HKUST IEMS
INSTITUTE FOR EMERGING MARKET STUDIES

# A Better Vision for Development: Eyeglasses and Academic Performance in Rural Primary Schools in China

Paul GLEWWE, Albert PARK, Meng ZHAO

HKUST IEMS Working Paper No. 2015-37

June 2016



HKUST Institute for Emering Market Studies
The Hong Kong University of Science and Technology
Clear Water Bay, Kowloon, Hong Kong
T: +852 3469 2215 E: iems@ust.hk W: iems.ust.hk

# A Better Vision for Development: Eyeglasses and Academic Performance in Rural Primary Schools in China

Paul GLEWWE, Albert PARK, Meng ZHAO

HKUST IEMS Working Paper No. 2015-37
June 2016

## Abstract

About 10% of primary school students in developing countries have poor vision, but very few of them wear glasses. Almost no research examines the impact of poor vision on school performance, and simple OLS estimates could be biased because studying harder may adversely affects one's vision. This paper presents results from a randomized trial in Western China that offered free eyeglasses to rural primary school students. Our preferred estimates, which exclude township pairs for which students in the control township were mistakenly provided eyeglasses, indicate that wearing eyeglasses for one academic year increased the average test scores of students with poor vision by 0.16 to 0.22 standard deviations, equivalent to 0.3 to 0.5 additional years of schooling. These estimates are averages across the two counties where the intervention was conducted. We also find that the benefits are greater for under-performing students. A simple cost-benefit analysis suggests very high economic returns to wearing eyeglasses, raising the question of why such investments are not made by most families. We find that girls are more likely to refuse free eyeglasses, and that parental lack of awareness of vision problems, mothers' education, and economic factors (expenditures per capita and price) significantly affect whether children wear eyeglasses in the absence of the intervention.

## Authors' contact information

Paul Glewwe
Department of Applied Economics
University of Minnesota
E: pglewwe@umn.edu
Click here to enter text.

Albert Park
Division of Social Science, Department of Economics
Hong Kong University of Science and Technology
E: albertpark@ust.hk

Meng Zhao
Faculty of International Social Sciences
Gakushuin University
E: meng.zhao@gakushuin.ac.jp

# A Better Vision for Development: Eyeglasses and Academic Performance in Rural Primary Schools in China

Paul Glewwe
University of Minnesota

Albert Park
Hong Kong University of Science and Technology, CEPR, and IZA

Meng Zhao
Gakushuin University

May, 2016

**Abstract**:

About 10% of primary school students in developing countries have poor vision, but very few of them wear glasses. Almost no research examines the impact of poor vision on school performance, and simple OLS estimates could be biased because studying harder may adversely affects one's vision. This paper presents results from a randomized trial in Western China that offered free eyeglasses to rural primary school students. Our preferred estimates, which exclude township pairs for which students in the control township were mistakenly provided eyeglasses, indicate that wearing eyeglasses for one academic year increased the average test scores of students with poor vision by 0.16 to 0.22 standard deviations, equivalent to 0.3 to 0.5 additional years of schooling. These estimates are averages across the two counties where the intervention was conducted. We also find that the benefits are greater for under-performing students. A simple cost-benefit analysis suggests very high economic returns to wearing eyeglasses, raising the question of why such investments are not made by most families. We find that girls are more likely to refuse free eyeglasses, and that parental lack of awareness of vision problems, mothers' education, and economic factors (expenditures per capita and price) significantly affect whether children wear eyeglasses in the absence of the intervention.

Data collection for the Gansu Survey of Children and Families was supported by grants from The Spencer Foundation Small and Major Grants Programs (wave 1), by NIH Grants 1R01TW005930-01 and 5R01TW005930-02 (wave 2), and by a grant from the World Bank (wave 2). Travel was supported by the Center for International Food and Agricultural Policy at the University of Minnesota. We would like to thank numerous seminar participants, as well as the co-editor and two anonymous reviewers, for very helpful comments on earlier drafts of this paper.

## 1. Introduction

Many prominent economists agree that higher education levels increase economic growth (Barro, 1991; Mankiw et al., 1992; Krueger & Lindahl, 2001; Hanushek & Woessmann, 2008). Yet school enrollment may not increase economic growth and individuals' incomes if children acquire few skills while they are in school. Recent research has produced valuable evidence on the effect of specific interventions on student learning (see, *inter alia*, Glewwe et al., 2013). Most interventions have focused on improving the quality of schools and teachers: the supply side of education. Less attention has been given to increasing students' capacity to learn, which often reflects the decisions of parents. Researchers have found that health interventions – such as school meals, deworming programs, and iron tablets to reduce anemia – increase enrollment (Afridi, 2011; Miguel and Kremer, 2004; Vermeersch and Kremer, 2004) and learning (Luo et al., 2012).

This paper examines a health-related intervention that has received little attention in developing countries: providing eyeglasses to students with vision problems. About 10% of primary school students in developing countries have vision problems (Bundy et al. 2003). In almost all cases their vision can be corrected with properly fitted eyeglasses, but very few of them have eyeglasses. This paper presents results of a randomized trial in Western China that offered free eyeglasses to children in grades 4, 5 and 6. It estimates the impact of being offered eyeglasses and, because one third of those offers were turned down, the impact of wearing eyeglasses.

Due to program implementation problems in six townships, we present two sets of estimates: our preferred estimates, which include only the 12 township pairs (and an unpaired township) where the program was correctly implemented, and a set in the online appendix that uses all 18 township pairs (and the unpaired township). Averaging over both counties, our preferred estimates indicate that the intent-to-treat effect of offering free glasses to students with poor vision increases their average test scores by 0.11 to 0.16 standard deviations. Instrumental variable estimates of the impact of wearing glasses for one academic year are 0.16 to 0.22 standard deviations, which is equivalent to 0.3 to 0.5 more years of schooling. The estimated present discounted value of higher life cycle wages due to 0.3 more years in school easily exceeds the cost of the glasses.

