

## **Value-added in a Virtual Learning Environment: An Evaluation of a Virtual Charter School**

**MARTIN LUEKEN, PH.D.**

*Friedman Foundation for Educational Choice, United States*  
marty@edchoice.org

**GARY RITTER, PH.D.**

*University of Arkansas, United States*  
garyr@uark.edu

**DENNIS BECK, PH.D.**

*University of Arkansas, United States*  
debeck@uark.edu

This paper evaluates an online charter school that serves children in grades K-8 in a southern state in the United States. We compare growth in math and literacy learning on state standardized assessments between students enrolled in this school and “matched twin” students enrolled in traditional public school students statewide each year from 2010 to 2012. We also examine the impact of the school subgroups that include minorities, students with special needs, and students in the bottom and top quartiles within their schools. We then use propensity score matching to match students who enrolled in the online school at least three consecutive years with an individually matched comparison group of public school students and estimate the differential value-added in math and literacy between the two groups. We find evidence of negative transition effects in the first year for students who enroll in the online school. These effects, however, dissipate after the first year and, in some instances, turn positive.

## INTRODUCTION

Distance learning is increasingly becoming an integral component of K-12 education in the United States and has experienced robust growth at all levels of education. A national survey of degree-granting tertiary institutions found that a large majority offer some kind of distance learning course (Parsad & Lewis, 2008). There are many ways that distance education is provided, and online schools provide one kind.

Many authors have considered the specific factors that characterize an online school. For example, geographic distance between teacher and student is insisted upon by some, while others suggested that this is not a mandatory component (Moore, 1991; Moore & Kearsley, 2011). Others chose to categorize online schooling based on whether the instruction is received inside or outside of a brick and mortar school building (Levenberg & Caspi, 2010) or whether it is delivered in real time or separated by time (Moore & Kearsley, 2011).

In the United States, online schools provide Internet-based courses and are state sanctioned (Barbour & Reeves, 2009). Online schools provide part of or an entire curriculum online for students and do not conduct any operations in a traditional brick-and-mortar environment. In this article, we will delineate the differences between two types of online schools at the K-12 level - a cyber school and a virtual school. We make this distinction in terms of whether their curriculum is supplemental to face-to-face instruction, or in a completely stand-alone online environment. A virtual school typically provides curriculum online for students in a supplemental manner. On the other hand, a cyber school offers curriculum completely online with students taking all courses online, and it does not conduct any operations in a traditional brick-and-mortar environment. The Florida Virtual School (FLVS), opened in 1997, is the first statewide online public school in the United States. This school provides its educational program for K-12 students online, and all of its courses are aligned with Florida's content standards. Today, FLVS Part Time offers about 120 courses, and FLVS Full Time grants diplomas for high school graduation (<https://www.flvs.net/>).

Since the opening of FLVS, distance education in the United States has experienced rapid growth at the K-12 level, although most of that growth is in supplemental online course delivery. Even so, thirty states have fully online schools that enroll about 275,000 students (Gemin, Pape, Vashaw, & Watson, 2015), and 55 percent of public school districts nationwide reported students enrolled in distance education courses in 2009-10, most at the high school level (Queen & Lewis, 2011). Proponents often advocate for online learning on grounds of efficiency (Hill & Roza, 2010) while opponents argue the need for increased regulation (Glass & Welner, 2011).

Moreover, a survey of high school and community college students indicate that students choosing coursework in distance learning environments value control over scheduling and pace of learning, whereas students choosing face-to-face courses value interaction with the instructor and classmates (Roblyer, 1999), therefore, suggesting a tradeoff between decisions about learning environments.

Determining the impact of online learning on student outcomes is important for policy. In the case of online charter schools, the need for answers is magnified due to the vibrant policy debate over the usefulness and effectiveness of public charter schools. In this study, we examine a fully online cyber charter school that serves K-8 students statewide in the Southern United States (we refer to it throughout this paper as “SVA”). Students enroll in the school and take its curriculum, which is aligned with state standards. SVA is a standalone school where students enroll and take an entire curriculum online in their private homes. This study attempts to ascertain the impact of SVA on its students’ learning relative to other traditional public brick-and-mortar K-8 schools across the state.

The remainder of the paper includes a review of the evidentiary record on virtual schools delivering online courses, a discussion of the challenging transition that occurs when students transition from brick-and-mortar schools to cyber school settings, and a concluding discussion of the research of the effectiveness of cyber schools. In the next section, we provide a comprehensive description of the cyber school of interest in this study, SVA, focusing on where its students come from and the background characteristics of its students.

## LITERATURE REVIEW

A substantial portion of the literature on K-12 distance learning in general is qualitative in nature and focuses on participants’ expectations or perceptions (Matthew & Varagoor, 2001; Shuldman, 2004; Oliver, Osborne, & Brady, 2009; Karp & Woods, 2003); the needs of schools considering developing online courses (Oliver et al., 2010); teacher development and integration of technology in urban classrooms (Hughes & Ooms, 2004; Staples, Pugach, & Himes, 2005; Mouza, 2011); and delivery and implementation of online courses (Cavanaugh et al., 2008). Cavanaugh (2009) noted that most of the literature on cyber charter schools, in particular, focuses on policy and management rather than on students’ academic outcomes. Although these areas are important for policy effectiveness, impact evaluations provide an important complement by assessing the extent to which these policies produce desirable outcomes related to student learning.

Despite a large body of research on supplemental online learning, a limited portion quantitatively analyzes differential impacts on student outcomes in full-time cyber schooling environments. For instance, Molnar et al. (2015) state in their annual report on Virtual Schooling in the U. S., “While there has been some improvement in what is known about supplemental K-12 online learning, there continues to be a lack of reliable and valid evidence to guide full-time online practice and policy” (pg. 31). The authors detail studies in Arizona, Colorado, Minnesota, Ohio, Pennsylvania, and Wisconsin that provide mixed evidence on whether full-time online students perform as well as brick and mortar students (Colorado Department of Education, 2006; Zimmer et al., 2009; Joint Legislative Audit Committee, 2010; Hubbard & Mitchell, 2011; Office of the Legislative Auditor, 2011; Ryman & Kossan, 2011; Ohio Alliance for Public Charter Schools, 2009; Innovation Ohio, 2011; Center for Research on Education Outcomes, 2011). Even in studies on student performance outcomes in online learning, there is a perceived shortage of rigorous evidence. For example, the U.S. Department of Education’s Office of Planning, Evaluation and Policy Development conducted a meta-analysis to compare online learning environments with face-to-face instruction (2010) and noted a dearth of research exists on K-12 cyber learning. It found only a small number of studies that employed sufficiently rigorous research methods to draw meaningful conclusions about the effectiveness of online learning compared to that of face-to-face instruction at the K-12 level. None of the studies in the meta-analysis employed experimental or rigorous quasi-experimental evaluation designs, and most of the studies on K-12 online learning focused on blended (virtual), not fully online, learning.

At minimum, impact studies should control for baseline outcome measures to account for differential ability that students bring into the classroom. Because online education is still a fairly novel method for many students, transition effects are another important consideration for evaluating the impact of these methods students using them for the first time. Much has been written on the negative transition effects of students transferring from one school to another. The literature appears to be focused particularly on the elementary to middle/junior high transition (Eccles & Midgley, 1989), the areas in which transitions occur (e.g., learning goals, peers, learning environment, roles, assessments), and the factors that influence how smoothly the transition occurs such as self-efficacy (Friedel, Cortina, Turner, & Midgley, 2010), peer acceptance (Kingery, Erdley & Marshall, 2011), socioemotional adjustment (Martínez, Aricak, Graves, Peters-Myszak & Nellis, 2011), and teacher perceptions (Ross & Brown, 2013).

A primary concern of most of these researchers is tendency of these transitions to have negative effects on student achievement. Most studies suggest that students transitioning to charter schools, innovative or not, often

struggle in terms of student achievement in the first year or two after the transition (Sass, 2006). Consequently, during the initial years of operation, charter schools generally underperform traditional public schools (Hanush-ek, Kain, Rivkin, & Branch, 2007). The question is, will these negative transition effects be even more problematic when the students move to an entirely different style of schooling, such as from a brick-and-mortar school to a cyber school? Our study of SVA will provide some insight into this question.

### **Studies on Online Courses**

In this section, we review studies that examine the effect of online courses on student outcomes (sometimes referred to as “virtual schools”). Cavanaugh, Gillan, Kromrey, Hess and Blomeyer (2004) reviewed 14 studies published between 1999 and 2004 that estimate effects of virtual schooling on student outcomes in K-12 education. Overall, the authors found no significant differences in outcomes between classroom-based learning and web-based distance learning.

O’Dwyer, Carey, and Kleiman (2007) found that high school students in virtual classes during the 2004-05 school year performed as well as their peers enrolled in traditional algebra classes on a post-test aligned with Louisiana’s Grade Level Expectations for Algebra I. Karp and Woods (2003) evaluated the effectiveness of a virtual wellness course on student learning and student perceptions by comparing students in virtual and face-to-face classes taught by the same instructor. Both groups experienced significant gains during the classes, though there were no significant differences between the two groups. Because virtual schools may well work hand-in-hand with traditional brick-and-mortar schools, it’s also important to examine the effectiveness of cyber schools, which are distinct in that they offer a fully online experience.

