

# Achievement and Behavior in Charter Schools: Drawing a More Complete Picture

Scott A Imberman<sup>1</sup>

University of Houston

*Abstract: Using a long panel with broad grade coverage, I assess how charter schools affect test scores, attendance, and discipline in order to establish whether these schools affect cognitive and non-cognitive skill formation. Schools that begin as charters generate large improvements in discipline and attendance but not test scores, with the exception of math scores in middle schools. I interpret this as reflecting improvements in non-cognitive skills but not cognitive skills. These improvements do not persist if students return to regular public schools. Charters that convert from regular public schools have little impact on cognitive or non-cognitive skill formation. These results are robust to potential bias from selection off of pre-charter trends, attrition and persistence.*

---

<sup>1</sup>204 McElhinney Hall, Houston, TX 77204-5019. simberman@uh.edu. I would like to thank the Maryland Population Research Center for their financial support. I extend my sincerest gratitude to employees and administrators of an anonymous school district for providing me with data and assistance and for making this project possible. I am especially grateful for guidance and assistance provided by my dissertation advisor Mark Duggan and the immeasurable help from Steven Craig. In addition, I'd like to give special thanks to Judy Hellerstein, Bill Evans, and Jeff Smith. I would also like to thank Rajashri Chakrabarti, Ken Chay, Aimee Chin, Jose Galdo, Jonah Gelbach, Ginger Jin, Beom-Soo Kim, Melissa Kearney, Adriana Kugler, Jordan Matsudaira, Jennifer King Rice, John Rust, Seth Sanders, John Shea, Barbara Sianesi, Alex Whalley, Ye Zhang, Ron Zimmer and seminar participants at Georgia State University, Stanford Institute for Economic Policy Research, Mathematica, RAND, University of Houston, University of Maryland, UNC-Chapel Hill, Urban Institute, Virginia Tech, APPAM, the North American Summer Meetings of the Econometric Society, and SEA for their helpful advice and comments. This work was done as part of my dissertation at the University of Maryland. All errors remain my own. c 2007 by Scott Imberman. All rights reserved. Short sections of text not to exceed two paragraphs may be quoted without the author's permission if properly cited.

# 1 Introduction

One of the fastest growing education reforms in the US today is the charter school movement. Charter schools operate under a contract, called a charter, with a government agency. These schools are provided a degree of autonomy from local school boards and freedom from some regulations in return for additional accountability requirements. Despite often being managed by private organizations, charters are public schools and receive almost all of their funding from government sources. Since 1997 the number of charter schools in the US has increased almost six fold, and the number of charter students has more than doubled since 1999, as is shown in Figure 1. As of 2006, 1.15 million students nationwide attended charter schools.

One of the largest questions in the charter literature is how charter schools affect the outcomes of students who attend them. It is unclear whether charters are beneficial or detrimental to students. On one hand, charters have fewer regulatory burdens and are at higher risk of being shut down if they under-perform, thus providing incentives to increase effort. On the other hand, charters have high levels of student turnover and eliminating some regulations may be detrimental to students. In addition to this theoretical ambiguity, the empirical evidence has been mixed. Of the papers which use more advanced econometric techniques, some researchers find insignificant or negative impacts of attending a charter school (Zimmer, Blanc, Gill and Christman, 2008; Hanushek, Kain, Rivkin and Branch, 2007; Betts, Rice, Zau, Tang and Koedel, 2006; Bifulco and Ladd, 2006; Sass, 2006; Zimmer and Buddin, 2006; Zimmer and Buddin, 2003), while others find positive impacts (Hoxby and Murarka, 2008; Booker, Gilpatric, Gronberg and Jansen, 2007; McClure, Strick, Jacob-Almeida and Reicher, 2005; Hoxby and Rockoff, 2004; Solmon and Goldschmidt, 2004; Solmon, Paark and Garcia, 2001).<sup>2</sup> Thus, the effect of charter schools on student outcomes is unclear.

One of the potential reasons for the wide variation in results is that some charter schools

---

<sup>2</sup>Of these all but Hoxby (2008), McClure (2005), and Hoxby and Rockoff (2004) utilize fixed-effects or some similar identification strategy. These papers instead use admission lotteries into charters as natural experiments.

may not see test scores as their primary output. Many charters focus on students with special needs such as those who are over-age for their grade, new immigrants, or students who have difficulty behaving well in a normal school environment. These students often need instruction not just in their academic ability - cognitive skills - but in motivation, self-esteem, and self-discipline which are types of non-cognitive skills.

The distinction between how cognitive skills and non-cognitive skills are affected by education interventions has become increasingly important in light of recent research showing that non-cognitive skills have substantial influence on labor market outcomes and degree attainment (Heckman, Stixrud and Urzua, 2006; Heckman and Rubinstein, 2001). Thus, if charters improve non-cognitive skills while having a limited effect on cognitive skills, these schools could still be effective tools for improving students' later-life outcomes.

Unfortunately, it is very difficult to measure non-cognitive skills directly. While testing of achievement is common in schools, other skills are rarely, if ever, tested. However, Heckman, Urzua, and Stixrud (2006) establish that both cognitive and non-cognitive skills improve behavioral outcomes. Thus if charters improve behavior and attendance rates, it could be indicative of improvement in non-cognitive skills even if cognitive skills - measured by achievement - do not improve. In particular, attendance is a logical proxy for self discipline and motivation. In addition, behavior and attendance are potentially important outcomes on their own since parents consider behavior to be an important factor in the decision on whether to send their children to charters. For example, in a survey of Texas charter parents, Weiher and Tedin (2002) show that only one-quarter of parents list test scores as the primary reason for sending their children to charters while more than two-thirds cite moral values, discipline, or safety.

Thus, I utilize a unique dataset from a large urban school district in the southwest (LUSD-SW) to provide a broad look at achievement, student discipline, and attendance and establish whether there is evidence of charter schools affecting cognitive or non-cognitive skills. To my knowledge, no study has considered how charters affect both of these skill sets.

This panel is also useful in that it provides wider coverage over time and grades than any previous study of charter schools. In total I have test score data for grades 1 - 11 over 9 years, during which charter schools were operating in each year. Attendance and discipline data covers grades 1 - 12 over 13 years. Using test scores with wide grade coverage appears particularly important, as limiting to the gradespans covered in Hanushek, Kain, Rivkin (2007) and Branch, Gilpatric, Gronberg and Jansen (2007) provide different results. In addition, I am able to look in detail at long-term impacts of charter schools while students are enrolled in charters and after they return to regular public schools, both of which have only rarely been studied in prior work while these studies used samples that did not include high school students.<sup>3</sup> Whether or not charter schools generate lasting impacts is particularly important. For the foreseeable future, the stock of charter schools in the US will be small relative to non-charters. Thus most students who enter charters in elementary and middle school will return to non-charter schools before leaving the public school system. If charters provide short-term benefits but no long-term benefits, the usefulness of these schools for generating human capital improvements will be limited. Hence, through studying multiple outcomes, using a broad base of students over a long time frame, and analyzing long-term effects I provide a comprehensive and wide-ranging analysis of charter school impacts on charter students.

I separate my analysis by two types of charters - startups and conversions. Previous work has shown these schools to differ in their impacts (Sass, 2006; Bifulco and Ladd, 2006; Zimmer, et. al., 2003) thus considering them together could lead to aggregation bias.<sup>4</sup> In addition, identifying whether these schools provide different impacts may have policy implications, since states and districts could allow only one type when starting a charter

---

<sup>3</sup>Booker et. al. (2007) find that student performance improves as time in charters increase and also when they leave charters. Bifulco and Ladd (2006) also look at how students perform in their first year in a charter and later years separately and find that first-year results were considerably worse. My strategy also differs from these two studies as I am able to instrument for students' charter exit decision which affects both the time-in-charter and persistence results.

<sup>4</sup>Hanushek, et. al. (2007) and Booker et. al. (2007) do not disaggregate by these charter types. However, in Texas only a few charters are conversions.

program, as is the case in Iowa, Mississippi, and Nevada. Startup charters are schools which begin as charters and enrollment is voluntary. Conversions in LUSD are schools that were previously regular public schools which convert to charter status. They keep the same staff, location, and attendance zones, thus most of their students are assigned based on location of residence like any normal school. Both types of charters benefit from exemptions from some regulations. Nonetheless, we would expect little impact from conversions since the change in the structure of the school is minimal. My results show this to generally be true.

Startup charters, on the other hand, generate impacts on student outcomes. While I find no statistically significant effect overall from attending a startup charter on test scores, there is an improvement in math for students in middle school grades of 0.07 to 0.18 standard deviations. Nonetheless, despite these test score finding, students gain large and statistically significant improvements in attendance and discipline from attending startup charters. On average, attendance rates increase by 2.4 percentage points. This is 23% of the absence rate in the year prior to charter entry. Startups also reduce annual disciplinary infractions by 0.5 to 0.8 instances. This is a very large impact relative to the average of 1.1 infractions in the year prior to entry. However, these results do not persist after students return to non-charter schools. In particular, attendance and discipline impacts disappear immediately after students return to regular public schools. Thus, while the impact of startup charters on cognitive skills is small, they generate a substantial improvement in non-cognitive skills. However, the drop-off after students return to regular public schools shows that these skills require continual reinforcement in students.

Nonetheless, one should be cautious in this interpretation as there are other potential explanations for the behavior results. One possibility is that charters differ from regular public schools in how they enforce or report discipline. This is an important concern and, as such, while it is not possible to completely rule out this explanation, I provide a series of tests and arguments that show that it is unlikely this is driving the results. Of particular importance is that the strong attendance results, which are much harder for the charters

to manipulate than discipline, serve to reinforce the discipline results and the fact that both measures show large and significant improvements provide evidence that there are non-cognitive skill improvements.

In addition to having multiple outcomes and the ability to assess long-term charter impacts, I also address some econometric issues. One potential econometric problem is that the assumptions underlying fixed-effects are invalid if students choose to attend charter schools based on changes in outcomes. If this occurs then the estimates of charter impacts may be contaminated by mean reversion. This phenomenon has been widely noted in the job-training literature (Heckman and Smith, 1999; Ashenfelter, 1978) while, in education, mean-reversion has been shown to occur in standardized exams (Chay, McEwan and Urquiola, 2005). Hanushek, et. al. (2007) use interrupted panel estimates to argue that this selection does not pose a problem. While I find some graphical evidence of this type of selection in startup charters, my interrupted panel estimates also show little change from baseline. Thus, while this selection does appear to exist, it does not substantially change the impact estimates.

Another potential problem is non-random attrition. This could create bias if the reason charter students leave the district for other schools at a rate that differs from public school students. While LUSD is a central city school district, it is bordered by 10 school districts. It also has many state charter schools and private schools within its boundaries. Thus there are a lot of educational options to parents and hence we need to be concerned about the attrition bias. Other papers have also shown that attrition bias in charter research is a potential problem. For example, Hanushek, et. al. (2007) find that charter students leave Texas public schools at more than 2.5 times the rate of non-charter students. To address this, in addition to tests for differential attrition, I utilize a unique semi-parametric attrition adjustment procedure for fixed-effects analyses proposed by Kyriazidou (1997). This procedure has not been previously used to assess charter impacts. These results suggest that non-random attrition does not have a substantial effect on the charter impact estimates.

## 2 LUSD Characteristics and Data Description

LUSD-SW was one of the first school districts in the US to institute a charter program. The program began in 1996 with two schools but in 1997 and 1998 expanded to half its current size. By 2006 there were 32 charter schools, 23 startups and 9 conversions.<sup>5</sup> Students from nearly 300 non-charter schools are also observed in the panel. During the time period studied, approximately 50 state charter schools were also in operation for which I do not have data. Figure 2 shows the evolution of the charter program in LUSD by examining the fraction of enrollment in startup and conversion charters. In 1997 and 1998 all of the conversions obtained charter status after which their enrollment shrank relative to total growth. Most of the startup charters opened in 1998 and 2001, but their population steadily increased over time. As of the 2006-2007 school year four percent of students in LUSD attended a charter school. Table 1 provides some summary information about charter schools and students. Startup students are more likely to be minority, poorer, and more at-risk than non-charter. The schools themselves are smaller on average and spend more per-student on instruction but less on other expenses. Conversions are also more heavily minority and poorer than regular public schools but are otherwise similar.<sup>6</sup>

In this paper I utilize a unique set of administrative records from a large urban school district in the southwest. This dataset includes information on scores from the Stanford Achievement Test, disciplinary records, attendance rates, and a number of student characteristics. By combining these outcomes I assess to what extent charter schools affect both cognitive and non-cognitive skill development. The data cover the 1994-1995 to 2006-2007

---

<sup>5</sup>One charter existed under contract with LUSD prior to the enactment of the state's charter laws and then promptly switched. Since enrollment is voluntary I define it as a startup. Some startups reside on the campus of an existing school but are considered independent schools and have voluntary enrollment. One conversion charter maintains a large gifted and talented magnet program. In order to prevent the impact of this program from influencing the charter impact estimates I drop any student who attends that school from the analysis, leaving 8 conversions in the final sample.

<sup>6</sup>Startup charters tend to be spread across all grade levels. There are thirteen startups covering at least one elementary grade, twelve covering at least one middle, and eight covering at least one high school grade. Since charters often diverge from the standard grade structure, these numbers only refer to 23 schools. Seven of the eight conversions I study cover at least one elementary grade. Three cover at least one middle school grades, but only one of these include grades 7 or 8. There are no high school conversions

academic years and I am able to follow individual students for as long as they attend school in LUSD, providing a long time-series on many students.<sup>7</sup> Hence, I am also able to look at how these skills develop in the long-term both during and after charter exposure.

LUSD started giving the Stanford Achievement test in the fall of 1997 and then subsequently winter of 1999 and every year thereafter in February or March. Since the 1997 test was given early in the school year I use the test scores from the 1998-99 school years and later. Thus the data includes nine years of test score data. The exam is norm-referenced, thus it reflects achievement relative to a national sample of students. Three exam subjects - math, reading and language - were given in grades 1 - 11 and thus I analyze these. For each subject I standardize the scaled scores to be mean zero, standard deviation one within grade and year across LUSD. Hence, impacts are measured in standard deviation units relative to the district average. In order to ensure that the impacts are analyzed on the same sets of students across all tests, after standardizing the scores I limit the sample to students who have scores for all three exams.<sup>8</sup> The final sample for test scores includes approximately 1.14 million student-year observations including 15,000 startup and 20,000 conversion charter observations. In Table 1 we see that for both types of charters the test score means are not statistically significantly different from regular public schools.

