positive: Magnet school advantages in reading and social studies	4	3	3	2	3	1	1	2	1	1	21	1
<b>E</b>	Heebner, A.L. (1995). The impact of career magnet high schools: Experimental and qualitative evidence. <i>Journal of Vocational Education Research</i> , 20(2), 27-35.	Uses data from five schools in one city to compare lottery winners and nonwinners on pretest and posttest <i>Limitations:</i> Incomplete methods sections (years of data obtained are unclear), not generalizable to other programs	Slightly positive: Lottery winners had higher math scores; students with medium reading scores benefited from winning the lottery	10	3	2	2	1	0	1	1	1	1	22	1
<b>F</b>	Institute for Assessment and Evaluation. (2006). <i>Knox County magnet schools evaluation</i> . Knoxville: Author, University of Tennessee.	Uses county data to track consecutive cohorts over four years; gains compared with national norms <i>Limitations:</i> No demographic controls used, no data on high schools, school level data	Strongly negative: Magnet schools perform more poorly than in Knox County and the state mean	1	4	1	2	1	1	1	2	1	0	14	-2
<b>G</b>	Christenson, B., Eaton, M., Garett, M.S., Miller, L.C., Hikawa, H., & DuBois, P. (2003). <i>Evaluation of the magnet schools assistance program, 1998 grantees</i> . Washington: U. S. Department of Education, Office of the Under Secretary.	Using national school level data, schools are matched based on student demographics and gains compared for matched magnet and traditional public schools <i>Limitations:</i> Multiple state tests used, school level data, data limited to elementaries only	Mixed: When controls for the composition of the schools used, gains of MSAP-sponsored schools were not significantly different than others.	1	3	4	2	3	0	0	2	1	1	17	0
<b>H</b>	Betts, J.R., Rice, L.A., Zau, A.C., Tang, Y.E., & Koedel, C.R. (2006). <i>Does school choice work? Effects on student integration and achievement</i> . San Francisco: Public Policy Institute of California.	Compares three forms of intradistrict choice in San Diego district using natural lottery to compare winners and nonwinners <i>Limitations:</i> Incomplete methods section (no sample size), comparison limited to one district	Slightly positive: Magnet enrollees showed higher scores in high school math in the second and third year of school placement	10	3	5	2	3	2	1	1	1	1	29	1
<b>I</b>	Eagle, N., & Ridenour, G. (1969). Differences in academic performance and report card grades between "open enrollment" and "matched home" elementary school children, after one and two years. <i>Urban Education</i> , 4, 115-123.	Focuses on effect of desegregation on academic achievement <i>Limitations:</i> Old data, small sample size, few demographic controls utilized, limited generalizability	Slightly positive: magnet enrollees did not show differences after one year of treatment; but as grade level increased, so did a statistically significant achievement level	8	3	2	2	2	0	1	0	0	0	19	1

## APPENDIX D CHARTER SCHOOL STUDIES

Code Used in the Chart	Title of Study/Evaluation	Description of the Study (include details about the design, comparison groups, test and outcome measure used, and scope of study)	Key Findings (Include rating and then bulleted summary of key findings)	Study Design	Duration of Study	Controls Used	Measure of Performance	Scope of the Study			Completeness of the Technical Report			Quality Rating	Impact Rating
								Scope of study 0-3	Grades levels covered 0-2	Subjects covered 0-1	Clear and complete methods section 0-2	Complete set of findings 0-1	Limitations of study included 0-1		
AZ1	Mulholland, L. (1999, March). <i>Arizona charter school progress evaluation</i> . Tempe: Morrison Institute for Public Policy, Arizona State University.	Analysis of consecutive cohorts with comparison group and statistical controls; stratified sample of individual gain scores from 82 out of 137 charter schools open in Arizona at the time <i>Limitations</i> : Low matching rate in high schools (32%-66%); rate is higher in charter schools	Mixed: No difference overall	1	3	0	2	3	2	1	2	1	1	16	0
AZ2	Solmon, L.C., & Goldschmidt, P. (2004). <i>Comparison of traditional public schools and charter schools on retention, school switching and achievement growth</i> . Policy Report: Goldwater Institute. No. 192.	Three-level hierarchical linear model used to measure achievement growth trajectories; used 158,000 test scores of more than 60,000 Arizona students attending 873 charter and traditional public schools statewide over a three-year period <i>Limitations</i> : None addressed, controls included may not address all differences in students	Slightly positive: Achievement growth varies by grade level; elementary charter school students' growth was higher; in middle grades, traditional and charter growth comparable; higher grades, traditional public school achievement growth was higher; overall charter school students gained faster	8	3	2	2	3	2	1	2	1	0	24	1