### **Studies on Cyber Schools**

Zimmer et al. (2009) used non-experimental methods to analyze charter schools, of all types, in eight states. They estimated models with fixed effects to compare students in Ohio enrolled in 40 cyber charter schools with students enrolled in classroom-based charters. They found that students in cyber schools performed significantly lower in both math and reading relative to when they were enrolled in classroom-based charter schools (by 0.44 and 0.25 standard deviations, respectively). Though these effects are large, the authors exercise caution in interpreting their results because of external validity concerns due to a significant portion of the cyber school population that was left out of the analysis.

In an earlier study, Zimmer et al. (2003) studied the effectiveness of charter schools in California. The authors used student-level data to estimate academic outcomes from 1997-98 through 2001-02. One important limitation of their analysis was an inability to control for baseline scores, though they controlled for other important differences. They found that students in startup and conversion non-classroom-based charter schools score, on average, significantly lower than students in comparable classroom-based charters (these differences range between 5 and 9 percentile points on measures of the Stanford 9 math and reading exams).

Regarding relatively objective measures of academic gains, different studies of cyber schools have found markedly different outcomes. For example, an independent evaluation of California's Rocketship Education showed sizable math gains among participating kindergarten and first grade students (Wang & Woodworth, 2011). In sharp contrast, full-time cyber students in Minnesota were more likely to drop out than their peers, and those in grades 4-8 made half the progress on state math tests as their traditional counterparts (Minnesota Office of the Legislative Evaluator, 2011). Cyber schools in Colorado produced three times more dropouts than graduates over a four-year period, with students' scores averaging well below the state average in reading, writing, and math (Hubbard and Mitchell, 2011). It is important to note, however, that these last two studies did not include any controls in their analysis. In Ohio, only 3 of 27 cyber schools were rated "effective" or "excellent" on the state's accountability scale in 2010, and on-time graduation rates were well under 50 percent (Tucker, Dillon & Jambulapati, 2011). Finally, the CREDO 2011 study of Pennsylvania charter schools found that all eight cyber charter schools in the study performed significantly worse in reading and math than their traditional school counterparts in terms of student gains, with none of the cyber charters making AYP and only one exceeding the state's average graduation rate (Woodworth & Raymond, 2013).

Results from the Center for Research on Educational Outcomes (CREDO) 2015 online charter school study, among others, appear to indicate that students in brick and mortar schools overall perform better than those in cyber charter schools (Woodworth et al, 2015; Woodworth & Raymond, 2013). Nationally, K12 Inc. is the biggest player in the K-12 cyber schooling market. Only 27.4 percent of cyber schools run by K12 Inc. met AYP in 2010-2011 as compared to 51.8 percent of brick and mortar charter schools. Also, in math, K12 Inc. students scored far lower than students in their host states, and the on-time graduation rate for K12 Inc. schools is 49.1% compared with 79.4% in the states in which K12 Inc. operates (Miron & Urshel, 2012). Because they did not include controls, however, their analysis may not pick up differences in students' difficulties and surely did not account for differences in student abilities or backgrounds.

Taken as a whole, the research on student achievement in cyber schools shows mixed to negative results (Molnar et al., 2013, 2014, 2015), though there is some evidence that cyber charter schools serve disproportionate numbers of children who had serious academic or emotional problems in traditional public schools, and thus may have academic disadvantages that are not easily captured by statistical controls (Beck, Maranto, & Lo, 2013). Carnahan and Fulton (2013) examined enrollment of special education students in Pennsylvania public cyber schools, defined as schools that deliver instruction without any face-to-face interaction, and showed that the number of special needs students enrolled in cyber schools increased at a greater rate than in public schools statewide during 2006-2009.

A more recent report by Molnar et al. (2014) acknowledged that “there continues to be a deficit of empirical, longitudinal research to guide the practice of K-12 online learning, particularly full-time learning” (p. 34). The authors compared cyber schools’ performance data and enrollment by characteristics with all public schools in the United States and found that cyber schools, on average, serve lower proportions of disadvantaged student populations and score lower on performance indicators like Adequate Yearly Progress (AYP) and on-time graduation. This conclusion is different than that drawn by Carnahan and Fulton (2013) in Pennsylvania.

In sum, the studies that evaluate virtual and cyber charter schools find mixed effects on student academic outcomes. It is unclear, however, what portion of these effects are due to students’ first years in these new learning environments and whether students who remain recover from such possible transition effects. Our study addresses this issue by evaluating students in SVA by years enrolled and comparing those who remain for at least three years with a group of students with similar background characteristics. Thus, we are able to identify the existence of negative transition effects and gauge whether students recover from them.

## QUESTIONS, DATA, AND METHODS

In this section, we provide an introduction to SVA by presenting details about the school, its operations, student demographics, and from where its students enroll. We then lay out our research questions and describe the data and research methods.

### About SVA

SVA is a full-time virtual charter school that enrolls K-8 students. Classes are taken online, and class materials such as books and CDs are delivered to the students’ homes. Core courses are English and Language Arts, math, science, history, and art. Resource-constrained families who qualify can

receive a computer and printer on loan plus a stipend for Internet service. As a public charter school, SVA is held to the same accountability standards and state laws as are other public schools.

During the period studied, state law limited SVA's enrollment to 500 students at any given time. As a public charter school, SVA does not charge tuition to students, and revenue that it receives from the state is based on the state's funding formula. Because of its enrollment cap, the number of pupils used to calculate its total state revenue does not exceed 500 students. While SVA's student body does not resemble the overall state's composition, it cannot selectively admit its students. Instead, when there are more applicants than there are available seats, students are admitted through a randomization process, and students who do not win a spot are placed on a waiting list. Thus, the makeup of the school's student body is a function of who chooses to enroll, which may partly depend on SVA's recruiting and counseling efforts as well. In every year of its existence, SVA has received more applicants than spots available and, therefore, has conducted a lottery.

Students enrolled in SVA take the same exams as other public school students statewide. SVA students in grades 3-8 take the state benchmark exams while all students in grades K-8 take the norm-referenced test each year. Because students conduct most of their coursework at home, a lack of exposure to peers in a classroom setting is one perceived disadvantage of cyber schooling. To address this issue, SVA provides opportunities for students to interact via social outings and activities that include trips to museums, skate parks, and zoos.

### **Study Sample and Research Questions**

Of course, the challenge inherent in doing any school-based program evaluation is answering the question "*compared to what?*" Given our setting, the ideal research method is random assignment. But because we are unable to do this, we rely on two strategies, both of which are imperfect but nonetheless provide very reasonable approaches. First, we compare SVA students with other students statewide by estimating academic growth models. Second, we use propensity score matching techniques to match SVA students with their "twins." We then estimate value-added models to compare student growth for SVA students with the student growth for their matched peers. Using these strategies, we ask the following two broad research questions:

1. How does SVA's academic growth for various groups of students compare to the academic growth for other similar public school students across the state?
  - a. For all students.

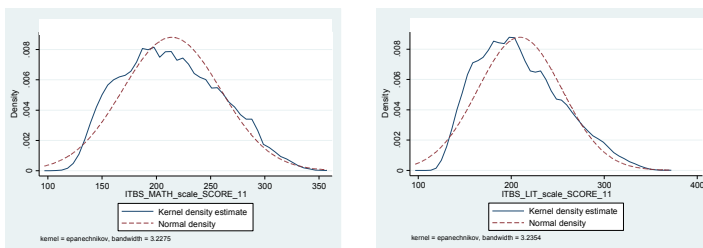


- b. For students grouped by number of years enrolled in the school (students enrolled one year, two years, and at least three years).
  - c. For specific subgroups of students, such as special needs, minority, top-quartile, and bottom-quartile students.
2. For students from the 2008 cohort who were enrolled in SVA at least three consecutive years,
    - a. How does their value-added compare to other students with similar background characteristics?
    - b. To what extent do these students recover from low outcomes experienced in their first couple years?

## Data

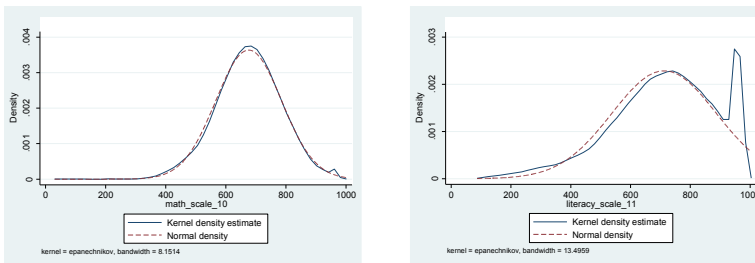
Our analysis is based on two longitudinal student-level data sources: criterion-referenced test (CRT) and norm-referenced test (NRT) data over 2007-2012. Each of these exams is administered as part of the state benchmark system, which mandates that statewide accountability exams are given every year to students in grades 3-8. The state-developed criterion-referenced exam assesses the extent to which students master the skills and content that is outlined in the state's curricular frameworks. The CRT exam is most commonly used as a source for reporting on student achievement because it is aligned with the state standards and because it is a high stakes exam used to satisfy NCLB requirements. We normalized CRT test scores so that the mean and standard deviation for each grade in each year is zero and one, respectively. Thus, the unit of analysis for CRT scores is given by standard deviations.