Discipline and attendance records of students cover 1994 through 2006 for all students in grades 1 - 12. The attendance rate provides the percent of days the student attends school while he or she is enrolled. Discipline records provide information on the type and length of punishment for any infraction that results in an in-school suspension or more severe

---

<sup>7</sup>After dropping observations for early education, pre-kindergarten, and kindergarten, 56% of students who are first observed in the data prior to ninth grade have at least four observations. In addition, only 27% of startup charter and 21% of conversion charter students have neither pre nor post-charter observations.

<sup>8</sup>Some students who are not proficient enough in English in grades 1-8 took a separate Spanish language exam called Aprenda. While I have data on these exam results, the scores are not directly comparable to those of students taking the English exam so I do not include them in the analysis. Almost all students who take Aprenda are in grades 1 - 5 and account for 24% of all test-takers in those grades. One concern is that, since startups include fewer LEP or special education students, and thus only a handful of Aprenda takers, the Stanford Achievement Test results may be biased. However, test score regressions limited to students who are not classified as LEP or special education for the duration of the test sample show similar results to the baseline regressions. Hence, this is not a substantial problem.

punishment.<sup>9</sup> I use the total number of disciplinary infractions per year in most of the analyses. Across LUSD, 17% of student-year observations have at least one infraction and 9% have multiple infractions. For startup charters those figures are 8% and 3%, respectively and for conversions they are 14% and 6%. In total the sample for discipline and attendance includes 2.23 million observations of which 20,000 are students in startups and 40,000 are students in conversions. Once again returning to Table 1, we see that on average startup charters have significantly lower overall infraction rates and fewer fighting infractions than non-charters. Conversions have fewer substance abuse and violent crime infractions. Means for attendance are not statistically significantly different from regular public schools for either charter type. Thus, the summary statistics provide suggestive evidence of there being little impact on test scores for either charter type but some improvement in behavior for startup charters. This is consistent with charters improving non-cognitive skills but not cognitive skills. Thus, to confirm this result we turn to regression analysis.

### 3 Empirical Strategy

#### 3.1 Baseline Model

In order to identify whether charters improve cognitive and non-cognitive skills, I utilize the following individual fixed-effects model:

$$(1) \quad y_{it} = \alpha + \theta_C \text{Conversion}_{it} + \theta_S \text{Startup}_{it} + \text{Demog}_{it} \Gamma + \text{Switch}_{it} \Phi + \text{Gradear}_{it} \Psi + \phi_i + \epsilon_{it}$$

where  $y_{it}$  is some outcome measure for student  $i$  at time  $t$ ,  $\text{Conversion}_{it}$  and  $\text{Startup}_{it}$  are indicators for the type of charter the student is enrolled in status,  $\text{Demog}_{it}$  is a vector of time-variant observable demographic characteristics,  $\text{Gradear}_{it}$  is a set of grade-by-year

---

<sup>9</sup>Unfortunately, infractions that generate less harsh punishments, such as detention, are not in the data. In addition, the records provide only limited information on the type of infraction, since 80% of infractions are unspecified student code violations. Nonetheless, a few severe infractions such as substance abuse, criminal behavior and, for 2002 and later, fighting are identified in the data.

indicator variables which account for changes in outcomes over time and grade level,  $\phi_i$  is defined as above, and  $\epsilon_{it}$  is i.i.d. error.<sup>10</sup>  $Switch_{it}$  is a set of variables that define whether a student changes schools in year  $t$  due to a structural change from normal grade progression, a non-structural change for some other reason, or if it's the student's first year in the district. All of these variables are interacted with indicators for the student's grade level.<sup>11</sup>

A problem with this strategy is that prior test scores play a role in current achievement. To address this, researchers often use a value added (or gains) version of the fixed-effects model where the dependent variable is the annual change in outcomes.<sup>12</sup> This is equivalent to assuming that if lagged achievement is an explanatory variable then its coefficient equals one. Since the role of lagged achievement likely decays, we may prefer a model that explicitly includes lagged achievement as an explanatory variable, as in Hanushek, et. al. (2007) and Sass (2006)

$$(2) \quad y_{it} = \alpha + \beta y_{i,t-1} + \theta_C Conversion_{it} + \theta_S Startup_{it} + Demog_{it} \Gamma + Switch_{it} \Phi + Gradeyear_{it} \Psi + \phi_i + \epsilon_{it}.^{13}$$

Since lagged scores are endogenous, these papers instrument with twice-lagged scores. However, recent research has suggested that factors in children's distant youth play important roles in later achievement (Todd and Wolpin, 2004) suggesting that twice-lagged scores are unlikely to be exogenous and thus the estimates in these papers may be biased.

---

<sup>10</sup> $Demog_{it}$  includes free-lunch status, reduced-price lunch status, other economic disadvantage, whether the student is a recent immigrant, and whether a parent is a migrant worker. Other economic disadvantage indicates the student does not receive free or reduced-price lunch but does receive other poverty assistance. More detailed definitions are in the online appendix.

<sup>11</sup>Hanushek, et. al. (2007), Booker, et. al. (2007), Bifulco and Ladd (2006) suggest that properly controlling for student switches is important for separating charter impacts from the effects of switching schools. I follow Bifulco and Ladd and define a non-structural switch as switching into a school where less than 10% of a student's previous class switches into the same school. Conversely, a student undergoes a structural switch when more than 10% of his or her previous class switch into the same school. Thus, 12% of student-years undergo non-structural switches, 11% of student-years undergo structural switches and 12% are students moving from outside of the district.

<sup>12</sup>Alternatively one could use a random trend model where both sides of the estimating equation are differenced then demeaned. While this allows for individual time trends, it substantially reduces precision and could exacerbate bias if the trends are non-linear. Nonetheless, estimates from this model are similar to the baseline levels models.

As an alternative, I use both levels and value-added models. In expectation these two models bound the true estimate which would be identified by equation (2). I provide the proof of this in the appendix. Thus, I am able to identify charter impacts within a range of values while avoiding biases that are introduced by endogenous instruments.

## 3.2 Potential Biases from Selection

While attending any school is a choice, for most charters parents are not restricted by having to live in a specific attendance zone or meet some transfer qualification. Thus, selection into charters may be a more substantial problem than selection into regular public schools. The student fixed-effects analysis used in this paper corrects for selection into charters based off of unobserved characteristics that do not change over time, such as innate ability. Nonetheless, we may be concerned of selection due to time-varying factors. While I cannot eliminate all types of selection that could play a role, I am able to check for a few specific types that would be particularly important in the charter context.<sup>14</sup> My main findings that there is little improvement of cognitive skills but substantial improvement in non-cognitive skills from attending a startup charter are robust to these potential biases.

The first issue I consider is whether entry into charter schools is based on pre-entry trends in the dependent variable (Booker, Gilpatric, Gronberg and Jansen, 2007; Hanushek, Kain, Rivkin and Branch, 2007; Bifulco and Ladd, 2006; Sass, 2006). We may suspect that students enter charter schools due to changes in test scores or behavior, or a change in some strong correlate with these outcomes. Such a situation has been widely noted

---

<sup>14</sup>Another solution is to use oversubscription lotteries as natural experiments (Hoxby and Murarka, 2008; McClure, Strick, Jacob-Almeida and Reicher, 2005; Hoxby and Rockoff, 2004). While a lottery based strategy has substantial advantages over fixed-effects, there are three potentially undesirable aspects. First, oversubscribed schools are likely of higher quality than schools with spaces available, hence a comparison of lottery winners and losers will only identify the impacts for the best charter schools. Second, lottery studies may be subject to more attrition bias than panel studies, since parents who lose lotteries may be more likely to send their children to private schools or other districts than those who win. Third, often lotteries are limited to a very small number of schools - four in the case of Hoxby and Rockoff (2004) and one in the case of McClure, et. al. (2006). Nonetheless, Hoxby and Murarka (2008) are able to address all three of these concerns well as their sample of NYC charter schools has 40 schools with lotteries and they appear to have little attrition bias.

in the job-training literature and is commonly called “Ashenfelter’s dip” (Heckman and Smith, 1999; Ashenfelter, 1978). Since a parent may see a drop in performance as an indicator that the current school does not meet his or her child’s needs, it is reasonable to believe that students change schooling environments in response to poor performance. This violates the strict exogeneity assumption for fixed-effects, which requires that future and past outcomes are not affected by current charter status. While much of this problem is dealt with by appropriately controlling for students’ switching status, I provide evidence that some small “dips” remain for startup charters. Thus, I follow Hanushek, et. al. (2007) in the use of interrupted panel estimates to check for bias from endogenous charter entry. In addition, regressions that look at how charter impacts vary with time spent in a charter are used to check for mean reversion, which would be a sign that the estimates suffer from this bias.

Another problem is that some parents may choose to leave LUSD altogether if students perform poorly in charter schools. If this occurs at a different rate than in regular public schools conditioning on observables and student fixed-effects then there are time periods when these students should be observed but are not. This could lead to attrition bias. For LUSD this is particularly important as there are many other options available to students including private schools, state charters, and more than ten suburban school districts. Figure 3 provides some suggestive evidence that endogenous attrition is a problem. While about 16% of non-charter students exit LUSD each year, 26% of startup charter students attrit. Though not shown in the figure, the differences are more dramatic over longer time periods. For example, 39% of non-charter third graders in 1998 - 2000 are no longer in LUSD five years later while that number is 43% for conversion students and 63% for startup students. These statistics may simply reflect different characteristics of the schools such as different grade levels covered or the types of students who attend. Indeed, regressions of attrition propensity on charter status including all of the covariates in equation (1) show no statistically significant relationship between charter status and attrition propensity at all grade levels.<sup>15</sup> However, a

---

<sup>15</sup>In the model with a pooled estimate across all grade levels the coefficient on attrition from startups is 0.037 (s.e. 0.025) and conversions is 0.003 (s.e. 0.010).

model that interacts charter status with test scores show that startup students with higher reading scores and conversion students with higher math scores are less likely to attrit than comparable non-charter students. These results can be found in the online appendix.

Thus, while differential attrition appears to only be a minor problem the evidence so far is not overwhelming. As such, I use a non-parametric procedure for correcting sample selection in individual fixed-effects models proposed by Kyriazidou (1997) to further check for attrition bias. Her insight is that, since fixed-effects correct for attrition based on time-invariant factors, one can correct for endogenous attrition due to time-variant factors by weighting towards observations where there is no change in attrition propensity.

To produce Kyriazidou’s (1997) estimator, one must first define the selection equation,

$$(3) \quad s_{it} = W_{it}\Omega + \zeta_i + \mu_{it}$$

where  $W_{it}$  is a set of variables that are observed whether or not the individual has attrited,  $\zeta_i$  is an individual specific fixed-effect, and  $\mu_{it}$  is random i.i.d. error.  $W_{it}$  need not contain all (or any) of the variables in the outcome equation, but it does need to contain an exclusion restriction. In this paper I use a model that includes the student’s last observed startup or conversion charter status, free lunch status interacted with grade-level, reduced-price lunch interacted with grade level, other economic disadvantage interacted with grade level, recent immigration status, whether a parent is a migrant worker, grade-by-year indicators, and whether the student is ineligible to attend his or her previous school due to his or her predicted grade not being offered, which serves the exclusion restriction.<sup>1617</sup> This model is estimated using a conditional “fixed-effects” logit. The outcome equation in my model is the baseline charter impact equation.

After removing the fixed-effect in the outcome equation through first-differencing, Kyr-

---

<sup>16</sup>The idea behind this exclusion restriction is that a student would be more likely to leave the district if she has to switch schools anyway; that is the relative costs of leaving the district falls if students are forced to switch schools.

<sup>17</sup>For attrited observations I use the grade the student would have been in assuming normal grade progression

iazidou argues that in observations where  $(W_{it} - W_{is})\Omega = 0$  for  $s < t$ , the individual has not had a change in circumstances which affect attrition. Since a student’s innate tendency to switch schools is captured by fixed-effects we can generate consistent estimates of  $\theta$ , the charter effect, by using only those observations where this holds true.

Since limiting to observations where  $(W_{it} - W_{is})\Omega = 0$  would reduce power substantially, Kyriazidou proposes using kernel weights to focus the analysis on observations that are close to  $(W_{it} - W_{is})\Omega = 0$ . Thus, in the second stage I run a first-differenced version of (1) weighted by

$$(4) \quad \hat{\psi}_{it,n} = \frac{1}{h_n} K\left(\frac{(W_{i,t} - W_{i,t-1})\hat{\Omega}}{h_n}\right)$$

where  $K$  is a kernel function with bandwidth  $h_n$  and  $(W_{it} - W_{i,t-1})\hat{\Omega}$  is the first-differenced linear prediction from the selection model estimation.<sup>18</sup>

A third problem that could arise is bias from students leaving the charters and returning to the regular public schools. Both the possibility that students leave because they perform poorly in charters and that charters have long-term impacts on outcomes can contribute to the bias. In the first case, if students leave charters prematurely due to poor performance then they will reduce the number of charter period observations reducing the influence of bad charters. In the second case, charters that have long-term impacts that persist when students return to regular public schools will also affect periods when students are not in charters, thus biasing impacts towards zero.<sup>19</sup> To solve both of these problems, I conduct analyses that include indicators for whether a student is in a “post-charter” period which allows us to compare charter impacts directly to pre-charter periods while identifying whether charter

---

<sup>18</sup>The appropriate bandwidth is found using the mean-squared error (MSE) minimization procedure described in Kyriazidou (1997). Since the MSE minimizing bandwidth is sensitive to the initial value, I follow Dustman and Rochina-Barrachina (2000) in using the initial value that is as close as possible to the constant for the MSE minimizing bandwidth as asymptotic theory says the two should converge. I use the Gaussian kernel to generate the weights.

<sup>19</sup>This is a violation of the strict exogeneity assumption for fixed-effects which implies that outcomes in future and past periods cannot be correlated with current charter status conditioning on observables and the fixed-effect.

impacts persist after students return to regular public schools.<sup>20</sup>

$$\begin{aligned}
 (5) \quad y_{it} = & \alpha + \theta_{S,0}Startup_{it} + \sum_{k=1}^4 \theta_{S,k}PostStartup_k \\
 & + \theta_{C,0}Conversion_{it} + \sum_{k=1}^4 \theta_{C,k}PostConversion_k \\
 & + Demog_{it}\Gamma + Switch_{it}\Phi + Gradyear_{it}\Psi + \phi_i + \epsilon_{it}
 \end{aligned}$$

where  $PostStartup_k$  and  $PostConversion_k$  denote the student being in year  $k$  after leaving the charter. For  $k = 4$ , I include any time period after the third year. To account for students endogenously exiting charters, I instrument for being in a post-charter period with the whether the student is the listed number of predicted grades past his or her previous charter’s highest grade covered. I use predicted grades where I calculate normal grade progression starting from the year of charter entry.<sup>21</sup> This avoids potential complications from the charter’s retention policies. For example, the estimator for being one year post startup is instrumented using whether the student is one predicted grade level higher than the highest grade covered in the charter school he or she attended. Thus, we are able to isolate the effects from students who are forced out of charters due to being grade-ineligible, instead of those who leave voluntarily, possibly because they do not perform well in the charter

## 4 Results

### 4.1 Charter School Impacts on Student Outcomes

I now turn to my main results. We start with test scores to see how charter schools affect cognitive skill development. First, in Figure 4 I provide graphs that trace out the residuals

---

<sup>20</sup>It is possible to be in a conversion charter and also in a post-startup period at the same time (and conversely for startup and post-conversion), however only 2% of charter students ever attend both types.

