The Impact of Providing Vision Screening and Free Eyeglasses on Academic Outcomes: Evidence from a Randomized Trial in Title 1 Elementary Schools

Paul Glewwe*, Kristine West**, Jongwook Lee*

September 2014

Abstract

More than 20% of all school aged children in the United States have vision problems, and low-income and minority children are even more likely to have unmet vision care needs. We use a randomized control trial to evaluate the impact of enhanced vision services provided by a local non-profit organization to Title 1 elementary schools in three large central Florida school districts. That organization provides state-of-the-art screening, comprehensive vision exams and free eyeglasses for low-income children. We find that providing additional/enhanced screening alone is generally insufficient to improve student achievement in math and reading, yet in two of the three counties studied providing free vision exams and eyeglasses significantly improved student achievement in math and reading in grade 5 (but for the most part not in grade 4). The magnitude of the impact ranges from 0.07 to 0.16 standard deviations of the distribution of students' test scores. The impact on English Language Learner (ELL) students is particularly large, increasing math and readings scores by about 0.15 standard deviations (averaged across grades 4 and 5). Most impacts faded out completely by the second year, but the impact on ELL students' math scores (0.08 standard deviations in year 2) did not fade out. A simple calculation suggests that providing screening and free eyeglasses has benefits that may be two orders of magnitude higher than the costs.

* Department of Applied Economics, University of Minnesota
** Department of Economics, St. Catherine University

We would like to thank James Cooper, Nicole Davis and Tung Le for excellent research assistance, and Meta Brown and other participants at the Association for Education Finance and Policy conference for helpful comments on an earlier draft. We are indebted to the staff members in each of the three districts for welcoming the study into their schools and for gathering and sharing student-level data. Lastly, this study would not have been possible without the excellent and tireless work of Nancy Jeppesen and the entire staff at Florida Vision Quest.

I. Introduction

More than 20% of school aged children in the United States have vision problems (Ethan et al. 2010, Basch 2011, Zaba 2011). Low-income and minority children have a greater than average risk of under-diagnosis and under-treatment of vision problems (Ganz et al. 2006, 2007, Basch 2010). For example, Title 1 students are two to three times more likely than non-Title 1 students to have undetected or untreated vision problems (Johnson et al. 2000).[1] Given this situation, it is surprising that there has been little attempt to quantify the impact of vision interventions on student outcomes such as test scores, attendance rates, and discipline incidents. This study attempts to fill this void.

We evaluate the impact of a randomized control trial in which a non-profit, Florida Vision Quest (FLVQ), offered enhanced vision services to a randomly selected group of Title 1 elementary schools in three large central Florida school districts during the fall of the 2011-12 school-year. FLVQ provides state-of-the-art screening, comprehensive vision exams and free eyeglasses for low-income children in central Florida. Despite a long record of partnership with area school districts, thus far FLVQ's work had been motivated solely by compelling anecdotal evidence; this is the first independent evaluation of FLVQ and the services it offers.[2]

The results of this analysis will be useful to policy makers and practitioners in Florida and elsewhere. There is a long tradition in the United States of public schools providing basic screening for health problems such as hearing and vision impairment. This study investigates whether school districts can improve student outcomes by upgrading their vision screening technologies and/or collaborating with local non-profits such as FLVQ to provide comprehensive

[1] Title 1 is a federally funded program to assist low-income students. The funding is based on the number of students who qualify for a free or reduced-price lunch. "Title 1 schools" are schools where at least 40% of their students are from low-income families. All schools in this study are Title 1 schools and have at least 70% low-income students.
[2] For more information on Florida Vision Quest, see its website: http://www.flvq.org/.

vision exams and free eyeglasses. Vision interventions may be a very cost-effective way to improve student outcomes; if districts can identify and treat vision problems early, they may be able to avoid costly remediation in subsequent years.

The economic theory for such an intervention is simple – identifying and treating vision problems should increase students' acquisition of human capital. If students cannot see, they cannot read (be it their textbooks and/or the writing on the board at the front of the classroom), and if they cannot read they have little hope of keeping pace with the demands of school and will likely underperform relative to their full potential. By identifying and treating vision problems at an early age, students will acquire human capital at a faster rate, which will yield both private and social benefits (see Cunha et al. 2006; Heckman et al., 2006, Lange and Topel, 2006, Grossman, 2006, and Lochner, 2011, for reviews of the private and social benefits of education).

We find that providing additional/enhanced screening alone is generally insufficient to improve student math and reading skills as measured by test scores. However, in two of the three counties studied, the intervention that included provision of free vision exams and eyeglasses significantly improved student achievement in math and reading in grade 5 (but for the most part not in grade 4) as measured by the Florida Comprehensive Achievement Test (FCAT). The magnitude of the impact ranges from 0.07 to 0.16 standard deviations of the distribution of students' test scores. The impact on English Language Learner (ELL) students is in particular large, increasing math and readings scores by about 0.15 standard deviations (averaged across grades 4 and 5). Most impacts faded out completely by the second year, but the impact on ELL students' math scores (0.08 standard deviations in year two) did not fade out.

The rest of this paper is organized as follows. The next section reviews the relevant literature, after which the design and implementation of the experiment are explained. Section IV describes the data and the estimation method, and the following section presents the results. The final section concludes.

II. Literature Review

This study investigates whether diagnosing and providing treatment to students with poor vision enables them to acquire human capital at a faster rate than would occur in the absence of diagnosis and/or the offer of treatment. We contribute to three strands of the existing literature.

First, as noted above, there is some evidence on the prevalence of undiagnosed and untreated vision problems among school aged children in the U.S. (Ethan et al 2010, Basch 2011, Zaba 2011). Common refractive errors such as myopia (nearsightedness), hyperopia (farsightedness) and astigmatisms can be corrected with eyeglasses, but many children either do not know that they have problems, do not have glasses, or do not wear their glasses. Research also shows that the rates of undiagnosed and untreated vision problems vary by race and ethnicity (Kleinstein et al 2003) and socio-economic status (Ganz et al 2006, 2007). Yet this evidence is sparse and incomplete. We contribute new data to this literature.

Second, there is some evidence linking vision problems and academic outcomes. This literature is largely correlational in nature. Correlations, unfortunately, cannot yield a clear answer as to whether providing screening and/or treatment to students who need glasses will improve their academic performance. Walline and Johnson Carder (2012) find that U.S. children who receive special education services have significantly higher rates of myopia, hyperopia and astigmatism than the general population. Gomes-Neto et al. (1997) find that primary school aged children in Brazil with vision problems are 10 percentage points more likely to drop out, 18 percentage points more likely to repeat a grade, and score 0.2 to 0.3 standard deviations (of the distribution of test scores) lower on achievement tests relative to their peers without vision problems. Both of these are examples of studies that compare students with vision problems to those without vision problems. These studies do not necessarily imply that treating vision problems will improve educational outcomes for students with vision problems because those students may be fundamentally different from students without vision problems in some unobserved way. If a third variable is causing both the poor vision and the low academic performance – for example, low birth weight could lead to vision problems and learning disabilities – then correcting the vision impairment may do little to improve academic outcomes.

Even among students whose vision problems are detected, there may be unobservable differences between those who actually go on to wear glasses and those who do not follow up with any treatment. Hannum and Zhang (2012) use a propensity score matching design and find that, for 13-16 year old children in China with vision problems, wearing glasses increases math and literacy scores by 0.26 and 0.34 standard deviations, respectively. They cannot, however, address self-selection into wearing glasses and they find that, among all students, wearing glasses is positively correlated with socio-economic status and overall academic achievement.

Sonne-Schmidt (2011) applies a more rigorous methodology, regression discontinuity, to estimate the impact of eyeglasses on middle school and high school students in the United States. Using data from the National Health Examination Survey of Youth (NHES III, 1966-70), he finds that wearing glasses increases test scores by approximately 0.1 standard deviations (of the distribution of test scores) across all eyeglasses wearers and by as much as 0.3 standard deviations for students with myopia. Yet this methodology focuses on students whose vision problems are not particularly severe, those on the borderline of being diagnosed with a vision problem, for which the impact of providing eyeglasses may be relatively small; this would underestimate the impact of providing eyeglasses to children with more severe vision problems.

It is common for authors of reports that call for better screening and better access to treatment to make impassioned claims without strong evidence that these interventions will improve student outcomes. For example, Zaba (2011) claims that vision interventions will increase both academic outcomes and eventual labor market outcomes, but these assertions are not backed by rigorous research. Instead he simply states: "*Is there any doubt that children must have the vision care and vision skills required in order to perform successfully in school and workplace environments?"* It may seem intuitive that providing vision screenings, eye exams and eyeglasses will improve student outcomes, but startlingly little research supports this claim.

The gold standard for identifying the causal impact of an intervention is a randomized control trial. We know of only one other such study, Glewwe et al. (2014). They conduct a study in Gansu Province, China, in 2004-05 in which students in grades 4-6 in a randomly selected group of townships were provided vision exams and, if needed, eyeglasses. They find that wearing glasses increases average test scores for students with poor vision by 0.15 to 0.22 standard deviations. They also find that not all students accepted the glasses, therefore the impact of being diagnosed with a vision problem is lower. We contribute the first evidence from a randomized control trial conducted in a developed country, the United States.

Perhaps the study closest to ours is Kimel (2006). She documents an intervention in Rockford, Illinois that provided students with two free pairs of glasses if they failed a vision screening. However, she did not assess the ultimate impact on student outcomes. Kimel finds that students who received free eyeglasses were more likely to wear glasses in the months after the intervention than students who were given only the information that they failed the screening.

Finally, we contribute to a third strand of literature which attempts to identify the most cost-effective ways to raise student achievement. We can do only rough back-of-the-envelope calculations, yet even these improve upon the existing literature because, to date, no one has tried to measure the economic return to providing vision screenings and/or glasses in a developed country.[3] One calculation compares the impact of FLVQ's services to a different use of school district dollars, reducing class sizes, and another is an overall comparison of costs and benefits

III. Experimental Design and Implementation

A. Experimental Design. To rigorously evaluate the impact of FLVQ's provision of vision exams and eyeglasses, we conducted a randomized control trial targeting 4th and 5th grade students in Title 1 elementary schools in central Florida. The randomization was done at the school level. Schools were randomly assigned to one of two intervention groups or to a control group. The benefit of the randomized design is that it provides a valid counterfactual. That is, it allows for an estimate of the impact of an intervention by comparing two groups of students who received FLVQ's interventions to another group that provides estimates of what would have happened in the absence of the interventions.

FLVQ uses a state-of-the-art screening tool called the "Spot," which is manufactured by PediaVision.[4] The Spot, essentially an infrared camera, uses auto-refraction and video-retinoscopy technologies to screen for refractive errors, amblyopic precursors and pupil abnormalities. The person doing the screening stands about one meter away from the student and

[3] Glewwe et al. (2014) provide such calculations for China. We are aware of only one other attempt to conduct any cost-benefit analysis, but this was not done within the context of an educational intervention. White (2004) focuses on amblyopia (lazy eye) and subsequent blindness, conducting a cost-benefit analysis using QALYs (quality-adjusted life years). This is quite different from our benefit measure.

[4] For more information on the Spot technology, see the PediaVision website: http://www.spotvisionscreening.com.

takes a digital photograph of the student's eyes. The information acquired yields an automatic assessment of a student's vision. With the Spot, screening is very accurate, and very quick.

Using the Spot, FLVQ screened all 4th grade and 5th grade students in the intervention schools. No screening was done for 4th and 5th grade students in the control schools. For the first of the two groups of intervention schools, which we refer to as the "screen-only" schools; this was the only service provided. Students who failed the screening were sent home with a note (in English or Spanish) for parents indicating that they should follow-up with a doctor of their choosing. For the other group of intervention schools, which we call "full-treatment" schools, students who failed the screening were offered comprehensive vision exams aboard the FLVQ mobile vision clinic, which is a bus that has been equipped with all the tools usually available in an optometrist's office, and is staffed with licensed eye care professionals. If the onboard optometrist prescribed glasses, FLVQ provided two pairs of glasses to the student at no charge.

FLVQ did not have sufficient resources to screen all of the Title 1 elementary schools in the three school districts, nor did they have sufficient resources to provide follow-up exams and glasses to all of the schools that were screened. Rather than have FLVQ choose which schools to serve, we persuaded them to randomize the choice, thus using the resource constraint as an opportunity to provide a rigorous evaluation of the two levels of services provided by FLVQ.