The NRT focuses on more general skills and knowledge, can be compared to a national norming sample of students, and has results that are reported as national percentile units or normal curve equivalents. NRT scores are reported as normal curve equivalents (NCE's) and do not require us to standardize the values.



**Figure 1.** Kernel density graphs for 2011 scale scores from norm-referenced exams with normal density curve (dashed line), all state public school students.

It has been documented that data based on CRT exams are susceptible to so-called ceiling effects and may produce biased value-added estimates (Koedel & Betts, 2010). This should be evident in Figure 1 and Figure 2, which display kernel density plots of 2011 CRT and NRT outcome measures for all public school students in the state. A normal density curve (dashed line) is superimposed on each graph. Observe in panel B that state-wide scores in literacy on the CRT “bunch up” in the upper tails. Students scoring at the top do not have much room to improve, and students with perfect scores have reached the “ceiling.” This bunching up may lead to imprecise literacy value-added estimates, particularly for students scoring at the upper end of the distribution. We do not observe bunching for CRT math and any NRT exams, where the tails are relatively thin.



**Figure 2.** Kernel density graphs for 2011 scale scores from criterion-referenced exams with normal density curve (dashed line), all state public school students.

Given potential sensitivity issues with our academic performance estimates from CRT data, we estimate the same models using the NRT data to check for consistency in the results. The advantage of using NRT data is that such tests are less sensitive to ceiling effects, thus providing more stable estimates of the impact of SVA on student academic growth. Furthermore, these tests are low-stakes and subsequently provide a broader sense of student learning because school leaders and teachers may feel less compelled to engage in “teaching to the test” than with high-stakes exams. In addition, NRT data provides test score information for grades 1-8 while CRT data gives outcome information only for grades 3-8. NRT exams, however, may not be aligned as well with state standards as the CRT exams. Overall, it is fair to see that there are tradeoffs associated with each exam; thus, we investigate our research questions using both tools.

### Analytic Strategy

In the first part of our analysis, we estimate academic growth models to compare SVA students with other students statewide. We also make comparisons within certain subgroups of students. In the second part of our analysis, we use a propensity matching technique to build a more targeted sample. Whereas in our first strategy we examined all students in the state, in our second strategy we create a more targeted (and more comparable) sample of students. To do so, we pair each SVA student who has been enrolled there at least three years with two similar students based on a vector of observable characteristics. Then we estimate value-added models in which we predict year-end scores as a function of student ability, student characteristics, and, most importantly, whether or not the student was in SVA. We then compare SVA's value added in math and literacy with those same scores for the comparison groups of students.

**Statewide academic growth comparisons.** The first part of our analysis relies on the following academic growth model, using a statewide sample of students:

$$(1) \quad Y_{it} = \alpha + \beta_1 Y_{i(t-1)} + \beta_2 Y_{i(t-2)} + \delta_1 Y'_{i(t-1)} + \delta_2 Y'_{i(t-2)} + \psi SVA_{it} + \varepsilon_{it}$$

where  $Y_{it}$  is student  $i$ 's score in math or literacy in year  $t$  ( $t=2010, 2011, \text{ or } 2012$ );  $Y'$  indicates the non-outcome subject;  $SVA_{it}$  is a binary indicator that takes on 1 if student  $i$  was enrolled in SVA in year  $t$  and 0 otherwise; and  $\varepsilon_{it}$  represents a normally distributed stochastic error term. The coefficient of interest is  $\psi$  and gives the average difference in measured growth between students enrolled in SVA and non-SVA students. In these models, the comparison group includes all non-SVA students across the state.

**Value-added analysis with matched peer comparison group.** In the second part of our analysis, we seek to answer the straightforward question: relative to their matched comparison peers, how well do SVA students who have enrolled in SVA for at least three consecutive years perform in math and literacy when they are in their first, second, third, fourth, and fifth year, and do they recover from any initial "dip" that might occur? We require at least three years of test scores in order to control for two years of prior test scores. We compute residual value-added estimates by subject for each student in each year and use these estimates as the outcome measures in our comparisons between SVA students and their matched peers. We examine the 2008 cohort of students in SVA and focus on those who remained for at least three years over the period 2008-2012. Because we control for two years of prior test scores, our outcomes are for years 2010-2012. To get a sharper sense of whether SVA students recover from any transitional dip and how fast, we also estimate matching models with value-added outcomes that control for baseline, pre-SVA 2007 test scores rather than 2-years lagged scores.

Because it is important to control for baseline test scores to improve the accuracy of estimates on the determinants of student achievement (Hedges, Laine, & Greenwald, 1994; Boardman & Murnane, 1979; Hanushek, 1979), ideally our analysis would focus on students whose data on pre-SVA test scores are available. Unfortunately, a substantial proportion of SVA students come from non-public school environments, and this approach would force us to omit a significant number of students and potentially important group. Therefore, we match on first-year test scores in order to increase power. The drawback is that we do not have data on SVA students' geographic origin. Matching on these dimensions would substantially reduce the size of our sample. To reduce bias, we match on same-subject test scores observed in a student's first year in SVA and their grade. We then obtain a propensity score for each student by estimating a probit model:

$$(2) \quad \Pr(D=1|\mathbf{X}) = \Phi(\mathbf{X}'\boldsymbol{\beta})$$

where the left-hand side term is the conditional probability of enrolling in SVA,  $\Phi$  is the cumulative distribution function of the standard normal distribution, and  $\mathbf{X}$  is a vector of observable characteristics that include the student's lagged or baseline test scores in math and literacy, race, IEP status, and gender. After matching grade and same-subject scores in the first year, we match each SVA student to two non-SVA students based on their propensity score. This matching analysis is conducted separately for math and literacy outcomes.

Nearest neighbor (NN) matching pairs each treatment observation to a comparison based on the propensity score. One-to-one NN matching will match each treatment with one comparison while one-to- $k$  NN matches  $k$  comparisons to each treated observation. Given the large pool of potential matches with both data sets, we match without common support and without replacement. We use 1-to-2 nearest neighbor matching in our analysis. Thus, the comparison group is roughly twice as large as the treatment.

Once we create the full sample of the SVA students and the matched comparison pool, we calculate the average difference in outcome measure between the two groups. First, the outcome measure is a variation on value-added (it is a residual value-added indicator). We obtain value-added estimates for all students for each outcome year by estimating a similar equation to model (1) above where we omit the SVA indicator and save the residuals. After we obtain these estimates for every student, we use them to estimate the difference in value-added between students enrolled in SVA for at least three years and a matched comparison group assembled via propensity score matching methods. After we compute VA scores for all students in the state and in our restricted sample (treatment and comparison), we then estimate the following regression model on our restricted sample:

$$(3) \quad V_{it} = \alpha + \gamma X_{it} + \lambda Treat_{it} + \varepsilon_{it}$$

where  $V_{it}$  is the residual value-added estimate for student  $i$  in year  $t$ ,  $X_{it}$  is the vector of matching covariates,  $Treat_{it}$  is an indicator for a student who entered SVA in 2008 and enrolled for at least three years, and  $\lambda$  represents the adjusted estimate. Because the matching dimensionality is reduced to one, it is possible that some matches are not made exactly on all matching covariates. This procedure allows us to control for any remaining differences after matching and reduces potential bias. Finally, we obtain standard errors with a bootstrap method.

### Student Characteristics

Although students are drawn from all parts of the state, SVA students are not a representative sample of the statewide public school student population. Table 1 summarizes the distribution of background characteristics for SVA and non-SVA students across the state. SVA enrolls a larger percentage of white students and fewer minority students than the state average. The proportion of minority students is less than half that for the entire state. The school does not enroll any English language learners, but it enrolls a significantly larger percentage of special needs students. About 14 percent of students enrolled in SVA have special needs, compared with 10 percent statewide.

**Table 1**  
Distribution of background characteristics for SVA students and non-SVA public school students

	SVA	State
Free and reduced price lunch <sup>†</sup>	22.2%	60.5%
Special needs	14.4%	9.4%
English language learners	0.0%	6.9%
Racial/ethnic background		
African American	4.5%	20.2%
Asian	1.2%	1.3%
Hispanic	2.2%	9.4%
Native American	0.7%	0.6%
White	86.6%	62.1%
Two or more	0.2%	1.8%
Female	45.7%	46.8%
Male	54.3%	53.2%
Number	418	285,607

<sup>†</sup> The statistic for SVA indicates students in SVA who subscribed to a FRL program while at a different public school over the period 2008-2011. Because the state databases do not report FRL enrollment for SVA students, this statistic is likely underestimated.

Because SVA's delivery of its curriculum differs from traditional methods of instruction and is relatively new, it may not be a good fit for some while effective for others. After all, evaluating goodness of fit between a novel learning environment and students is a process rather than an event. Thus, we might have good reason to expect high student turnover as this process proceeds. This is a limitation of a study of a school like ours that allows for students to transfer in and out of the school.