<sup>21</sup>While I do not include Kindergarten in the analysis, as there are no test scores for this grade and enrollment is optional in LUSD, I am able to identify whether a student enters a charter in Kindergarten.

of student outcomes from a fixed effects regression including all of the covariates in model (1) except charter status. Thus, we are able to look at how outcomes change as students enter charters net of fixed student characteristics, economic status, immigration status, grade level, and time. The top row shows startup charters, which are the type of charter we would most expect to see charter impacts. While math scores seem to increase after entry, reading and language show far less improvement. The graphs also show some evidence of selection into the charter off of pre-charter trends since test scores drop in the year immediately prior to entry (-1), which could potentially bias the estimates. For conversions, test scores appear to fall off slightly after entry. Thus the graphs suggest that, with the possible exception of math in startups, cognitive skill improvement is small at best.

Table 2 provides the main test score results for this paper along with the interrupted panel and Kyriazidou (1997) estimates. The standard errors for each regression are robust to heteroscedasticity and clustered by school. The baseline results for startups confirm the graphical analysis as there is no statistically significant impact of startup charters on test scores in levels or value-added models. With the exception of math in levels all of the point estimates are relatively close to zero. Thus, these results suggest that there is little improvement in cognitive skills in startup charters on average. For conversions while math in levels is statistically significant at 0.07 standard deviations, the value added estimate is insignificant and thus we cannot say that conversions have a math impact.<sup>22</sup> In addition, this impact will disappear when we consider the role of persistence. One should note that these results differ considerably from those restricting grade span coverage to 3 - 8 in levels and 4 - 8 in value-added models. These are the grade spans used by Hanushek, et. al. (2007) and Booker, et. al. (2007). These regressions show positive and statistically significant impacts on math and language in startups and thus underscore the importance of having data on a

---

<sup>22</sup>At-risk, LEP, gifted & talented, and special education status could be useful covariates but they are also potentially influenced by school quality or charters may not classify students the same way as non-charters. As such, these variables would also be inappropriate as outcomes. Nonetheless adding LEP, at-risk status, gifted & talented status, and special education status to the baseline model has very little effect on the estimates.

wide range of grades.<sup>23</sup>

The next two rows show the interrupted panel estimates.<sup>24</sup> These estimates account for the pre-entry dips seen in Figure 4. These estimates are close to the baseline results regardless of whether one or two years prior are dropped. A possible exception is math for conversions in the value-added models where the impact increases to 0.07 standard deviations, but is only significant at the 10% level. The last three rows provide estimates using the Kyriazidou (1997) attrition correction procedure. Since this relies on a first-differenced estimator rather than a fixed-effects estimator I use these to examine how much estimates change when the attrition corrections are applied rather than as true impact estimates. The unweighted model is the first-differences corollary to the baseline model in the first row. If attrition is a problem then we would expect there to be large changes in the estimates as the weights are added. I show both models using mean-squared error minimizing weights and weights where the bandwidths are 1/4 of the MSE minimizing bandwidths. While startup math impacts increase and become statistically significant in value-added models, levels models are still statistically insignificant and change little. There is also some increase for language in conversions. Nonetheless, these results at worst suggest that we may be slightly underestimating math scores for startups and language scores for conversions, but in general the baseline results hold up well.

Thus, we see that startup charters have no significant effect on cognitive skills. However, it is still unclear how charters affect non-cognitive skills. Since many startup charters in LUSD target students with behavioral problems and those who are classified as “at-risk” of dropping out of school, we may expect them to focus more on non-cognitive than cognitive skills. While some cognitive skill improvement spills into discipline and attendance, the

---

<sup>23</sup>Results are provided in the online appendix.

<sup>24</sup>In value-added models Hanushek, et. al. (2007) keep the gain in the first charter year as the difference between year  $t$  and  $t - 1$ . Thus, while this reduces the bias from the pre-charter gain measures it does not reduce bias from the excessive gain in the first charter year. I modify the procedure such that the dependant variable in the value-added models is the average gain over the dropped years. Thus, when year  $t-1$  is dropped, I use the difference in test scores in  $t$  and  $t - 2$  divided by 2. Both strategies provide similar results for all outcomes.

lack of test score impacts imply that improvements in these outcomes would likely be due to non-cognitive skill enhancement. Figure 5 provides the same graphs for discipline and attendance as Figure 4 provides for test scores. While there is little change in either of these measures after charter entry for conversions, the figures show drops in disciplinary infractions and increases in attendance rates immediately after entry. In both cases there is some reversion after the first year, but both outcomes remain at levels that improve on pre-entry performance. As in the case with test scores, however, there is evidence of selection into charters off of worsening attendance, so once again we need to use interrupted panel estimates to address this potential bias.

Panel A of Table 3 shows the baseline, interrupted panel, and attrition adjusted results for discipline and attendance. While there are no statistically significant estimates in any model for conversion charters, startup charters have large and statistically significant improvements in discipline and attendance. Using the levels and value-added model estimates as bounds, baseline estimates in the first row show annual disciplinary infractions falling by between 0.5 and 0.8 incidences when students attend startup charters. These estimates are statistically significant at the 1% level. Average infractions in the year prior to charter entry is 1.1 so the startup charter impact is between 45% and 73% of the pre-charter mean. Startups also have statistically significant impacts at the 1% level on attendance. Baseline estimates show an increase of 2.3 percentage points. In the year prior to entry, attendance rates average 91.0% so the attendance impact accounts for 26% of the pre-charter absence rate. In addition, despite the graphical evidence for the Ashenfelter dips, discipline and attendance results hold in the interrupted panel and attrition adjusted estimates. Thus, it appears that startup charter improve non-cognitive skill formation.

Since LUSD audits attendance by checking teacher's logs with reported attendance, any systematic misreporting would require participation of administrators and teachers, which would be very difficult. Thus, we can be confident that the attendance results reflect actual behavioral improvements. Nonetheless, it is possible that the discipline results are

due to differences in enforcement or reporting rather than actual behavior improvements. To address this concern, I provide a few pieces of evidence that at least part of the impact reflects behavioral improvements. First, the large impact from attendance reinforces the discipline results since they are highly correlated and both reflect non-cognitive skill improvement.<sup>25</sup>

Second, regressions that use severe infractions - substance abuse, violent criminal activity, and fighting - as outcomes show statistically significant drops from attending a startup charter. In levels models, substance abuse infractions fall by 0.015 (s.e. 0.006), violent crimes by 0.008 (s.e. 0.002) and fighting by 0.024 (s.e. 0.009).<sup>26</sup> It is unlikely that principals punish students for these infractions with punishments that are less severe than in-school suspensions, thus improvements in these outcomes are most likely due to behavioral improvements.<sup>27</sup>

Third, at four to seven times the standard error, the enforcement or reporting bias would need to be very large to make the discipline estimates statistically insignificant. Fourth, a multinomial logit regression of type of punishment on charter status shows that startup charter students were more likely to receive out-of-school suspensions than in-school suspensions, suggesting that if anything punishments in startup charters are harsher than in non-charters.<sup>28</sup>

Finally, in panel B of Table 3 the first three columns provide results for some alternative measures of discipline. Since these are all binary outcomes I only report level models. All regressions include the same covariates as in the baseline regressions. I consider the effect of attending a charter on whether the student has any disciplinary infractions during the year,

---

<sup>25</sup>An OLS regression of attendance rate on the number of infractions, free-lunch, reduced-price lunch, other economic disadvantage, recent immigrant, parents' migrant status, gender, race, and grade-year interactions provides a point estimate on infractions of -1.12 and a standard error (clustered by school) of 0.05.

<sup>26</sup>All standard errors are clustered by school. Value added models were not statistically significant. However, since these are low-frequency events it is actually very rare to get more than one infraction in subsequent years, thus it would be very difficult for value-added models to identify impacts on these outcomes

<sup>27</sup>In addition, for substance abuse and criminal infractions the principal is legally obligated to notify the police department.

<sup>28</sup>The unit of observation is at the student-infraction level and the left-out category is in-school suspension. The regression includes covariates for gender, race, free lunch, reduced-price lunch, other economic disadvantage, immigration status, whether student is a recent immigrant, whether the parent is a migrant worker, infraction type - substance abuse, violent crime, non-violent crime, truancy, other - and grade-by-year dummies. The estimate of attending a startup charter on out-of-school suspensions is statistically significant at the 5% level. These results are available upon request.

whether the student is expelled, and whether the student is placed in alternative education. Startup charters provide statistically significant drops in likelihood of all three of these outcomes. Conversion charters also show drops in expulsion rates and alternative education placement, but these are substantially smaller than for startups. Thus, the discipline results for startups are robust across multiple margins. The last column of panel B looks at retention. Whether or not students are retained more in charters is potentially interesting but provides an unclear interpretation. The results show a slight increase in retentions for startups, but this is only significant at the 10% level. Even if we accept this as a retention impact, the interpretation is unclear. Higher retention could indicate that students perform worse, but it could also indicate a policy difference where charters are more likely to hold marginal students back or that charters are better at identifying students who need to be held back.

So far, I have established that on average startup charters provide little improvement in cognitive skills while generating large improvements in non-cognitive skills. Since this is an average result, these impacts may vary by school and student characteristics. Table 4 provides some results on how charter impacts differ by grade-level and age of the charter. Each regression includes main effects and interactions. Panel A looks at variation by grade level. Since there is only one conversion charter in the sample with grades 7 and 8 and none with grades 9 - 12, a pooled estimate for conversion charters is included in the regression but not shown. In columns (1) - (3) we see that while elementary and high-school startups have no statistically significant impact on any test score, startups with grades 6 - 8 (middle school grades) fare quite well. All three of the estimates in levels models show that students in these grades perform significantly better than those in elementary grades and tests of the sums of the main effects and the middle school interaction are also statistically significant at the 5% or higher level. However, for value-added models only math has a statistically significant combination of main and interaction effects. Thus middle school startups increase math scores by 0.7 to 0.18 standard deviations while reading and language increase by at most 0.07 and 0.08 standard deviations but I cannot rule out a zero effect

for those. Hence, while on average there is little evidence of cognitive skill improvements from startup charters, there is some weak evidence of improvements for middle schools. Discipline and attendance results show evidence of non-cognitive skill improvements at all grade levels. Elementary schools show improvement in discipline but not attendance while other grade levels show improvements in both measures. Nonetheless, average attendance rates are higher in elementary grades so there is less room for improvement. The district-wide average is 96.4% for grades 1 - 5 and 92.6% for grades 6 and higher.

Panel B looks at how charter impacts vary by the age of the charter in levels models. Value-added models provided similar results and are available in the online appendix. Previous work on charter schools have generally found that as charters age their test score impacts improve (Booker, Gilpatric, Gronberg and Jansen, 2007; Hanushek, Kain, Rivkin and Branch, 2007; Sass, 2006; Bifulco and Ladd, 2006). While I caution that since I have 31 charters there is less variation in this school level characteristic than in previous papers, I nonetheless find little evidence of improvement in test scores for startups or conversions as the schools age. I do find some evidence of improvement in discipline in startups as those that are five or more years old have 0.3 fewer annual disciplinary infractions per student than first-year startups. However, this is offset by lower attendance improvements after the first year. It is unclear why attendance rates would drop after the first year. One possibility is that schools that are new need to maintain a close knit community in order to succeed and thus they try harder to induce students to attend.

In table 5 I consider how charter school impacts vary by student characteristics. Since I do not find substantial evidence of impacts for conversion charters, I report only the estimates for startup charters.<sup>29</sup> Each column in section A or B is a separate regression with a main effect and interaction effects. No estimate for race is significantly different from the main effect in both levels and value-added models so we cannot say for sure if there is a racial difference. There are, however some notable differences by gender. Women in startups

---

<sup>29</sup>Full results including conversion estimates are provided in the online appendix.

appear to perform better on math tests than men by 0.03 to 0.06 standard deviations. While somewhat tenuous, as the value-added estimate is only significant at 10%, this is an intriguing result as there has been considerable focus in education policy on how to improve math scores for girls. Disciplinary infraction impacts, on the other hand, are smaller for girls, though they still show large improvements. This is not surprising, however, as boys tend to have more discipline problems on average. For economic status, there is no statistically significant difference for any outcome. Immigrants do seem to have lower improvements in attendance than non-immigrants, but the combined main effect and interaction is still positive.<sup>30</sup>

In Table 6 I investigate some mechanisms the charter impacts may work through. We saw in Table 1 that startup charters have more instructional expenditures, less expenditures on other functions, and are generally smaller than non-charter schools. Thus, seeing how the impact estimates change as we control for these characteristics can provide insight into the paths through which charters affect students. Nonetheless these covariates are potentially endogenous, and thus we should be careful of interpreting any change in the charter impact estimate as being causally determined by the added covariates. Tests score estimates for all charters and discipline and attendance estimates for conversions are generally unaffected by adding these covariates so I leave those results in the online appendix.

The first row repeats the baseline estimates from Tables 2 and 3. The second through fifth rows show what happens when we add different categories of per-student expenditures. Adding these factors does not generate substantial changes in the impact estimates. In row 6 I control for the percent of each school that is white, Hispanic, black, LEP, special education, gifted, and economically disadvantaged. In this case there is a drop in the discipline estimate in value-added models to the point where it is only significant at the 10% level. However, all of the other estimates change little.

---

<sup>30</sup>An additional element of heterogeneity that one may expect to see is that students with pre-charter behavioral problems improve more in startup charters than others since they drive the discipline results. However regressions that interact charter status with infractions in the year prior to charter entry show that test score improvements fall with the number of infractions, though not to the point where a large number of students would have significantly negative test score impacts. These results are available upon request.

Row 7 controls for a quadratic in total school enrollment. Since startups tend to be smaller than non-charters, administrators may find it easier to maintain control over students and to spend more time dealing with discipline problems. The results in row 7 are consistent with this theory. When enrollment is added the disciplinary infraction estimate rises by 0.145 in levels and .373 in value-added models. While the levels estimate remains statistically significant, the value-added estimates becomes insignificant. Thus, it appears that school size and peer characteristics may play roles in the discipline improvements found in the startup charters. In particular, part of the discipline improvement may be due to administrators having smaller schools that they can closely monitor.