There are two main mechanisms through which these interventions may affect student outcomes. First, perhaps there is an "information problem." That is, there may be students (parents) who do not know that they (their children) have vision problems. In Florida, students are routinely screened for vision problems in Kindergarten, 1st grade, 3rd grade, and 6th grade. This intervention targets 4th and 5th graders, thus adding an extra screening. Also, the Spot may be a more effective screening tool than standard technology. FLVQ claims that the commonly used Snellen chart may miss over 60% of vision problems (Indian River Study, 2010). Both the fact that students are screened in grades that schools do not usually screen and the fact that the screening is done with an arguably superior technology should identify students who are missed by the district's standard screenings. If students in the screen-only schools outperform students in the control group, this suggests that simply providing information will increase student learning.

Second, if the main barrier is not identifying vision problems, but rather obtaining glasses, then the screen-only intervention will be insufficient. If the real issue is an "access problem" there will be no difference between students in the screen-only schools and students in the control group. Students (parents) may know about vision problems but lack the resources needed to obtain access to a doctor and/or eyeglasses to remedy those problems. In the full-treatment schools any student who is identified with vision problems is offered a vision exam and two pairs of eyeglasses, all free of charge. Not only are the exam and eyeglasses free, they are brought to the students at their schools. There is no need for students (parents) to invest any resources other than the time and effort needed to return a permission form, and to use and care for the glasses. The difference, if any, between the screen-only and the full-treatment schools will isolate the importance of providing onsite follow-up exams and free eyeglasses.

B. Implementation. Three Florida school districts agreed to participate in the study: we will refer to them as County 1, County 2 and County 3.[5] Only the Title 1 elementary schools in each district were eligible to participate. County 1 has 46 elementary schools, of which 37 are Title 1 schools. The respective numbers are 24 and 16 for County 2 and 125 and 65 for County 3.

Randomization. In each school district, we ranked the Title 1 schools by their students' academic proficiency. Specifically, we used the average of each school's "points" over the

[5] In Florida, each county is a separate school district. At the request of the three districts, we do not use their names.

preceding three years. The "points" measure, designed by the Florida Department of Education, includes pass rates as well as gains on the state-mandated Florida Comprehensive Achievement Tests (FCATs). The schools were then grouped (stratified) by academic proficiency and the randomization was conducted within these groups (strata). This stratification provides additional assurance that the treatment and control groups will have comparable levels of academic proficiency prior to the intervention.

There were eleven strata in County 1, five in County 2 and seven in County 3. The number of strata was determined by the number of schools that FLVQ estimated their resources would cover. One full treatment school and one screen only school were selected from each strata. The remaining schools serve as the control. Strata ranged in size from three to six schools. Because the strata contain different numbers of schools, and because schools differ in size, the final sample has between 268 and 1,069 students in each strata. The average is 657. Controls (dummy variables) for each strata are included in the analysis.

County 1 had 37 Title 1 elementary schools in 2010-11. We randomly assigned 11 to the screen-only group, 11 to the full-treatment group and 15 to the control group. After doing this, we learned that one school assigned to the full-treatment group was not part of the district but rather a charter school, so it was dropped, leaving ten full-treatment schools. We later learned that FLVQ had a significant presence in 2009-10 at two of the full-treatment schools. These two schools are kept in the analysis but the results are robust to excluding these two schools.

In County 2 there were 16 Title 1 elementary schools in 2010-11. Prior to randomization, we learned that FLVQ had a significant presence at one school in 2009-10, so this school was manually placed in the full-treatment group along with five randomly chosen schools, for a total of six full-treatment schools.[6] Five schools were assigned to the screen-only group. One school in the screen-only group refused to participate, but the data from this school are included in the study so the screen-only results for County 2 should be interpreted as intent-to-treat (ITT) effects.[7] The remaining five schools were placed in the control group.

County 3 had 65 Title 1 elementary schools in 2010-11. FLVQ had a significant presence in 28 of them over the prior two years, so our analysis uses only the remaining 37. We assigned seven to the screen-only group, seven to the full-treatment group and 23 to the control group. The control group was much larger than the treatment groups because funding constraints limited the number of schools that FLVQ could serve.

Table 1 provides information on the number of schools and students in the treatment and control groups, by county. We show numbers of students both for the full sample and for the final sample that is used for the regression analysis. Students were excluded from the final sample only if data were missing for one or more variables used in the regression analysis. The variable that was missing most frequently was the 2011 FCAT score (i.e. pre-intervention test scores) which primarily reflects students moving between districts. Student mobility also led to incomplete matching of names on any two lists of students compiled at different times for the same school.

Balance tests are presented in Table 2; they provide evidence that the randomization was successful. Students in full-treatment, screen-only, and control schools appear similar on all observable characteristics. The only differences that are statistically significant are that screen-

[6] The logic for placing this school in the full-treatment group is that it would provide a compromised control because students and parents at the school already knew of FLVQ and the services it offers. Another option would be to drop this school from the study. Results are robust to dropping this school from the sample.
[7] ITT is the most appropriate estimate for policy makers who may want to replicate this intervention. Offering the treatment is the intervention. Part of offering the treatment is the possibility that some schools may decline.

only schools in County 1 have slightly fewer multi-race students than other schools in County 1 and full-treatment schools in County 2 have slightly more girls than other schools in County 2. Given the number of hypotheses tests (74), finding one difference that is significant at the 1% level and another that is significant at the 5% level is consistent with the null hypothesis of no differences between the three groups of schools. Overall, we conclude that, within each county, the randomization created three groups for which there are no systematic differences of any kind.

Delivery of Vision Services. FLVQ arranged screening dates with each school in the two intervention groups during the fall of 2011. County 1 schools were screened in September and, in the full-treatment schools, exams were conducted a few weeks later, with glasses distributed to students in October and November. County 2 schools were screened in October, and glasses were distributed in November and December. County 3 schools were screened in November and December, and glasses were distributed in December, January and February.[8]

Table 3 summarizes the results of the screenings and, where applicable, the follow-up exams. As seen in the top panel, in the full-treatments schools, we have screening data for 80% of students. Some students were not screened because they were absent on the day of the screening (about 4 percentage points). Matching problems due to differences in spelling of names on any two lists of students collected by two different organizations further reduced the screening rate. The bottom panel of Table 3 reports similar results when the sample is limited to the students who constitute the final sample. Screening data are more common in the final sample due to mobility; students with high mobility are both more likely not to be screened and more likely not to have the 2011 FCAT score and thus to be excluded from the final sample.

---

[8] One student in County 2 did not receive glasses until February. In County 3, one student did not receive glasses until March and two students did not receive glasses until April.

Note that students who have glasses with them at the time of the screening are instructed to wear their glasses during the screening, and thus are not identified as having untreated vision problems (unless their glasses are inadequate). In a few cases, FLVQ gave students who already had glasses (and passed the screening with those glasses) a new pair. Specifically, there were 55 students who passed the vision screening but were issued new glasses by FLVQ. Results are robust to dropping these students from the analysis.

In the full-treatment schools, students who failed the screening were offered a comprehensive vision exam aboard the mobile clinic, and 75% of those identified with untreated vision problems (914 out of 1,222) were seen by an optometrist.[9] This rate is less than 100% because students must complete a parent permission form to see the optometrist and they must be present on the day(s) the mobile clinic is scheduled for their school. Almost all of the students seen aboard the mobile clinic were given glasses.[10] In the end, 16% of students in the full-treatment schools (784 out of 4,806) were provided glasses by FLVQ. Among the 1,222 children in the full-treatment schools who failed the screening, 64% (784) were provided eyeglasses.

There were three major problems with the implementation in County 1, which was the first of the three counties in which the program was implemented. First, the notices that were sent home to the screen-only group indicated that they would be receiving the full treatment. That is, they mistakenly indicated that free eyeglasses would be provided. Parents were then informed of this error via the district's automated phone messaging system. Second, vision screening data from the screen-only group was not recorded properly. As indicated in Table 3,

---

[9] There were some students who did not fail the screening but were still seen aboard the mobile clinic. This happened for two reasons: (1) the Spot device was unable to get an accurate reading; and (2) in County 1 a problem interpreting the Spot results caused some students to be incorrectly passed/failed (see next paragraph).

[10] The exceptions were mostly in County 1 where the error misidentified some students. The students who mistakenly failed the screening were often not prescribed glasses (although, interestingly, sometimes they were prescribed glasses which indicates that the threshold for failing the screening may have been different from the threshold that the optometrists used for prescribing glasses.)

there are no data for this group on how many were screened and how many failed the screening. Third, there was a problem interpreting the output of the Spot device at most of the full-treatment schools and some of the screen-only schools. This problem affected the screening results for at least 100 students, and possibly many more. Some students were incorrectly identified as having vision problems (false positive) and some students who had vision problems were missed (false negative). When this issue was discovered, FLVQ gave the schools where this occurred a list of the affected students. FLVQ offered to see all these students aboard the mobile clinic. The majority of the false negatives were not seen aboard the mobile clinic because of insufficient time to collect parent permission forms. Thus many students who otherwise would have been given an exam, and likely glasses, were missed in County 1.

The implementation in Counties 2 and 3 was done after that in County 1, and was much smoother. Due to this variation in implementation, all estimates are shown separately by county.

IV. Data and Estimation Method

A. Data. The data for the study come from three sources: the Spot screenings; the records kept by FLVQ; and administrative records from each school district. We constructed a student-level panel that includes vision data for the intervention schools and demographic, attendance, discipline and test score records for both intervention and control schools. Table 4 provides summary statistics for the students in the final sample.

When a student is screened using the Spot device, it stores detailed readings for each of the student's eyes. We primarily use the summary result that indicates whether a student passed or failed the screening. We also have data on the device's preliminary diagnosis; the most common diagnoses are myopia and astigmatism. After a student is seen aboard the mobile clinic, FLVQ records whether the student is prescribed glasses as well as when the glasses were given to the student. We have no data on students' prior vision services. For instance, we do not know whether they had previously failed vision screenings and/or been prescribed glasses. Also, we have no vision data for students in any of the control schools.

The primary outcomes used to assess the impact of the intervention are reading and math scores on the Florida Comprehensive Achievement Test (FCAT). The FCAT is given in April of each year, near the end of the academic year; thus, for example, the 2011 FCAT occurs near the end of the 2010-11 academic year. We obtained three years of FCAT data (one pre-intervention and two post-intervention) from Counties 1 and 3, and two years of FCAT data (one pre-intervention and one post-intervention) from County 2.[11] High rates of student mobility mean that we do not have pre-intervention FCAT scores for all students. Approximately 10% of students do not have a 2011 test score available. This may bias results if students who transfer districts are more(less) likely to have undiagnosed or untreated vision problems and/or benefit more(less) from the vision intervention than non-transferring students. However, we find little difference in mobility (as measured by who remains in the same district in from 2011-12 to 2012-13) between students who pass and students who fail the screening.

We have FCAT scale scores (a continuous measure) for each student in County 1 and County 3, but not for County 2. [12] For all three counties we have FCAT achievement level scores (a categorical measure), which range from 1 to 5. Level 3 is defined as "demonstrating a satisfactory level of success." Levels 4 and 5 are more than satisfactory and levels 1 and 2 are

[11] Despite an agreement in principle to provide the data, and repeated requests, the County 2 school district has thus far not provided the second year of post-intervention data, citing that they simply did not have the staff do this.
[12] In 2011-12 the state transitioned to the FCAT 2.0. The 2011 scale scores range from 100 to 500, while the 2012 scale scores range from 140 to 302 in reading and 140 to 298 in math. For all regression estimates scale tests are normalized to have means of 0 and standard deviations of 1 within grade level, separately by county and by year.

less than satisfactory. Table A in the appendix shows the cutoffs for each level for the grades relevant to this study.

Additional outcome data include attendance rates and discipline records (office referrals and suspensions). It may be that students with undiagnosed and/or untreated vision problems miss school more often than their peers, or are more likely to misbehave in class and be referred to the principal, or even suspended. All of these are likely to reduce academic achievement. Attendance and discipline variables allow us to test whether these are mechanisms through which the intervention has an impact.