## RESULTS

This section reports the results for our comparisons of SVA to all students statewide, followed by results of our analysis comparing SVA students with their matched peers.

### Statewide Analysis

Table 2 displays the academic performance estimates for math and literacy over 2010-2012. Row (1) reports growth by year by subject for all SVA students enrolled in those years with at least two years of prior test scores. We first look at CRT results. In 2010, students in SVA experienced, on average, 0.15 and 0.14 standard deviations lower growth than non-SVA students in math and literacy, respectively. Both results are statistically significant ( $p < 0.01$ ). These gaps shrunk and became statistically indistinguishable for math in 2011 and 2012 and literacy in 2011. In 2012, the coefficient on literacy became positive and statistically significant ( $p < 0.01$ ). All SVA students in 2012 experienced, on average, 0.13 standard deviations higher growth on the CRT literacy exam than the comparison group ( $p < 0.01$ ). Estimates from NRT data indicate similar patterns. In 2010, all SVA students experienced 1.69 NCE points lower in math growth ( $p < 0.05$ ). Estimates for math in 2011 and 2012 became statistically insignificant. The estimates for literacy are positive but insignificant for each year.

Row (2) provides the growth estimates for first-year SVA students. Students who transferred to SVA experienced significantly lower growth in both math and reading in their first year. This effect is stronger in math than literacy. Again, we examine CRT results first. In 2010, first-year SVA students experienced, on average, 0.44 standard deviations lower math growth than the statewide average. This effect is also evident for first-year students during 2011 and 2012, although the magnitude of the gap diminished slightly for first-year students during these years. First-year SVA students during 2010 and 2011 experienced 0.36 and 0.29 standard deviations lower literacy growth than non-SVA students, respectively. The estimate for first-years in 2012 declined to -0.15 standard deviations and became statistically insignificant. The average gain in math and literacy on the NRT by first year

**Table 2**  
**SVA math and literacy growth by years enrolled in SVA, 2010-2012**

	Math			Literacy			
	CRT	NRT		CRT	NRT+		
<b>(1) All SVA students</b>	2010	2011	2012	2010	2011	2012	
	-0.15***	0.00	-0.06	-1.69**	0.44	0.70	
	(0.05)	(0.04)	(0.05)	(0.86)	(0.91)	(1.00)	
	number in treatment group	137	133	129	209	169	131
total sample size	125,519	125,582	124,316	182,655	154,441	127,226	
<b>(2) 1st year students</b>	-0.44***	-0.34**	-0.37***	-7.63***	-4.07	-3.70	
	(0.13)	(0.14)	(0.12)	(2.19)	(2.65)	(2.76)	
	number in treatment group	18	13	17	32	20	17
	total sample size	125,400	125,482	124,204	182,478	154,292	127,003
<b>(3) 2nd year students</b>	-0.29**	-0.07	-0.22	-4.78**	3.71	2.52	
	(0.12)	(0.14)	(0.17)	(2.38)	(2.59)	(3.60)	
	number in treatment group	20	14	9	27	21	10
	total sample size	125,402	125,463	124,196	182,473	154,293	126,996
<b>(4) 3rd, 4th &amp; 5th year students</b>	-0.07	0.06	0.02	0.21	0.60	1.26	
	(0.05)	(0.05)	(0.05)	(1.01)	(1.05)	(1.12)	
	number in treatment group	99	106	102	149	128	103
	total sample size	125,481	125,555	124,289	182,595	154,400	127,089
	2010	2011	2012	2010	2011	2012	
	-0.14***	-0.07	0.13***	0.58	0.03	1.27	
	(0.04)	(0.04)	(0.05)	(0.87)	(0.86)	(0.86)	
	number in treatment group	137	133	130	210	169	131
total sample size	125,519	125,448	125,171	184,508	154,635	127,371	
<b>(2) 1st year students</b>	-0.36***	-0.29**	-0.15	-8.66***	-3.51	-0.05	
	(0.12)	(0.14)	(0.13)	(2.24)	(2.51)	(2.39)	
	number in treatment group	18	13	17	32	20	17
	total sample size	125,400	125,328	125,058	184,330	154,486	127,148
<b>(3) 2nd year students</b>	-0.21*	-0.05	0.03	-0.06	1.94	0.34	
	(0.11)	(0.14)	(0.17)	(2.44)	(2.45)	(3.11)	
	number in treatment group	20	14	9	27	21	10
	total sample size	125,402	125,329	125,050	184,325	154,487	127,141
<b>(4) 3rd, 4th &amp; 5th year students</b>	-0.09*	-0.04	0.19***	2.79***	0.27	1.51	
	(0.05)	(0.05)	(0.05)	(1.03)	(0.99)	(0.97)	
	number in treatment group	99	106	103	150	128	103
	total sample size	125,481	125,421	125,144	184,448	154,594	127,294

Standard errors in parentheses (\*\* p<0.01, \* p<0.05, \* p<0.1). Academic growth models control for twice-lagged test scores in both math and literacy-related tests. Test scores for the CRT are normalized year-by-grade such that the mean is zero and standard deviation is one and are from the state's benchmark (criterion-referenced) exams. Scores for NRT are reported as normal curve equivalents (NCE's).

SVA students in 2010 was 7.63 and 8.66 NCE's less than non-SVA students, both statistically significant estimates. Unlike the CRT estimates, however, these deficits decreased for other first-year SVA students in 2011 and 2012 by about half for math and became statistically insignificant. The 2012 NRT literacy coefficient became practically zero and insignificant. These results are based on small samples and should be interpreted with caution.

Row (3) suggests that second-year students fared somewhat better than first-year students. While SVA students who were first-years in 2010 experienced large negative effects, the size of these coefficients significantly decreased in their second year in 2011 (seen in the down-right-diagonal cells from the 2010 first-year estimates) and turned positive for NRT estimates, becoming statistically insignificant for all estimates. We observe a similar pattern for the group of SVA students who were in their first year in 2011. While math and literacy estimates for 2011 first-years were -0.34 and -0.29 standard deviations, respectively, estimates for their second year in 2012 became -0.22 and 0.03 standard deviations. Academic growth estimates in both subjects for second-year SVA students in 2012 were statistically indistinguishable from their non-SVA counterparts. Based on these results, students who remained in SVA for a second year performed at least as well as other students statewide. As with first year results, these results are based on a small sample of SVA students and should be interpreted cautiously.

Row (4) displays growth estimates for students enrolled in SVA for at least three consecutive years. This group's estimate for math was similar to the statewide average and statistically indistinguishable. Its 2010 and 2011 literacy growth estimates were -0.09 standard deviations ( $p < 0.1$ ) and -0.04 standard deviations 2011 ( $p > 0.1$ ). In 2012, students enrolled in SVA for at least three consecutive years gained, on average, 0.19 standard deviations above the state average in literacy ( $p < 0.01$ ). Almost all NRT estimates for math and literacy were positive, and the coefficient for 2010 literacy was positive and statistically significant ( $p < 0.01$ ).

Our results provide evidence that SVA students, on average, experienced a negative transition effect in their first year. Examination of the results in Table 2 along the diagonals suggests that students who remained in SVA improved relative to non-SVA public school students statewide. Estimates across columns suggest that average growth experienced by SVA students in years after 2010 improved relative to the average growth experienced by students in other public schools statewide.

The analysis discussed above compares all students in SVA relative to other public school students statewide. Certain groups of students, however, may benefit in particular from non-traditional learning environments, especially those that allow flexibility to learn at their own pace. Conversely, these students may struggle with learning in a new environment and experience problems adapting to a new format or learning new technologies. Thus, we next examined subgroups of students to shed light on these questions.



### Subgroup Analysis

We conducted the same analysis described above on the following subgroups: students with minority ethnic or racial backgrounds, students with special needs, and students in the bottom and top quartiles of test performance. We compared SVA students for each group with non-SVA students statewide having the same characteristics. Quartile groups were formed on same-subject test scores during the pre-SVA year. Table 3 summarizes the results.

**Minority students.** In 2010, the average CRT math growth for minority students enrolled in SVA was 0.40 standard deviations below non-SVA minority students statewide in 2010 ( $p < 0.01$ ) and statistically insignificant in 2011 and 2012. Average growth in CRT literacy was also negative (-0.27) but only marginally significant ( $p < 0.10$ ). The coefficient decreased by half and turned insignificant in 2011. In 2012, minority students enrolled in SVA gained 0.37 standard deviations more on CRT literacy than minority students statewide ( $p < 0.01$ ). The only statistically significant results from NRT data were 2010 math and 2011 literacy – the SVA group gained 11.6 and 6.3 NCEs less than the comparison group, respectively, though the literacy estimate was marginally significant. Overall, minority students experienced significantly lower growth in 2010, and the gap diminished over time and turned significantly positive in the case of literacy. These results are based on a very small sample, however, and should be interpreted cautiously.