We also find evidence of heterogeneous treatment effects. Most notably, in one county the intent-to-treat effect on test scores was 0.26 to 0.39 standard deviations, while there was little or no impact in the other county. We also find evidence that eyeglasses are less effective for children likely to have Vitamin A deficiencies and children with more educated parents, and that these differences can explain about half of the differential effects across the two counties.

Our results imply that many households fail to make what appears to be a high-return investment. What explains this failure? We study the determinants of children accepting the offer of free eyeglasses, and we also use a richer dataset on rural children in the same province to examine the determinants of wearing glasses absent the intervention. We find that information failures, such as lack of awareness of vision problems, and credit constraints appear to be major factors.

The rest of the paper is organized as follows. Section 2 introduces relevant aspects of primary education in rural China, and reviews the literature on vision problems among primary school students in developing countries, and how those problems affect their academic performance. Sections 3 and 4 describe the randomized trial and the data collection, and the methodology used to estimate program impacts. Section 5 examines whether the treatment and control townships are similar and investigates the possibility of selection bias when township pairs that did not implement the program correctly are excluded from the estimates. The next two sections present the results and investigate whether they vary by student characteristics, respectively. Section 8 explores why some children did not accept the free eyeglasses, and more generally why most children with poor vision do not wear eyeglasses. A final section summarizes the results and makes recommendations for further research.

## 2. Background and Literature Review

**A. Primary Education in Rural China.** China has achieved nearly universal primary school enrollment. In 2000, only 4% of adults aged 25 to 29 had no formal schooling (Hannum et al., 2008). The 1986 Law on Compulsory Education mandates that all children complete six years of primary school and three years of lower secondary school, yet the rural poor and some minority groups face difficulties in meeting this goal (Hannum, Park, and Cheng, 2007).

In rural areas of Western China, nearly all children attend the nearest public primary school, in their village or a nearby village. The county Educational Bureau allocates teachers to schools and pays their salaries. Thus, school quality disparities within counties tend to be modest (Li et al., 2009), reducing incentives to bypass the local school. Each county's Health Bureau conducts physical exams of all students, including eye exams. These exams should be annual, but budgetary and staff constraints cause many schools to conduct them only once every two or three years. The exam results are given to teachers, who are expected to convey them to parents.

**B. Vision Problems and School Performance.** Little data exist on children's vision problems in developing countries. Bundy et al. (2003) report that about 10% of 5-15 year old children have refraction errors (myopia, hypermetropia, strabismus, amblyopia, and astigmatism), which constitute 97% of their vision problems. Almost all refraction errors can be corrected with properly fitted eyeglasses, but most children with these problems in poor countries do not have glasses. Zhao et al. (2000) found that, in one district in Beijing, 12.8% of children age 5-15 years had vision problems, and 90% were refraction errors. Of these children, only 21% had glasses. He et al. (2007) report that 36.8% of 13-year-olds and 53.9% of 17-year-olds in middle schools in a county in Guangdong, a wealthy province, had myopia, and that less than half had glasses. Children with vision problems in poor areas are even less likely to have glasses, as seen below. In China, a common (but mistaken) belief is that wearing glasses worsens children's visual acuity.

Only two published studies examine the impact of poor vision on student academic performance in developing countries. Gomes-Neto et al. (1997) found that primary school children in Northeast Brazil with vision problems had higher probabilities of dropping out (10 percentage points) and repeating a grade (18 percentage points), and scored 0.2 to 0.3 standard deviations lower on achievement tests. Yet these estimates could be biased; if some of these children had glasses, their vision may be correlated with unobserved factors that affect learning, such as parent preferences for education. Even if none had glasses, their vision can be affected by their home environment (e.g. lighting quality) and daily activities, such as time studying or doing homework. Thus their vision may be correlated with unobserved factors that directly affect learning (e.g. hours studying), leading to biased estimates. Second, Hannum and Zhang (2012), using household survey data from Gansu province (described below) and propensity score matching, find that, for children with poor vision aged 13-16, wearing glasses sharply increases math and literacy test scores (by 0.27 and 0.43 standard deviations) but not language scores. Unfortunately, they could not fully address self-selection into wearing glasses; indeed, they show that wearing glasses is positively associated with socio-economic status and academic achievement and engagement.

We add to this small literature by providing experimental estimates of the impact of offering students eyeglasses in one of China's poorest provinces. It may seem obvious that providing glasses raises student learning, but the size of this impact is unknown. The "obvious" pathway is that students who cannot see, cannot learn, so providing glasses allows them to see the blackboard, read their textbooks, and study at home, reducing eye strain and possibly headaches. Yet complications can arise. First, some students or parents may refuse glasses due to worries that wearing them increases vision problems. Second, the impact of glasses may vary by parental characteristics, for example more educated parents may better compensate for children's undiagnosed vision problems, and by the type of vision problem. Third, implementation problems may arise when providing eyeglasses. Indeed, all three of these arose in this evaluation, so it is possible that offering eyeglasses does not always increase student learning.

## 3. Project Description and Data Available

The lack of evidence on the impact of offering glasses to students in developing countries led to the Gansu Vision Intervention Project. This section describes the project and the data collected.

**A. The Gansu Vision Intervention Project (GVIP).** In 2004, a team of Chinese and international researchers implemented a randomized trial to examine the impact of providing glasses to students with poor vision in two counties (hereafter, County 1 and County 2). The project covered nearly all grade 4-6 students in primary schools in these two counties.