Demographic variables from school administrative records include grade, age (in months), race/ethnic group, gender, free/reduced lunch status, and receipt of English Language Learner (ELL), special education, and gifted services. We use these variables, and prior FCAT scores, to control for observable student differences. This is not necessary to avoid bias, since the randomization should remove all bias, yet controlling for these variables should yield more precise estimates of the impact of the intervention. These variables also allow us to investigate the possibility of heterogeneous effects by student characteristics.

B. Estimation Method. Although randomization provides a convincing counterfactual, econometric methods can estimate the effect of the intervention more precisely than comparisons of group means. The simplest regression model that one could estimate is:

$$Y = \alpha + \beta P + u \qquad (1)$$

where Y is the outcome of interest, such as student test scores, P equals one if a student's school was randomly assigned to the program (i.e. the intervention) and equals zero if that school was randomly assigned to the control group, u represents all other factors (observed or unobserved) that could affect test scores, and the α and β are parameters (coefficients) to be estimated.

Given that assignment of schools to the program was random, the variable P will be uncorrelated with u, so ordinary least squares estimates of β will be unbiased estimates of the program's impact. However, greater statistical precision can be obtained if other variables that affect test scores are added to the regression. This can be depicted as:

$$Y = \alpha + \beta P + \gamma_1 X_1 + \gamma_2 X_2 + \ldots \gamma_k X_k + u \qquad (2)$$

where the X variables indicate these other factors. The main X variables used in the analysis are students' prior year achievement, the students' age in months, and dummy variables for racial and ethnic groups, gender, free or reduced lunch status, receipt of ELL, gifted, or special education services, and, in some regressions, grade and/or diagnosis of specific vision problems.

Another extension is to allow the program to vary according to student characteristics, although this must be done with caution to avoid finding "significant" results by doing this for a large number of student characteristics; this should be done only for a few variables, those for which there is a clear reason to expect a differential effect. For example, suppose that X_1 indicates a type of student who would most likely benefit from the program, for example a student with vision problems or a student from a very poor family that perhaps cannot afford eyeglasses (which could be measured by the variable indicating eligibility for a free lunch). The following regression allows for separate impacts of students with and without this status:

$$Y = \alpha + \beta P X_1 + \delta P(1 - X_1) + \gamma_1 X_1 + \gamma_2 X_2 + \ldots \gamma_k X_k + u \qquad (3)$$

In this regression, β indicates the effect of the program for students with $X_1 = 1$ and δ is the effect for students for whom $X_1 = 0$.

Finally, recall that there are two programs, screen-only and full-treatment. The effects of both of these programs are estimated in the following regression:

$$Y = \alpha + \beta_1 P_1 + \beta_2 P_2 + \gamma_1 X_1 + \gamma_2 X_2 + \ldots \gamma_k X_k + u \qquad (4)$$

where P_1 equals one for students assigned to the first program (screen-only) and equals zero otherwise, and P_2 is analogously defined for students assigned to the second program (full-treatment). In this case, β_1 and β_2 are the impacts of the first and second programs, respectively.

Finally, when program impacts are allowed to vary by a student characteristic, X_1, equation (4) can be similarly modified:

$$Y = \alpha + \beta_1 P_1 X_1 + \beta_2 P_2 X_1 + \delta_1 P_1 (1 - X_1) + \delta_2 P_2 (1 - X_1) + \gamma_1 X_1 + \gamma_2 X_2 + \ldots \gamma_k X_k + u \quad (5)$$

where β_1 and β_2 are the impacts of the first and second programs, respectively, on students with $X_1 = 1$, and δ_1 and δ_2 are analogously defined for students with $X_1 = 0$.

V. Results

This section presents the results of our analysis. The first subsection presents cross-sectional evidence on the prevalence of untreated vision problems and the second examines correlations between untreated vision problems and academic performance. The last subsection presents our experimental evidence on the impact of providing vision services on students' academic outcomes.

A. Prevalence of Untreated Vision Problems. We begin by documenting the prevalence of untreated vision problems in schools serving low-income students. We find that a startlingly high percentage of students in these schools need glasses but either do not have them or have them but do not wear them regularly (as proxied by not having glasses on the screening day).

We have vision screening data from 5,747 grade 4 and 5 students in Title 1 elementary schools in central Florida.[13] As seen in the top panel of Table 3, 1,760 of the 5,747 students failed the screening and thus are likely to have untreated vision problems. This is a 31% failure rate. Of the students who failed the screening and were seen by an optometrist aboard the mobile unit, 86% (784 out of 914), were prescribed glasses. Recall that the main reason why some were not prescribed glasses is that some County 1 students who would have passed the screening were mistakenly classified as having a vision problem (false positives). Still, assuming that only 86% of the students who fail a screening in the screen-only schools need glasses, our data show that approximately one in four students in low-income schools have untreated (or undertreated) vision problems. This suggests that lack of information and/or lack of access to vision care are very common among low-income students in central Florida.

Figure 1 shows statistics by demographic subgroups. We find that, among race/ethnic categories, Asian students are most likely to fail the screening and be prescribed glasses. About 38% of Asian students failed the screening compared to 26-30% in other race/ethnic categories, and FLVQ dispensed glasses to 29% of Asian students compared to 14-20% for other race/ethnic categories. These differences are statistically significant at the 5% level.

Students who are eligible for free and reduced lunch (FRL), a proxy for household income, are slightly more likely to fail the screening (27% vs. 25%) and be prescribed glasses

[13] As seen in Table 3, 3,862 full-treatment students and 1,885 (1,036 + 849) screen-only students were screened.

(17% vs. 14%) than their non-FRL peers, but these differences are not statistically significant. English language learner students (ELL) are slightly more likely to fail the screening (31% vs. 27%) than their non-ELL peers, a difference that is statistically significant at the 5% level, but are equally likely to be provided glasses (18% for both). Perhaps most striking is the difference between special education and non-special education students; special education students are less likely to fail the screening than their non-special education peers (25% vs. 30%), and consequently are less likely to be provided with glasses (17% vs. 21%). Both differences are significant at the 5% level.

There are also some differences by demographic group on the specific vision problem identified by the screening, which are summarized in Table 5. Myopia is much more common among Asian students (relative to other students), and this is highly significant ($p = 0.008$). Astigmatism is more prevalent for Black, Asian, Native American and special education students, but this is significant only for Black students ($p = 0.046$), which reflects the relatively small samples of the other three groups. Finally, there are more "other" vision problems among Black and Asian students, but again this is significant only for Black students ($p = 0.001$). Other vision problems include Anisocoria, Anisomertopia, Gaze, and Hyperopia (farsightedness).

B. Correlation between Vision Problems and Academic Outcomes. Interestingly, we do not find that students with vision problems tend to have lower test scores. Figure 2 compares the distributions of the math and reading pre-test (2011 FCAT) scale scores for students who passed the screening exam to the same distributions for those who failed the exam, combining Counties 1 and 3.[14] For both math and reading, the distribution of scores for students who failed the screening is very similar to the distribution for students who passed the screening; children

[14] Separate estimates by county are very similar. Recall that we do not have 2011 FCAT scale scores for County 2.

with vision problems do not have lower baseline test scores than children without vision problems.

The fact that students with and without vision problems have similar scores may indicate that the vision problems identified by the intervention developed in the previous year or two. The students in this study are young, and their eyes are still developing. It could be that the intervention detected emerging problems that had not yet impacted learning.[15] If this is the case, some of the students identified by the intervention would likely be identified by conventional screenings in 6th grade, and the main benefit of the intervention is that the problem is identified and/or treated one or two years earlier than it would have been in the absence of the intervention.

An alternative explanation for the lack of a difference in baseline test scores between students with and without vision problems is that students accommodate for their vision problems. Their vision problems may not be severe enough to make it impossible for them to be successful. For example, a student who has trouble seeing the blackboard may adapt by sitting closer to the front of the classroom or by asking classmates to tell him/her what is written on the board. If this is the case, students who are successful in spite of their vision problems will benefit from the intervention because they do not have to make extra efforts to adapt to their poor vision.

A third possibility is that students who exert extra effort are more likely to have vision problems. Glewwe et al (2014) discuss research that suggests that more frequent studying may increase the likelihood of myopia. If this is the case, conditional on effort levels, the students who failed the screening would have lower test scores than those who passed the screening.

It is worth noting that policy makers interested in closing the gap between low and high achieving students may find that providing vision services will not contribute to this goal. The

[15] Glewwe et al. (2014) note that in China's Gansu province only 3.0% of Grade 1 students have vision problems, but this rate increases to 7.7% by Grade 3 and 16.5% by Grade 5.

fact that vision problems exist across the achievement distribution indicates that vision interventions will have an impact not only on students who are performing below grade level but also on those who are performing at, or even above, grade level. We do see, however, in Table 5 that vision problems are not uniformly distributed across demographic groups. Therefore, gaps between demographic groups may be affected by vision interventions.

C. Experimental Estimates of the Impact of Vision Services. Although students with vision problems are not concentrated at the bottom of the achievement distribution, at virtually any point on the distribution such students may be underperforming relative to their full potential. The randomized control trial that we conduct offers the first evidence from the United States (or any developed country) on the impact of providing vision services (screening only or screening with free follow-up exams and eyeglasses) on student achievement. Tables 6 and 7 present estimates of equation (4), and Tables 8 and 9 do the same for equation (5). In general, all these estimates are intent to treat (ITT) estimates, in two distinct senses. First, they are estimates of the impact of offering services, and some students did not obtain the services because, for example, they were absent on the days of the screening or did not return a permission slip to be seen on the mobile eye clinic. Second, except for Table 9 the estimates compare all students, both those with and without vision problems, and so the estimated impacts are for offering vision services to the average student, not just to students who need vision services.

Average Impacts on FCAT scores. To begin, Table 6 presents estimates of the average impacts of the two programs for each of the three counties, separately for grades 4 and 5. The dependent variable is the 2012 FCAT achievement level, the only learning outcome measure that we have for all three counties. All our estimates are of learning gains over one year, since they are conditional on the previous year's test score.[16] One should also bear in mind that this impact is an ITT effect averaged over all students, not just those who have vision problems; later we present separate estimates by the vision status of students in the intervention schools.

The coefficients on the non-program variables are as expected.[17] Reading and math skills in the prior year strongly predict current reading and math achievement. Demographic patterns are consistent with the literature on student achievement. Girls made larger gains in reading and smaller gains in math than boys. In all three counties, black students almost always had smaller gains in both subjects than white (the omitted category) and Asian students. In Counties 1 and 3, Hispanic students had somewhat lower gains than baseline (white) students, but this is not the case in County 2. Students who are eligible for free or reduced lunch made smaller gains than non-eligible students. English language learner students made smaller gains on reading tests in two of the three counties than their peers who have English as their first language, however they made larger gains in County 3 in the 4th grade and they generally made larger gains in math. Within grade, older students made fewer gains than younger students. This may be because students who are struggling are more likely to be held back and thus be older than their peers.[18]

Turning to the variables of interest, the full-treatment schools almost always made larger gains than the screen-only schools, though this difference is usually not significant. This result is strongest for reading scores, especially in County 2. In regressions combining all three counties (not shown), F-tests for grade 4 show no significant differences in impacts on reading or math scores between full-treatment and screen-only schools ($p = 0.156$ for reading, 0.733 for math),

[16] In fact, they would be estimates of learning gains even without conditioning on the previous year's test scores since randomized assignment implies that the treatment and control groups were virtually identical at baseline, so any differences after the intervention must be due to gains that occurred after the intervention was implemented.
[17] The variables in the regressions differ slightly by county. County 1 did not provide free/reduced lunch status and County 2 did not provide special education or gifted status.
[18] We find that just over 22% of students are one to two years older than the typical age for their grade level and just over 2% of students are two or more years older than the typical age for their grade level.

but the grade 5 impacts are significantly larger for full-treatment schools (p = 0.001 for reading, and 0.090 for math). This indicates that there is an "access problem;" it is not simply a problem of lack of information concerning vision problems – some students were constrained by lack of access to follow-up services. Notably, even when FLVQ significantly lowered barriers to access, not everyone was treated. We find that attending school on the day the mobile unit came to the school and having a signed permission slip proved formidable barriers. Only 75% of the students in the full-treatment schools who failed the screening were seen aboard the mobile unit.