AZ3	Garcia, D.R. (2008). <i>Growing pains: Revisiting academic achievement in the earliest years of the charter school movement</i> . Manuscript submitted for publication.	Compares the academic achievement of charter and traditional public elementary students while controlling for prior achievement, grade, student demographics, school mobility, and student entrance into a first-year charter school <i>Limitations</i> : Differences may not be adequately controlled for	Slightly positive: Charter schools outperform traditional public schools in total scores; advantages largely attributable to greater achievement gains relative to traditional public schools in the basic skills areas of reading vocabulary and mathematics procedures	8	4	3	2	3	1	1	2	1	0	25	1
CA1	EdSource. (2007). <i>California's charter schools: Measuring their performance</i> . Mountainview, CA: Author.	Cross-sectional analysis with statistical controls used to compare charter schools scores with noncharter school scores; 60% of charter schools in operation in 05-06 and 79% of noncharter schools in operation in same year <i>Limitations</i> : Doesn't account for motivation or differences in funding; cross-sectional, school level data	Mixed: Negative for elementary charters, positive for middle school charters, positive but inconsistent for charter high schools	0	0	4	2	3	2	1	2	1	1	16	0
CA2	Rogosa, D. (2003). <i>Student progress in California charter schools, 1999-2002</i> . Palo Alto, CA: Stanford University.	Controls for API and Stanford 9 test scores; all students in 93 charter schools and 6,584 noncharter schools in most complete analysis; uses consecutive cohort and same cohort designs <i>Limitations</i> : School level data, controls may not be adequate	Mixed: More comparable gains than in Rogosa (2002)	4	4	1	2	3	2	1	2	1	1	21	0
CA3	Raymond, M.E. (2003). <i>The performance of California charter schools</i> . Palo Alto, CA: CREDO: Hoover Institution, Stanford University.	Multivariate regression models were constructed for each year of API scores from 1999 to 2002, regressing school scores on student body characteristics, family education characteristics and school attributes <i>Limitations</i> : Shortcomings of the API, school level data	Slightly positive - Against all other California schools, the changes in charter schools' API scores at the elementary and middle school levels are not statistically different, but slightly lower. Compared with other California high schools, California charter high schools on average have growth in API scores that is positive and statistically significant. Charter elementary and middle schools were found to create equivalent gains for students as their conventional peer schools. Charter high schools produced significantly more positive changes in API scores	1	3	2	2	3	2	1	1	1	1	17	1
CA4	Zimmer, R., Buddin, R., Chau, D., Gill, B., Guarino, C., Hamilton, L., Krop, C., McCaffrey, D., Sandler, M., & Brewer, D. (2003). <i>Charter school operation and performance: Evidence from California</i> . Santa Monica: RAND.	Approach III: Longitudinally links student-level data → value-added estimate of the contribution of charter schools to student achievement. <i>Limitations</i> : Availability of data in only a few districts; no comparison between different types of charter schools possible	Mixed: Slightly negative for math score comparisons on primary and secondary school level; reading—positive for secondary school level in comparison with public schools, but neutral for primary school	8	4	3	2	3	1	1	1	0	1	24	0
CO	Colorado Department of Education. (2006). <i>The state of charter schools in Colorado in 2004-05: The characteristics, status, and performance record of Colorado charter schools</i> . Denver: Author.	Comparison of average charter school % meeting standards and noncharter school students meeting standards <i>Limitations</i> : No use of gain score or controls; cut score is used	Mixed: Charter schools scored better in lower grades; noncharter school students scored better in high school grades	0	0	0	1	3	2	1	2	1	0	10	0
CT	Miron, G. (2005). <i>Evaluating the performance of charter schools in Connecticut</i> . Kalamazoo: The Evaluation Center, Western Michigan University.	Looks at changes in average scaled scores for same and consecutive cohorts <i>Limitations</i> : School level data, CAPT had weaker design	Slightly positive: 3 of 4 cohorts in lower grades made much larger gains than comparison groups, but 10th grade results mixed to negative	4	3	0	2	1	2	1	2	1	1	17	1