Gansu province is in northwest China. Its geography is diverse, including the Gobi desert, mountains and vast grasslands. Its population was 25.4 million in 2004, three fourths of whom lived in rural areas. It ranked 30th out of 31 provinces in rural per capita disposable income (National Bureau of Statistics, 2005). Using official poverty lines, the World Bank (2000) found a 22.7% poverty rate for Gansu's rural population in 1996, compared to 6.3% for China as a whole.

The two counties were chosen for study because they are typical rural counties in Gansu, are near Lanzhou (the provincial capital), which eased monitoring by Gansu's Center for Disease Control (CDC), and had CDC staff to implement the project. County 1 is a Tibetan minority autonomous district. Its population was 217,000 in 2004, 85% of whom were in rural areas. In the 2000 census, 63% of its population was Han Chinese and 30% were Tibetan. County 2 is more populous and in a different municipality, but has a similar land area.[1] Its 2004 population was over 500,000, of whom 87% were in rural areas. Nearly all were Han Chinese. Both are typical Gansu counties in terms of GDP per capita in 2004, although County 1 was somewhat poorer.

[1] Municipalities are groups of counties in a province. County 1 census data are from searching the county's name in Wikipedia (accessed Nov. 23, 2011). All other figures in this paragraph are from Gansu Statistical Yearbook (2005).

County 1 has 22 townships; 19 of them, with 101 primary schools, participated in the program. Ten of these 19 townships were randomly assigned to the program, and the other nine were the control group. County 2 has 23 townships; 18 of these, with 155 primary schools, participated. Nine of these 18 were randomly assigned to the program, and the other nine were the control group. In both counties, the excluded townships were the county seat (the main urban center, where glasses are easy to obtain) and a few townships in sparsely populated remote areas.

Random assignment was conducted in 2004 as follows. In each county, all participating townships were ranked by 2003 per capita income. Starting with the two wealthiest, one was randomly assigned to be a treated township and the other to the control group; this was repeated for all subsequent township pairs. In County 1, the 19$^{th}$ township (the poorest) was not paired with another township; it was randomly assigned to the treatment group. In each township primary schools were either all assigned to the treated group or all to the control group.[2]

Baseline data were collected in June of 2004 (end of the 2003-04 school year) on student characteristics, exam scores, and visual acuity. Data were collected from treatment and control schools for all students finishing grades 1-5 in June of 2004. Treatment school students slated to enter grades 4-6 in the fall of 2004 who had poor vision were offered free eyeglasses. In each county, an optometrist hired for the summer visited all townships to conduct formal eye exams for students who accepted the glasses. If poor vision was confirmed, they were prescribed appropriate lenses. Students had a limited choice of colors and styles for their eyeglasses. All the eyeglasses were ordered from a reputable company. The 2004 fall semester began on August 26; most students who accepted the offer received glasses by mid-September. Teacher monitoring and field visits by project staff found high rates of wearing eyeglasses. At the end of the 2004-05 school year (late June/early July of 2005), fall and spring semester exam scores were collected.

---

[2]Primary schools with less than 100 students were excluded to avoid high travel costs to a few very remote schools.

Unfortunately, in 5 of the 18 control townships some students were given eyeglasses; after providing eyeglasses in the treatment townships, local officials used the remaining funds to buy them for students with poor vision in the paired control township. This occurred in two control townships in County 1,[3] and three in County 2. In another township pair in County 1, there was a "role reversal"; no one in the treatment township was offered glasses, while many children with poor vision in the control township were offered glasses. In contrast, the randomization was correctly implemented for six pairs of townships in County 2 and six pairs (plus the poorest township, the one randomly assigned to be treated) in County 1. To check for selection bias due to the exclusion of township pairs that deviated from random assignment, we conducted several robustness checks; they provide very little evidence of such bias.

**B. Data**. We use four sources of data: 1) school records of pupil characteristics and exam scores before and after the intervention; 2) results of health exams and vision tests conducted by the county CDC in each primary school before glasses were provided; 3) optometrists' records (for students fitted for eyeglasses); and 4) the Gansu Survey of Children and Families (GSCF), a longitudinal study of children in rural areas of 20 counties in Gansu. The school records include students' grade in school for the 2003-04 and 2004-05 school years, their sex, ethnicity, and birthdate, and the occupation and education of the head of their household (usually the father). Scores on exams (Chinese, math and science) given at the end of each semester were also collected.[4]

One characteristic of the exams has major implications for analysis: many schools design their own exams, so test scores may not be comparable across schools. Given random assignment of townships to treatment and control groups, this noncomparability of exams across schools does not cause bias, but it does add noise to the data. Section 4 discusses implications for estimation.

[3] In a third County 1 control township, four children received glasses, yet only one had poor vision. We retained this township. Excluding it and its matched pair (or excluding only these four students) has little effect on the results.
[4] In some schools, these semester exam scores are averages of several exams, including an end of semester exam.

The school health data include whether a student wears glasses (and if so, the grade when glasses were first worn), students' height, weight and hemoglobin count, and at least one vision measurement for each eye (students provided glasses have more measurements due to the fitting process). In China, eye exams are usually conducted by asking patients to read (with one eye covered) an eye chart from five meters away. The chart has 12 rows of the letter E facing different directions; the top row has large E's, and subsequent rows have smaller E's. If a patient cannot read the first row, the worst possible eyesight, his or her vision is coded as 4.0. If he or she can read the first row but not the second, his/her vision is coded as 4.1, and so on. A patient who can read the 10th row, the normal level, is coded as 5.0. Anyone who can read all 12 rows is coded as 5.2. There is also information from optometrists, only for children offered eyeglasses; it includes whether a child was fitted for glasses and, if not, the reason glasses were declined.