The key question posed by the randomization, however, is whether the full-treatment schools and screen-only schools made larger gains the control schools. We begin with evidence from the full-treatment schools for the reading exams. In County 3, both 4th and 5th grade students in the full-treatment schools made larger gains than their counterparts in the control schools for the reading exam, these differences are significant only at the 10% level; the magnitude of these effects was 0.082 and 0.084, respectively. These coefficients are difficult to interpret because they describe movement between achievement levels on the FCAT. They correspond to approximately 0.07 standard deviations of the distribution of student test scores.[19] In County 2, grade 5 students in the full-treatment schools made larger gains than students in the control schools in reading. The magnitude of the coefficient is relatively large – 0.126, which corresponds to 0.11 standard deviations – and is significant at the 5% level. Grade 4 students in County 2, however, saw no statistically significant impact from the full treatment nor did grade 5 students in County 1. Surprisingly, grade 4 students in County 1 did significantly *worse* than the control schools. Perhaps the intervention was ineffective in County 1 in part because of the error that led to false positives and false negatives in some schools, but it could also reflect random

[19] The within county standard deviations for the FCAT variables can be found in Table 4.

chance given that there are six estimates for grade 4 in Table 6 and the true impact for that grade may be close to zero.

Turning to results in math, in County 3 grade 5 the full-treatment students made much larger learning gains than did control school students. The coefficient, 0.197, is significant at the 1% level and corresponds to an impact of 0.16 standard deviations. Grade 4 students, however, did no better than those in the control schools. In County 2, the effect of the full treatment intervention on math is also concentrated in grade 5, with a coefficient of 0.168 that is significant at the 1% level and corresponds to 0.14 standard deviations. The same is true for both grades in County 1.

The results in Table 6 also indicate that the full-treatment intervention was more effective for grade 5 students than for grade 4 students; in regressions combining all three counties (not shown), F-tests show that the impacts in grade 5 are significantly larger than those for grade 4 for both readging ($p = 0.027$) and math ($p = 0.016$) There are two reasons to expect higher impacts for grade 5 students. First, recall that Florida schools screen all children for vision problems in grades 3 and 6, so two years have elapsed since grade 5 students have had their vision checked, compared to only one year for grade 4 students, so the former have had twice as much time for their vision to deteriorate since their last screening, or to have lost or broken any glasses they acquired in response to the grade 3 screening. Second, the curriculum becomes more reading intensive (and less reliant on oral instruction) with each grade, which likely increases the need for glasses in higher grades.

Turning to the screen-only schools, their estimated impacts were almost always smaller than the full-treatment school estimates, and only one of the 12 screen-only estimates is significantly positive: grade 5 in County 3. Not only are none of the other 11 estimates significantly positive, reading scores are significantly *lower* than the control schools for both grades in County 1 and grade 4 in County 2 , and the same is true for grade 5 math scores in County 1 (though only at the 10% level). Thus there is very little evidence that the screen-only intervention increased test scores beyond what we would have seen in the absence of the intervention, which again implies that an "information problem" is not the primary reason why children in central Florida have uncorrected vision problems.[20] It does not appear that providing students and parents information about vision problems will lead to improved student learning.

Indeed, there may even be reason to worry that the screen-only intervention was harmful. Students (parents) could have been upset that they are told about a vision problem but not offered help to address that problem. This may make a student more inclined to give up and attribute his or her academic difficulties to the vision problem. This is consistent with evidence given below that discipline referrals and suspensions increased in the County 3 screen-only schools, even after controlling for prior referrals and suspensions. This may have been particularly likely in County 1, where screen-only schools were mistakenly told that they would be provided glasses.[21]

Another possible explanation is that the control schools were somehow compromised. For example, if principals in the control group heard that their peers' schools were getting extra vision services and did something (vision related or otherwise) to attempt to "keep up," then the control group may not be a valid counterfactual. This is most plausible in the districts with fewer control schools (i.e. County 1 and County 2). In these districts the feeling of being left out of the intervention would be more acute. In County 3 there were many more control schools than treatment schools so principals who were not offered the intervention may have been less likely

[20] Unfortunately, we have no data on whether students in screen-only schools acquired glasses after being screened.
[21] Note that there are no referral or suspension data available for County 1.

to feel the need to "keep up" with the treatment schools.[22] The behavior results mentioned in the previous paragraph could be consistent with this hypothesis if principals in control schools did something that affected behavior outcomes.

The estimates in Table 6 may be somewhat imprecise because the FCAT level scores "throw away" variation within each of the five levels. To avoid this, Table 7 presents estimates of the average impact of the two interventions on the standardized scale scores (rather than achievement level scores) in 2012 and 2013 for Counties 1 and 3 (recall that County 2 provided only scale scores for 2012, and provided no data for 2013). The scores were standardized to have a mean of 0 and a standard deviation of 1 within each district, subject and grade. For brevity, the control variable coefficients are not shown, but they are very similar to those in Table 6.

The top panel in Table 7 uses 2012 scores as the outcome. As in Table 6, the full-treatment schools in County 3 saw larger gains in reading scores than the control schools, on the order of 0.08 standard deviations for both 4th and 5th grade students. Math scores in County 3 full-treatment schools increased by approximately 0.15 standard deviations more than in control schools for 5th graders but were no better for 4th graders. As in Table 6, the full treatment does not appear to have been beneficial in County 1 and may have even lowered grade 4 reading scores, and the screen-only schools did not outperform the control schools in 2012, except in County 3 for 5th grade math scores. The negative coefficients in County 1 are no longer statistically significant, except (marginally) for 4th grade reading scores. It could be that concerns that the screen-only intervention was detrimental may be an artifact of the categorical

[22] Further, recall that in County 3, FLVQ had a significant presence at 28 Title 1 schools in prior years (these schools were excluded from the study). This indicates that knowledge of FLVQ was likely already higher across principals in County 3. A similar explanation for why we see the control schools outperform the screen only schools in County 1 and County 2 but not in County 3, could be that district officials, rather than principals, sought to compensate the schools in the control group. Any resources used to compensate the control group would have been spread out over more schools, and thus would have had a smaller impact, in County 3.

cutoffs in the achievement levels or that the improvement in the control schools was systematically at the cutoffs.[23]

Recall that the 2012 tests were administered during the last two weeks of April. Students who received glasses from FLVQ had them for between 1.5 and 6.5 months before taking the 2012 tests. In County 3, the last county to receive the intervention, students had their glasses for an average of only three months prior to the tests. For students who had glasses for a relatively short period, FCAT gains may mostly be due to being more able to read the test, as opposed to increased acquisition of human capital; that is, the test became a more accurate measure of their existing human capital.[24] Having glasses for a longer period of time should lead to additional acquisition of human capital that is reflected in higher test scores, though it is possible that the benefit of glasses could erode over time if students break, lose, or stop using them.

Table 7 also reports the results for 2013 tests, which were taken about 1.5 years after the intervention. The results suggest that virtually all of the positive impacts found in 2012 faded out by 2013.[25] Fade out in education interventions, especially when the outcome measure is a test score, is common and so may not be cause for alarm. For example, Duncan and Magnuson (2013) find that impacts from pre-school programs such as Head Start fade out rather quickly. Another example of fade out is Taylor (2014), who finds that gains from an extra math course quickly fade out for middle school students. In this case fade out may indicate that students are losing or breaking their glasses, and/or not persisting in wearing them regularly. Unfortunately the available data contain no information on whether students who were given eyeglasses still had them in 2013, or whether they were wearing them regularly. A final point is that these

[23] Visual inspection using density estimation shows no patterns of any kind at the categorical cutoff points.
[24] Yet some vision problems are unlikely to affect test taking skills; for example, myopia, the most common vision problem, impairs vision only for distant objects.
[25] Notably, the negative impact of the screen-only intervention in County 1 persists into 2013.

estimates are averages over all students; as will be seen below, for some types of students significantly positive effects persist into the 2013 school year.

Impacts on FCAT Scores for Different Types of Students. Next, we examine variation in the impact of the interventions by student characteristics. We focus entirely on County 3; the intervention appears to have had the largest impact in that county, so we want to explore this result further. Also, of the three counties the sample size is largest in County 3 (see Table 1), and although the sample in County 1 is almost as large recall that there were several implementation problems in County 1. For brevity, we combine 4^{th} and 5^{th} grade students (and add a control variable for grade). Results separately by grade yield no noteworthy patterns.

A priori, one would expect that providing vision screening services and free eyeglasses should have a larger impact on children from low income families, who presumably have limited medical care options and are less able to afford eyeglasses. On the other hand, low-income families may be better served by social safety nets such as Medicaid, leaving lower middle income students with fewer healthcare options. The top panel of Table 8 presents estimates of equation (5), which allows the impact of each program to vary by whether students in the treatment schools receive a free or reduced price lunch (FRL). We find that the impact appears stronger for students who were *not* receiving free or reduced price lunch, particularly on the reading test. This is a small group, fewer than 11% of the students in these schools were not receiving free or reduced price lunch, so the standard errors are larger. In fact, the differences between FRL and non-FRL are only marginally significant ($p = 0.08$). There is also a positive and statistically significant impact of the screen-only intervention for students who were *not* receiving a free or reduced lunch. This, however, occurs only in 2013 and only for math.

The bottom panel of Table 8 shows that the full treatment had a large and statistically significant impact on students who were receiving English Language Learner (ELL) services. The average impact on both reading and math is about 0.15 standard deviations. For students who receive ELL services, the positive impact appears to persist into 2013; their math scores are 0.08 standard deviations higher than those of ELL students in the control schools more than one year after the intervention, and this effect is significant at the 5% level. For students who receive ELL services, there is also a positive and statistically significant impact of the screen-only intervention on math scores in 2013 (but not in 2012).[26]

The finding that the impact is strongest for ELL students, coupled with the result that the impact does not differ by FRL status, leads us to posit that barriers to access may not be entirely economic. If ELL students are more likely to be undocumented immigrants, or to have parents who are undocumented, it may be that this population has little access to standard healthcare services. Nearly two thirds of undocumented Latino immigrants are uninsured (Rodriguez et. al. 2009) and undocumented immigrants are excluded from Medicaid except under extreme emergency circumstances – which would not include vision care (Sommers 2013).

Perhaps the most obvious distinction to make among students is between those who passed the screening – and therefore were not offered glasses or any other services – and those who failed the screening and so were either notified and advised to seek treatment (screen-only schools) or offered a free eye exam and free eyeglasses (full-treatment schools). We expected most, and perhaps all, of the benefits to accrue to the students identified as having unmet vision

[26] We also have results by gender and race (not shown here but available from the authors.). For reading, the full treatment had a stronger impact on girls. For math, it had a similar impact on girls and boys. Consistent with the fact that most ELL students in Florida are Hispanic, the interventions had the largest impact on Hispanic students.

problems. This is examined in Table 9, which reports regressions that allow the impact of the program to vary by whether students passed or failed the screening.[27]

Surprisingly, it is not the case that students identified as having unmet vision problems benefitted most from the full treatment in 2012. The coefficients on the interaction of full-treatment and passed-screening (0.078 and 0.086 for reading and math, respectively) are almost as large as those on the interaction of full-treatment and failed-screening (0.095 and 0.101, respectively), and in fact are more statistically significant (which reflects in part that most students passed the screening, which raises the statistical precision for that group). This suggests that the full treatment generates sizable spillovers. There do not appear to be sizable spillovers in the screen-only schools; in these schools the point estimates for students who failed the screening are generally larger than those for their peers who passed the screening, and none is significant.

While we find no statistically significant effects in the screen-only schools, it is worth noting that the size of the coefficient on the interaction between screen-only and failed-screening is similar to that of the interaction between full-treatment and failed-screening, so perhaps the screening-only intervention had an impact. In particular, one could argue that we also find no significant effect for full-treatment interacted with failed-screening, so either both interventions had an effect on students with poor vision, albeit imprecisely estimated, or neither had an impact. However, the fact that there is a statistically significant spillover effect in the full-treatment schools suggests that the glasses must have had some impact on students with poor vision in

[27] In principle, this specification should include a dummy variable indicating poor vision for students in both sets of treatment schools and in the control schools. Unfortunately, this is not possible because there are no data on vision problems for control school students. Thus the coefficients on the interaction between the treatment and "failed screening" should be interpreted as the impact of the treatment on students who failed the screening relative to an average control group student (i.e. not the impact on students with bad vision in the treatment group relative to a control group student with bad vision). The analogous interpretation applies to the "passed screening" interaction.

those schools; there must have been an effect on the students who received eyeglasses for there to be an impact on their classmates who did not need, and thus did not receive, eyeglasses.