Code Used in the Chart	Title of Study/Evaluation	Description of the Study (include details about the design, comparison groups, test and outcome measure used, and scope of study)	Key Findings (Include rating and then bulleted summary of key findings)	Study Design	Duration of Study	Controls Used	Measure of Performance	Scope of the Study			Completeness of the Technical Report			Quality Rating	Impact Rating
								Scope of study 0-3	Grades levels covered 0-2	Subjects covered 0-1	Clear and complete methods section 0-2	Complete set of findings 0-1	Limitations of study included 0-1		
DC	Henig, J.R., Holyoke, T.T., Lacireno-Paquet, N., & Moser, M. (2001, February). <i>Growing pains: An evaluation of charter schools in the District of Columbia; 1999-2000</i> . Washington, DC: The Center for Washington Area Studies, The George Washington University.	Comparison of poorly performing public and charter schools with similar proportions of needy students; also a comparison of stability of test scores between the two types of school over time <i>Limitations</i> : Group level data	Strongly negative: <u>Consecutive cohorts</u> : DCPS schools more likely to have improved, less likely to have declined than charter schools. <u>Cross-sectional analysis</u> : more charter schools scored "below basic" than DCPS schools; differences hold up under statistical elaboration	1	3	1	2	3	2	1	2	1	1	17	-2
DE	Miron, G., Cullen, A., Applegate, E.B., & Farrell, P. (2007). <i>Evaluation of the Delaware charter school reform: Final report</i> . Kalamazoo: The Evaluation Center, Western Michigan University.	Students matched on 4 student-level characteristics; 4x4 factorial ANCOVA; for group or school level analysis, residual gains analysis was used <i>Limitations</i> : Cannot be generalized to other states' programs; controls may not adequately account for differences	Positive: Charter schools at secondary level gaining more as compared with traditional public school students	8	4	5	2	1	2	1	2	1	1	27	1
FL1	Florida Department of Education. (2006). <i>Florida's charter schools: A decade of progress</i> . Tallahassee: Author.	Examines change in FCAT Development Scale Score (DSS) from grade to grade for charter and traditional students from 2001-2002 to 2005-2006 <i>Limitations</i> : No demographic controls, no statistical significance tests	Mixed: No consistent pattern	4	4	0	2	3	2	1	1	1	0	18	0
FL2	Sass, T. R. (2006). <i>Charter schools and student achievement in Florida</i> . Gainesville, FL: American Education Finance Association.	Longitudinal data, control for student level fixed effects, uses econometric model of student achievement <i>Limitations</i> : Those who leave one form for another may have unobservable characteristics not controlled for	Slightly positive: Achievement initially lower in charters; but by fifth year of operation, achievement is on par and reading achievement scores are higher than traditional school counterparts	8	4	5	2	3	2	1	2	1	0	28	1
GA	Plucker, J., Eckes, S., Rapp, K., Ravert, R., Hansen, J., & Trotter, A. (2006, April). <i>Baseline evaluation of Georgia's charter schools program</i> . Atlanta: Georgia Department of Education.	Cross-sectional series of analyses of covariance (ANCOVA) were conducted, reliance on both statistical significance and effect size interpretation, controls for student ethnicity and gender <i>Limitations</i> : Incomplete methods section, cross-sectional, no control for SES	Mixed: Charter schools are achieving at similar levels as their peers statewide and in comparison schools, with significant variations by subject area, grade, and length of time attending charter schools; most differences favor charter schools, but not universal	0	0	2	1	3	1	1	1	1	1	11	0
ID	Ballou, D., Teasley B., & Zeidner, T. (2006). <i>Charter schools in Idaho</i> . Nashville, TN: National Center on School Choice. Prepared for the National Conference on Charter School Research at Vanderbilt University on September 29, 2006.	Student gain scores were calculated for students' math scores in grades 2-10; virtual schools (5) dropped from sample, and those students who switched during year dropped from sample; models created using ordinary least squares and controls for grade level, ethnicity, and special education <i>Limitations</i> : Fixed effects model and no fixed effects model produce completely different results, school level data	Mixed: Analysis of switchers favors CS, while simpler gains analysis does not. Elementary students in CS have made greater gains than they would have made had they remained in traditional public schools (though the difference in higher grades is reversed or insignificant). The smallest drop in gain scores occurred among students who moved from the district schools to CS. The largest drop occurred among students who moved in the opposite direction.	4	3	2	2	0	2	0	1	1	1	16	0