The GSCF was implemented in rural areas of Gansu. It was first conducted in 2000 for a random sample of 2000 children aged 9-12. A second wave (GSCF2) was conducted in 2004; 1869 of the original 2000 children were re-interviewed,[5] as were 886 oldest younger siblings of the original 2000 children, if 8 years old or older. The GSCF2 collected detailed information on vision and wearing eyeglasses from both sets of children and their parents, and data on lighting conditions at home and at school, the cost and availability of eyeglasses, and many household and village variables. The GSCF2 contains self-reported vision data and measurements of each child's eyesight via an eye exam, for both sets of children, conducted by Gansu CDC staff.

**C. Descriptive Statistics.** Table 1 presents descriptive statistics for the 25 "compliant" townships (township pairs for which both townships complied with their random assignment) and for all 37 townships. The former group had 18,902 students in grades 4-6 in 2004-05. Of these

[5] The reasons for no reinterview include: 108 had moved out of the counties where they had resided in 2000; 8 died; 4 were seriously ill; 2 had parents who divorced; 1 household refused; and for 8 children the reason is unknown.

students, 13.4% (2,529) had poor vision in the sense that either the left eye or the right eye (or both) had a visual acuity score below 4.9.[6] Only 2.3% of the students who had vision problems (59 out of 2,529) already had glasses. Those with vision problems had test scores almost identical to those of students without problems (78.2% vs. 78.9% for Chinese, 78.5% vs. 79.1% for mathematics, and 80.6% vs. 80.7% for science) at the end of the spring 2004 semester (before the program began). Very similar patterns are also seen for the full sample of 37 townships.

Simple t-tests show that none of these small differences in test scores is significant. But this does not imply that vision problems do not affect student learning because study habits may affect eyesight. Several studies show that doing "near-work" – spending many hours on activities with the eyes focused on objects about 1 meter away – can cause myopia (Angle and Wissmann, 1980; Mutti et al., 2002).[7] Thus, students who study more may tend to develop myopia, the most common vision problem. If so, simple comparisons of test scores of students with and without vision problems can underestimate the impact of vision problems on learning as they ignore the possibility that more studious students (who will have higher scores) have more visions problems.

Table 2 shows how the GVIP was implemented for the students with poor vision, for both the compliant (25 township) and full (37 township) samples. Of the 2,529 students with poor vision in the compliant sample, 1,528 were in the program schools. Of these, 1,066 (69.8%) accepted, while the other 462 declined, the eyeglasses. The main reasons for declining were objection of the household head (145) and refusal by the child (80). Similar patterns hold for the full sample of 37 townships; in particular, 70.0% of those offered eyeglasses accepted them.[8]

[6] Children with a visual acuity of 4.9 in one or both eyes were offered glasses, but only 6.8% (17 of 249) accepted. In contrast, 56.5% (109 of 193) of those with a visual acuity of 4.8 in one or both eyes accepted glasses. Since the definition of poor vision is somewhat arbitrary, this suggests defining poor vision as below 4.9, instead of below 5.0. Also, the low take-up for children with an acuity of 4.9 prevents estimation of the program impact on those children.

[7] However, the evidence is not unanimous; Lu et al. (2009) find no relationship between near-work and myopia.

[8] There are 703 students with bad vision in the full sample of County 1 control schools but only 112 in the compliant sample, as one of the three County 1 control townships that improperly implemented the intervention was very large.

## 4. Methodology

Almost all primary school age children in Gansu province are in school; the GSCF data show that only 1.4% of children age 9-12 in 2000 were not in school. Thus, providing eyeglasses cannot increase enrollment; the sole impact is on academic performance. The random assignment of schools to participate or not participate in the GVIP greatly simplifies analysis of that project's impact on student learning. To ease interpretation, all estimates in this paper use test scores that are standardized by subtracting the control schools' mean and dividing by the control schools' (student level) standard deviation, separately for each subject and grade.[9]

The simplest estimate of the program's impact on students with poor vision is a t-test that compares the mean test scores of students with poor vision in the program schools with the same mean for their counterparts in the control schools. This estimates the impact of *offering* eyeglasses (intent to treat effect), not the impact of receiving them. This can be done by regressing the (standardized) test score ($T_{is}$) on a constant and a binary variable for enroling in a program school ($P_s$):

$$T_{is} = \alpha + \beta P_s + u_{is} \qquad (1)$$

for student i in school s, where the residual $u_{is}$ is uncorrelated with $P_s$ due to random program assignment. Reflecting the sample design, all regressions include a dummy variable for each pair of townships within which randomization was done (not shown in equation (1)). Bruhn and McKenzie (2009) provide a justification for adding strata dummy variables.

Equation (1) uses only students with poor vision. More precise estimates of β can be obtained by adding students with good vision. This "double difference" method compares the difference in test scores between students with good vision and poor vision in treatment schools to

[9] This may be misleading since each school administered a different test. Yet the test score data range from 10 to 100 points for each test, and school level medians are close to 80 (between 72 and 87 for 80% of the schools, with smallest being 61). Also, school level standard deviations are about 10 (between 7 and 14 for 80% of the schools). Thus our normalization is not forcing schools with very different distributions onto a similar scale. To check robustness, we normalized test scores using within-school means and standard deviations; the results are very similar.

the same difference in control schools. Another advantage of adding these students is that it compares only students who took the same test, as it is based on within-school comparisons.