Spillovers could come in a number of forms. Perhaps students who were already wearing glasses and passed the screening tended not to wear their glasses prior to the intervention. In the full-treatment schools, 10-20% of the students were given glasses, which may have been a sufficient critical mass to change those schools' "glasses-wearing culture". If students now view it is normal, or even desirable, to wear glasses, this could spill over onto students who already had glasses but who were not wearing them regularly. Similarly, teachers may decide to encourage wearing of glasses. More generally, teachers' classroom behavior may have changed in response to some of their students having a learning impediment removed, and this may have benefited all students. Another possibility is that students who failed the screening used to ask their peers to help them, and now their peers can focus on their own assignments. Lastly, it may be that the students who received glasses improved their behavior, which reduced disruptions in class. We find no support in the data, however, for this last hypothesis (see the discussion below of Table 10) and the other hypotheses cannot be tested with the data we have.

A final point regarding the surprising finding that the spillovers effects are almost as large as the effect on students who failed the screening exam is that although we cannot reject the hypothesis that these two effects are equal, we also cannot reject the hypotheses that the latter impacts are twice, or even three times, as large as the former impacts (due to the imprecision of the estimates of the latter).[28] In the end, the estimated impacts by students' visual acuity are insufficiently precise to distinguish whether the spillovers are of equal size or only half or one third of the size of the effects on the students who failed the screening.

[28] The p-value for the hypotheses that the impacts are equal, twice and three times are p=0.76, p=0.39 and p=0.16 respectively. These values are for the 2012 reading results.

Next, we consider estimates of program impacts separately by the type of vision problem detected by the "Spot". Traditional screenings, such as the Snellen chart, provide little or no information on the type of vision problem. In contrast, the Spot provides a preliminary diagnosis that must be confirmed by an optometrist.[29] Panel B of Table 9 presents results that interact the intervention with the most common conditions detected. Turning first to the 2012 results, the students given a preliminary diagnosis of myopia (nearsightedness), the most common diagnosis (see Table 4), appear to benefit from both the screen-only and the full-treatment interventions, but the estimated effects are relatively small and only one of the four is significant.

For students with a preliminary diagnosis of astigmatism (irregular shape of the cornea or lens), the second most common vision problem, there is no impact on reading or math scores for the screen-only intervention but there is a large (0.15 standard deviations) impact of the full treatment intervention – one that appears to persist into 2013 (although large, at 0.12 standard deviations, the 2013 coefficient is not statistically significant). In contrast, the effect of the full-treatment intervention on math scores is smaller and statistically insignificant.

These estimates by type of diagnosis are imprecise due to splitting the students who failed the screening into three distinct categories, but it may be that the large effect of the full treatment on students with astigmatism reflects the fact that astigmatism affects both near and far vision, while myopia affects only far vision. We see some evidence that students diagnosed with "other" vision problems (i.e. other than myopia or astigmatism) made gains in math in 2012 in full-treatment schools and in 2013 in screen-only schools. Further analysis, not shown here, reveals that these results are largely driven by anisocoria (unequal pupil size).

[29] We do not have detailed data from the optometrist. We know what prescription the students were given but not the optometrist's assessment of their condition (i.e. myopia, astigmatism, etc).

Impacts on Non-Academic Outcomes. Finally, Table 10 shows estimated impacts of both interventions on two types of non-academic outcomes, namely attendance and behavior outcomes. All estimates control for the prior year's data; for example, the regression for absences includes as an explanatory variable the prior year's absences. All estimates also control for 2011 FCAT achievement levels. We see no evidence that the intervention reduced student absences. In fact, the only statistically significant result on absenteeism is only significant at the 10% level and has an unexpected sign: students in the full treatment schools appear to have been absent *more* often than control school students even after conditioning on prior year's absences.

There is also no evidence that either intervention reduced discipline problems. In fact, in County 3 students in the screen-only schools were *more* likely to have behavior issues than students in the control schools even after conditioning on prior year's behavior issues. More generally, using data one year before the interventions we find no correlation between bad vision and behavior problems. Neither do we find that students who failed the screening in the fall of 2011 had more absences or behavior referrals in the academic year prior to the intervention (2010-11 school year) than did students who passed the screenings.

VI. Conclusion

This is the first study in the United States to use a randomized trial to estimate the impact of vision services on student outcomes. We find that providing additional/enhanced screening alone is generally insufficient to improve student learning as measured by test scores. However, in two of the three counties studied, the full-treatment intervention, which included free vision exams and free eyeglasses, significantly improved student achievement in math and reading skills in grade 5 (but for the most part not in grade 4) as measured by the Florida Comprehensive

Achievement Test (FCAT). The magnitude of the impact ranges from 0.07 to 0.16 standard deviations of the distribution of student test scores. The impact on English Language Learner (ELL) students in 2012 is particular large, about 0.15 standard deviations for both math and reading (averaged across grades 4 and 5). Although most impacts faded out after the first year, the impact on ELL students' math scores in the second year was 0.08 standard deviations and statistically significant.

Several other findings are of interest, and some merit further study. First, there is clear evidence that the full-treatment intervention has positive spillover effects onto children who passed the screening and thus did not receive eyeglasses, yet the mechanism by which this occurred is unclear. A second surprising result is that the impact of the full-treatment intervention on middle-income students (measured by ineligibility for a free or reduced-price lunch) is as large, if not larger, than its impact on low-income students; this does not support the hypothesis that low income is a barrier to obtaining vision care services. Third, there is no evidence that either intervention reduced student absenteeism or discipline problems.

Returning to the finding that the full treatment increased student learning by 0.07 to 0.16 standard deviations of the distribution of test scores, one way to put this in context is to note that Krueger (1999) estimated that reducing class size in the early elementary grades by about eight students will raise student test scores by 0.22 standard deviations. Reducing class size, however, is much more expensive than the vision intervention.

Another way to assess the merits of this intervention is to calculate a simple back-of-the-envelope estimate of the benefit/cost ratio. FLVQ estimates that it costs $2 per student to screen and $100 per student to provide an exam and eyeglasses. Assuming a class size of 20 students, and that FLVQ screens about 85% of students and identifies about 20% for exams (see Table 3) this implies that the cost of the full intervention is about $434 per class. The full social value of a 0.20 of a standard deviation achievement gain for a kindergarten class has recently been estimated conservatively at $200,000 in net present value (Hanushek 2010; Chetty et al. 2011). This includes both private (i.e. increased earnings for students) and social (i.e. lower crime rates) benefits. The full-treatment schools in County 3 experienced an average gain (averaging over both grades and both tests in Table 7) of 0.08 standard deviations yielding a present value of $80,000 per class. This is about 184 times the cost of FLVQ's intervention, which suggests that the intervention has a very high benefit-cost ratio. Granted, school districts do not realize the majority of these benefits. In order to assess the benefits to the school district we would need to follow the students in this intervention and assess potential benefits such as decreased rates of special education referrals.

While these estimates generally support vision service programs that include provision of free eyeglasses, we note at least four caveats. First, the implementation in County 1 (but not in Counties 2 and 3) experienced some difficulties, which may explain why the program did not raise student learning in that county, and may even have disrupted their learning. Second, it is not clear why the program increased the learning of grade 5 students in Counties 2 and 3 but had little or no effect on grade 4 students in those counties; we provide two plausible explanations but could not test either with the data available. Third, the estimates that attempt to measure differences in program impacts by the type of vision problem (e.g. myopia, astigmatism) are tentative at best. Fourth, the increases in test scores do not appear to have lasted for more than one year for most groups of students, so it may be that the very high benefit-cost ratio calculated in the previous paragraph is greatly overestimated -- although even if the true ratio were only one tenth of that calculation it would still be a very high benefit-cost ratio.

Future research on the impact of provision of vision services on students' educational outcomes should address issues that could not be resolved in this study. First, further research on the size of spillover effects, and on the mechanisms that generate them, may have very important policy implications. Second, more needs to be learned about barriers that prevent many students with poor vision from obtaining eyeglasses. Third, the results appear to fade out quickly from the first to the second year; additional research on why this happens – and what can be done to minimize it – would be very valuable. Finally, it would be useful to conduct similar research at the secondary school level to see whether vision services programs can improve the educational outcomes of older students. While much remains to be learned, it appears that vision interventions have the potential to improve student learning at a relatively low cost, and thus may be a wise investment for scarce resources in education in the United States, and presumably in other countries as well.

References

Basch, C. (2010). *Healthier Students Are Better Learners: A Missing Link in School Reforms to Close the Achievement Gap.* Equity Matters: Research Review No. 6. Teachers College, Columbia University.

Basch, Charles (2011) "Vision and the Achievement Gap Among Urban Minority Youth." *Journal of School Health*, 81(10): 593-598.

Chetty, R., J. Friedman, N. Hilger, E. Saez, D. Schanzenbach, and D. Yagan, (2011) "How Does Your Kindergarten Classroom Affect Your Earnings?," *Quarterly Journal of Economics*, 126(4):1593-1660.

Cunha, Flavio, James Heckman, Lance Lochner and Dimitriy Masterov. 2006. "Interpreting the Evidence on Life Cycle Skill Formation", in in E. Hanushek and F. Welch, eds., *Handbook of the Economics of Education*. North Holland.

Dirani, Mohamed, Xiaoe Zhang, Liang Ke Goh, Terri Young, Paul Lee and Seang Mei Saw (2010) "The Role of Vision in Academic School Performance." *Ophthalmic Epidemiology*, 17(1): 18-24.

Duncan, Greg, and Katherine Magnuson. (2013) "Investing in Preschool Programs". *Journal of Economic Perspectives* 27(2):109-132.

Ethan, Danna, Charles Basch (2008) "Promoting Healthy Vision in Students: Progress and Challenges in Policy, Programs and Research." *Journal of School Health*,78(8): 411-416.

Ethan, Danna, Charles Basch, Roger Platt, Elizabeth Bogen and Patricia Zybert (2010) "Implementing and Evaluating a School-Based Program to Improve Childhood Vision". *Journal of School Health*, 80(7): 340-345.

Ferebee, Annette. (2004). *Childhood vision: Public challenges and opportunities. A policy brief*. Washington, DC: Center for Health and Health Care in Schools, School of Public Health, George Washington University Medical Center.

Florida Department of Education, Office of Assessment Publication. "Understanding FCAT 2.0 Reports Spring 2012," Retrieved online 6/18/2013 as document s12uf2r.pdf.

Ganz, M., Xuan, Z., & Hunter, D.G. (2006). "Prevalence and correlates of children's diagnosed eye and vision conditions." *Ophthalmology, 113*(12), 2298-2306.

Ganz, M., Xuan, Z., & Hunter, D.G. (2007). "Patterns of eye care use and expenditures among children with diagnosed eye conditions." *Journal of the American Association for Pediatric Ophthalmology and Strabismus,11*(5), 480-487.

Glewwe, Paul, Albert Park and Meng Zhao. (2014) “A Better Vision for Development: Eyeglasses and Academic Performance in Rural Primary Schools in China”. Draft. University of Minnesota and Hong Kong University of Science and Technology.

Gomes-Neto, João., Eric Hanushek, Raimundo Leite and Roberto Frota-Bezzera. 1997. “Health and Schooling: Evidence and Policy Implications for Developing Countries.” *Economics of Education Review*. 16(3):271-283.

Grossman, Michael. 2006. “Education and Nonmarket Outcomes”, in E. Hanushek and F. Welch, eds., *Handbook of the Economics of Education*. North Holland.

Hannum, Emily, and Yuping Zhang. 2012. “Poverty and Proximate Barriers to Learning: Vision Deficiencies, Vision Correction and Educational Outcomes in Rural Northwest China”. *World Development* 40(9):1921-1931.

Hanushek, E.A., (2010) “The economic value of higher teacher quality," *Economics of Education Review* 30(3):466-479.

Heckman, James, Lance Lochner and Petra Todd. 2006. “Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond”, in in E. Hanushek and F. Welch, eds., *Handbook of the Economics of Education*. North Holland.

Indian River Study (2010). Unpublished research report. Florida Vision Quest.

Johnson, R.A.; Blair, R.; Zaba, J. (2000). “The Visual Screening of Title 1 Reading Students.” *Journal of Behavioral Ophthalmology*, 4 (1): 3-6.

Kimel, Linda (2006) “Lack of Follow-up Exams After Failed School Vision Screenings: An investigation of contributing factors.” *Journal of School Nursing*, 22(3): 156-162.