IL1	Hoxby, C.M., & Rockoff, J.E. (2004). <i>The impact of charter schools on student achievement</i> . Nashville: Working Paper Series, National Center on School Choice.	Compares gains for lottery winners and lottery losers; student level analysis for lottery applicants to 3 CICS schools in 2000, 2001, and 2002 <i>Limitations</i> : Not generalizable to nonapplicants; private school students can't be compared	Strongly positive: After 2 years in a charter school, average of 6 percentile points higher on standardized tests	10	3	5	2	3	1	1	2	1	1	29	2
IL2	Nelson, C., & Miron, G. (2002). <i>The evaluation of the Illinois charter school reform: Final report</i> . Report submitted to the Illinois State Board of Education. Kalamazoo: The Evaluation Center, Western Michigan University.	Compares percentages passing state tests in charter schools and demographically similar schools statewide <i>Limitations</i> : Cross-sectional, small sample of schools	Mixed: Statewide, charter schools perform slightly below demographically similar schools; in Chicago, charter schools have higher proportions scoring at or above national norms than do demographically similar schools	1	3	3	1	2	1	1	2	1	1	16	0
IL3	Chicago Public Schools. (2007). <i>Charter schools: 2005/2006 annual performance report</i> . Chicago: Author.	Compares percentage of high, middle, and low ratings received by 21 charter schools and district schools on absolute student and operational performance measures; looks at changes from 2002-2006 <i>Limitations</i> : Aimed at charter school supporters, school level data, use of general rating as measurement	Strongly positive: Charter schools had higher percentage of high and middle ratings than did district schools	0	1	0	1	2	2	1	0	1	0	8	2
MA	Massachusetts Department of Education. (2006). <i>Massachusetts charter school achievement comparison study: An analysis of 2001-2005 MCAS performance</i> . Boston: Author.	HLM growth models for each charter school and its corresponding comparison sending district <i>Limitations</i> : School level data, concerns about MCAS scaled scores and interpretation across 5-year period, length of charter school operation not taken into account	Slightly positive: HLM data show some charter scores as highest of all schools	4	4	3	2	3	2	1	1	1	1	22	1

Code Used in the Chart	Title of Study/Evaluation	Description of the Study (include details about the design, comparison groups, test and outcome measure used, and scope of study)	Key Findings (Include rating and then bulleted summary of key findings)	Study Design	Duration of Study	Controls Used	Measure of Performance	Scope of the Study			Completeness of the Technical Report			Quality Rating	Impact Rating
								Scope of study 0-3	Grades levels covered 0-2	Subjects covered 0-1	Clear and complete methods section 0-2	Complete set of findings 0-1	Limitations of study included 0-1		
MI1	Eberts, R.W., & Hollenbeck, K.M. (2002). <i>Impact of charter school attendance on student achievement in Michigan</i> . Kalamazoo, MI: Upjohn Institute Staff Working Paper. No. 02-080.	Pairs charter schools with public school districts, used fixed effects to control for factors in the areas common to both types of schools <i>Limitations:</i> No use of gain scores, cross-sectional only, analysis explains only small proportion of variance	Strongly negative: With student, building, and district controls, students attending charters have lower test scores	0	4	2	2	3	0	1	2	1	1	16	-2
MI2	Michigan Department of Education (December, 2007). <i>Public school academies: Michigan Department of Education report to the legislature</i> . East Lansing: Author.	Comparison of proficiency levels for PSAs, host districts, and non-PSAs for MEAP and other measures; broken down by age of PSA, economically disadvantaged students, ethnicity, students with disabilities, and correlation of proficiency level with percentage of free and reduced price lunch students (all controls/subgroups analyzed separately) <i>Limitations:</i> None addressed, cross-sectional, cut scores used, emphasis on elementaries and middle schools performing well	Slightly positive: Elementary and charter middle schools consistently have a higher percentage of proficient students on MEAP than do counterparts in geographical districts in which PSAs are located; charter high schools "are struggling"	0	0	3	1	3	2	1	1	1	0	12	1