**Students with special needs.** We found some evidence that students with special needs in SVA outgained their counterparts statewide. While estimates for 2010 and 2011 were statistically insignificant, all 2012 estimates for math and literacy for CRT and NRT exams were positive, large, and statistically significant. Math and literacy growth estimates were 0.20 and 0.41 standard deviations on the CRT and 5.64 and 3.92 NCE's on the NRT, respectively. The CRT math coefficient was marginally significant. Thus, students with special needs enrolled in SVA over time appeared to experience significantly higher growth than their counterparts statewide, particularly in literacy.

**Ability quartiles.** SVA students in the bottom quartile outgained other bottom-quartile students statewide in CRT literacy by 0.36 standard deviations ( $p < 0.01$ ). This was a significant improvement from 2011, when the SVA group gained on average 0.18 standard deviations below their comparison group's average ( $p < 0.1$ ). The coefficient on NRT 2010 literacy was -4.54 NCE's and statistically significant. SVA students in the top quartile were outgained by other top-quartile students statewide, on average, in math and literacy in 2010. Only the 2010 CRT math coefficient was statistically significant at the 95 percent level. These estimates became statistically insignificant in 2011 and 2012. Bottom-quartile students in SVA experienced significantly higher growth in literacy over time than their counterparts statewide. Growth estimates for top-quartile students in SVA became statistically indistinguishable from their counterparts statewide.

**Table 3**

Differences in academic growth between SVA and non-SVA schools by subgroup, 2010-2012, using CRT and NRT data

	Math						Literacy					
	CRT			NRT			CRT			NRT		
	2010	2011	2012	2010	2011	2012	2010	2011	2012	2010	2011	2012
<b>(1) Minority</b>	-0.40**	0.16	-0.08	-11.56***	-5.01	-0.78	-0.27*	-0.13	0.37***	0.18	-6.31*	2.29
standard error	(0.16)	(0.18)	(0.13)	(2.98)	(3.86)	(2.86)	(0.15)	(0.18)	(0.14)	(2.97)	(3.62)	(2.47)
number in treatment group	11	8	14	17	9	15	11	8	14	17	9	15
total sample size	39,457	40,057	39,699	55,747	48,202	40,610	39,457	39,990	40,219	56,394	48,285	40,664
<b>(2) Special needs</b>	-0.13	-0.04	0.20*	2.09	1.19	5.64**	-0.14	-0.06	0.41***	-0.18	-0.70	3.92**
standard error	(0.13)	(0.11)	(0.11)	(2.79)	(2.52)	(2.32)	(0.13)	(0.12)	(0.12)	(2.62)	(2.29)	(1.98)
number in treatment group	18	23	24	23	28	29	18	23	25	23	28	29
total sample size	15,567	16,090	15,343	19,945	18,918	17,011	15,567	15,987	15,939	20,459	18,942	17,039
<b>(3) Bottom quartile</b>	-0.06	0.02	0.03	1.98	2.79	1.67	-0.14	-0.18*	0.36***	-4.54**	-8.12	-0.03
standard error	(0.10)	(0.09)	(0.08)	(2.07)	(2.04)	(2.30)	(0.12)	(0.10)	(0.09)	(1.96)	-1.84	(1.91)
number in treatment group	28	37	36	39	39	27	25	35	42	38	51	28
total sample size	30,027	30,216	28,450	43,722	38,511	31,174	30,293	30,158	31,101	45,371	33167	30,239
<b>(4) Top quartile</b>	-0.28**	-0.06	-0.14	-2.55*	1.07	1.14	-0.11*	-0.06	0.03	0.05	0.89	0.47
standard error	(0.12)	(0.10)	(0.10)	(1.46)	(1.82)	(1.94)	(0.06)	(0.07)	(0.06)	(1.45)	(1.28)	(1.58)
number in treatment group	22	27	34	62	44	37	38	30	32	76	78	42
total sample size	31,757	32,264	32,023	47,211	35,838	31,606	31,365	31,562	31,548	46,916	51,035	29,800

NOTES: Standard errors in parentheses (\*\*\* p<0.01, \*\* p<0.05, \* p<0.1). Academic growth models control for twice-lagged test scores in both math and literacy-related tests. Test scores for the CRT are normalized year-by-grade such that the mean is zero and standard deviation is one and are from the state benchmark (criterion-referenced) exams. Scores for NRT are reported as normal curve equivalents (NCE's).

## Matching

One weakness of the above analysis is that SVA students are compared with all students statewide, although the students who entered SVA are not representative of all students across the state. Thus, here we refine our analysis by forming comparison groups based on observable characteristics for students enrolled in SVA for at least three consecutive years. This analysis will help us understand in finer detail the nature of the effects we observed for this group in the previous analysis as our estimate of the counterfactual here may be better. Table 4 summarizes our matching results using both CRT and NRT data. Outcomes are presented as residual value-added estimates calculated from models that control for two years of lagged test scores. The columns pertain to value-added during the cohort's second year, third year, fourth year, and fifth year.

Based on CRT data, the SVA group value added scores were 0.29 standard deviations lower in math than their matched comparisons during their second year. This difference decreased and turned insignificant for math in their third year and became positive, large, and significant in the fourth and fifth years. Students who enrolled in SVA during their fourth and fifth years gained 0.17 and 0.27 standard deviations higher in math than their matches. The pattern for literacy was more favorable to the SVA group than for math. None of the coefficients were statistically significant except for the fifth year (0.22 standard deviations).

Results for both subjects from NRT data convey a somewhat similar story as CRT math. Math and literacy value-added experienced by the SVA group were 6.5 and 6.1 NCE's lower than their matches during their second year and turned statistically insignificant in their third year. In their fourth year, the SVA group earned value added scores that were, on average, approximately 4 NCE's greater than the scores of their matched peers. In the fifth year, these magnitudes remained positive but decreased and became insignificant.

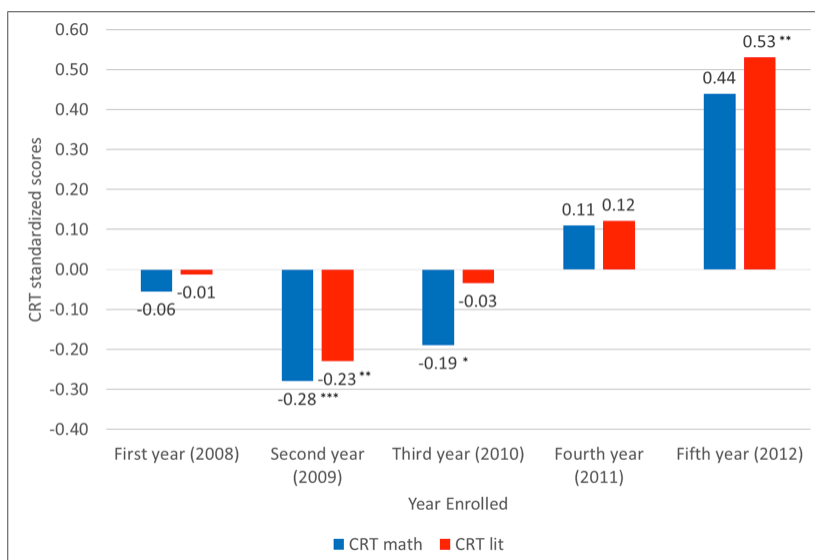
Cumulative effects. Because these models are based on value-added estimates that control for lagged scores, they simply present individual annual scores and do not provide a sense about the extent to which the SVA group recovers from their initial slump. Instead, the coefficients represent gains experienced between the lagged periods and the evaluation period, rather than between the beginning of the sample period and the evaluation period. To gain a better sense of whether and to what extent students recovered from their first year, we ran matching models based on residual value-added estimates, which controlled for pre-SVA test scores (i.e., from 2007) and match on this baseline year. That is, the estimates for each year indicate the gain or loss that the student experiences from the baseline year to the evaluation period relative to their matched comparisons. The results for these models are represented by Figure 3 and Figure 4. Table A.1 in the appendix provides more detailed information of these results.

**Table 4**  
**One-to-two nearest neighbor propensity score matching results for sample of 2008 cohort of SVA students enrolled at least 3 consecutive years**

Years enrolled in SVA	First year (2008)	Second year (2009)	Third year (2010)	Fourth year (2011)	Fifth year (2012)
<b>CRT data:</b>					
Math	0.03	-0.29***	-0.07	0.17**	0.27***
standard error†	0.11	0.08	0.06	0.07	0.10
number SVA students	32	53	99	67	40
total sample size	96	159	297	201	120
Literacy	0.08	0.07	-0.04	0.07	0.22**
standard error†	0.09	0.09	0.05	0.08	0.09
number SVA students	32	53	99	67	40
total sample size	96	159	297	201	120
<b>NRT data:</b>					
Math	-4.09	-6.52***	-0.88	4.24***	0.59
standard error†	2.75	1.61	1.20	1.39	1.60
number SVA students	30	52	149	112	77
total sample size	90	156	429	318	225
Literacy	-3.25	-6.06***	-1.20	3.85***	1.08
standard error†	2.65	1.60	1.25	1.41	1.49
number SVA students	30	52	149	112	77
total sample size	90	156	435	324	225

NOTES: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Outcomes are residual value-added measures for CRT math and literacy scores as function of 2-years of lagged test scores. Estimates are based on regression-adjusted OLS models. Outcome measures are residual value-added estimates from data over the period 2006-2012 derived from models that control for two years of lagged test scores in both subjects. Each student in the treatment group is paired with two non-SVA students based on grade and other covariates observed during the student's first year in SVA. The pool of matches comprises public school students over the sample period who were observed in the same school at least three years. Students are matched on propensity scores from probit models based on a set of covariates that include math and literacy test scores, race, gender, and IEP status. FRL data on students enrolled in SVA is not collected by the state and is not included in analysis. Treatment and control groups were balanced on observables (difference in means were statistically insignificant) during matching.