Kleinstein, R., et al (2003) “Refractive Error and Ethnicity in Children.” *Archives of Ophthalmology*, 121(8): 1141-7.

Krueger, Alan. 1999. “Experimental Estimates of Education Production Functions”. *Quarterly Journal of Economics* 114(2): 497-532.

Lange, Fabian, and Robert Topel. 2006. “The Social Value of Education and Social Capital”, in in E. Hanushek and F. Welch, eds., *Handbook of the Economics of Education*. North Holland.

Lochner, Lance. 2011. “Nonproduction Benefits of Education: Crime, Health, and Good Citizenship”, in E. Hanushek, S. Machin and L. Woessmann, eds., *Handbook of the Economics of Education: Volume 4*. North Holland.

Maples, WC (2003) “Visual Factors that Significantly Impact Academic Performance”. *Optometry*, 74(1): 35-49.

Maples, WC (2001) “A Comparison of Visual Abilities, Race and Socio-economic Factors and Predictors of Academic Achievement”. *Journal of Behavioral Optometry*, 12(3): 60-65.

National Commission on Vision and Health, “Vision Exams for Children Prior to Entering School,” Retrieved online 6/19/2013 http://www.visionandhealth.org/documents/FactsheetVisionexams123008MAS26.pdf

Rodriguez MA, Bustamante AV, Ang A. (2009) “Perceived Quality of Care, Receipt of Preventive Care, and Usual Source of Health Care among Undocumented and other Latinos.” *Journal of General Internal Medicine* 24(3):508-513

Sommers, Benjamin (2013) “Stuck between Health and Immigration Reform – Care for Undocumented Immigrants,” *New England Journal of Medicine* 369: 593-595.

Sonne-Schmidt “Wearing Glasses during Classes: Evidence from a Regression Discontinuity Design.” Retrieved online 6/18/13 http://www.sole-jole.org/12229.pdf

Taylor, Eric (2014) “Spending more of the school day in math class: Evidence from a regression discontinuity in middle school.” *Journal of Public Economics* 117(1):162-181.

Walline, J., and Johnson Carder, E. “Vision Problems of Children with Individualized Education Programs.” Journal of Behavioral Optometry, 23(4): 87-93.

White, Alan J. (2004) “Eye Exams for Children: Their Impact and Cost Effectiveness,” Abt Associates Inc.

Zaba, Joel (2011) “Children’s Vision Care in the 21st Century and its Impact on Education, Literacy, Social Issues and the Workplace: A Call to Action. *Journal of Behavioral Optometry*, 22(2): 22-39.

Table 1: Summary of Treatment and Control Groups

Schools			
	County 1	County 2	County 3
Control Group	15	5	23
Screen Only	11	5	7
Full Treatment	10	6	7
Total	36	16	37
Students - Full Sample			
	County 1	County 2	County 3
Control Group	2,823	1,408	4,356
Screen Only	2,143	1,329	1,325
Full Treatment	1,601	1,836	1,369
Total	6,567	4,573	7,050
Students - Final Sample			
	County 1	County 2	County 3
Control Group	2,544	1,101	3,684
Screen Only	1,904	903	1,111
Full Treatment	1,384	1,302	1,184
Total	5,832	3,306	5,979

Note: The number of schools is the same in both full sample and final sample. The final sample includes only students for whom we have the complete set of variables used in the regression analysis.

Table 2: Balance Tests

	County 1		County 2		County 3	
	Full Treatment	Screen Only	Full Treatment	Screen Only	Full Treatment	Screen Only
Reading 2011	-0.078	-0.025	-	-	-0.082	-0.050
	(0.089)	(0.081)			(0.076)	(0.096)
Math 2011	-0.069	0.029	-	-	-0.053	-0.067
	(0.105)	(0.087)			(0.082)	(0.096)
Reading 2011 (level)	-0.083	0.005	0.037	-0.054	-0.091	-0.096
	(0.100)	(0.097)	(0.094)	(0.106)	(0.103)	(0.092)
Math 2011 (level)	-0.088	0.042	-0.039	0.006	-0.050	-0.064
	(0.113)	(0.097)	(0.083)	(0.111)	(0.091)	(0.102)
Grade	0.009	0.002	0.010	0.007	0.000	-0.005
	(0.018)	(0.014)	(0.023)	(0.017)	(0.015)	(0.016)
Age (Months)	0.419	0.267	0.927	0.904	0.032	-0.453
	(0.386)	(0.479)	(0.572)	(0.528)	(0.697)	(0.701)
Asian	-0.001	-0.004	0.006	0.001	0.000	-0.003
	(0.005)	(0.004)	(0.008)	(0.008)	(0.009)	(0.009)
Black	0.000	-0.042	0.020	0.022	0.048	-0.001
	(0.078)	(0.070)	(0.043)	(0.030)	(0.151)	(0.159)
Hispanic	0.046	0.095	-0.031	-0.022	0.018	0.053
	(0.065)	(0.074)	(0.042)	(0.070)	(0.121)	(0.132)
Multiple race	-0.008	-0.023***	0.002	-0.012	0.002	-0.001
	(0.008)	(0.008)	(0.015)	(0.012)	(0.008)	(0.006)
Girl	-0.005	-0.003	0.030**	0.022	-0.020	-0.004
	(0.016)	(0.015)	(0.013)	(0.022)	(0.012)	(0.014)
Special education	-0.005	-0.004	-	-	0.003	0.012
	(0.013)	(0.014)			(0.015)	(0.018)
Gifted	-0.015	-0.040	-	-	0.006	0.001
	(0.049)	(0.032)			(0.014)	(0.017)
Free or reduced lunch (FRL)	-	-	-0.021	0.053	0.036	-0.006
			(0.035)	(0.037)	(0.029)	(0.038)
ELL	0.036	0.096	-0.022	0.086	0.039	-0.014
	(0.042)	(0.066)	(0.054)	(0.049)	(0.064)	(0.092)

Note: This table reports the results from 74 separate regressions. Each cell represents an OLS regression in the following format: row variable = α + β×column variable + ε. For example, the first number, -0.082 is the $\hat{\beta}$ from: $reading2011 = \alpha + \beta Full\ treatment + \varepsilon$. These regressions use the students in the final sample.
*** p<0.01, ** p<0.05, * p<0.1

Table 3: Screening and Exam Results by County

Panel A: Full Sample								
Full Treatment								
	County 1		**County 2**		**County 3**		**All counties**	
	Count	*Share*	*Count*	*Share*	*Count*	*Share*	*Count*	*Share*
Total students	1,601	1.00	1,836	1.00	1,369	1.00	4,806	1.00
Screened	1,317	0.82	1,470	0.80	1,075	0.79	3,862	0.80
Failed screening	456	0.28	426	0.23	340	0.25	1,222	0.25
Exam given	408	0.25	246	0.13	260	0.19	914	0.19
Glasses dispensed	326	0.20	226	0.12	232	0.17	784	0.16
Screen Only								
	County 1		**County 2**		**County 3**		**All counties**	
	Count	*Share*	*Count*	*Share*	*Count*	*Share*	*Count*	*Share*
Total students	2,143	1.00	1,329	1.00	1,325	1.00	4,795	1.00
Screened	N/A	N/A	849	0.64†	1,036	0.78	N/A	N/A
Failed screening	N/A	N/A	263	0.20†	275	0.21	N/A	N/A

Panel B: Final Sample								
Full Treatment								
	County 1		**County 2**		**County 3**		**All counties**	
	Count	*Share*	*Count*	*Share*	*Count*	*Share*	*Count*	*Share*
Total students	1,384	1.00	1,302	1.00	1,184	1.00	3,870	1.00
Screened	1,222	0.88	1,186	0.91	997	0.84	3,405	0.88
Failed screening	431	0.31	348	0.27	324	0.27	1,103	0.29
Exam given	388	0.28	199	0.15	250	0.21	837	0.22
Glasses dispensed	309	0.22	181	0.14	223	0.19	713	0.18
Screen Only								
	County 1		**County 2**		**County 3**		**All counties**	
	Count	*Share*	*Count*	*Share*	*Count*	*Share*	*Count*	*Share*
Total students	1,904	1.00	903	1.00	1,111	1.00	3,918	1.00
Screened	N/A	N/A	656	0.73†	947	0.85	N/A	N/A
Failed screening	N/A	N/A	193	0.21†	255	0.23	N/A	N/A

Note: This table reports the number and share of students who take part in each element of the intervention. Data from the screen-only schools in County 1 are unavailable. *Share* is share of total students. N/A indicates that screening data are not available for County 1 screen-only schools.

† These include one school that refused the screening.

Table 4: Summary Statistics

		County 1			County 2			County 3		
		2010-11	**2011-12**	**2012-13**	**2010-11**	**2011-12**	**2012-13**	**2010-11**	**2011-12**	**2012-13**
Demographic Data										
Total Students		-	5,832	5,323	-	3,306	N/A	-	5,979	5,386
Share Free/Red Lunch		-	N/A	N/A	-	0.85	N/A	-	0.89	0.89
Share Female		-	0.49	0.49	-	0.49	N/A	-	0.50	0.50
Share White			0.54	0.54	-	0.12	N/A	-	0.15	0.14
Share Black		-	0.17	0.17	-	0.16	N/A	-	0.43	0.43
Share Hispanic		-	0.24	0.24	-	0.45	N/A	-	0.37	0.37
Share Asian		-	0.01	0.01	-	0.03	N/A	-	0.02	0.03
Share ELL		-	0.13	0.14	-	0.53	N/A	-	0.36	0.37
Share Special Ed		-	0.16	0.15	-	N/A	N/A	-	0.12	0.11
Share Gifted Ed		-	0.05	0.05	-	N/A	N/A	-	0.05	0.05
Age (month) at time of intervention		-	133	133	-	132	N/A	-	133	133
Outcome Data										
Reading	Scale score	317.96 (53.10)	214.87 (20.12)	221.03 (20.34)	N/A	N/A	N/A	306.88 (53.85)	212.75 (19.75)	218.64 (19.74)
	Level	3.02 (1.14)	2.78 (1.17)	2.73 (1.18)	2.83 (1.15)	2.73 (1.14)	N/A	2.80 (1.13)	2.67 (1.15)	2.59 (1.15)
Math	Scale score	322.73 (57.19)	215.45 (21.33)	220.57 (20.00)	N/A	N/A	N/A	320.10 (59.80)	214.07 (21.40)	219.08 (20.13)
	Level	3.01 (1.07)	2.62 (1.22)	2.49 (1.18)	2.87 (1.11)	2.43 (1.15)	N/A	2.96 (1.11)	2.55 (1.23)	2.40 (1.18)
Absences	Total	8.66 (7.53)	8.50 (8.07)	9.14 (9.11)	5.83 (10.00)	9.37 (8.83)	N/A	7.69 (7.33)	7.74 (7.50)	8.35 (8.67)
	Unexcused	4.71 (5.78)	4.86 (6.31)	5.39 (7.44)	2.25 (4.05)	4.63 (4.96)	N/A	5.10 (5.54)	5.05 (5.45)	5.08 (5.92)
Behavior	Referrals	N/A	N/A	N/A	N/A	N/A	N/A	0.31 (1.43)	0.34 (1.19)	0.74 (2.38)
	Suspensions	N/A	N/A	N/A	N/A	N/A	N/A	0.31 (1.43)	0.48 (1.88)	1.38 (5.21)
Vision Problems										
Failed Screening		-	442	-	-	544		-	585	-
- Myopia diagnosis		-	366	-	-	388		-	390	-
- Astigmatism diagnosis		-	85	-	-	132		-	144	-
- Other diagnosis		-	105	-	-	155		-	190	-

Note: Mean (standard deviation) for the final sample. Scale scores are the continuous measures, and levels are the categorical measures. Here we report FCAT scale scores in their raw form. For the analysis in the paper, we convert these to standardized z-scores within district-grade. ELL = English Language Learner. N/A indicates data that are not available.