MI3	Bettinger, E.P. (2005). The effect of charter schools on charter students and public schools. <i>Economics of Education Review</i> , 24(3), 133-147.	Estimates charter school achievement for charter schools opening in 1996/97; difference in difference estimator for consecutive cohorts; second model controls for ethnicity and free and reduced lunch <i>Limitations:</i> Group-level data, limited to charters opened in 1996-1997 school year	Slightly negative: charter schools' scores "may" decline; results are negative	8	3	2	2	3	0	1	2	1	1	23	-1
MI4	Miron, G., & Nelson, C. (2002). <i>What's public about charter schools? Lessons learned about choice and accountability</i> (pp. 134-147). Thousand Oaks, CA: Corwin.	Compares changes in school-level passing rates between charter schools and districts <i>Limitations:</i> School level data, passing rates as measure of performance	Slightly negative: Host districts' passing rate gains exceed charter school rate gains for all subjects and grades except 4th grade math	1	4	1	1	3	1	1	2	1	1	16	-1
MO	Metis Associates. (2004). <i>A study of the Kansas City, Missouri, charter public schools 2000-2003</i> . New York: Author.	Compares change in average charter school score with average change in district and state score <i>Limitations:</i> No controls used, group-level data	Slightly positive: Charter school students start out behind but close gap	1	4	0	2	2	2	1	1	1	1	15	1
NC1	Noblit, G.W., & Corbett, D. (2001). <i>North Carolina charter school evaluation report</i> . Raleigh: North Carolina State Board of Education.	Compares percentage of traditional public school students proficient with % of charter school students proficient <i>Limitations:</i> Percentage of students proficient used as measure of performance, cross-sectional	Strongly negative: Charter school students start with higher prior achievement scores, but lose ground to their peers in all grades and subject areas	8	4	1	2	2	1	1	2	1	1	23	-2
NC2	Billulco, R., & Ladd, H.F. (2006). <i>School choice, racial segregation and test-score gaps: Evidence from North Carolina's charter school program</i> . Paper presented at the annual meeting of Allied Social Science Associations, Boston.	Models include grade/year fixed effects and are estimated using "within" student estimator. Dependent variable is annual gain in end of grade development scale scores transformed into standard scores <i>Limitations:</i> not applicable to other states' charter schools; students who switch sectors may have unobservable characteristics that are not adequately controlled for, introducing sampling bias	Strongly negative: Charter schools have produced larger achievement gaps between Caucasian and African-American students	8	4	3	2	3	2	1	2	1	1	27	-2
NJ	Barr, J. (2007). <i>Charter school performance in New Jersey</i> . (Working Paper #2007-006). Newark: Rutgers University.	Regression analyses done on panel data of fourth graders from 1999 to 2006; 35 charter schools in 18 districts included; comparisons made only to those districts that have a charter school; first regression analysis looks at each school's passing rate on 4th grade standardized language arts and mathematics exams <i>Limitations:</i> School level data, cut score used as measure of performance	Slightly negative: Charter schools have lower performance than public schools in the same districts on fourth grade standardized tests for language and math, but performance improves with experience. The estimated time to close the gap between charter and traditional schools is about a decade	1	4	3	1	3	0	1	2	1	0	16	-1
NY1	New York Board of Regents. (2003). Report to the governor, the temporary president of the senate, and the speaker of the assembly on the educational effectiveness of the charter school approach in New York State.	Compares percentage of students passing from 2002-2003 between charter schools and their districts <i>Limitations:</i> Cross-sectional, school level data, no use of gain score	Slightly negative: No real aggregate results/conclusions presented, but for some charter schools, greater % classified with serious deficiencies	0	0	4	0	2	1	1	1	0	0	9	-1
NY2	Hoxby, C.M., & Murarka, S. (2007). <i>Charter schools in New York City: Who enrolls and how they affect their students' achievement</i> . Cambridge, MA: National Bureau of Economic Research.	Comparison of students who are lotteried-in and lotteried-out of charter schools using instrumental variables regression <i>Limitations:</i> Known underreporting of special education status	Strongly positive: For every year in charter schools, students gain 3.8 scale score points in math (12% of performance level), 1.6 scale score points in reading (3.5% performance level)	10	4	2	2	3	2	1	2	1	1	28	2