† Standard errors are estimated with a bootstrap procedure (1,000 replications)



**Figure 3.** Difference between SVA students (2008 cohort enrolled at least 3 years) and matched comparison group.

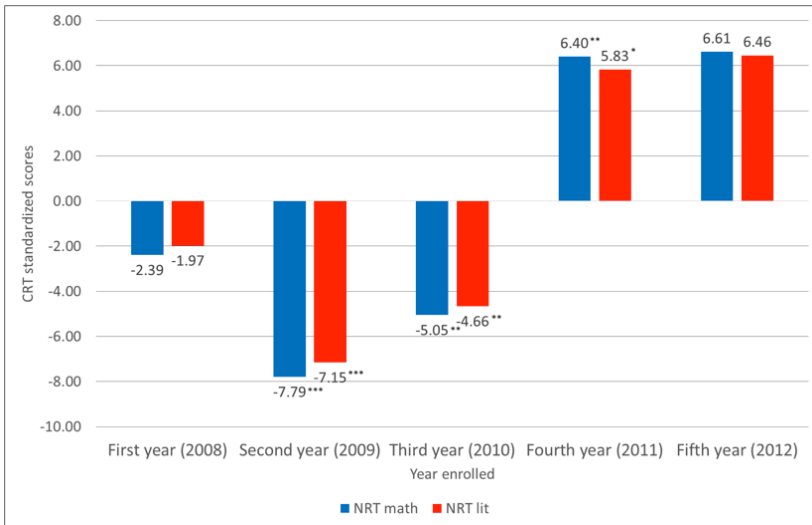
Note: outcomes are CRT value-added that controls for pre-SVA 2007 test scores

Figure 3 suggests that SVA students recovered and even significantly outperformed their matched comparisons after their third year. We first examine CRT data. During their second year, the SVA group experienced a loss of 0.28 and 0.23 standard deviations in math and literacy, respectively, relative to the comparison group. They recovered slightly in math during their third year, though cumulative math value-added over the baseline year remained negative and marginally significant. Recovery from the initial dip in literacy was more rapid, with the coefficient becoming almost zero and insignificant by the third year. Thus, these students by their third year in SVA recovered to neutral in literacy but did not recover in math.

For about half the students from this group who remained for their fourth year, cumulative value-added from the baseline year became positive and statistically insignificant for both subjects. For the small number of students who continued in SVA for a fifth year, cumulative learning became large and positive for both subjects and statistically significant for literacy. By five years, this group experienced 0.53 standard deviations higher value-added in literacy than their comparisons, or an annualized difference of about 0.11 standard deviations.

NRT results in Figure 4 tell a similar story. Relative to their comparison peers, SVA student gains, on average, were negative but statistically insignificant in their first year. They experienced a dip during their transition

year and recovered slightly in the subsequent third year, though not fully. Students who continued in SVA for their fourth and fifth years, however, fully recovered and experienced significantly larger gains than their matched comparisons. Estimates for cumulative math and literacy gains during the fourth year were statistically significant and marginally statistically significant. Fifth year estimates were also positive and large, though statistically insignificant because of the small sample size.



**Figure 4.** Difference between SVA students (2008 cohort enrolled at least 3 years) and matched comparison group.

Note: outcomes are NRT value-added that controls for pre-SVA 2007 test scores

In summary, students who remained in SVA at least three years experienced a significant negative transition effect in their first few years in terms of math and literacy value-added. Following their second year in SVA, although the students recovered somewhat, the net value-added for the whole group remained negative in the third year. Students who remained longer than three years completely recovered. They experienced significantly greater value-added in their fourth and fifth years relative to their matched comparison group. This recovery was more marked in literacy than math. Thus, the negative transition effects experienced by these students disappeared and reversed themselves over time.

### Limitations

We acknowledge important limitations with this analysis. First and foremost, our analysis was based on non-random selection. We addressed this problem somewhat by controlling for two years of prior achievement in our value-added analysis. Second, the analyses we conducted include students that enter SVA from non-public school settings such as private schools or home-school environments; for these students, we cannot observe their pre-SVA levels and thus have no true “baseline” score for many of the students. Third, both the CRT and NRT data are imperfect, though we argue that we are better off with the knowledge generated by this data than not having this knowledge at all. The former are susceptible to ceiling effects while results from the latter should be interpreted with caution as the state switched exams in 2011. Because of these concerns, we conducted analyses using two assessment measures and multiple models in an attempt to check the robustness of our results. Nevertheless, in this study we were simply unable to draw the types of strong causal conclusions that are drawn in random assignment designs.

## DISCUSSION AND POLICY IMPLICATIONS

Our study addressed a perceived shortage in empirical, longitudinal research on cyber schooling (e.g., Rice, 2006; Molnar et al., 2013, 2014, 2015) by using a rich panel data set to examine the impact of a statewide fully online cyber charter school in a southern state on student academic outcomes. Although some studies that examined the impact of cyber schools on student academic performance did not control for baseline outcome measures (e.g., Molnar et al., 2014, 2015; Minnesota Office of the Legislative Evaluator, 2011; Hubbard and Mitchell, 2011), our panel data allowed a more nuanced comparison by allowing us to control for these important factors. Moreover, studies such as Zimmer et al. (2009), though longitudinal, omitted switchers from their analysis. Our analysis included both switchers and non-switchers.

Our results suggest that students transitioning to a non-traditional cyber, fully online learning environment experience an initial dip in performance. We found evidence of transition effects mostly in students' first year in SVA. These effects decreased, however, for students who remained enrolled in the school past one year. Indeed, we observed a large, positive, and significant estimate on 2012 literacy for the group of SVA students who were in their third, fourth, or fifth year. Furthermore, SVA itself may have improved in effectiveness over time as negative effects tend to be found in 2010 while positive effects were largely found in 2012. Observed negative effects tended to dissipate after 2010.

Although we were unable to ascertain the reasons for the initial negative transition effects, there are several plausible explanations. Students may need time to adjust to using technology in a new learning environment. They may also need to adjust to learning at home after being acclimated to learning in a classroom-based environment. Finally, it may be the case that parents choose to enroll their children in SVA because of problems or “shocks” experienced in their previous school that might be connected to drops in student performance. Unfortunately, our data did not allow us to tackle these questions.

Overall, the somewhat unsurprising conclusion is that SVA students experience an initial drop in academic achievement but then improve to the point that students in SVA three or more years generally do as well as or better than their matched peers. Importantly, we also considered the experience of various subgroups of students. These sub-analyses uncovered several interesting patterns.

First of all, we considered student scores in 2010, 2011, and 2012. A review of these three samples revealed our first positive trend for SVA students: our data suggest that SVA is becoming more effective over time. Compared with their matched student peers, first and second year SVA students experienced significantly lower math and literacy value-added than their comparisons in 2010. CRT math and literacy estimates for first year students, for instance, were  $-0.44$  and  $-0.36$ , respectively. NRT estimates from 2010 were similar. Estimates for third to fifth year students were also negative, though statistically insignificant in most cases. Overall, outcomes for SVA students improved in 2011, and by 2012 differences were either positive or neutral, suggesting that SVA improved over time.

Second, we examined a cohort of SVA students from 2008 who enrolled at least three years and matched them to similar students based on observable student characteristics. In this way, we were able to observe the possibility of improvements over time at SVA. Here again, albeit with small samples, we found clear evidence of “veteran” SVA students outperforming their matched peers across the state. For instance, students enrolled in SVA for five consecutive years outgained their comparison group on CRT math and literacy by  $0.27$  and  $0.22$  standard deviations, respectively. These patterns suggest that although SVA may not provide positive value-added for all students who enroll, some students may benefit from its offerings.

Finally, our data allowed us to consider subgroups of students from minority backgrounds, students with special needs, and students from the top and bottom quartiles of the ability distribution. For the most part, due to the small samples of students in each group, clear patterns did not emerge – with one exception. The 25-30 students identified as special needs students in SVA in 2012 experienced very positive scores on both math and



literacy on both the NRT and CRT exams. This finding suggests that, for special needs students, relative to their peers in traditional public schools, SVA was as effective in the earlier years and more effective in 2012. This finding is important in light of the recent findings that some cyber schools indeed serve disproportionate numbers of special needs students (Egalite, Beck, & Maranto, 2014; Beck, Maranto, & Lo, 2013; Carnahan and Fulton, 2013). Moreover, this pattern is also consistent with the improving SVA performance over time.