## 4.2 Evolution of Charter Impacts Over Time

One of the key advantages of the LUSD-SW dataset is that the length and breadth of the panel allows me to investigate some long-term impacts of charters on cognitive and non-cognitive skills that most of the previous research is not able to do. First, in Table 7 I provide results from regressions where charter impacts are allowed to vary by time spent in a charter. We may suspect that initially a student may not perform well in a charter due to the shock of having to change educational environments, but over time the student could acclimate to the new environment and improve. On the other hand, it's possible that the new environment of attending a charter could temporarily motivate the student and improve performance but as time goes on the student may return to old habits.

In order to address the potential for endogenous exogenous exit I instrument for charter attendance in any year after the first. In this case I use the number of potential years a student could have been in the charter. That is I subtract the first grade offered in the charter from the the grade the student is predicted to be in based on date of birth and assuming normal grade progression. For test scores (columns 1 - 3) startup charters show some evidence of improvement when students have been in a charter for four or more years, but only in value-added models. Thus, there appears to be some weak evidence that students

who spend long periods of time in startups improve but those who spend only short periods of time in startups do no better or worse. For conversion charters it appears that test score impacts remain roughly constant as time in the charter increases.<sup>31</sup>

Looking at columns (4) and (5), for discipline and attendance there appears to be improvement as time in startup charters increase. For example in the first year annual disciplinary infractions are down by 0.8 in levels models. After the third year they are down by 1.6. An F-test of this difference is statistically significant at the 5% level. Attendance also shows a pattern of improvement. The pattern for discipline does not hold for startups in value-added models. This suggests that discipline improves substantially in the first year, but further improvements in subsequent years are small. Nonetheless, discipline impacts do increase as time spent in the charter increases. These results also show that there is no mean reversion in charter impacts, and thus provides further evidence that the pre-entry dips in test scores and attendance do not impose substantial bias.

I now turn to how charter impacts persist after students return to regular public schools. Analyzing persistence informs us as to whether the skills students learn in charters remain without the need for reinforcement. Persistence can also generate bias in the impact estimates since some of the impact would be attributed to non-charter periods. First, figures 6 and 7 provide an initial look at whether charter impacts persist by graphing residuals from the regressions used for figures 4 and 5 for students before and after they leave charters. In figure 6, we see that after charter exit math test scores drop for startups and all subjects drop for conversions. In figure 7, we also see that the discipline and attendance gains achieved while students are enrolled in startups drop off after they leave. These measures are relatively stable after students leave conversions. Thus these results suggest that charter impacts do not persist after students leave.

In table 8 I provide regression results that account for the persistence of charter impacts when students return to regular public schools. The table shows 2SLS results using the

---

<sup>31</sup>The 2SLS models are somewhat different from OLS models which show improvement in startup charters over time. These results are available in the online appendix.

instrument for “post-charter” periods described at the end of section 3.2. Since two and three years after attending a charter provide results similar to one and four-plus years, I only show the latter combination, though the F-tests are for all four categories.<sup>32</sup> Test scores are provided in columns (1) - (3), discipline in column (4) and attendance in column (5). Each column in panel A and panel B refer to a single regression.

First, for conversion charters, any significant positive impacts found in the baseline models become statistically insignificant. This provides additional confirmation that conversion charters are ineffective at improving student outcomes. Turning to the results for startups, we see in columns (1), (2) and (3) that after leaving the charter students perform well on test scores in the first year, with statistically significant and positive impacts in the levels models for math and language. Nonetheless almost all of the other estimates for post-startup test scores are either negative or statistically insignificant. Thus, the evidence for persistence in test scores for startups is weak. As such, the estimates for the startup impact on test scores remain statistically insignificant in all models.

In columns (4) and (5) we see that after startup students return to regular public schools disciplinary infractions increase and attendance fall to the point where they do not differ from non-charter students<sup>33</sup> While it is true that in the post-charter periods the students are older and thus we would expect these outcomes to rise, the inclusion of grade-by-year indicators corrects for this problem. Thus, there is no persistent impact of startup charters on non-cognitive skills.

The most likely explanation for these results is that the non-cognitive skill improvements from startup charters need continual reinforcement or else they dissipate. Hence, while the startups provide an environment that helps the student improve behavior and attendance, once they return to regular public schools they lose the benefit of that environment. An alternative explanation is that the estimated impacts reflect reporting and enforcement bias

---

<sup>32</sup>OLS and complete 2SLS results are available in the online appendix.

<sup>33</sup>The F-Test of the post-startup estimates in column 4 is significant because the two year and three year estimates are both positive 0.3 and statistically significant.

in charters. However, as discussed in section 4.1 it is highly unlikely that attendance rates are manipulated while the discipline results appear to be, at least partially, due to real behavioral changes.

## 5 Conclusion

In this paper, I investigate the effect of charter schools on students who attend them using data from a large urban school district in the southwest (LUSD-SW) with an extensive charter program. I provide a broad outlook on charter schools by focusing on multiple outcomes and using a panel with wide grade coverage and long time frames to provide analyses of long-term impacts both while in and after attending charters. By combining analyses of test scores with discipline and attendance impacts, I am able to identify whether charter schools affect both cognitive and non-cognitive skill formation.

I find that, while charters that convert from regular public schools to charter status have little effect on cognitive or non-cognitive skills, schools that begin as charters, called startup charters, generate large improvements in non-cognitive skills as proxied by discipline and attendance. Disciplinary infractions drop by 0.5 to 0.8 instances per year compared to a pre-charter mean of 1.1 while attendance rates rise by 2.3 percentage points which is equivalent to 26% of the pre-charter mean. These impacts are statistically significant at the 1% level across many models and specification checks. While discipline and attendance impacts could also be caused by cognitive skill enhancements, I can only establish improvements in test scores for math in middle school startups whereas estimates for elementary and high schools are close to zero on all test score measures. Hence, I interpret the discipline and attendance results as showing that startup charters generate non-cognitive skill improvements. These individual fixed-effects results are robust to potential biases from selection into charters off of trends in outcomes and attrition.

The long and broad nature of my panel also allows me to follow students for a number

of years while they are enrolled in charters and after they leave charters. While impacts for either type of charter are relatively stable as time in the charter increases, the discipline and attendance impacts found in startups do not persist after students return to regular public schools so there does not appear to be long-term behavioral improvements. This result suggests that any non-cognitive skills that have improved in startup charters require constant reinforcement.

Nonetheless, I caution that while the discipline and attendance results are intriguing they may reflect reporting or enforcement differences. The lack of persistence increases this concern as such a result is consistent with reporting bias. Since attendance is difficult to manipulate or misreport, this is mainly a concern for discipline. Thus, I provide multiple pieces of indirect evidence that suggests the discipline results, at least in part, reflect real behavioral changes in the startup charter students. In particular, the large impacts on attendance buttress the discipline improvements and combined, they provide substantial evidence that startups generate non-cognitive skill improvements in spite of the lack of cognitive skill improvement.

## References

- Ashenfelter, Orley, “Estimating the Effect of Training Programs on Earnings,” *The Review of Economics and Statistics*, 1978, 60 (1), 47–57.
- Betts, Julian R., Lorien A. Rice, Andrew C. Zau, Y. Emily Tang, and Cory R. Koedel, “Does School Choice Work? Effects on Student Integration and Achievement,” Technical Report, Public Policy Institute of California, San Francisco, CA 2006.
- Bifulco, Robert and Helen F. Ladd, “The Impacts of Charter Schools on Student Achievement: Evidence from North Carolina,” *Education Finance and Policy*, 2006, 1 (1), 123–138.
- Booker, Kevin, Scott M. Gilpatric, Timothy Gronberg, and Dennis Jansen, “The Impact of Charter School Attendance on Student Performance,” *Journal of Public Economics*, 2007, 91 (5/6), 849–876.
- Buddin, Richard and Ron Zimmer, “Student Achievement in Charter Schools: A Complex Picture,” *Journal of Policy Analysis and Management*, 2005, 24 (2), 351–371.
- Chay, Kenneth Y., Patrick J. McEwan, and Miguel Urquiola, “The Central Role of Noise in Evaluating Interventions that Use Test Scores to Rank Schools,” *American Economic Review*, 2005, 95 (4), 1237–1258.
- Hanushek, Eric A., John F. Kain, Steven G. Rivkin, and Gregory F. Branch, “Charter School Quality and Parental Decision Making With School Choice,” *Journal of Public Economics*, 2007, 91 (5/6), 823–848.
- Heckman, James J. and Jeffrey A. Smith, “The Pre-programme Earnings Dip and the De-

terminants of Participation in a Social Programme. Implications for Simple Programme Evaluation Strategies,” *The Economic Journal*, 1999, 109 (457), 313–348.

— and **Yona Rubinstein**, “The Importance of Noncognitive Skills: Lessons from the GED Testing Program,” *The American Economic Review - Papers and Proceedings*, 2001, 91 (2), 145–149.

— , **Jora Stixrud**, and **Sergio Urzua**, “The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior,” *Journal of Labor Economics*, 2006, 24 (3), 411–482.

**Hoxby, Caroline M. and Jonah E. Rockoff**, “The Impact of Charter Schools on Student Achievement,” *Department of Economics, Harvard University, mimeo*, 2004.

— and **Sonali Murarka**, “Charter Schools in New York City: Who Enrolls and How They Affect Their Students’ Achievement,” *NBER*, 2008.

**Kyriazidou, Ekaterini**, “Estimation of a Panel Data Sample Selection Model,” *Econometrica*, 1997, 65 (6), 1335–1364.

**McClure, Larry, Betsy Strick, Rachel Jacob-Almeida, and Christopher Reicher**, “The Preuss School at UCSD: School Characteristics and Students’ Achievement,” *CREATE-UCSD*, 2005.

**Sass, Tim R.**, “Charter Schools and Student Achievement in Florida,” *Education Finance and Policy*, 2006, 1 (1), 123–138.

**Solmon, Lewis and Pete Goldschmidt**, “Comparison of Traditional Public Schools and Charter Schools on Retention, School Switching, and Achievement Growth,” policy report, Goldwater Institute 2004.

— , **Kern Paark, and David Garcia**, “Does Charter School Attendance Improve Test Scores? The Arizona Results,” occasional report, Goldwater Institute 2001.

**Todd, Petra E. and Kenneth I. Wolpin**, “The Production of Cognitive Achievement in Children: Home, School and Racial Test Score Gaps,” *University of Pennsylvania, PIER Working Paper*, 2004.

**Zimmer, Ron and Richard Buddin**, “Academic Outcomes,” in “Charter School Operations and Performance,” RAND, 2003, pp. 37–62.

— **and** — , “Charter School Performance in Two Large Urban Districts,” *Journal of Urban Economics*, September 2006, *60* (2), 307–326.

— , **Suzanne Blanc, Brian Gill, and Jolley Christman**, “Evaluating the Performance of Philadelphia’s Charter Schools,” *Rand Working Papers*, 2008.

## Appendix

*Proof of Expected Value of Level and Value-Added Fixed Effects Estimates Bounding the Lagged-Dependent Variable Model with Fixed Effects*

Let us first simplify notation and denote  $\mathbf{X}$  as a  $k \times nt$  vector of demeaned covariates while  $\mathbf{Y}$  is a  $1 \times nt$  vector of the demeaned student outcome variable and  $\mathbf{Y}_{t-1}$  is the  $1 \times nt$  vector of demeaned once-lagged outcome variables. Our true model becomes

$$(6) \quad \mathbf{Y}_t = \mathbf{X}\beta + \mathbf{Y}_{t-1}\gamma + \epsilon$$

In a levels framework, the lagged outcomes enter into the error term such that we have composite error

$$(7) \quad \mu = \gamma\mathbf{Y}_{t-1} + \epsilon.$$

This provides us with

$$(8) \quad \mathbf{E}(\hat{\beta}^L) = \beta + \gamma[\mathbf{X}'\mathbf{X}]^{-1}[\mathbf{X}'\mathbf{E}(\mathbf{Y}_{t-1})]$$

For a value added model we subtract  $\mathbf{Y}_{t-1}$  from each side of (6) to get

$$(9) \quad \mathbf{Y}_t - \mathbf{Y}_{t-1} = \mathbf{X}\beta + (\gamma - \mathbf{1})\mathbf{Y}_{t-1} + \epsilon$$

which will provide us with an estimate of  $\beta$  such that

$$(10) \quad \mathbf{E}(\hat{\beta}^{VA}) = \beta + (\gamma - \mathbf{1})[\mathbf{X}'\mathbf{X}]^{-1}[\mathbf{X}'\mathbf{E}(\mathbf{Y}_{t-1})]$$

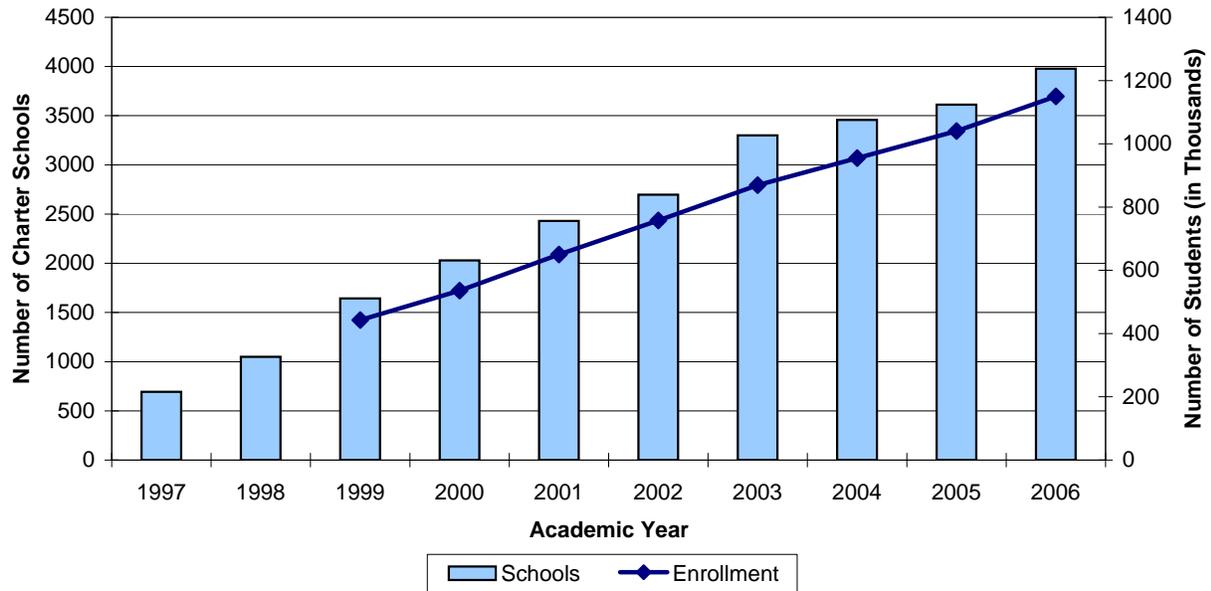
Let us further define the matrix  $\mathbf{A} = [\mathbf{X}'\mathbf{X}]^{-1}[\mathbf{X}'\mathbf{E}(\mathbf{Y}_{t-1})]$  and the  $k^{th}$  row of  $\mathbf{A}$  as  $\mathbf{A}_k$ , hence

$$(11) \quad \mathbf{E}(\hat{\beta}_k^L) = \beta_k + \gamma \mathbf{A}_k$$

$$(12) \quad \mathbf{E}(\hat{\beta}_k^{VA}) = \beta_k + (\gamma - 1) \mathbf{A}_k$$