The equation to be estimated for this specification is:

$$T_{is} = \alpha + \pi PV_{is} + \tau P_s + \beta PV_{is}*P_s + u_{is} \qquad (2)$$

where $PV_{is}$ is a variable indicating poor vision. In this specification the program's impact on students with good vision ($PV_{is} = 0$) is $\tau$, which should be zero unless there are spillover effects onto these students, and the program's impact on students with poor vision is $\tau + \beta$, which equals $\beta$ if $\tau$ equals zero. If no spillovers exist, the $\tau$ coefficient is a check on the randomization; if the schools assigned to the program were better (worse) than average, then $\tau$ would be positive (negative). Finally, $\pi$ measures the impact of poor vision on test scores, which should be negative. Yet this estimate will be biased toward zero since students who study more tend to have worse vision. Fortunately, neither correlation between u and PV nor random measurement error in PV lead to bias in the estimate of the program impact ($\beta$). (This is explained in Online Appendix I.)

Adding other explanatory variables to equations (1) and (2) could lead to more precise estimates. Several child and parent variables were tried, but none increased precision. In contrast, adding students' test scores in the spring of 2004, before eyeglasses were provided, greatly increases the precision of the estimated program impacts.

A final issue is the correct standard errors for the estimated program effects. They should allow for heteroscedasticity of unknown form, and for correlation in u across students in the same schools, and across students in different schools in the same townships. Indeed, schools typically use their own tests, not county or province tests; Online Appendix 2 shows that this generates a school level random effect but does not lead to an inconsistent estimate of $\beta$.

The best approach to address this correlation is to use covariance matrices that allow for clustering of the error terms. Yet for this paper the standard clustering formula has two disadvantages. First, estimation of equations (1) and (2) that allows for correlation of unknown form at the township level loses information, leading to less precise estimates, because the covariance matrices do not distinguish between students in the same school and students in different schools in a given township. Unobserved school effects imply that the error terms are likely to be more strongly correlated for the first set of students. To allow for this differential correlation, we estimate specifications with school random effects, which distinguish between students in the same school and those in different schools, *and* we allow for correlation of unknown form for the error terms of students in the same township. This yields correct inference even if the errors in equations (1) and (2) do not follow the classical random effects form (Wooldridge, 2010, pp.866-67).

The other disadvantage is that covariance matrices that allow for clustered errors are valid only as the number of clusters (townships) goes to infinity. Our estimates that drop township pairs with a township that did not follow its random assignment are based on 25 townships. Such covariance matrices can be misleading if there are 30 or fewer clusters (Cameron et al., 2008). For robustness, we present p-values estimated using the wild bootstrap, as Cameron et al. suggest.

The discussion so far has focused on the impact of being offered glasses, not the impact of receiving them. Instrumental variable (IV) estimation can yield consistent estimates of the impact of receiving glasses. This is done for equations (1) and (2) by replacing $P_s$ (offer of glasses) with "$G_{is}$", actually receiving glasses.[10] $G_{is}$ may be correlated with the residual, but $P_s$ can be used to instrument $G_{is}$; $P_s$ is, by definition, uncorrelated with $u_{is}$, and has strong predictive power for $G_{is}$.[11]

[10] Technically, the IV estimates are local average treatment effect (LATE) estimates: the impact of wearing glasses for those students the program induced to wear them. Yet since few students had glasses before the program, LATE estimates are very close to the impact of glasses on those who received them: average treatment on the treated (ATT).

[11] $G_{is} = 1$ for program school students who accepted glasses *and* for any student in any school who already had glasses.

There is one complication with IV estimates of equation (2); replacing $P_s$ with $G_{is}$ yields $T_{is} = \alpha + \pi PV_{is} + \tau G_{is} + \beta PV_{is}*G_{is} + u_{is}$. Although one can be in a program school if one does not have poor vision, it makes little sense for such people to wear glasses, thus $G_{is} = 0$ when $PV_{is} = 0$, and so $G_{is}$ and $PVis*G_{is}$ are perfectly correlated.[12] Thus IV estimates must exclude the $\tau G_{is}$ term. Note also that IV estimation is valid even if the randomization was incorrectly implemented; the *planned* randomization still satistifies the exclusion restriction. A final limitation of IV estimation is that the performance of the wild bootstrap has not been verified for IV estimation.

**5. Checks for Treatment/Control Balance and for Selectivity into the Compliant Sample**

Before presenting estimated program impacts, we check whether the treatment and control townships in the compliant sample are similar, which one would expect since the treatment was randomly assigned. We then check whether the compliant sample shows signs of selection bias.

Table 3 assesses whether the treatment and control townships are well balanced for the 25 compliant townships. Results are presented separately by county; combining both counties yields similar results. The treatment and control means for 10 key variables are in the first and second columns, respectively, and the third reports the difference, with asterisks indicating statistical significance. For County 1 students, none of the ten differences is significantly different at the 5% level, although two are for the bootstrapped p-values. While this is more significance than expected, which is one variable significant at the 10% level, a joint test based on regressing the treatment variable on these 10 variables cannot reject the null hypothesis that none of these variables has explanatory power (p-value of 0.452). Limiting the sample to students with poor vision in County 1 yields similar results (see Online Appendix Table A.1).

[12] This correlation is not exactly equal to one in the data (it is 0.86); this occurs because a very small percentage of students report wearing eyeglasses even though they have good vision.

The bottom half of Table 3 presents similar comparisons for County 2. None of these 10 differences is significant, indicating a well-balanced random assignment, although the analogous joint test suggests some differences (p-value of 0.019). Focusing only on children with poor vision also suggests some differences (see Online Appendix Table A.1). Overall, the Table 3 results are consistent with random assignment in County 1, but perhaps less so in County 2. This is also the case for balance checks including all 37 townships; see Online Appendix Table A.2.