It is important to note here that these outcomes are consistent with a large number of educational media comparison research studies – that the online environment or technology in general does not *by itself* dictate student performance. This was most notably summarized by Richard Clark in his (1983) *Review of Educational Research* article in which he compared the impact of the medium of student learning to the impact that a delivery truck has on the nutritional value of the groceries that it carries. As a result, the question for researchers and policymakers should not be what impact does the technology have on student learning outcomes, but what actually caused these impacts? We know that SVA delivers an entire curriculum in a way not offered anywhere else in the state among public schools. This unique curriculum delivers mixed results for SVA students, mostly negative at first and mostly positive after several years; this finding is generally consistent with the literature on distance learning. So, what does this all mean to policymakers who must decide whether to expand or limit seats at cyber schools across the nation?

First, based on the student turnover and the fact that “veteran” SVA students consistently improve over time, the flexibility inherent in the SVA cyber school environment may work well for some students and not well for others. Therefore, policies might focus on ascertaining the “goodness of fit” between SVA and potential enrollees. For instance, policy might facilitate good matches with information drives and help families decide whether this type of schooling is suitable for their children. SVA school leaders may want to address this issue by allocating more resources, time, or focus onto first-year students in an effort to more efficiently acclimate them. Given its unique learning environment, however, SVA may not be a good match for everyone. Thus, we would encourage policymakers, to the extent possible, to create policies that ease the transitions for students and make it possible for students to easily exit the cyber sector if it does not appear to be a good fit.

Second, not only do SVA students do better over time, but it appears that the school itself is improving over time as evidenced by the changes from 2010 to 2012. This suggests that policymakers who have decided to award charters to cyber schools should be willing to allow a few years for a school to reach some level of maturity in curricular and administrative decision making.

Third, this analysis suggests, based on our admittedly small sample of students, that some special needs students may well find the flexibility of the SVA cyber environment more conducive to learning than a traditional environment. In our view, it may be the case that the availability of cyber schools such as SVA can provide additional options for some students for whom the traditional system is not working. An important follow-up for research is to investigate the conditions that enable SVA's success educating some students. Perhaps for some students struggling with challenges as intense as bullying or as simple as an inability to concentrate fully in a crowded classroom, the option to receive online instruction via cyber schooling may well be helpful.

While many avenues for research remain open, this analysis adds to the current body of knowledge on cyber schools by shedding light on the extent that some students recover from the transitions they experience early on. Further analysis that investigates the reasons that parents seek cyber schooling options and the reasons for these challenges with the transition is warranted, however. For instance, a detailed analysis of time-on-task and the role of both teachers and parents in supervising students might inform school and public policy. In any event, as educators and policymakers alike continue to experiment with various forms of technology to improve our ability to educate all students, researchers should continue to monitor and report on this work.

## References

- Barbour, M. K., & Reeves, T. C. (2009). The reality of virtual schools: A review of the literature. *Computers & Education, 52*(2), 402-416.
- Beck, D., Egalite, A., & Maranto, R. (2014). Why they choose and how it goes: Comparing special education and general education cyber student perceptions. *Computers & Education, 76*, 70-79.
- Beck, D. E., Maranto, R., & Lo, W. J. (2014). Determinants of student and parent satisfaction at a cyber charter school. *The Journal of Educational Research, 107*(3), 209-216.
- Boardman, A. & Murnane, R. (1979). Using panel data to improve estimates of the determinants of educational achievement, *Sociology of Education, 52*(2), 113-121.
- Booker, K., Gilpatric, S., Gronberg, T., & Jansen, D. (2007). The impact of charter school attendance on student performance, *Journal of Public Economics, 91*(5/6), 849-876.
- Carnahan, C. & Fulton, L. (2013). Virtually forgotten: Special education students in cyber schools, *TechTrends, 57*(4), 46-52.
- Cavanaugh, C., Gillan, K. J., Kromrey, J., Hess, M., & Blomeyer, R. (2004). The effects of distance education on K-12 student outcomes: A meta-analysis. *Learning Point Associates/North Central Regional Educational Laboratory (NCREL)*.

- Cavanaugh, C., Gillan, K. J., Bosnick, J., Hess, M., & Scott, H. (2008). Effectiveness of interactive online algebra learning tools. *Journal of Educational Computing Research, 38*(1), 67-95.
- Cavanaugh, C. (2009). A Review of Research Learnings. *TechTrends, 53*(4), 29.
- Clark, R. E. (1983). Reconsidering research on learning from media. *Review of educational research, 53*(4), 445-459.
- Clausen, J. M. (2007). Beginning teachers' technology use: First-year teacher development and the institutional context's effect on new teachers' instructional technology use with students. *Journal Of Research On Technology In Education, 39*(3), 245-261.
- Colorado Department of Education (2006). Report of the State Auditor: Online education. Denver, CO. Retrieved March 5, 2015, from <http://www.cde.state.co.us/onlinelearning/download/2006%20Report%20of%20the%20State%20Auditor.pdf>
- Dehejia, R. & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *The Review of Economics and Statistics, 84*(1), 151-161.
- Eccles, J. S., & Midgley, C. (1989). Stage-environment fit: Developmentally appropriate classrooms for young adolescents. *Research on Motivation in Education, 3*, 139-186.
- Friedel, J. M., Cortina, K. S., Turner, J. C., & Midgley, C. (2010). Changes in efficacy beliefs in mathematics across the transition to middle school: Examining the effects of perceived teacher and parent goal emphases. *Journal of Educational Psychology, 102*(1), 102.
- Florida Tax Watch (2007). *Final report: A comprehensive assessment of Florida Virtual School*. Tallahassee, FL: Florida Tax Watch Center for Educational Performance and Accountability. Retrieved December 7, 2015 from <http://www.scribd.com/doc/47743217/Florida-Virtual-School-Report>.
- Gemin, B., Pape, L., Vashaw, L., & Watson, J. (2015). Keeping pace with K-12 digital learning: An annual review of policy and practice. Durango, CO: Evergreen Education Group. [http://www.kpk12.com/wp-content/uploads/Evergreen\\_KeepingPace\\_2015.pdf](http://www.kpk12.com/wp-content/uploads/Evergreen_KeepingPace_2015.pdf)
- Glass, G. V & Welner, K.G. (2011). *Online K-12 Schooling in the U.S.: Uncertain Private Ventures in Need of Public Regulation*. Boulder, CO: National Education Policy Center. Retrieved December 7, 2015 from <http://nepc.colorado.edu/publication/online-k-12-schooling>.
- Hanushek, E. (1979). Conceptual and empirical issues in the estimation of educational production functions. *The Journal of Human Resources, 14*(3), 351-388.
- Hanushek, E. A., Kain, J. F., Rivkin, S. G., & Branch, G. F. (2007). Charter school quality and parental decision making with school choice. *Journal of public economics, 91*(5), 823-848.
- Hedges, L., Laine, R., & Greenwald, R. (1994). Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes. *Educational Research, 23*(3), 5-14.
- Hill, P., & Roza, M. (2010). Curing Baumol's disease: In search of productivity gains in K-12 schooling, CRPE White Paper n 2010\_1. *Center for Reinventing Public Education. Seattle, WA: University of Washington*. Retrieved from [http://www.crpe.org/cs/crpe/download/csr\\_¿les/whp\\_crpe1\\_baumols\\_jul10.pdf](http://www.crpe.org/cs/crpe/download/csr_¿les/whp_crpe1_baumols_jul10.pdf).
- Hoxby, C. M., & Murarka, S. (2009). *Charter schools in New York city: Who enrolls and how they affect their students' achievement* (No. w14852). National Bureau of Economic Research.
- Hubbard, B. & Mitchell, N. (2011). Troubling questions about online education. *EdNews Colorado*.