Code Used in the Chart	Title of Study/Evaluation	Description of the Study (include details about the design, comparison groups, test and outcome measure used, and scope of study)	Key Findings (Include rating and then bulleted summary of key findings)	Study Design 0-10	Duration of Study 0-4	Controls Used 0-6	Measure of Performance 0-2	Scope of the Study			Completeness of the Technical Report			Quality Rating 0-32	Impact Rating -2 to +2
								Scope of study 0-3	Grades levels covered 0-2	Subjects covered 0-1	Clear and complete methods section 0-2	Complete set of findings 0-1	Limitations of study included 0-1		
OH1	Carr, M., & Staley, S. (2005). <i>Using the Ohio proficiency test to analyze the academic achievement of charter school students: 2002-2004</i> . Columbus, OH: The Buckeye Institute.	Compares gains of percentage of students passing Ohio Proficiency Tests made by low-performing charter and district schools, controlling for family income, race, poverty <i>Limitations:</i> School level data, not generalizable to all community schools in Ohio, sample restricted to lowest performing districts, cross-sectional study	Strongly positive: In all cases and both analyses, charter schools performed as well as or better than traditional schools	1	3	3	1	3	2	1	1	0	0	15	1
OH2	Legislative Office of Education Oversight. (2003). <i>Community schools in Ohio: Final report on student performance, parent satisfaction, and accountability</i> . Columbus, OH: Author.	Compares scores on Ohio Proficiency Test and the percentage proficient through matching of schools based on grades served and demographics <i>Limitations:</i> School level data, cross-sectional, method for matching schools is incomplete	Slightly negative: District schools generally outperformed community schools, but small differences; when there were statistically significant differences, generally favored district schools	0	0	0	2	3	1	1	2	1	1	11	-1
OR	Bates, M., & Guile, D. (2006). <i>Oregon charter schools 2004-2005 evaluation report</i> . Salem: Oregon Department of Education.	Examines AYP general ratings for charter and traditional public schools at the elementary, middle school, and high school levels <i>Limitations:</i> General rating used as measurement, cross-sectional, no use of demographic controls, complete set of findings not presented	Mixed: Charter schools outperform at elementary benchmark levels; traditional public schools outperform charters at middle and high school benchmark levels	0	0	0	0	2	2	1	1	0	1	7	0
PA	Miron, G., Nelson, C., & Risley, J. (2002). <i>Strengthening Pennsylvania's charter school reform: Findings from the statewide evaluation and discussion of relevant policy issues</i> . Kalamazoo: The Evaluation Center, Western Michigan University	Compares charter school scores with similar district schools using regression analysis; determines how charter school scores change in conjunction with length of operation <i>Limitations:</i> School level data, cross-sectional study	Slightly positive: Pennsylvania charter schools appear to be attracting students with lower-than-average achievement levels and producing small relative gains (15 points per year, on average) in their achievement level	1	4	4	2	3	2	1	2	1	1	21	1
TX1	Maloney, C., Sheehan, D., Huntsberger, B., Caranikas-Walker, F., & Caldera, S. (2007). <i>Texas open-enrollment charter schools: 2005-06 evaluation</i> . Austin: Texas Center for Educational Research	Cross-sectional comparisons for each year, each grade, each subject; patterns for different ethnicities also determined <i>Limitations:</i> No controls used, cross-sectional study, no use of gains	Strongly negative: Accountability ratings are negative for charter schools at each year; TAKS scores: all subjects, all years, negative for charter schools; differences in magnitude of negative change by ethnicity, but Caucasian and African-American students both have lower scores in charter schools	0	0	0	2	3	2	1	2	1	0	11	-2
TX2	Gronberg, T., & Jansen, D.W. (2005). <i>Texas charter schools: An assessment in 2005</i> . Austin: Texas Public Policy Foundation.	Comparing gains in z scores for 2003 and 2004 for charter school students and predicted gain in z scores if those students had continued to attend TPS; matched student design employed <i>Limitations:</i> Concerns over attrition patterns, longitudinal but only 2 years of study	Slightly positive: Gains for students in lower grades who stay in charter schools are higher than matched students in district schools; at-risk charter school students do better than their matches at district schools; students in charter high school score lower than their matches	8	3	3	2	3	2	1	2	1	1	26	1