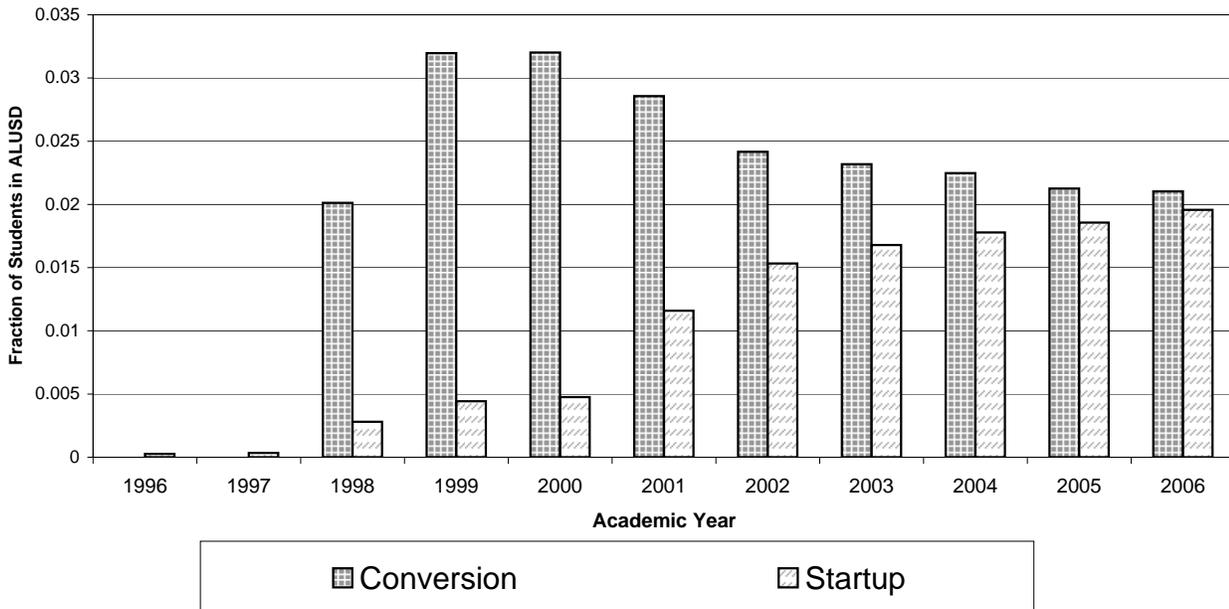
Thus, assuming that  $0 \leq \gamma \leq 1$ , if  $\mathbf{A}_k > \mathbf{0}$  then  $\mathbf{E}(\hat{\beta}_k^L) > \beta > \mathbf{E}(\hat{\beta}_k^{VA})$  while if  $\mathbf{A}_k < \mathbf{0}$  then  $\mathbf{E}(\hat{\beta}_k^L) < \beta < \mathbf{E}(\hat{\beta}_k^{VA})$ . In either case, the levels model and value added models bound  $\beta$ .

**Figure 1: Charter Growth In the US**



Sources: 1997 - 1998, US Dept. of Education National Charter School Reports. 1999 - 2003, US Dept. of Education Common Core of Data. 2005, National Alliance for Public Charter Schools. 2006, Center for Education Reform. 2004 data are unavailable so a linear interpolation is provided

**Figure 2 - ALUSD-SW Charter Enrollment by Year (Grades 1 - 12)**



**Figure 3: Transitions Between School Types 1998 - 2005, Grades 1 - 11**

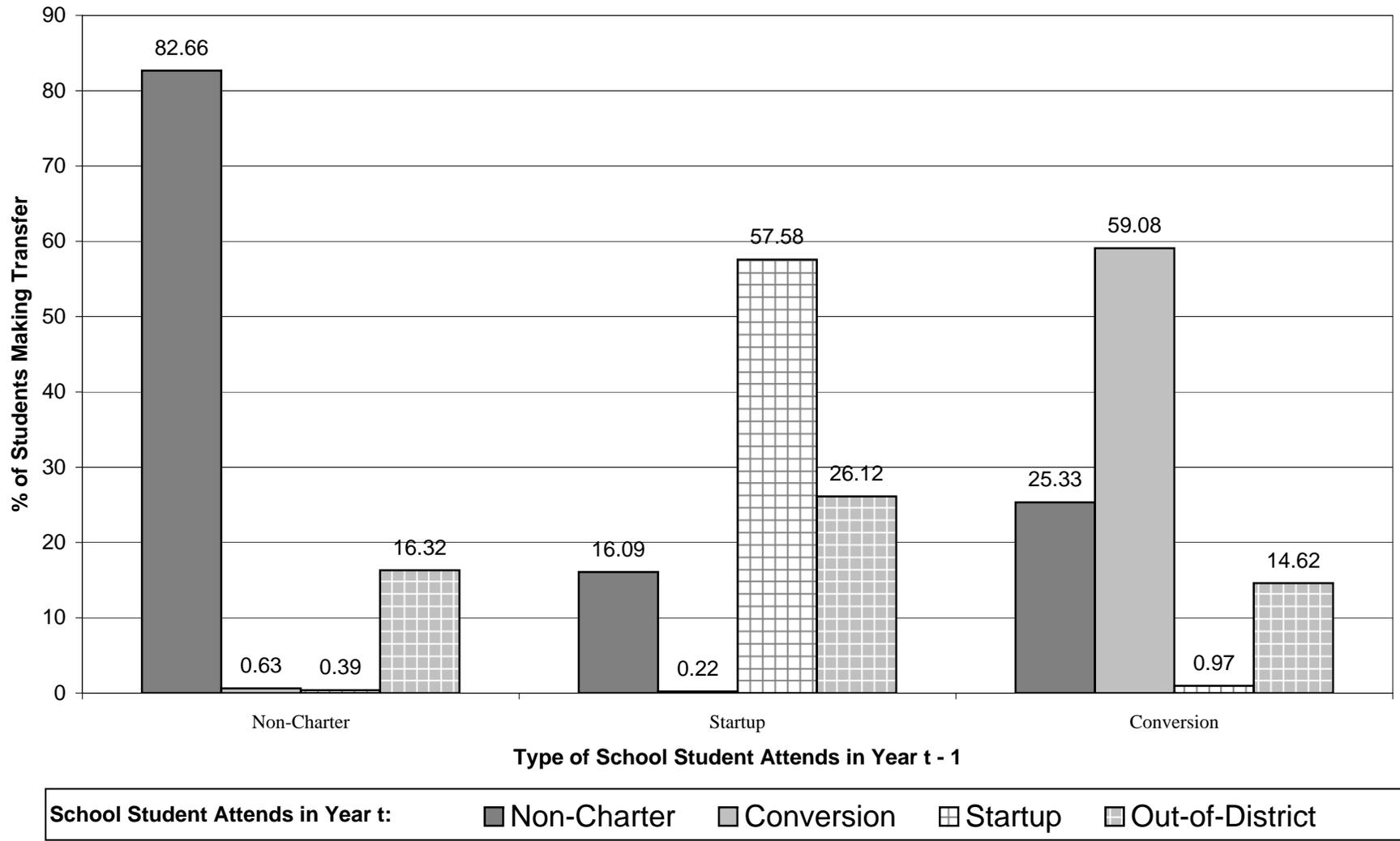
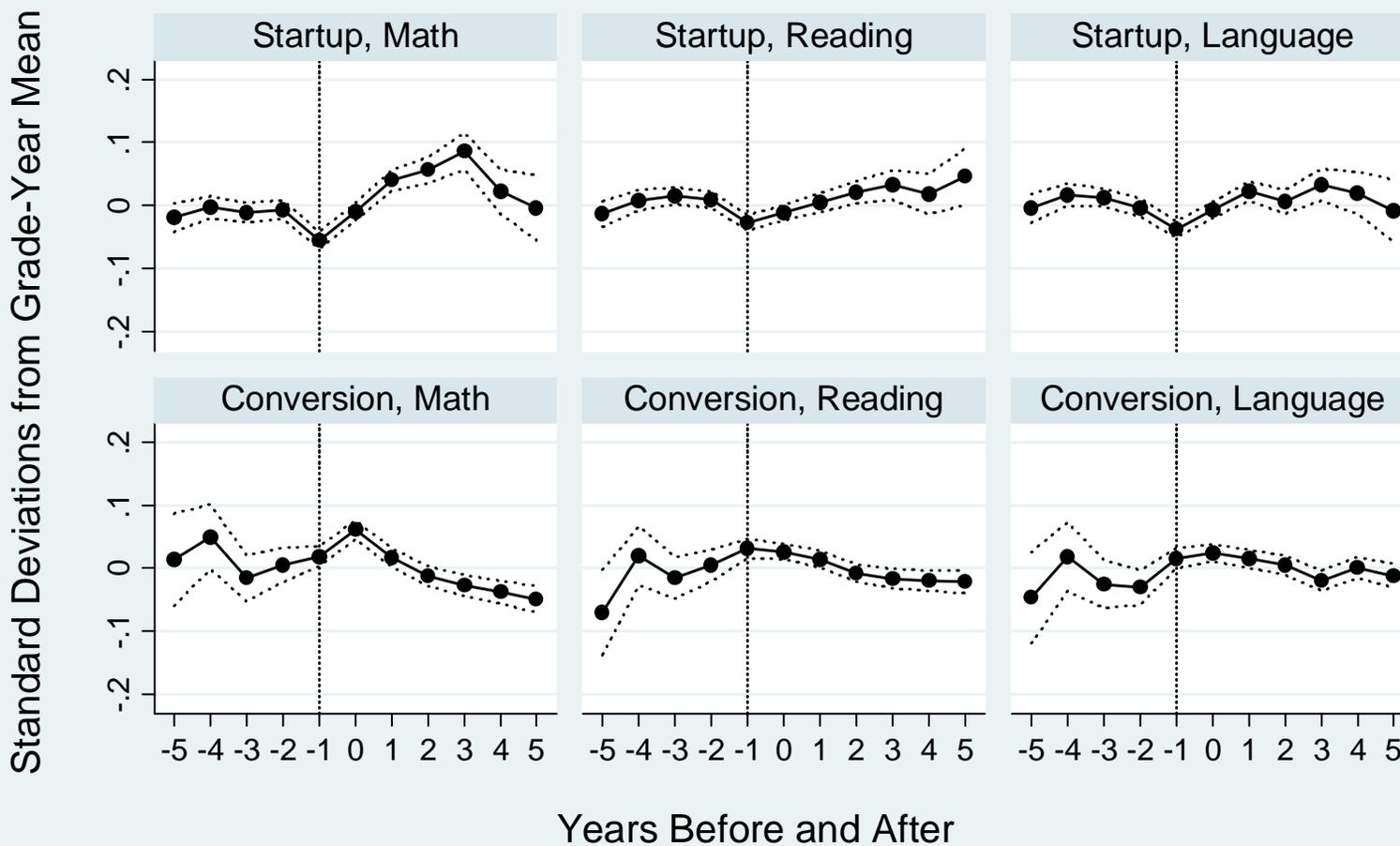
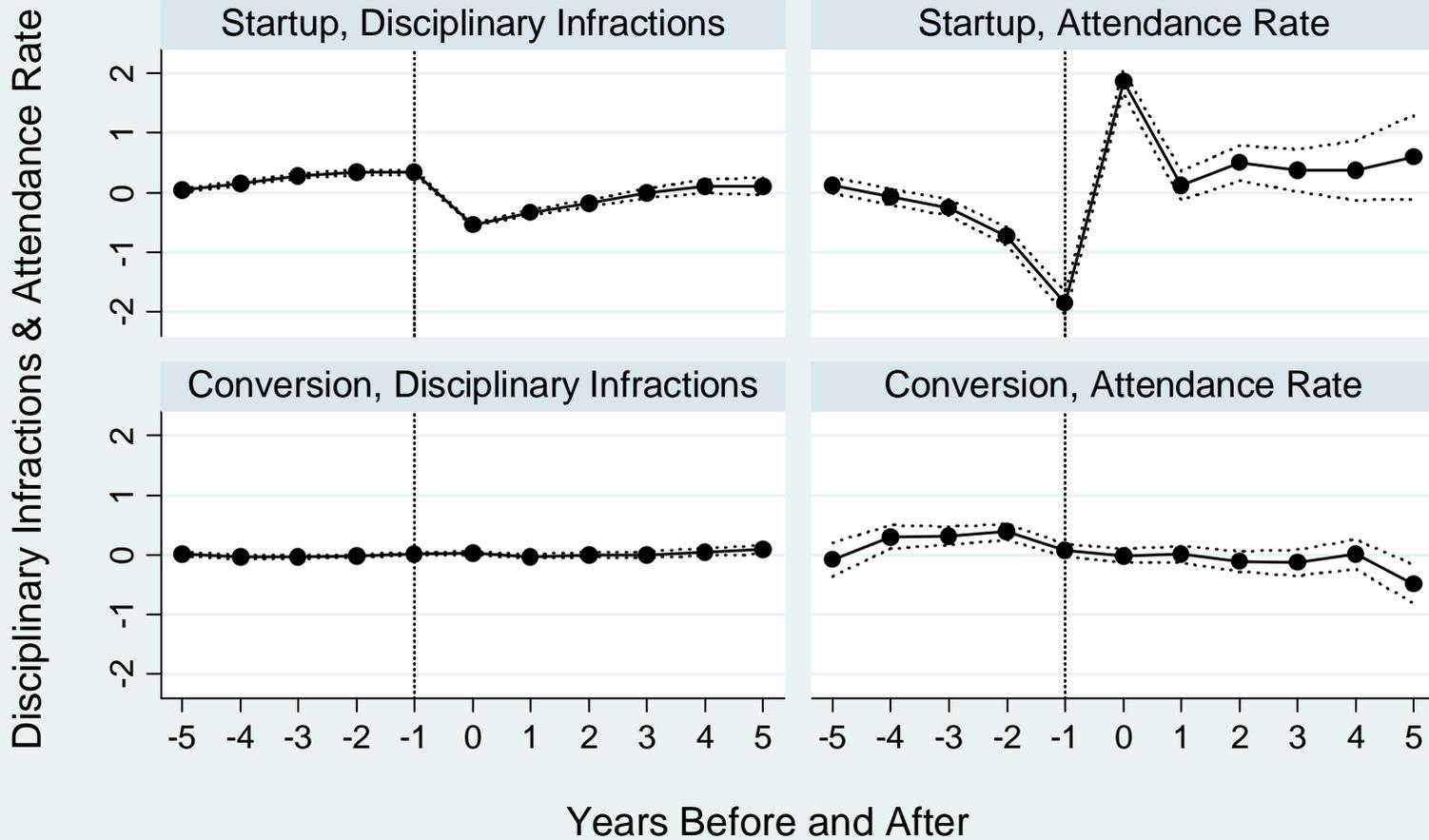