- Hughes, J. E., & Ooms, A. (2004). Content-Focused Technology Inquiry Groups: Preparing Urban Teachers to Integrate Technology to Transform Student Learning. *Journal Of Research On Technology In Education*, 36(4), 397-411.
- Innovation Ohio (2011). Ohio e-schools: Funding failure; Coddling contributors. Columbus, OH: Author. Retrieved March 5, 2015, from <http://innovationohio.org/2011/05/12/ohio-e-schools-funding-failure-coddling-contributors-2/>.
- Joint Legislative Audit Committee (2010). *An evaluation: Virtual charter schools*. Madison, WI: Legislative Audit Bureau. Retrieved March 5, 2015, from <http://legis.wisconsin.gov/lab/reports/10-3full.pdf>.
- Karp, G. & Woods, M. (2003). Wellness NutriFit online learning in physical education for high school students, *The Journal of Interactive Online Learning*, 2(2), 1-19.
- Kingery, J. N., Erdley, C. A., & Marshall, K. C. (2011). Peer acceptance and friendship as predictors of early adolescents' adjustment across the middle school transition. *Merrill-Palmer Quarterly*, 57(3), 215-243.
- Koedel, C. & Betts, J. (2010). Value added to what? How a ceiling in the testing instrument influences value-added estimation. *Education Finance and Policy*, 5(1), 54-81.
- Levenberg, A., & Caspi, A. (2010). Comparing perceived formal and informal learning in face-to-face versus online environments. *Interdisciplinary Journal of E-Learning and Learning Objects*, 6(1), 323-333.
- Martínez, R. S., Aricak, O. T., Graves, M. N., Peters-Myszak, J., & Nellis, L. (2011). Changes in perceived social support and socioemotional adjustment across the elementary to junior high school transition. *Journal of Youth and Adolescence*, 40(5), 519-530.
- Matthew, K. I., & Varagoor, G. (2001). Student Responses to Online Course Materials. *Journal of Research On Computing In Education*, 33(5), 1-20.
- Minnesota Office of the Legislative Auditor, September 2011, K-12 Online Learning Evaluation Report Summary.
- Miron, G., & Urschel, J. L. (2012). *Understanding and Improving Full-Time Virtual Schools: A Study of Student Characteristics, School Finance, and School Performance in Schools Operated by K12 Inc*. National Education Policy Center, July.
- Molnar, A., Miron, G. Huerta, L. Rice, J. K., Cuban, L., Horvitz, B., Gulosino, C., Shafer, S. R. (2013). *Virtual Schools in the U.S. 2013: Politics, Performance, Policy, and Research Evidence*. Retrieved January 31, 2014 from <http://nepc.colorado.edu/publication/virtual-schools-annual-2013>
- Molnar, A. (Ed.); Rice, J.K., Huerta, L., Shafer, S. R., Barbour, M.K., Miron, G., Gulosino, C, Horvitz, B. (2014). *Virtual Schools in the U.S. 2014: Politics, Performance, Policy, and Research Evidence*. Boulder, CO: National Education Policy Center. Retrieved August 25, 2014 from <http://nepc.colorado.edu/publication/virtual-schools-annual-2014>.
- Moore, M. G. (1991). Distance education theory. *The American Journal of Distance Education*, 5(3), 1-6.
- Moore, M. G., & Kearsley, G. (2011). *Distance education: A systems view of online learning*. Retrieved from CengageBrain.com.
- Mouza, C. (2011). Promoting Urban Teachers' Understanding of Technology, Content, and Pedagogy in the Context of Case Development. *Journal of Research on Technology In Education*, 44(1), 1-29.
- Mueser, P., Troske, K., & Gorislavsky, A. (2007). Using state administrative data to measure program performance, *The Review of Economics and Statistics*, 89(4), 761-783.

- O'Dwyer, L., Carey, R., & Kleiman, G. (2007). A study of the effectiveness of the Louisiana Algebra I Online Course, *Journal of Research on Technology in Education*, 39(3), 289-306.
- Office of the Legislative Auditor, State of Minnesota (2011). *Evaluation report: K-12 online learning*. Retrieved February 26, 2014 from <http://www.auditor.leg.state.mn.us/ped/pedrep/k12oll.pdf>
- Ohio Alliance for Public Charter Schools (2009). *E-schools show superior results: Analysis of state value added data confirms e-schools students' progress*. Columbus, OH: Author. Retrieved March 5, 2015, from [http://www.oapcs.org/files/ESchoolStudy\\_final6-24-09.pdf](http://www.oapcs.org/files/ESchoolStudy_final6-24-09.pdf).
- Oliver, K., Osborne, J., & Brady, K. (2009). What are secondary students' expectations for teachers in virtual school environments? *Distance Education*, 30(1), 23-45.
- Oliver, K., Kellogg, S., Townsend, L., & Brady, K. (2010). Needs of elementary and middle school teachers developing online courses for a virtual school, *Distance Education*, 31(1), 55-75.
- Parsad, B., & Lewis, L. (2008). *Distance Education at Degree-Granting Postsecondary Institutions: 2006-07* (NCES 2009-044). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.
- Queen, B., & Lewis, L. (2011). *Distance Education Courses for Public Elementary and Secondary School Students: 2009-10* (NCES 2012-008). U.S. Department of Education, National Center for Education Statistics. Washington, DC: Government Printing Office.
- Rice, K. (2006). A Comprehensive Look at Distance Education in the K-12 Context. *Journal of Research on Technology In Education*, 38(4), 425-448.
- Roblyer, M. D. (1999). Is choice important in distance learning? A study of student motives for taking Internet-based courses at the high school and community college levels. *Journal of Research on Computing In Education*, 32(1), 157.
- Ross, H., & Brown, J. (2013). *Teachers' perceptions of pupil active citizenship and the transition from primary to secondary school*. Edinburgh: University of Edinburgh. Retrieved February 27, 2014 from [http://www.research.ed.ac.uk/portal/files/12564933/RossBrown2013\\_GCFProjectReport\\_AsDelivered.pdf](http://www.research.ed.ac.uk/portal/files/12564933/RossBrown2013_GCFProjectReport_AsDelivered.pdf)
- Ryman, A., & Kossan, P. (2011). The race to online: Arizona experiments with virtual K-12 schools. Will they work for your child? *Arizona Republic*. Retrieved March 5, 2015, from <http://www.azcentral.com/news/education/online-school/>.
- Sass, T. R. (2006). Charter Schools and Student Achievement in Florida. *Education Finance and Policy*, 1(1), 91-122.
- Shuldman, M. (2004). Superintendent Conceptions of Institutional Conditions That Impact Teacher Technology Integration. *Journal Of Research On Technology In Education*, 36(4), 319-343.
- Solomon, L. C., & Goldschmidt, P. (2004). Comparison of traditional public schools and charter schools on retention, school switching, and achievement growth. *Policy Report*, 192.
- Staples, A., Pugach, M. C., & Himes, D. (2005). Rethinking the Technology Integration Challenge: Cases from Three Urban Elementary Schools. *Journal Of Research On Technology In Education*, 37(3), 285-311.
- Tucker, B., Dillon, E. & Jambulapati, P. (2011). Ohio E-Schools: Learning from their experience. *Education Sector*. Retrieved June 26, 2013 from <http://www.educationsector.org/publications/ohio-e-schools-learning-their-experience>

- U.S. Department of Education, Office of the Deputy Secretary (2004). *Evaluation of the Public Charter Schools Program: Final Report*. Washington, D.C.
- U.S. Department of Education, Office of Planning, Evaluation, and Policy Development (2010). *Evaluation of Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies*, Washington, D.C., 2010.
- Wang, H., & Woodworth, K. (2011). *Evaluation of Rocketship Education's Use of Dream-Box Learning's Online Mathematics Program*, SRI International.
- Woodworth, J. L., Raymond, M. E., Chirbas, K., Gonzalez, M., Negassi, Y., Snow, W. & Van Donge, C. (2015). Online charter school study. Retrieved November 9, 2015 from <http://credo.stanford.edu/pdfs/Online%20Charter%20Study%20Final.pdf>.
- Woodworth, J. L., & Raymond, M.E. (2013). *Charter School Growth and Replication, Volume II*. (Stanford: CREDO, January 30 2013), accessed March 31 at <http://credo.stanford.edu/pdfs/CGAR%20Growth%20Volume%20II.pdf>.
- Zimmer, R. et al. (2003). *Charter School Operations and Performance: Evidence from California*. MR-1700, Rand Corporation.
- Zimmer, R., Gill, B., Booker, K., Lavertu, S., Sass, T. R., & Witte, J. (2009). *Charter Schools in Eight States: Effects on Achievement, Attainment, Integration, and Competition*. Santa Monica, CA: RAND Corporation.

## APPENDIX

Table A.1

One-to-two nearest neighbor propensity score matching results for sample of 2008 cohort of SVA students enrolled at least 3 consecutive years

Years enrolled in SVA	First year (2008)	Second year (2009)	Third year (2010)	Fourth year (2011)	Fifth year (2012)
<b>CRT data:</b>					
Math	-0.06	-0.28***	-0.19*	0.11	0.44
standard error†	0.09	0.10	0.12	0.14	0.29
number SVA students	53	53	53	30	12
total sample size	159	159	159	90	36
Literacy	-0.01	-0.23**	-0.03	0.12	0.53**
standard error†	0.08	0.10	0.11	0.14	0.26
number SVA students	53	53	53	30	12
total sample size	159	159	159	90	36
<b>NRT data:</b>					
Math	-2.39	-7.79***	-5.05**	6.40**	6.61
standard error†	2.27	1.66	2.31	3.07	4.40
number SVA students	52	52	52	29	12
total sample size	156	156	156	87	36
Literacy	-1.97	-7.15***	-4.66**	5.83*	6.46
standard error†	2.26	1.67	2.28	3.05	4.52
number SVA students	52	52	52	29	12
total sample size	156	156	156	87	36

Notes: \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$ . Outcomes are residual value-added measures for CRT math and literacy scores as function of baseline pre-SVA 2007 test scores. Estimates are based on regression-adjusted OLS models. Outcome measures are residual value-added estimates from data over the period 2007-2012 derived from models that control for baseline test scores (2007) in both subjects. Each student in the treatment group is paired with two non-SVA students based on grade and other covariates observed during the pre-SVA year. The pool of matches comprises public school students over the sample period who were observed in the same school at least three years. Students are matched on propensity scores from probit models based on a set of covariates that include math and literacy test scores, race, gender, and IEP status. FRL data on students enrolled in SVA is not collected by the state and is not included in analysis. Treatment and control groups were balanced on observables (difference in means were statistically insignificant) during matching.