Unfortunately, six of the 37 townships implemented the program incorrectly, affecting six township pairs. In five pairs, 34% to 72% of students with poor vision in control group schools were offered, and accepted, glasses. In the sixth, glasses were not offered to students in the treated township but were offered, and accepted, by 33% of poor vision students in the control township.

This offer of glasses to a third of the control group causes underestimation of the intent to treat (ITT) effect (impact of offering glasses to students with poor vision) of the program for estimates that use all 37 townships because offering glasses to control group students raises their test scores (if eyeglasses increase learning). While one could argue that this estimates the (ITT) effect of the program as actually implemented, most violations of random assignment were offers of glasses to the control group, an error that cannot occur in a "real" implementation of the program, which has no control group. In contrast, estimates based on the 25 townships that followed random assignment are unbiased as long as there are no systematic differences between the full sample and the 25 townships that correctly implemented the program.

Online Appendix Table A.3 examines whether key variables have systematic differences between the 25 compliant and 12 "non-compliant" townships, separately for the two counties. (Estimates combining both counties are similar.) The results for County 1 (top half of table) show almost no statistically significant results, which suggests no difference between the compliant and non-compliant townships. Of the 19 differences for County 2, none is significant at the 1% level, but two are at the 5% level, close to what one can expect if the true values are all zero, and wild p-values indicate that the two significant differences may be significant only at the 10% level.

Another check of whether analysis of the 25 compliant townships leads to selection bias is to consider the 18 townships in the full sample that were randomly assigned to the control group. The issue is whether some school or student characteristics among the 105 schools in these 18 townships make some schools more likely to provide glasses (and thus violate random assignment) than other schools. For example, if control schools with wealthier or with better educated parents, or fewer ethnic minority students, are more likely to pressure officials to provide their children glasses, that may lead to selection bias. This was checked using school level regressions for these 105 schools. None of these has any influence on which control schools were (mistakenly) provided eyeglasses. Only two variables have predictive power: control schools with a high percentage of students with poor vision, and control schools paired (at the township level) with treatment schools that tended to have fewer children with poor vision, were more likely to be provided glasses. This is consistent with field reports that non-compliance was due to local officials using funds left over (after providing glasses to students with poor vision in the treatment schools) to provide glasses to those with poor vision in the control schools.

A last check for selection bias is to compare the difference in the baseline means of the test scores between the treatment and control townships for the compliant townships with the same difference for non-compliant townships. The concern is that the program impact could be underestimated if random assignment was violated mainly in township pairs where the treatment township had higher test scores than the control township (since this behavior implies that, in the compliant sample, the treated townships would tend to have lower scores than the control townships). In fact, the difference in these two differences is statistically insignificant for all test scores, and for the average test score (results available from the authors).

## 6. Estimates of Program Impact

This section presents estimates for the 25 compliant townships of the impact of the GVIP on the test scores of students in grades 4-6 in the spring of 2005 (results for all 37 townships are in the online appendix). These results measure the impact of the project after one academic year. As above, all test scores have been normalized separately for each subject and grade. To increase precision, all estimates condition on pre-intervention (spring 2004) test scores. Estimates without lagged scores as controls are similar but somewhat less precise (Online Appendix Table A.4).[13]

Table 4 presents estimates of equations (1) and (2) for the 25 townships that implemented the program correctly, for the pooled sample and by county.[14] All estimates include school random effects and allow for correlation (clustering) of unknown form at the township level for the individual level error term. Coefficients on lagged test scores, constants, and strata indicators are not reported to avoid clutter in the table.

Combining both counties, the estimated treatment effect on all test scores is 0.16 (0.11) standard deviations for the sample of poor vision children (all children). Both estimates are statistically significant using the non-bootstrapped standard errors, but the wild bootstrap yields significance at conventional levels only for the sample of all children (p-value of 0.046, compared to 0.118 for the sample of poor vision children). Surprisingly, the estimated impacts are much larger in County 1 than in County 2. For County 1, the estimated impact for the sample of children with poor vision is 0.39 standard deviations, which is highly statistically significant even when using the wild bootstrap (p-value of 0.026). Adding children with good vision, the estimated impact is 0.26 standard deviations, which is also highly significant. In contrast, the

[13] Estimates that add other controls (in addition to lagged scores), namely sex, child health (hemoglobin levels and height-for-age), ethnic group indicators, household head occupation and education, and grade dummy variables, did not affect the results.

[14] Estimates of equation (1) classify students whose worst eye has a visual acuity score of 4.9 as having good vision. Yet recall that such students were offered glasses, and 17 out of 249 accepted them. Those 17 are excluded from the regression. Dropping all 249 of these students from the sample does not affect the results.

estimate for County 2 is 0.08 (0.07) standard deviations and statistically insignificant for the sample of poor vision children (all children).

Given the strong impacts for County 1, we also investigate impacts for each subject (all County 2 subject-specific impacts are statistically insignificant). The impact is greatest for Chinese, 0.41 (0.29) standard deviations using the poor vision children (all children) sample, and smallest for science, 0.26 (0.15) standard deviations for the poor vision children (all children) sample. Yet these differences in the estimates across subjects are statistically insignificant. Using conventional clustered standard errors, all impacts on subject-specific scores are statistically significant; for the wild bootstrap, estimated impacts are significant for Chinese (for both samples, children with poor vision and all children) and science (only for the all children sample).

If one views the County 1 results as the true potential effect of offering glasses to children with poor vision, they can be regarded as the preferred estimates. Yet if one gives equal weight to the estimates in each county, the preferred estimates are those that pool the data from both counties. Much depends on why the County 2 results are insignificant; this is discussed in Section 8.