TX3	Hanushek, E.A., Kain, S.G., & Rivkin, S. (2002). <i>The impact of charter schools on academic achievement</i> . Unpublished manuscript.	Compares average test score gains of charter students with the same students' gains in district schools <i>Limitations:</i> Incomplete methods section (sample size not included); students who switch sectors may have different unobservable characteristics, controls employed may not be adequate	Slightly negative: Charter schools gains are initially lower, but no significant differences after 2 or 3 years of charter school	8	4	4	2		2	1	1	0	0	22	-1
TX4	Booker, K., Gilpatrick, S.M., Gronberg, T., & Jansen, D. (2004). <i>Charter school performance in Texas</i> . College Station: Texas A & M University.	Examines student gains for TAAS test in reading and math using student-level data and fixed effect method <i>Limitations:</i> Though overall sample is very large, paper does not indicate number of students in different categories of "movers," which is central to analysis; controls may not adequately account for unobserved differences in students	Strongly positive: After controlling for the mobility effect (the initial negative effect that transferring to a charter school causes), charter schools significantly improve the performance of students in both math and reading, with some evidence that school performance may improve as new charter schools progress beyond their first year in operation. African-American students in charter schools perform particularly well	8	4	1	2	3	1	1	2	1	1	24	0
UT	Was, C., & Kristjansson, S. (2006). <i>An analysis of charter vs. traditional public schools in Utah</i> . Salt Lake City: Utah State Charter School Board.	Cross-sectional, ANOVA used to compare standardized test scores in charter schools and traditional public schools, HLM used as well <i>Limitations:</i> Cross-sectional, school level data, no information on scope	Slightly positive: Charter schools outperform traditional public schools in lower grades, traditional public schools outperform high schools in grade 10	0	0	2	2		2	1	1	1	0	9	1
WI	Witte, J.F., Weimer, D.L., Schlomer, P.A., & Shober, A.F. (2004). <i>The performance of charter schools in Wisconsin</i> . Madison: Wisconsin Charter Schools Study.	Multichotomous logit group analysis, consecutive cohorts used to compare charter schools' and traditional schools' scores on Terra Nova test in grades 4 and 8 <i>Limitations:</i> School level data, does not examine charter high schools because 90% are aimed at high risk populations	Positive: For charters in elementary and middle grades across most comparison. High school results not shared due to concern that many of the charter schools at this level serve at-risk students.	1	3	4	0	2	1	1	2	1	1	16	1

Code Used in the Chart	Title of Study/Evaluation	Description of the Study (include details about the design, comparison groups, test and outcome measure used, and scope of study)	Key Findings (Include rating and then bulleted summary of key findings)	Study Design 0-10	Duration of Study 0-4	Controls Used 0-6	Measure of Performance 0-2	Scope of the Study			Completeness of the Technical Report			Quality Rating 0-32	Impact Rating -2 to +2
								Scope of study 0-3	Grades levels covered 0-2	Subjects covered 0-1	Clear and complete methods section 0-2	Complete set of findings 0-1	Limitations of study included 0-1		
US1	Finigan, K., et al. (2004). <i>Evaluation of the public charter schools program: Final report</i> . Prepared for U.S. Department of Education by SRI International, Washington, DC.	Logistical regression with background characteristics at school level controlled for <i>Limitations</i> : Cross-sectional, differences in standards and definitions of background characteristics from state to state	Strongly negative: Charter schools less likely to meet state standards than traditional public schools when background controls are taken into account	0	0	2	1	3	2	1	2	1	1	13	-2
US2	Hoxby, C.M. (2004). <i>Achievement in charter schools and regular public schools in the US: Understanding the differences</i> . Cambridge, MA: Harvard University and National Bureau of Economic Research.	Compares percentage proficient at charter school elementaries with those proficient at geographically closest elementary and with similar by race public school <i>Limitations</i> : Elementaries only, cross-sectional, various state standards used, single grade (4th) used	Strongly positive: Charter students are 5.2 percent more likely to be proficient in reading and 3.2 percent more likely to be proficient in math on their state's exams; stronger advantage for older charter schools, those with high minority populations, states with strong charter laws	0	0	1	1	3	0	1	2	1	1	10	2