Table 5: Screening and Exam Results by Group (percent)

Panel A

	White	Black	Hispanic	Asian	Native American	Multiple race
Myopia	22.1	21.7	22.7	33.6	21.5	23.6
Astigmatism	5.5	8.4	6.9	9.2	9.2	4.2
Other diagnosis	8.1	11.3	8.1	10.7	6.5	5.8

Panel B

	FRL	Non-FRL	ELL	Non-ELL	Special education	No special education
Myopia	20.5	19.0	23.6	21.9	21.1	23.8
Astigmatism	7.3	6.9	7.3	7.0	9.1	6.9
Other diagnosis	9.1	8.4	8.2	9.3	7.5	9.4

Note: This table reports the "Spot preliminary diagnosis" by demographic groups. It includes all students in the final sample who were screened, combining both the full treatment schools and the screen-only schools for which we have data.

Table 6: The Impact of Vision Intervention on 2012 FCAT Levels

	Reading						Math					
	County 1		County 2		County 3		County 1		County 2		County 3	
	Grade 4	Grade 5	Grade 4	Grade 5	Grade 4	Grade 5	Grade 4	Grade 5	Grade 4	Grade 5	Grade 4	Grade 5
Full treatment	-0.075**	-0.002	-0.049	0.126**	0.082*	0.084*	-0.077	-0.054	-0.083	0.168***	0.005	0.197***
	(0.033)	(0.034)	(0.029)	(0.054)	(0.047)	(0.045)	(0.098)	(0.062)	(0.066)	(0.044)	(0.103)	(0.069)
Screen only	-0.063**	-0.061**	-0.129***	0.006	0.039	0.029	-0.095	-0.077*	-0.031	0.050	-0.073	0.169***
	(0.027)	(0.028)	(0.031)	(0.058)	(0.037)	(0.040)	(0.082)	(0.041)	(0.073)	(0.063)	(0.119)	(0.058)
Reading 2011†	0.575***	0.590***	0.565***	0.568***	0.568***	0.527***	0.218***	0.273***	0.107***	0.192***	0.156***	0.216***
	(0.021)	(0.018)	(0.024)	(0.019)	(0.016)	(0.018)	(0.020)	(0.017)	(0.027)	(0.014)	(0.024)	(0.020)
Math 2011†	0.240***	0.215***	0.232***	0.259***	0.241***	0.190***	0.662***	0.643***	0.692***	0.669***	0.651***	0.632***
	(0.019)	(0.023)	(0.022)	(0.017)	(0.018)	(0.020)	(0.020)	(0.018)	(0.027)	(0.015)	(0.023)	(0.019)
Age (months)	-0.010***	-0.005***	-0.006	-0.008***	-0.007***	-0.009***	-0.004*	-0.007**	-0.009**	-0.004**	-0.007***	-0.009***
	(0.002)	(0.002)	(0.004)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)
Girl	0.074**	0.068**	0.015	0.096***	0.052*	0.086***	-0.029	-0.111***	-0.009	-0.024	-0.040	-0.027
	(0.028)	(0.030)	(0.039)	(0.030)	(0.029)	(0.025)	(0.027)	(0.022)	(0.031)	(0.023)	(0.038)	(0.022)
Asian	0.184	0.015	0.298***	0.096	-0.041	0.143**	0.415***	0.330***	0.242**	0.277***	0.095	0.240***
	(0.120)	(0.110)	(0.088)	(0.080)	(0.109)	(0.071)	(0.139)	(0.103)	(0.111)	(0.086)	(0.099)	(0.085)
Black	-0.069*	-0.095**	0.090	0.030	-0.160***	-0.082	-0.001	-0.156***	-0.104	-0.002	-0.178**	-0.103***
	(0.039)	(0.037)	(0.083)	(0.051)	(0.040)	(0.050)	(0.062)	(0.040)	(0.079)	(0.060)	(0.077)	(0.049)
Hispanic	-0.108***	0.072	0.088**	0.027	-0.145***	-0.044	-0.045	-0.053	0.027	0.010	-0.146**	-0.066*
	(0.036)	(0.057)	(0.037)	(0.033)	(0.035)	(0.044)	(0.040)	(0.039)	(0.048)	(0.035)	(0.072)	(0.039)
Multiple race	-0.031	0.011	-0.058	-0.038	-0.055	-0.003	-0.017	0.089	-0.033	-0.019	-0.192**	-0.093
	(0.058)	(0.071)	(0.094)	(0.074)	(0.088)	(0.089)	(0.081)	(0.066)	(0.076)	(0.066)	(0.094)	(0.083)
Special educ.	-0.110**	-0.159***	--	--	-0.024	-0.205***	-0.018	0.001	--	--	-0.014	-0.059
	(0.050)	(0.035)			(0.066)	(0.042)	(0.062)	(0.050)			(0.067)	(0.045)
Gifted	0.264***	0.388***	--	--	0.369***	0.453***	0.388***	0.380***	--	--	0.405***	0.496***
	(0.069)	(0.040)			(0.060)	(0.063)	(0.131)	(0.096)			(0.086)	(0.065)
Free lunch	--	--	-0.099**	-0.062	-0.159***	-0.084***	--	--	-0.084	-0.087**	-0.107	-0.109**
			(0.042)	(0.045)	(0.040)	(0.029)			(0.055)	(0.034)	(0.081)	(0.051)
ELL	-0.049	-0.102	-0.027	-0.124**	0.065*	-0.011	-0.101	0.087	0.108**	-0.034	0.026	0.007
	(0.042)	(0.061)	(0.045)	(0.050)	(0.034)	(0.051)	(0.071)	(0.056)	(0.039)	(0.050)	(0.039)	(0.040)
Obs.	2,744	3,088	1,668	1,638	2,905	3,074	2,744	3,088	1,668	1,638	2,905	3,074
R-squared	0.617	0.636	0.582	0.615	0.602	0.601	0.594	0.655	0.546	0.659	0.563	0.679
F-test††	0.15	2.84	9.48***	3.52*	0.66	1.16	0.03	0.13	0.71	2.75	0.35	0.14
(p-value)	(0.6976)	(0.1008)	(0.0076)	(0.0803)	(0.4216)	(0.2888)	(0.8540)	(0.7164)	(0.4118)	(0.1181)	(0.5577)	(0.7056)

Note: Final sample. Clustered standard errors at school level in parentheses. All regressions include controls for strata. †2011 FCAT level. †† F-test for the null hypothesis that full treatment and screen only have the same effect. *** p<0.01, ** p<0.05, * p<0.1

Table 7: The Impact of Vision Interventions on Standardized 2012 and 2013 FCAT Scale Scores

	2012							
	Reading				**Math**			
	County 1		**County 3**		**County 1**		**County 3**	
	Grade 4	**Grade 5**	**Grade 4**	**Grade 5**	**Grade 4**	**Grade 5**	**Grade 4**	**Grade 5**
Full treatment	-0.068**	0.029	0.082*	0.078*	-0.045	-0.056	0.018	0.149**
	(0.029)	(0.026)	(0.046)	(0.040)	(0.074)	(0.049)	(0.075)	(0.055)
Screen only	-0.045*	-0.006	-0.008	0.026	-0.065	-0.019	-0.081	0.161***
	(0.025)	(0.023)	(0.036)	(0.032)	(0.056)	(0.034)	(0.088)	(0.050)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,744	3,088	2,905	3,074	2,744	3,088	2,905	3,074
R-squared	0.694	0.694	0.667	0.660	0.673	0.726	0.635	0.728
F-test†	0.61	1.73	2.67	1.34	0.07	0.53	1.01	0.05
(p-value)	(0.4382)	(0.1971)	(0.1110)	(0.2546)	(0.7875)	(0.4734)	(0.3226)	(0.8323)
	2013							
	Reading				**Math**			
	County 1		**County 3**		**County 1**		**County 3**	
	Grade 4	**Grade 5**	**Grade 4**	**Grade 5**	**Grade 4**	**Grade 5**	**Grade 4**	**Grade 5**
Full treatment	-0.002	0.008	-0.021	-0.001	0.028	-0.060	0.062	-0.009
	(0.039)	(0.035)	(0.036)	(0.035)	(0.044)	(0.040)	(0.060)	(0.046)
Screen only	-0.065*	-0.076**	-0.004	-0.019	0.022	-0.122***	-0.005	0.028
	(0.032)	(0.030)	(0.028)	(0.027)	(0.041)	(0.042)	(0.044)	(0.040)
Control variables	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,557	2,766	2,665	2,721	2,554	2,762	2,659	2,709
R-squared	0.661	0.638	0.628	0.633	0.676	0.667	0.644	0.648
F-test†	2.54	5.44**	0.16	0.24	0.01	1.94	1.24	0.46
(p-value)	(0.1197)	(0.0255)	(0.6910)	(0.6262)	(0.9048)	(0.1727)	(0.2722)	(0.5030)

Note: Final sample. Clustered standard errors at school level in parentheses. All regressions include controls for strata. Additional controls not shown, same as in Table 6. County 2 did not provide FCAT scale scores for either year. † F-test for the null hypothesis that full treatment and screen only have the same effect. *** p<0.01, ** p<0.05, * p<0.1

Table 8: Result by Demographic Groups, County 3 Only

Panel A – By Free or Reduced Lunch (FRL) Status				
	2012		**2013**	
	Reading	**Math**	**Reading**	**Math**
Full treatment × FRL	0.072* (0.036)	0.088* (0.046)	-0.016 (0.029)	0.021 (0.032)
Full treatment×No FRL	0.181*** (0.050)	0.093 (0.067)	0.036 (0.086)	0.092 (0.110)
Screen only×FRL	0.016 (0.023)	0.047 (0.044)	-0.012 (0.023)	0.003 (0.029)
Screen only×No FRL	-0.006 (0.048)	0.045 (0.049)	0.021 (0.056)	0.110** (0.050)
Control variables	Yes	Yes	Yes	Yes
Observations	5,979	5,979	5,386	5,368
R-squared	0.661	0.675	0.628	0.643

Panel B – By English Language Learner (ELL) Status				
	2012		**2013**	
	Reading	**Math**	**Reading**	**Math**
Full treatment×ELL	0.145*** (0.025)	0.147*** (0.047)	0.006 (0.039)	0.079** (0.037)
Full treatment×Non-ELL	0.044 (0.042)	0.055 (0.048)	-0.017 (0.046)	-0.002 (0.045)
Screen only×ELL	0.014 (0.042)	0.063 (0.047)	0.024 (0.030)	0.067** (0.030)
Screen only×Non-ELL	0.009 (0.025)	0.036 (0.045)	-0.030 (0.027)	-0.015 (0.030)
Control variables	Yes	Yes	Yes	Yes
Observations	5,979	5,979	5,386	5,368
R-squared	0.662	0.676	0.629	0.643

Note: Final sample. Clustered standard errors at school level in parentheses. All regressions include controls for strata. Additional controls not shown, same as in Table 6.
*** p<0.01, ** p<0.05, * p<0.1

Table 9: Results by Screening Status, County 3 Only

Panel A	2012		2013	
	Reading	Math	Reading	Math
Full treatment × Failed screening	0.095 (0.057)	0.101 (0.060)	0.004 (0.029)	0.025 (0.044)
Full treatment × Passed screening	0.078** (0.034)	0.086* (0.048)	-0.014 (0.035)	0.030 (0.038)
Screen only × Failed screening	0.064 (0.041)	0.097 (0.064)	-0.013 (0.042)	0.058 (0.044)
Screen only × Passed screening	-0.003 (0.019)	0.030 (0.039)	-0.009 (0.022)	0.001 (0.033)
Control variables	Yes	Yes	Yes	Yes
Observations	5,979	5,979	5,386	5,368
R-squared	0.663	0.677	0.629	0.643

Panel B	2012		2013	
	Reading	Math	Reading	Math
Full treatment×Myopia	0.056 (0.055)	0.057 (0.056)	-0.005 (0.041)	-0.020 (0.039)
Screen only×Myopia	0.104** (0.050)	0.071 (0.056)	0.005 (0.049)	0.031 (0.067)
Full treatment×Astigmatism	0.152** (0.069)	0.031 (0.072)	0.117 (0.079)	0.072 (0.073)
Screen only×Astigmatism	-0.051 (0.054)	0.090 (0.075)	-0.091 (0.058)	-0.051 (0.075)
Full treatment×Other problems	0.008 (0.058)	0.158*** (0.054)	-0.054 (0.054)	0.039 (0.029)
Screen only×Other problems	0.007 (0.063)	0.048 (0.063)	-0.031 (0.072)	0.134** (0.059)
Full treatment × Passed screening	0.076** (0.034)	0.084* (0.048)	-0.015 (0.036)	0.029 (0.038)
Screen only × Passed screening	-0.005 (0.019)	0.029 (0.038)	-0.010 (0.022)	0.000 (0.034)
Control variables	Yes	Yes	Yes	Yes
Observations	5,979	5,979	5,386	5,368
R-squared	0.663	0.677	0.629	0.644