Figure 4: Test Scores Before and After Charter Entry



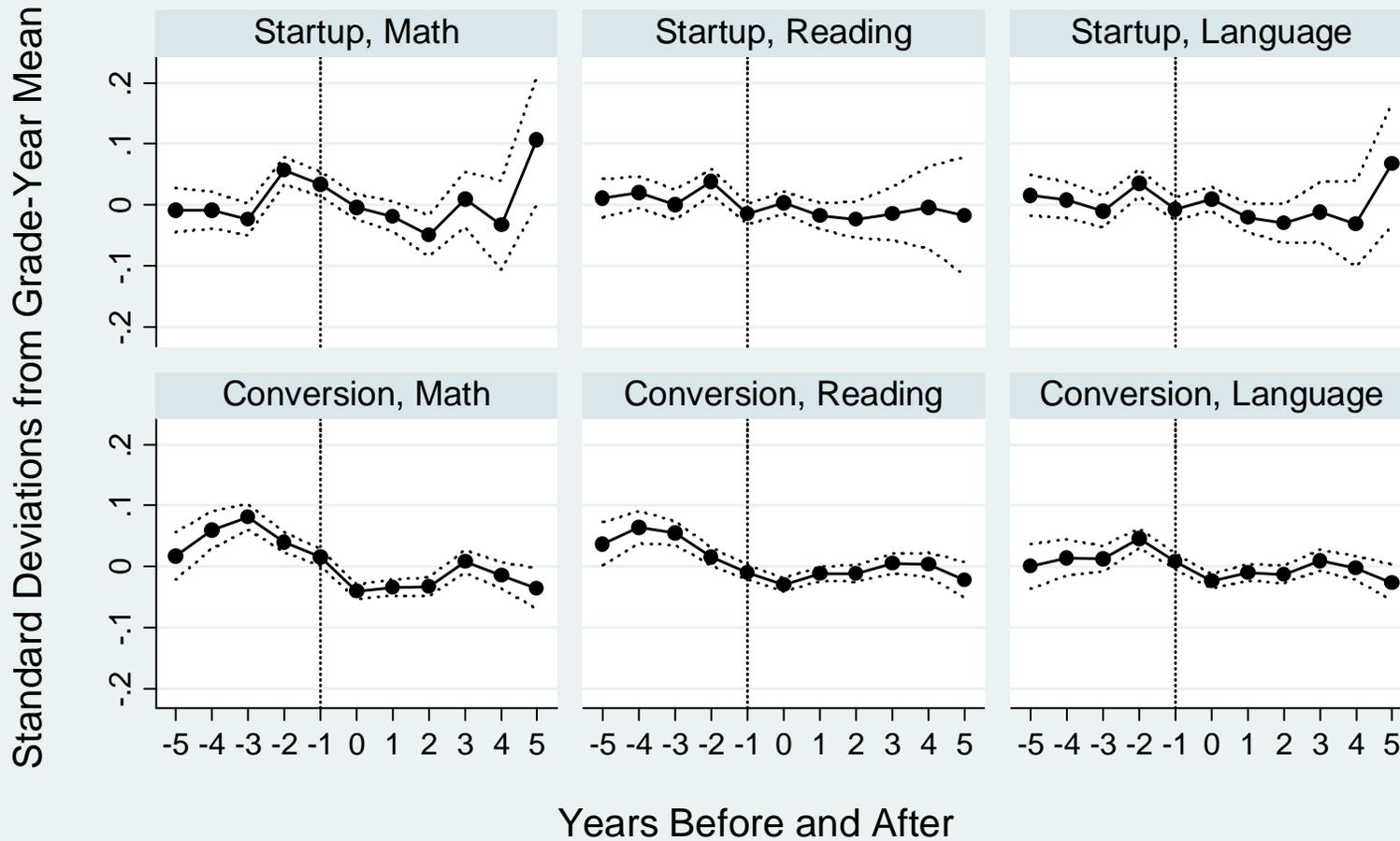
Graph shows residuals from a regression of test scores on observables included in the baseline model excluding charter status. All variables are demeaned within students to remove the individual fixed effect.

Figure 5: Discipline & Attendance by Before and After Charter Entry



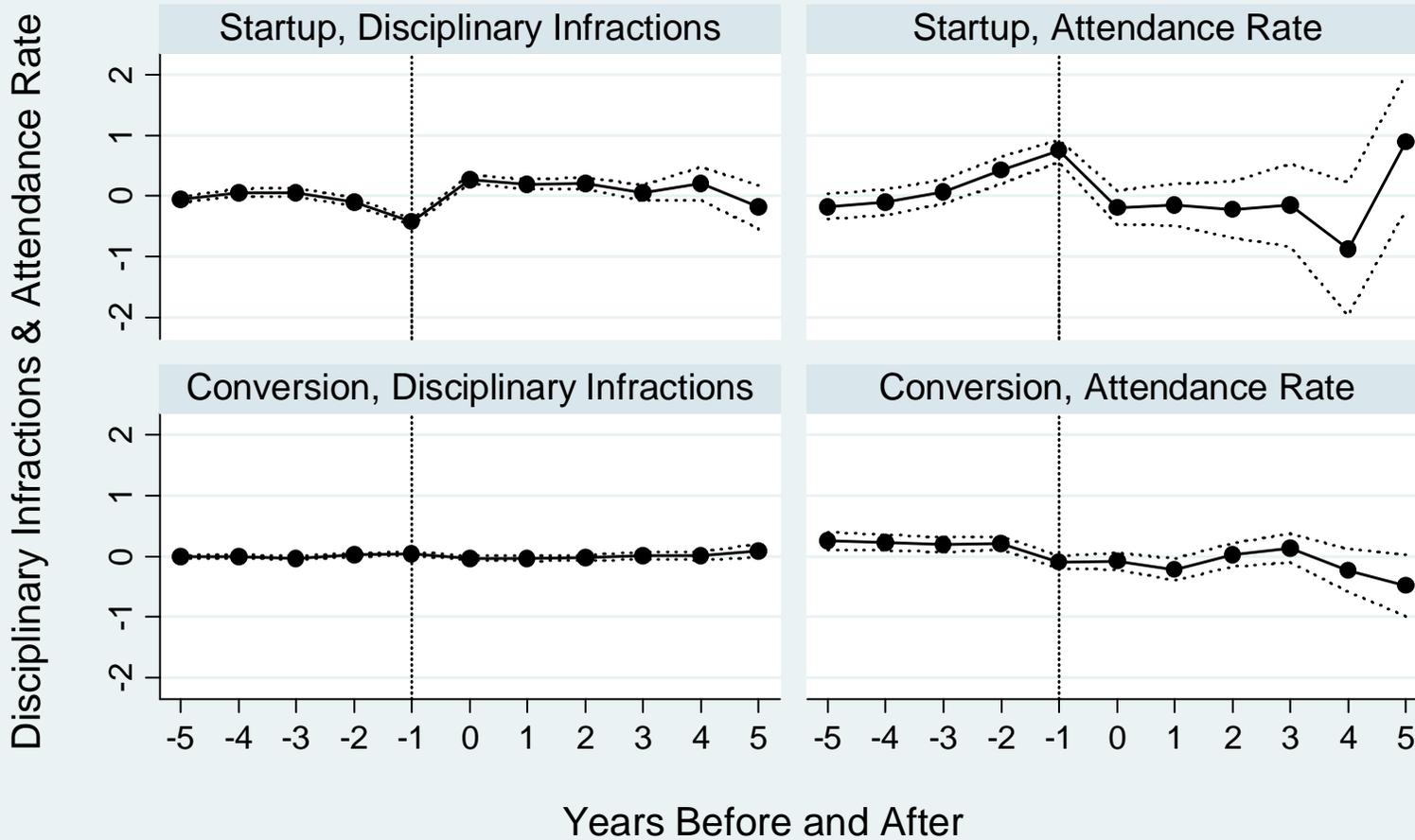
Graph shows residuals from a regression of test scores on observables included in the baseline model excluding charter status. All variables are demeaned within students to remove the individual fixed effect.

Figure 6: Test Scores Before and After Charter Exit



Graph shows residuals from a regression of test scores on observables included in the baseline model excluding charter status. All variables are demeaned within students to remove the individual fixed effect.

Figure 7: Discipline & Attendance by Before and After Charter Exit



Graph shows residuals from a regression of test scores on observables included in the baseline model excluding charter status. All variables are demeaned within students to remove the individual fixed effect.

Table 1: Summary Statistics By Charter Status

| Demographics (Student Level Observations) |                |                |                | Outcomes (Student Level Observations)              |                  |                  |                  |
|---|----------------|----------------|----------------|--|------------------|------------------|------------------|
| Variable                                  | Non-Charter    | Startup        | Conversion     | Variable   | Non-Charter      | Startup          | Conversion       |
| Female                                    | 0.49<br>(0.50) | 0.47<br>(0.50) | 0.48<br>(0.50) | Math   | -0.03<br>(0.97)  | -0.02<br>(0.93)  | -0.11<br>(0.95)  |
| White                                     | 0.10<br>(0.30) | 0.03<br>(0.16) | 0.01<br>(0.11) | Reading  | -0.03<br>(0.98)  | -0.06<br>(0.90)  | -0.11<br>(0.92)  |
| Black                                     | 0.33<br>(0.47) | 0.31<br>(0.46) | 0.50<br>(0.50) | Language   | -0.02<br>(0.98)  | -0.04<br>(0.91)  | -0.08<br>(0.94)  |
| Hispanic                                  | 0.54<br>(0.50) | 0.66<br>(0.47) | 0.48<br>(0.50) | Disciplinary<br>Infractions                        | 0.44<br>(1.37)   | 0.16<br>(0.73)   | 0.35<br>(1.29)   |
| Grade Level                               | 5.85<br>(3.33) | 6.53<br>(3.50) | 3.74<br>(2.07) | Substance Abuse<br>Infractions                     | 0.0074<br>(0.13) | 0.0073<br>(0.12) | 0.0007<br>(0.03) |
| Free Lunch                                | 0.60<br>(0.49) | 0.58<br>(0.49) | 0.79<br>(0.41) | Violent Crime<br>Infractions                       | 0.0047<br>(0.09) | 0.0024<br>(0.07) | 0.0019<br>(0.06) |
| Reduced Price<br>Lunch                    | 0.07<br>(0.25) | 0.11<br>(0.31) | 0.08<br>(0.27) | Fighting<br>Infractions                            | 0.0359<br>(0.22) | 0.0163<br>(0.15) | 0.0498<br>(0.27) |
| Other Economic<br>Disadvantage            | 0.05<br>(0.23) | 0.11<br>(0.32) | 0.07<br>(0.25) | Attendance Rate                                    | 94.4<br>(8.89)   | 93.2<br>(11.98)  | 95.7<br>(5.97)   |
| Limited English<br>Proficiency            | 0.25<br>(0.43) | 0.18<br>(0.39) | 0.29<br>(0.46) | School Characteristics (School Level Observations) |                  |                  |                  |
| At Risk Status                            | 0.58<br>(0.49) | 0.65<br>(0.48) | 0.59<br>(0.49) | Instructional<br>Spending                          | 4054<br>(5668)   | 5511<br>(5341)   | 3688<br>(743)    |
| Special Education                         | 0.11<br>(0.32) | 0.05<br>(0.21) | 0.11<br>(0.32) | Administrative<br>Spending                         | 519<br>(1670)    | 164<br>(350)     | 455<br>(158)     |
| Gifted and<br>Talented                    | 0.10<br>(0.30) | 0.01<br>(0.12) | 0.05<br>(0.23) | Other Spending                                     | 936<br>(1363)    | 254<br>(322)     | 918<br>(389)     |
| Immigrant                                 | 0.13<br>(0.33) | 0.13<br>(0.34) | 0.09<br>(0.29) | Enrollment   | 739<br>(509)     | 225<br>(263)     | 599<br>(277)     |
|   |                | [0.2]          | [1.7]          |  |                  | [7.9]            | [1.7]            |

Standard deviations in parentheses. Absolute T-statistics from a regression of the variable on startup and conversion status in brackets. Test scores are standard deviation units from scale scores normalized within grade and year. Spending figures are per-pupil. Demographics, attendance, and general infractions include 2.45 million student-year observations. Due to limited years of availability substance abuse and violence infractions include 1.79 million observations while fighting includes approximately 700,000 observations. Test scores include 1.2 million observations. School characteristics include approximately 3900 school-year observations.