US3	U.S. Department of Education, Institute for Education Sciences, National Center for Education Statistics. (2004). <i>The nation's report card: America's charter school report</i> , NCES 2005-456. Washington, DC: Author.	Compares NAEP national reading and math scores in charter schools and district schools <i>Limitations</i> : Cross-sectional, school level data	Slightly negative: Charter school students performed worse in math; free/reduced lunch students in charter schools performed worse; similar performance by ethnic groups	0	0	2	2	3	0	1	2	1	1	12	-1
US4	Loveless, T. (2003). <i>The 2003 Brown Center report on American education: Charter schools: Achievement, accountability, and the role of expertise</i> . Washington, DC: The Brookings Institution.	Compares changes in average charter school and district test scores in 10 states from 2000 to 2002. Brown Center researchers computed z-scores for charter schools, indexing charter schools' test scores relative to the mean and standard deviation of test scores within each state, and then examining z-scores nationally <i>Limitations</i> : School level data, tests vary from state to state, no controls used	Slightly positive: Charter schools have lower scores but larger gains	1	3	0	2	3	2	1	2	1	1	16	1
US5	Nelson, H.F., Rosenberg, B., & Van Meter, N. (2004). <i>Charter school achievement on the 2003 National Assessment of Educational Progress</i> . Washington, DC: American Federation of Teachers.	Comparison of NAEP scores for charter and traditional public schools <i>Limitations</i> : Cross-sectional, controls in separate analyses	Slightly negative: Charter school students worse in both fourth grade subjects, statistically significant	0	0	1	2	3	1	1	2	0	0	10	-1
US6	Greene, J.P., Forster, G., & Winters, M.A. (2003). <i>Apples to apples: An evaluation of charter schools serving general student populations</i> . (Education Working Paper No. 1). New York City: Center for Civic Innovation at the Manhattan Institute.	Regression analysis on two most recent years with year-to-year change reported <i>Limitations</i> : School level data, different tests used for different states, some states excluded from results	Strongly positive: Cross-sectional and longitudinal positive were overall positive for charter schools; TX and FL were most positive for charter schools	1	3	1	1	3	1	1	2	1	0	14	2
US7	Miron, G., Coryn, C., & Mackety, D. (2007). <i>Evaluating the impact of charter schools on student achievement: A longitudinal look at the Great Lakes states</i> . East Lansing, MI: Great Lakes Center for Education Research and Practice.	Linear regression models used to estimate student achievement patterns, producing three estimates: (1) actual scores, based on observed student achievement data provided by each school; (2) predicted scores, based on the performance of demographically similar public schools across the state; and (3) residual scores, based on the difference between predicted and actual charter school student achievement <i>Limitations</i> : School-level data, varied quality of achievement tests, missing or incomplete data for some schools	Slightly negative: Not currently outperforming demographically similar traditional public schools; scores lower than demographically similar traditional public schools with scores on achievement tests lower than TPS, especially for those with the newest charter school initiatives, IN & OH. IL has highest relative results, maybe because of effort to close low-performing charters? All states have some high performing charter schools	1	4	4	1	3	2	1	2	1	1	20	-1
US8	Braun, H., Jenkins, F., Grigg, W., & Tirre, W. (2006). <i>A closer look at charter schools using hierarchical linear modeling</i> . Washington: U.S. Department of Education	Phase 1: Charter schools are compared with all public noncharter schools, using a variety of models that incorporate different combinations of student and school characteristics (HLM); Phase 2: Charters classified into those who affiliated with public school districts and those not affiliated with public school districts; Phase 3: subset of public schools in urban areas with large minority populations are compared <i>Limitations</i> : Cross-sectional, self-selection bias may not be accounted for	Strongly negative: After adjusting for student characteristics, charter school mean scores in reading and mathematics were lower than public noncharters. Differences between public noncharter schools and charter schools affiliated with a public school district were not statistically significant, while charter schools not affiliated with a public school district scored significantly lower on average than public noncharter schools	4	0	3	2	3	1	1	2	1	1	18	-2