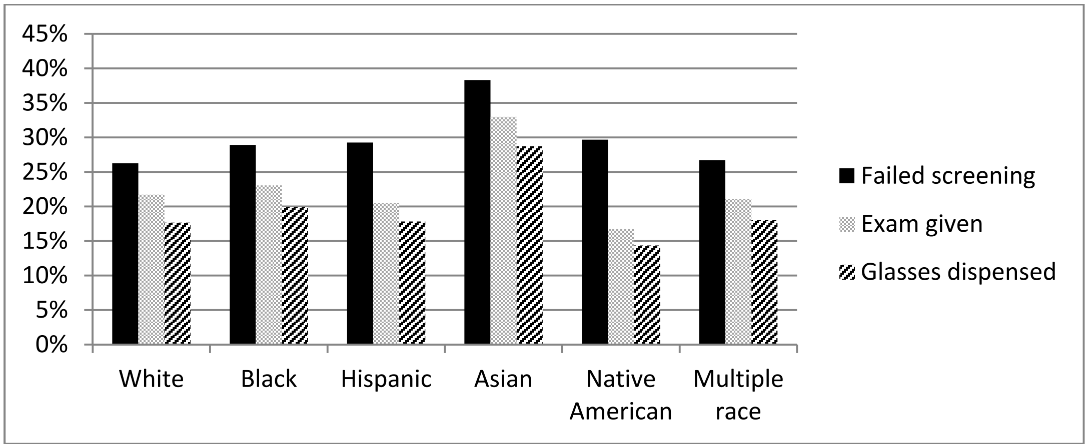
Note: Final sample. Clustered standard errors at school level in parentheses. All regressions include controls for strata. Additional controls not shown, same as in Table 6. *** p<0.01, ** p<0.05, * p<0.1

Table 10: Impacts of Vision Interventions on Attendance and Behavior Outcomes

	Total Absences 2012-13		Unexcused Absences 2012-13		Referrals 2012-2013	Suspensions 2012-2013
	County 1	County 3	County 1	County 3	County 3	County 3
Full treatment	0.072* (0.040)	-0.010 (0.043)	0.044 (0.053)	-0.027 (0.060)	0.264 (0.167)	0.042 (0.150)
Screen only	0.006 (0.030)	0.041 (0.036)	0.023 (0.050)	-0.042 (0.062)	0.317*** (0.104)	0.236*** (0.100)
FCAT reading 2011	0.036** (0.018)	-0.056*** (0.018)	0.007 (0.029)	-0.035 (0.023)	-0.122*** (0.046)	-0.108* (0.057)
FCAT math 2011	-0.072*** (0.026)	-0.045** (0.021)	-0.097*** (0.032)	-0.073*** (0.024)	-0.184*** (0.058)	-0.217*** (0.051)
Grade	-0.159*** (0.043)	-0.000 (0.033)	-0.237*** (0.069)	-0.158*** (0.055)	0.872*** (0.128)	1.224*** (0.118)
Age (months)	0.005** (0.002)	0.012*** (0.002)	0.009*** (0.003)	0.011*** (0.003)	0.033*** (0.006)	0.035*** (0.006)
Girl	-0.112*** (0.023)	-0.100*** (0.021)	-0.090*** (0.032)	-0.085*** (0.026)	-0.518*** (0.097)	-0.481*** (0.100)
Asian	-0.237 (0.158)	-0.342** (0.141)	-0.095 (0.235)	-0.493*** (0.187)	-0.615* (0.360)	-0.804 (0.535)
Black	-0.098** (0.040)	-0.127*** (0.034)	0.063 (0.061)	-0.077* (0.046)	1.024*** (0.123)	1.196*** (0.097)
Hispanic	-0.034 (0.039)	0.017 (0.035)	0.051 (0.055)	0.118*** (0.041)	0.204 (0.135)	0.378*** (0.126)
Multiple race	0.005 (0.074)	-0.013 (0.070)	0.179* (0.092)	0.000 (0.087)	0.171 (0.357)	0.262 (0.518)
Special Education	0.021 (0.039)	-0.215*** (0.042)	-0.035 (0.061)	-0.169*** (0.051)	-0.379*** (0.122)	-0.799*** (0.131)
Gifted	-0.091** (0.042)	-0.019 (0.068)	-0.242*** (0.090)	-0.126* (0.072)	-0.548* (0.299)	-0.171 (0.411)
Free lunch		0.149*** (0.037)		0.257*** (0.049)	0.919*** (0.151)	1.013*** (0.187)
ELL	-0.090** (0.045)	-0.157*** (0.033)	-0.105* (0.057)	-0.203*** (0.043)	-0.288*** (0.086)	-0.214** (0.107)
Total absences 2010-11	0.043*** (0.002)	0.047*** (0.002)				
Unexcused absences 2010-11			0.062*** (0.004)	0.058*** (0.004)		
Referrals 2010-11					0.138*** (0.026)	
Suspensions 2010-11						0.121*** (0.013)
Observations	5,697	5,506	5,697	5,506	5,506	5,979

Note: Final sample. Clustered standard errors at school level in parentheses. All regressions include controls for strata. *** p<0.01, ** p<0.05, * p<0.1

Figure 1: Screening and Exam Results by Group



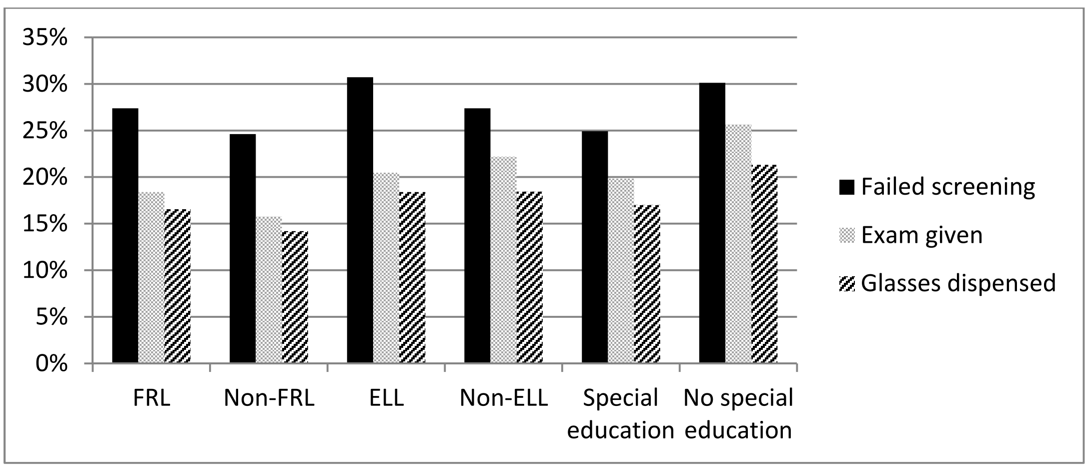
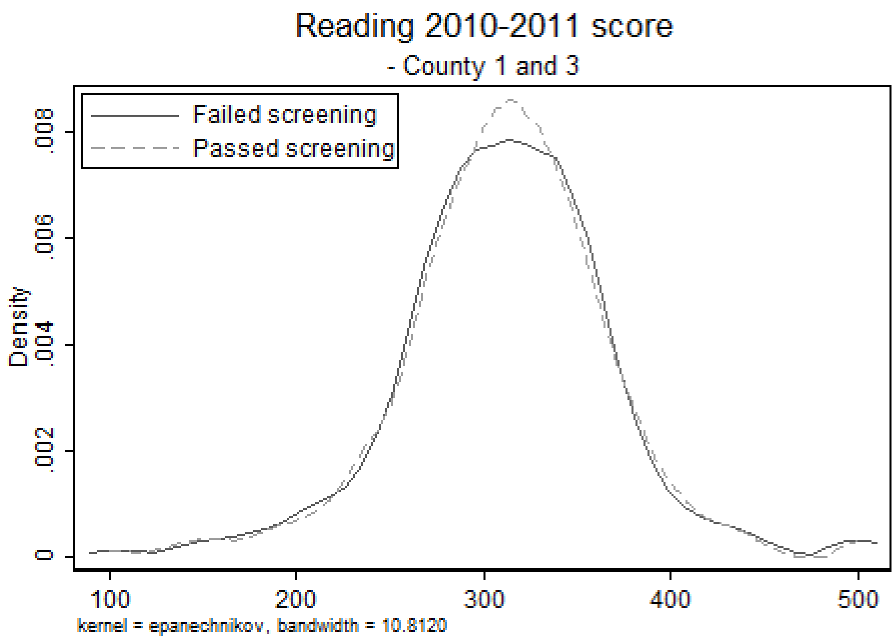
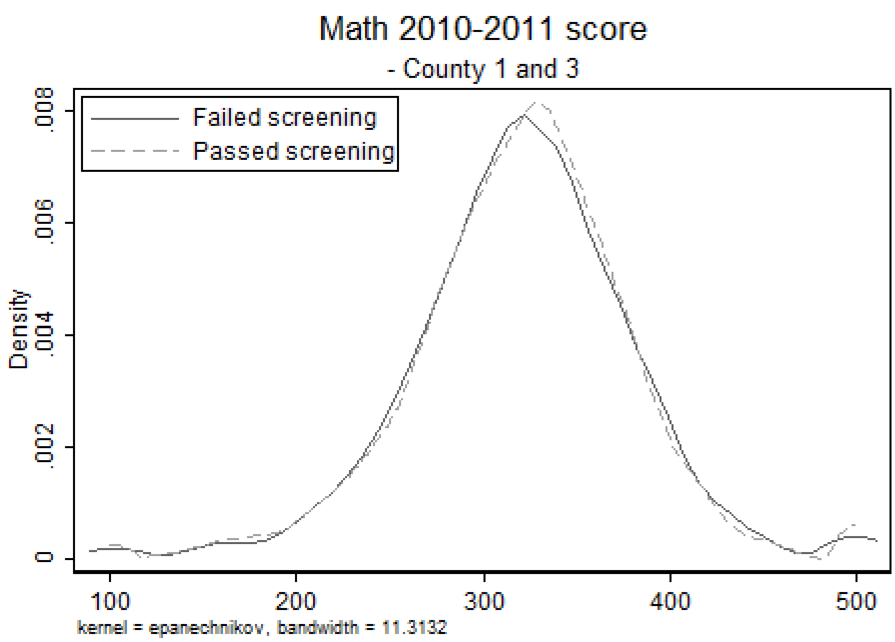


Figure 2


Reading 2010-2011 score
- County 1 and 3
Failed screening
Passed screening
Density
0
.002
.004
.006
.008
100
200
300
400
500
kernel = epanechnikov, bandwidth = 10.8120


Math 2010-2011 score
- County 1 and 3
Failed screening
Passed screening
Density
0
.002
.004
.006
.008
100
200
300
400
500
kernel = epanechnikov, bandwidth = 11.3132

Appendix

Table A: Achievement Levels for the FCAT and FCAT 2.0

Achievement Levels for the 2011 FCAT (2010-2011 school year)

Reading Scale Scores					
Grade	**Level 1**	**Level 2**	**Level 3**	**Level 4**	**Level 5**
3	100 - 258	259 - 283	284 - 331	332 - 393	394 - 500
4	100 - 274	275 - 298	299 - 338	339 - 385	386 - 500
Mathematics Scale Scores					
Grade	**Level 1**	**Level 2**	**Level 3**	**Level 4**	**Level 5**
3	100 - 252	253 - 293	294 - 345	346 - 397	398 - 500
4	100 - 259	260 - 297	298 - 346	347 - 393	394 - 500

Achievement Levels for the 2012 FCAT 2.0 (2011-2012 school year)

Reading Developmental Scale Scores (140 to 302)					
Grade	**Level 1**	**Level 2**	**Level 3**	**Level 4**	**Level 5**
4	154-191	192-207	208-220	221-237	238-269
5	161-199	200-215	216-229	230-245	246-277
Mathematics Developmental Scale Scores (140 to 298)					
Grade	**Level 1**	**Level 2**	**Level 3**	**Level 4**	**Level 5**
4	155-196	197-209	210-223	224-239	240-271
5	163-204	205-219	220-233	234-246	247-279

Source: Florida Department of Education