Table 2 - Effect of Attending a Charter School on Test Scores

|                                      |                                     |            | A. Levels           |                  |                     |           | B. Value-Added     |                   |                    |         |
|--------------------------------------|-------------------------------------|------------|---------------------|------------------|---------------------|-----------|--------------------|-------------------|--------------------|---------|
|                                      |                                     |            | Math                | Reading          | Language            | Obs       | Math               | Reading           | Language           | Obs     |
|                                      |                                     |            | (1)                 | (2)              | (3)                 |           | (4)                | (5)               | (6)                |         |
| OLS with Student<br>Fixed Effects    | Baseline                            | Startup    | 0.072<br>(0.052)    | 0.015<br>(0.038) | 0.022<br>(0.033)    | 1,141,480 | 0.026<br>(0.025)   | -0.033<br>(0.024) | -0.007<br>(0.023)  | 779,343 |
|                                      |                                     | Conversion | 0.070**<br>(0.033)  | 0.029<br>(0.026) | 0.041*<br>(0.024)   |           | 0.026<br>(0.034)   | -0.019<br>(0.021) | 0.014<br>(0.021)   |         |
| Interrupted Panel                    | Drop Year Prior to<br>Charter       | Startup    | 0.060<br>(0.055)    | 0.010<br>(0.041) | 0.013<br>(0.036)    | 1,135,119 | 0.062<br>(0.051)   | 0.002<br>(0.038)  | 0.015<br>(0.031)   | 772,568 |
|                                      |                                     | Conversion | 0.081**<br>(0.039)  | 0.040<br>(0.032) | 0.048*<br>(0.028)   |           | 0.070*<br>(0.041)  | 0.030<br>(0.035)  | 0.042<br>(0.027)   |         |
|                                      | Drop 2 Years Prior<br>to Charter    | Startup    | 0.056<br>(0.054)    | 0.009<br>(0.041) | 0.006<br>(0.036)    | 1,131,123 | 0.041<br>(0.051)   | -0.008<br>(0.038) | 0.001<br>(0.032)   | 767,916 |
|                                      |                                     | Conversion | 0.084**<br>(0.041)  | 0.042<br>(0.033) | 0.043<br>(0.029)    |           | 0.070*<br>(0.042)  | 0.029<br>(0.036)  | 0.038<br>(0.029)   |         |
| Kyriazidou (1997)<br>Attrition Model | First - Differences<br>(Unweighted) | Startup    | 0.049<br>(0.040)    | 0.003<br>(0.028) | 0.011<br>(0.022)    | 618,050   | 0.059<br>(0.036)   | -0.009<br>(0.033) | 0.006<br>(0.024)   | 526,203 |
|                                      |                                     | Conversion | 0.073***<br>(0.022) | 0.029<br>(0.022) | 0.054***<br>(0.017) |           | 0.063<br>(0.043)   | 0.001<br>(0.028)  | 0.028<br>(0.022)   |         |
|                                      | MSE Minimizing<br>Bandwidth         | Startup    | 0.053<br>(0.039)    | 0.004<br>(0.028) | 0.012<br>(0.022)    | 618,050   | 0.075**<br>(0.036) | -0.005<br>(0.034) | 0.007<br>(0.026)   | 526,203 |
|                                      |                                     | Conversion | 0.071***<br>(0.022) | 0.028<br>(0.022) | 0.054***<br>(0.017) |           | 0.051<br>(0.041)   | -0.008<br>(0.028) | 0.031<br>(0.021)   |         |
|                                      | 1/4*MSE<br>Minimizing<br>Bandwidth  | Startup    | 0.065<br>(0.042)    | 0.005<br>(0.026) | 0.018<br>(0.024)    | 618,050   | 0.100**<br>(0.046) | 0.014<br>(0.032)  | 0.043<br>(0.037)   | 526,203 |
|                                      |                                     | Conversion | 0.044*<br>(0.023)   | 0.018<br>(0.023) | 0.063***<br>(0.021) |           | 0.015<br>(0.044)   | -0.037<br>(0.037) | 0.066**<br>(0.031) |         |

All regressions include an individual fixed-effect, free lunch status, reduced-price lunch status, other economic disadvantage, recent immigration status, whether a parent is a migrant worker, and grade-by-year indicators. Standard errors clustered by school in parentheses. \*, \*\*, \*\*\* denote significance at the 10%, 5%, and 1% levels, respectively. Levels models cover 1998 - 2006 and grades 1 - 11. Value-added models cover 1999 - 2006 and grades 2 - 11. Because students are defined to be in the sample in the first grade and year of each sample, these observations are dropped from the Kyriazidou attrition analysis.

Table 3 - Effect of Attending a Charter School on Discipline, Attendance & Retention

A. Disciplinary Infrations and Attendance

|                                      |                                     |            | i. Levels                   |                     |           | ii. Value-Added             |                     |           |
|--------------------------------------|-------------------------------------|------------|-----------------------------|---------------------|-----------|-----------------------------|---------------------|-----------|
|                                      |                                     |            | Disciplinary<br>Infractions | Attendance          | Obs       | Disciplinary<br>Infractions | Attendance          | Obs       |
|                                      |                                     |            | (1)                         | (2)                 |           | (3)                         | (4)                 |           |
| OLS with Student<br>Fixed Effects    | Baseline                            | Startup    | -0.795***<br>(0.113)        | 2.268***<br>(0.607) | 2,233,050 | -0.498***<br>(0.151)        | 2.338***<br>(0.757) | 1,777,994 |
|                                      |                                     | Conversion | 0.023<br>(0.070)            | -0.044<br>(0.186)   |           | 0.005<br>(0.058)            | 0.047<br>(0.167)    |           |
| Interrupted Panel                    | Drop Year Prior to<br>Charter       | Startup    | -0.789***<br>(0.101)        | 2.149***<br>(0.545) | 2,221,517 | -0.789***<br>(0.085)        | 1.852***<br>(0.364) | 1,758,480 |
|                                      |                                     | Conversion | 0.028<br>(0.074)            | -0.045<br>(0.182)   |           | 0.024<br>(0.088)            | 0.199<br>(0.174)    |           |
|                                      | Drop 2 Years Prior<br>to Charter    | Startup    | -0.775***<br>(0.088)        | 2.058***<br>(0.491) | 2,211,570 | -0.797***<br>(0.084)        | 1.873***<br>(0.367) | 1,756,982 |
|                                      |                                     | Conversion | 0.026<br>(0.078)            | 0.025<br>(0.177)    |           | 0.024<br>(0.089)            | 0.181<br>(0.178)    |           |
| Kyriazidou (1997)<br>Attrition Model | First - Differences<br>(Unweighted) | Startup    | -0.809***<br>(0.171)        | 2.467**<br>(1.027)  | 1,457,716 | -0.849***<br>(0.208)        | 3.071***<br>(1.129) | 1,067,566 |
|                                      |                                     | Conversion | 0.029<br>(0.073)            | 0.017<br>(0.194)    |           | 0.034<br>(0.094)            | 0.032<br>(0.171)    |           |
|                                      | MSE Minimizing<br>Bandwidth         | Startup    | -0.818***<br>(0.174)        | 2.500**<br>(1.015)  | 1,457,716 | -0.842***<br>(0.216)        | 3.242***<br>(1.203) | 1,067,566 |
|                                      |                                     | Conversion | 0.034<br>(0.075)            | 0.017<br>(0.194)    |           | 0.062<br>(0.102)            | 0.061<br>(0.172)    |           |
|                                      | 1/4*MSE<br>Minimizing<br>Bandwidth  | Startup    | -0.879***<br>(0.190)        | 3.125**<br>(1.218)  | 1,457,716 | -0.865***<br>(0.302)        | 4.721***<br>(1.646) | 1,067,566 |
|                                      |                                     | Conversion | 0.109<br>(0.095)            | 0.013<br>(0.218)    |           | 0.271*<br>(0.161)           | 0.143<br>(0.212)    |           |

B. Additional Outcomes (Levels Only)

|         | Any<br>Infraction    | Expulsion              | Alternative<br>Education<br>Placement | Retention           |
|---------|----------------------|------------------------|---------------------------------------|---------------------|
| Startup | -0.250***<br>(0.036) | -0.0036***<br>(0.0013) | -0.020***<br>(0.004)                  | 0.0024*<br>(0.0014) |
| Convert | 0.002<br>(0.015)     | -0.0016**<br>(0.0007)  | -0.005**<br>(0.002)                   | -0.0001<br>(0.0001) |
| Obs     | (0.156)              | 1,740,282              | 1,740,282                             | 1,777,994           |

All regressions include an individual fixed-effect, free lunch status, reduced-price lunch status, other economic disadvantage, recent immigration status, whether a parent is a migrant worker, and grade-by-year indicators. Standard errors clustered by school in parentheses. \*, \*\*, \*\*\* denote significance at the 10%, 5%, and 1% levels, respectively. Levels models cover 1994 - 2006 and grades 1 - 12 and value-added models 1994 - 2006 and grades 2 - 12 for discipline and attendance. Because students are defined to be in the sample in the first grade and year of each sample, these observations are dropped from the Kyriazidou attrition analysis. Additional outcomes cover all grades 1 - 12 but retention, expulsion, and AEP placement are not available in all years.

Table 4: Charter Impacts by Grade Level and Age of Charter

|                                       | Math<br>(1)         | Reading<br>(2)      | Language<br>(3)     | Discipline<br>(4)    | Attendance<br>(5)    |
|---------------------------------------|---------------------|---------------------|---------------------|----------------------|----------------------|
| A. Impacts by Grade Level             |                     |                     |                     |                      |                      |
| i. Levels                             |                     |                     |                     |                      |                      |
| Startup                               | -0.010<br>(0.042)   | -0.005<br>(0.033)   | -0.011<br>(0.035)   | -0.182***<br>(0.035) | -0.029<br>(0.103)    |
| Startup * Grades 6 - 8                | 0.192***<br>(0.056) | 0.070**<br>(0.034)  | 0.087***<br>(0.028) | -0.751***<br>(0.099) | 1.636***<br>(0.235)  |
| Startup * Grades 9 - 12               | -0.011<br>(0.103)   | -0.034<br>(0.102)   | -0.019<br>(0.084)   | -0.734***<br>(0.156) | 3.591***<br>(0.617)  |
| Observations                          | 1,141,480           | 1,141,480           | 1,141,480           | 2,233,050            | 2,233,050            |
| ii. Value-Added                       |                     |                     |                     |                      |                      |
| Startup                               | -0.011<br>(0.056)   | -0.058<br>(0.036)   | -0.045<br>(0.038)   | -0.333***<br>(0.102) | 0.231<br>(0.224)     |
| Startup * Grades 6 - 8                | 0.082<br>(0.061)    | 0.063<br>(0.050)    | 0.056<br>(0.049)    | -0.344**<br>(0.152)  | 0.999***<br>(0.240)  |
| Startup * Grades 9 - 12               | -0.013<br>(0.074)   | -0.017<br>(0.065)   | 0.032<br>(0.050)    | -0.068<br>(0.212)    | 3.503***<br>(0.854)  |
| Observations                          | 779,343             | 779,343             | 779,343             | 1,777,994            | 1,777,994            |
| B. Impacts by Age of Charter (Levels) |                     |                     |                     |                      |                      |
| Startup                               | 0.046<br>(0.045)    | 0.040<br>(0.044)    | 0.031<br>(0.027)    | -0.653***<br>(0.098) | 5.941***<br>(2.161)  |
| Startup * 2 Years                     | 0.027<br>(0.055)    | -0.027<br>(0.056)   | 0.026<br>(0.050)    | 0.032<br>(0.072)     | -4.167***<br>(1.405) |
| Startup * 3 Years                     | -0.021<br>(0.065)   | -0.118**<br>(0.052) | -0.054*<br>(0.030)  | -0.038<br>(0.109)    | -4.229**<br>(1.922)  |
| Startup * 4 Years                     | 0.139<br>(0.088)    | 0.004<br>(0.048)    | 0.030<br>(0.031)    | -0.205*<br>(0.111)   | -2.547*<br>(1.302)   |
| Startup * 5+ Years                    | 0.010<br>(0.062)    | -0.010<br>(0.063)   | -0.019<br>(0.047)   | -0.269***<br>(0.083) | -4.842*<br>(2.757)   |
| Convert                               | 0.093<br>(0.071)    | 0.038<br>(0.034)    | 0.017<br>(0.043)    | 0.033<br>(0.062)     | -0.016<br>(0.208)    |
| Convert * 2 Years                     | -0.008<br>(0.031)   | 0.017<br>(0.023)    | 0.017<br>(0.016)    | -0.131*<br>(0.076)   | 0.134<br>(0.226)     |
| Convert * 3 Years                     | 0.056<br>(0.083)    | 0.039<br>(0.051)    | 0.100<br>(0.062)    | 0.037<br>(0.068)     | -0.192<br>(0.258)    |
| Convert * 4 Years                     | 0.030<br>(0.127)    | 0.047<br>(0.106)    | 0.064<br>(0.070)    | -0.048<br>(0.033)    | -0.041<br>(0.201)    |
| Convert * 5+ Years                    | -0.126<br>(0.097)   | -0.089*<br>(0.050)  | -0.022<br>(0.057)   | 0.060<br>(0.078)     | -0.146<br>(0.244)    |
| Observations                          | 1,141,480           | 1,141,480           | 1,141,480           | 2,233,050            | 2,233,050            |

Each column in each panel is a separate regression. All regressions include an individual fixed-effect, free lunch status, reduced-price lunch status, other economic disadvantage, recent immigration status, whether a parent is a migrant worker, and grade-by-year indicators. Standard errors clustered by school in parentheses. \*, \*\*, \*\*\* denote significance at the 10%, 5%, and 1% levels, respectively. Levels models cover 1998 - 2006 and grades 1 - 11 for test scores and 1994 - 2006 and grades 1 - 12 for other outcomes. Value-added models cover 1999 - 2006 and grade 2 - 11 for test scores and 1994 - 2006 and grades 2 - 12 for other outcomes. Columns in each panel show separate regressions. Grade-level regressions also contain an indicator for conversion charters, but since only one conversion covers grades 7 - 8 and no conversions cover grades 9 - 12, I do not separate that estimate by grade level. Value-added results for age of charter are similar to levels results and are available in the online appendix.

Table 5: Startup Charter Impacts by Student Characteristics

|                                      | Math<br>(1)         | Reading<br>(2)       | Language<br>(3)      | Discipline<br>(4)    | Attendance<br>(5)   |
|--------------------------------------|---------------------|----------------------|----------------------|----------------------|---------------------|
| A. Levels                            |                     |                      |                      |                      |                     |
| Startup                              | 0.286<br>(0.185)    | 0.086<br>(0.086)     | -0.024<br>(0.074)    | -0.948***<br>(0.120) | 2.601***<br>(0.956) |
| Startup * Black                      | -0.238<br>(0.166)   | -0.087<br>(0.076)    | 0.034<br>(0.058)     | 0.228*<br>(0.119)    | -1.094<br>(0.740)   |
| Startup * Hispanic                   | -0.305**<br>(0.154) | -0.133***<br>(0.045) | -0.006<br>(0.027)    | -0.098<br>(0.093)    | 0.434<br>(0.467)    |
| Startup * Female                     | 0.056**<br>(0.027)  | 0.005<br>(0.022)     | 0.013<br>(0.019)     | 0.238***<br>(0.074)  | 0.055<br>(0.244)    |
| Startup * Economically Disadvantaged | -0.028<br>(0.025)   | -0.024*<br>(0.013)   | -0.020<br>(0.017)    | -0.039<br>(0.043)    | -0.090<br>(0.268)   |
| Startup * Immigrant                  | 0.034<br>(0.030)    | 0.030<br>(0.019)     | 0.018<br>(0.025)     | 0.078**<br>(0.033)   | -1.358**<br>(0.565) |
| Observations                         | 1,141,480           | 1,141,480            | 1,141,480            | 2,233,050            | 2,233,050           |
| B. Value-Added                       |                     |                      |                      |                      |                     |
| Startup                              | 0.027<br>(0.061)    | -0.102*<br>(0.058)   | -0.158***<br>(0.048) | -0.745***<br>(0.219) | 3.381**<br>(1.404)  |
| Startup * Black                      | 0.054<br>(0.061)    | 0.093<br>(0.060)     | 0.196***<br>(0.044)  | 0.075<br>(0.161)     | -1.908<br>(1.506)   |
| Startup * Hispanic                   | -0.024<br>(0.038)   | 0.011<br>(0.056)     | 0.141***<br>(0.044)  | -0.052<br>(0.072)    | -0.060<br>(1.150)   |
| Startup * Female                     | 0.034*<br>(0.019)   | -0.034*<br>(0.019)   | -0.042*<br>(0.024)   | 0.381**<br>(0.151)   | 0.329<br>(0.429)    |
| Startup * Economically Disadvantaged | -0.041<br>(0.029)   | 0.014<br>(0.040)     | -0.003<br>(0.022)    | -0.071<br>(0.093)    | 0.021<br>(0.540)    |
| Startup * Immigrant                  | 0.053*<br>(0.027)   | -0.034<br>(0.027)    | -0.033<br>(0.026)    | 0.058<br>(0.042)     | -0.409**<br>(0.163) |
| Observations                         | 779,343             | 779,343              | 779,343              | 1,777,994            | 1,777,994           |

Each column in each panel is a separate regression. All regressions include an individual fixed-effect, free lunch status, reduced-price lunch status, other economic disadvantage, recent immigration status, whether a parent is a migrant worker, and grade-by-year indicators. Regressions also include the main effect and the same set of interactions for conversion charters along with startup and conversion interactions with other non-white students. These students make up less than 1% of startup and 4% of conversion charter students. Full results are available in the online appendix. Standard errors clustered by school in parentheses. \*, \*\*, \*\*\* denote significance at the 10%, 5%, and 1% levels, respectively. Levels models cover 1998 - 2006 and grades 1 - 11 for test scores and 1994 - 2006 and grades 1 - 12 for other outcomes. Value-added models cover 1999 - 2006 and grade 2 - 11 for test scores and 1994 - 2006 and grades 2 - 12 for other outcomes.