The compliant sample estimates are our preferred estimates. Online Appendix Table A.5 shows results for the full sample of 37 townships. As explained above, we expect that these underestimate the true program impacts; adding one "treatment" township that did not offer the treatment and six "control" townships that offered the treatment to some of their students will reduce the gap in test scores between the townships randomly assigned to be treated and those randomly assigned to be controls (if eyeglasses have an effect). The Online Appendix Table A.5 results are as expected; each intent to treat estimate is less than the corresponding estimate in Table 4. Nonetheless, it is worth noting that we still find some statistically significant impacts of the offer of eyeglasses in County 1. For non-bootstrapped standard errors, the estimated impacts using the sample of poor vision children are significant at the 5% level for Chinese and math scores. Wild-bootstrap p-values are significant at the 10% level for math test scores and very close to significant at the 10% level (p-value of 0.108) for Chinese scores for the sample of poor vision children.

A final point regarding the estimates in Table 4 is that they can underestimate the impact of providing eyeglasses if parents reallocate educational spending and efforts from a child who received glasses to their other children. Yet recent research by Shi (2012), who also uses data from Gansu, suggests that such intra-household substitution is unlikely. To the extent that there is intra-household substitution we estimate the impact on treated children after such behavioral effects, and thus we may be overlooking benefits that spilled over onto other family members.

IV estimates of the impact of wearing eyeglasses for one year on student test scores, for both the 25 compliant townships and the full set of 37 townships, are in Online Appendix Table A.6.[15] The (planned) randomization is a valid instrument for wearing glasses, so estimates for both the compliant and the full samples are consistent. Yet the compliant sample is still preferred because its estimates are more precise. As expected, all the estimated impacts are larger than the ITT estimates because they reflect the impact of actually receiving, rather than just being offered, eyeglasses.[16] Pooling both counties' compliant samples, the impacts on test scores are 0.22 (0.16) standard deviations for the sample of poor vision children (all children), and significant at the 5% level. Disaggregating by county, for County 1 all estimated impacts for the compliant sample are statistically significant. In contrast, as with the ITT estimates all of the County 2

[15] Some students had worn eyeglasses for more than one year; of the 1,245 children with glasses, 199 had obtained them on their own, of whom 94 obtained them one year ago, 85 obtained them two years ago, and 20 obtained them 3 or 4 years ago, so only 105 of the 1,245 children had them for more than one year. Recall that only 59 children in the sample with bad vision had glasses; thus 140 of the 199 children who report having obtained eyeglasses on their own do not appear to have had bad vision. This could reflect a misdiagnosis that led their parents to obtain glasses for them, or measurement error either in the variable indicating wearing eyeglasses. Measurement error in reported wearing eyeglasses does not imply inconsistency since that variable is instrumented.

[16] As noted above, the IV estimates' p-values cannot be corrected by the wild bootstrap. Yet for the (preferred) compliant sample note that ITT estimates' statistical significance is generally robust to using the bootstrap, and the statistical significance of the non-bootstrapped IV estimates and the non-bootstrapped ITT estimates is similar.

estimates are much smaller and statistically insignificant. In general, IV estimates are about 1.5 times larger than the corresponding ITT estimates; this is consistent with the experience that only about two thirds of the students who were offered glasses accepted them.

One can express the treatment effect in terms of an equivalent gain from additional time in school. The 2000 GSCF administered identical Chinese and math tests to students in grades 4, 5 and 6. Relatively few were in grade 6, so we focus on grades 4 and 5. The mean test scores of grade 5 students were 0.37 standard deviations higher in Chinese and 0.51 standard deviations higher in math than the grade 4 students' mean scores. Comparing the average gains on these two tests (0.44) with the estimated gains from wearing glasses of 0.16 to 0.22 in the two counties together, the impact of wearing glasses is equivalent to one third to one half of a year in school. Put another way, giving glasses to students with poor vision raised learning per year by 33-50%.

We end this section with two robustness checks of the Table 4 results. First, estimates of equation (2) in Table 4, which compare students with poor vision to those with good vision, assume that providing the former glasses does not affect the latter's test scores. This can be checked by examining the impact of being in a treatment township ($\tau$) in equation (2), as shown in the estimates for all students in Table 4. All estimates are very small and far from statistically significant, indicating that the program did not affect students with good vision.[17] Second, while unlikely, it may be that something else occurred in treatment schools around the same time that raised those schools' test scores. Online Appendix Table A.7 re-estimates the specification for equation (2) in Table 4, but does it using data from a year earlier. If "something else" were happening, one could find an "effect" even before the program began. Yet no evidence is found of such an effect; averaged over all three tests, the estimated "effect" is tiny (-0.02) and insignificant.

[17] We also estimated equation (1) for "good vision" students only; the program impacts are almost identical to the estimates of ($\tau$) in Table 4, as one would expect. These are available from the authors upon request.

## 7. Heterogeneous Treatment Effects

The impact of providing eyeglasses may vary over students. Indeed, we find evidence of heterogenous treatment effects by initial (2004) test scores, nutritional status, and parents' education. Differences in average student characteristics in terms of nutritional status and parents' education explain about half of the stronger impacts found in County 1, relative to County 2.

Perhaps the most obvious dimension along which the impact of glasses would vary is by students' visual acuity. The first column of Table 5 examines whether those with very bad vision benefit most from eyeglasses. Among students with poor vision (visual acuity < 4.9), we define very poor vision as visual acuity below 4.4; about 20% of students with poor vision have very poor vision. The first set of results uses only students with poor vision; it finds a positive program impact but no additional impact on students with very poor vision. Indeed, the additional impact is negative, though far from significant. Adding students with good vision to the regression (equation (2)) gives a similar result. Thus there is no evidence that students with very poor vision benefit more from the program. This could reflect compensatory behavior by some children with very poor vision, e.g. sitting in the front of the classroom, that reduces the impact of the degree of poor vision on learning, or it may be that eyeglasses do not fully correct the problems of those with very poor vision. This (lack of a) result may also simply reflect imprecision due to the relatively low number of students with very poor vision (only 20% of those with poor vision).