Table 6 - Mechanisms of Discipline and Attendance Impacts in Startup Charters

|     |  |              | A. Levels                |                     | B. Value-Added           |                     |
|-----|--|--------------|--------------------------|---------------------|--------------------------|---------------------|
|     |  |              | Disciplinary Infractions | Attendance          | Disciplinary Infractions | Attendance          |
| (1) | Baseline (from tables 2, 3)  | Coefficient  | -0.795***<br>(0.113)     | 2.268***<br>(0.607) | -0.498***<br>(0.151)     | 2.338***<br>(0.757) |
|     |  | Observations | 2,233,050                | 2,233,050           | 1,777,994                | 1,777,994           |
| (2) | Controlling for Per-Student Instructional Expenditures                             | Coefficient  | -0.799***<br>(0.121)     | 2.318***<br>(0.582) | -0.536***<br>(0.139)     | 2.406***<br>(0.791) |
|     |  | Observations | 2,221,995                | 2,221,995           | 1,769,243                | 1,769,243           |
| (3) | Controlling Per-Student School Leadership Expenditures                             | Coefficient  | -0.852***<br>(0.131)     | 1.908***<br>(0.583) | -0.586***<br>(0.183)     | 2.276***<br>(0.759) |
|     |  | Observations | 2,223,425                | 2,223,425           | 1,770,359                | 1,770,359           |
| (4) | Controlling Per-Student Other Expenditures   | Coefficient  | -0.849***<br>(0.129)     | 1.716***<br>(0.613) | -0.520**<br>(0.205)      | 2.073***<br>(0.756) |
|     |  | Observations | 2,221,995                | 2,221,995           | 1,769,243                | 1,769,243           |
| (5) | Controlling for Per-Student Instructional, School Leadership, & Other Expenditures | Coefficient  | -0.755***<br>(0.115)     | 1.733***<br>(0.641) | -0.354**<br>(0.144)      | 1.823**<br>(0.733)  |
|     |  | Observations | 2,221,995                | 2,221,995           | 1,769,243                | 1,769,243           |
| (6) | Controlling for Student Body Composition   | Coefficient  | -0.798***<br>(0.113)     | 2.343***<br>(0.582) | -0.341*<br>(0.184)       | 2.099***<br>(0.680) |
|     |  | Observations | 2,232,700                | 2,232,700           | 1,777,666                | 1,777,666           |
| (7) | Controlling for Enrollment   | Coefficient  | -0.550***<br>(0.129)     | 2.899***<br>(0.706) | -0.125<br>(0.200)        | 2.277***<br>(0.837) |
|     |  | Observations | 2,232,700                | 2,232,700           | 1,777,666                | 1,777,666           |

Each coefficient is from a separate regression. All regressions include an individual fixed-effect, free lunch status, reduced-price lunch status, other economic disadvantage, recent immigration status, whether a parent is a migrant worker, and grade-by-year indicators. Regressions also include a conversion charter estimate that is not shown here. Additional controls for potential charter mechanisms are added as described above. Each row/column combination is a separate regression with both conversion and startup charter estimates. Full results are provided in the online appendix. Standard errors clustered by school in parentheses. \*, \*\*, \*\*\* denote significance at the 10%, 5%, and 1% levels, respectively.

Table 7: Charter Impacts by Time in Charter

| A. Levels             |                    |                    |                    |                      |                     |
|-----------------------|--------------------|--------------------|--------------------|----------------------|---------------------|
|                       | Math               | Reading            | Language           | Discipline           | Attendance          |
|                       | (1)                | (2)                | (3)                | (4)                  | (5)                 |
| Startup - 1 Year      | 0.013<br>(0.060)   | -0.016<br>(0.045)  | -0.004<br>(0.042)  | -0.837***<br>(0.115) | 3.270***<br>(1.009) |
| Startup - 2 Years     | 0.008<br>(0.078)   | -0.062<br>(0.075)  | -0.008<br>(0.059)  | -1.042***<br>(0.200) | 2.197***<br>(0.412) |
| Startup - 3 Years     | -0.114<br>(0.098)  | -0.128<br>(0.106)  | -0.122<br>(0.082)  | -0.981***<br>(0.168) | 3.756***<br>(0.660) |
| Startup - 4+ Years    | 0.151<br>(0.270)   | -0.122<br>(0.150)  | -0.042<br>(0.130)  | -1.577***<br>(0.354) | 11.774<br>(10.629)  |
| Conversion - 1 Year   | 0.095*<br>(0.050)  | 0.056*<br>(0.029)  | 0.064*<br>(0.039)  | 0.006<br>(0.048)     | 0.029<br>(0.179)    |
| Conversion - 2 Years  | -0.013<br>(0.138)  | -0.074<br>(0.100)  | -0.007<br>(0.091)  | 0.075<br>(0.082)     | -0.175<br>(0.360)   |
| Conversion - 3 Years  | 0.062<br>(0.139)   | 0.031<br>(0.093)   | -0.020<br>(0.061)  | -0.182<br>(0.111)    | 0.564<br>(0.421)    |
| Conversion - 4+ Years | 0.019<br>(0.137)   | 0.000<br>(0.142)   | 0.095<br>(0.086)   | -0.207*<br>(0.118)   | 0.984*<br>(0.535)   |
| Observations          | 1,141,480          | 1,141,480          | 1,141,480          | 2,233,050            | 2,233,050           |
| B. Value-Added        |                    |                    |                    |                      |                     |
| Startup - 1 Year      | 0.053<br>(0.052)   | -0.006<br>(0.039)  | 0.046<br>(0.047)   | -0.792***<br>(0.222) | 3.800***<br>(1.408) |
| Startup - 2 Years     | -0.005<br>(0.068)  | -0.084*<br>(0.050) | 0.017<br>(0.048)   | -0.226*<br>(0.122)   | 1.700***<br>(0.442) |
| Startup - 3 Years     | 0.159<br>(0.163)   | 0.103<br>(0.112)   | 0.121<br>(0.089)   | -0.232<br>(0.359)    | 3.645***<br>(0.636) |
| Startup - 4+ Years    | 0.446**<br>(0.183) | 0.202*<br>(0.120)  | 0.375**<br>(0.148) | -0.035<br>(0.731)    | 7.327<br>(5.546)    |
| Conversion - 1 Year   | 0.069<br>(0.108)   | -0.021<br>(0.075)  | 0.053<br>(0.055)   | 0.004<br>(0.063)     | 0.148<br>(0.203)    |
| Conversion - 2 Years  | -0.163<br>(0.155)  | -0.179<br>(0.112)  | 0.018<br>(0.087)   | -0.186**<br>(0.093)  | 0.477<br>(0.329)    |
| Conversion - 3 Years  | 0.335**<br>(0.165) | 0.165<br>(0.139)   | 0.253<br>(0.223)   | -0.208**<br>(0.097)  | 0.740<br>(0.501)    |
| Conversion - 4+ Years | 0.012<br>(0.332)   | 0.001<br>(0.218)   | 0.255<br>(0.171)   | -0.150<br>(0.094)    | 0.992<br>(0.604)    |
| Observations          | 779,343            | 779,343            | 779,343            | 1,777,994            | 1,777,994           |

Each column in each panel is a separate regression. All regressions include an individual fixed-effect, free lunch status, reduced-price lunch status, other economic disadvantage, recent immigration status, whether a parent is a migrant worker, and grade-by-year indicators. In 2SLS models, all post first-year periods are instrumented by the potential number of years the student could have been in the charter school based on date of birth. Standard errors clustered by school in parentheses. \*, \*\*, \*\*\* denote significance at the 10%, 5%, and 1% levels, respectively. Levels models cover 1998 - 2006 and grades 1 - 11 for test scores and 1994 - 2006 and grades 1 - 12 for other outcomes. Value-added models cover 1999 - 2006 and grade 2 - 11 for test scores and 1994 - 2006 and grades 2 - 12 for other outcomes.

Table 8: Persistence of Charter Impacts

| A. Levels Models                        |                                 |                                  |                                 |                                    |                                    |
|---|---------------------------------|----------------------------------|---------------------------------|------------------------------------|------------------------------------|
|   | Math<br>(1)                     | Reading<br>(2)                   | Language<br>(3)                 | Discipline<br>(4)                  | Attendance<br>(5)                  |
| <b>Startup</b>                          | <b>0.134*</b><br><b>(0.071)</b> | <b>0.025</b><br><b>(0.054)</b>   | <b>0.043</b><br><b>(0.046)</b>  | <b>-0.749***</b><br><b>(0.125)</b> | <b>2.502***</b><br><b>(0.711)</b>  |
| 1 Year Post-Startup                     | 0.199***<br>(0.065)             | 0.038<br>(0.048)                 | 0.096**<br>(0.045)              | 0.100<br>(0.173)                   | 0.601<br>(0.555)                   |
| 4+ Years Post-Startup                   | 0.194<br>(0.172)                | -0.002<br>(0.167)                | -0.006<br>(0.171)               | -0.142<br>(0.166)                  | -0.060<br>(1.235)                  |
| <b>Conversion</b>                       | <b>-0.062</b><br><b>(0.044)</b> | <b>-0.071*</b><br><b>(0.043)</b> | <b>0.001</b><br><b>(0.038)</b>  | <b>0.093</b><br><b>(0.121)</b>     | <b>-0.745**</b><br><b>(0.290)</b>  |
| 1 Years Post-Conversion                 | -0.203***<br>(0.063)            | -0.169**<br>(0.070)              | -0.074<br>(0.057)               | 0.045<br>(0.189)                   | -1.103**<br>(0.438)                |
| 4+ Years Post-Conversion                | -0.180**<br>(0.070)             | -0.133*<br>(0.071)               | -0.069<br>(0.062)               | 0.031<br>(0.163)                   | -1.279***<br>(0.439)               |
| F-Test of All Post-Startup Estimates    | 3.48***                         | 0.92                             | 2.05*                           | 2.52**                             | 0.79                               |
| F-Test of All Post-Conversion Estimates | 3.14**                          | 1.58                             | 1.13                            | 1.23                               | 4.24***                            |
| Observations                            | 1,141,480                       | 1,141,480                        | 1,141,480                       | 2,233,050                          | 2,233,050                          |
| B. Value-Added Models                   |                                 |                                  |                                 |                                    |                                    |
| <b>Startup</b>                          | <b>0.011</b><br><b>(0.036)</b>  | <b>-0.065*</b><br><b>(0.035)</b> | <b>0.004</b><br><b>(0.033)</b>  | <b>-0.375**</b><br><b>(0.150)</b>  | <b>2.453***</b><br><b>(0.867)</b>  |
| 1 Year Post-Startup                     | -0.027<br>(0.074)               | -0.068<br>(0.047)                | 0.075<br>(0.054)                | 1.034***<br>(0.245)                | -0.954*<br>(0.505)                 |
| 4+ Years Post-Startup                   | -0.045<br>(0.214)               | -0.356**<br>(0.168)              | -0.068<br>(0.200)               | -0.065<br>(0.206)                  | 0.627<br>(0.905)                   |
| <b>Conversion</b>                       | <b>-0.122</b><br><b>(0.122)</b> | <b>-0.201</b><br><b>(0.130)</b>  | <b>-0.082</b><br><b>(0.070)</b> | <b>0.107</b><br><b>(0.124)</b>     | <b>-0.714***</b><br><b>(0.242)</b> |
| 1 Years Post-Conversion                 | -0.229<br>(0.175)               | -0.255<br>(0.190)                | -0.138<br>(0.098)               | 0.127<br>(0.279)                   | -1.440***<br>(0.380)               |
| 4+ Years Post-Conversion                | -0.183<br>(0.145)               | -0.250<br>(0.157)                | -0.141*<br>(0.077)              | 0.045<br>(0.128)                   | -1.123***<br>(0.316)               |
| F-Test of All Post-Startup Estimates    | 1.25                            | 1.74                             | 0.84                            | 7.02***                            | 2.62**                             |
| F-Test of All Post-Conversion Estimates | 0.51                            | 1.20                             | 1.49                            | 1.73                               | 5.31***                            |
| Observations                            | 779,343                         | 779,343                          | 779,343                         | 1,777,994                          | 1,777,994                          |

Each column in each panel is a separate regression. Two and three year post-charter estimates are similar to the one and four-year estimates so I do not show them here, though the full set of results are provided in the online appendix. All regressions include an individual fixed-effect, free lunch status, reduced-price lunch status, other economic disadvantage, recent immigration status, whether a parent is a migrant worker, and grade-by-year indicators. All post charter periods are instrumented by whether the student is the listed number of years beyond the last grade level covered by his or her prior charter school. Standard errors clustered by school in parentheses. \*, \*\*, \*\*\* denote significance at the 10%, 5%, and 1% levels, respectively. Levels models cover 1998 - 2006 and grades 1 - 11 for test scores and 1994 - 2006 and grades 1 - 12 for other outcomes. Value-added models cover 1999 - 2006 and grade 2 - 11 for test scores and 1994 - 2006 and grades 2 - 12 for other outcomes.