Another possible heterogeneity in program effects is by initial performance; students with poor vision *and* relatively low academic performance may experience greater learning than those with poor vision and average or above average academic performance. Alternatively, high ability students, as measured by initial test scores, may benefit more from better vision. This is examined in the second column of Table 5. When only students with poor vision are included, the impact is lower for those with higher initial (2004) test scores, but this negative interaction term is insignificant. Adding students with good vision to the regression yields a larger and more precisely estimated (triple) interaction effect, which is significant at the 1% level (with a bootstrapped p-value of 0.014). The average 2004 score was normalized to zero, so these estimates imply that average students experience a 0.11 standard deviation increase, while below average students (defined as those with a 2004 average test score one standard deviation below the mean) had a 0.27 standard deviation gain, and above average students (those with a 2004 average score one standard deviation above the mean) had a small loss of 0.06 standard deviations.[18] Thus providing eyeglasses appears to equalize educational outcomes among students with poor vision.

Next, we examine heterogenous impacts with respect to child nutrition and parents' education. Regarding child nutrition, Vitamin A deficiency (VAD) can lead to vision problems.[19] VAD is uncommon in most of China, but Greiner et al. (2001) report that 39% of preschool age children in Gansu suffer from VAD, by far the highest rate for the 10 provinces in that study (the others had rates from 4% to 18%). A more recent study (Zhang et al., 2011) of rural areas in six Chinese provinces found that Gansu had the highest VAD rate (25.5%).

Night blindness (nyctalopia) is the most common vision problem due to lack of Vitamin A. It reduces vision when lighting is dim, and wearing glasses does not alleviate it.[20] Night blindness can hamper studying in dwellings with inadequate lighting, a common problem in Gansu. Indeed, the 2004 GSCF data indicate that only 62.9% of children age 13-17 live in homes with "good" or "very good" lighting, while 33.1% live in homes with "so so" lighting and 4.0% are in homes with "bad" or "very bad" lighting. Also, 21.7% of GSCF children report "pain in eyes while studying because of dim light" (63.6% of whom are in homes with "so so", "bad" or "very

---

[18]These calculations use the fact that, for students with poor vision, the average 2004 test score was -0.16, and the standard deviation was 0.97. So, for example, the impact for an average student is $0.081 - 0.166\times(-0.16) = 0.108$.
[19] See http://en.wikipedia.org/wiki/Vitamin_A_deficiency (accessed April 14, 2015) for a general discussion of VAD and associated vision problems. We thank Nathan Congdon for advice on VAD, especially in the context of China.
[20] The most common vision problem addressed by our intervention is myopia; it is unlikely to affect studying at home since it is the inability to see relatively distant objects, not difficulty in seeing objects within arm's length.

bad" lighting). Despite this potential role for VAD, especially night blindness, we know of no studies that explicitly examined its impact on student learning.

Unfortunately, the GVIP collected no data on VAD or night blindness. Yet it did collect data on two conditions correlated with VAD: hemoglobin levels and height-for-age. Note that Semba and Bloem (2002) find evidence that VAD reduces hemoglobin levels, and Hu et al. (2001) find that children with VAD grow more slowly. There is no evidence that either low hemoglobin or slow growth (stunting) affect children's vision directly, so any relationship between these two and the impact of providing glasses may reflect their being an indicator of VAD, and especially night blindness.

Parental support can also affect the impact of providing glasses. In the absence of glasses, parents of children with vision problems who are more educated and wealthier are likely more able to help their children with their studies, so that providing glasses may have less impact on children of educated and/or wealthy parents. While an opposite effect is possible, that more educated parents better ensure that their children have and wear eyeglasses, Section 8 shows that educated parents are not more likely than less educated parents to accept free glasses.

Tables 6 and 7 present evidence that differences in child health and parents' ability to help their children with schoolwork may explain part of the differential effect across the two counties. Table 6 shows that students in County 1, where glasses had a large impact, have higher average hemoglobin levels and height-for-age Z-scores. This suggests less VAD in County 1, supporting the hypothesis that glasses had a larger impact on vision – and learning – in County 1 because of a lower incidence of VAD among students. Table 6 also shows that the household head's years of schooling is 1.2 years lower in County 1 than in County 2. If glasses are substitutes for parents' education (and for parents' income, which is correlated with their education) in the learning production function, this may also explain, in part, the larger impact of glasses provision in County 1.

Regression analysis in Table 7 indicates that variation in hemoglobin levels, height-for-age and parental education explains about half of the differential program impact across the two counties. The results for all students in Table 4 show a difference in the estimated program impact of 0.184 (0.257 - 0.073). The first column in Table 7 verifies this by adding to equation (2) interactions of three variables with a County 1 dummy variable: poor vision, treated township, and their interaction. The triple interaction coefficient, 0.205, is both significant and close to 0.184.

The next column of Table 7 adds three variables – hemoglobin, height-for-age Z-score and the household head's years of education – to the regression, as well as their interactions with the poor vision and treatment program dummy variables, and their interaction. The three "triple interactions" measure whether the program's impact on students with poor vision varies by these three variables. The signs support the above hypotheses: the program's impact on students with poor vision is larger for students with higher hemoglobin levels and higher height-for-age Z-scores, and lower for students whose household head (usually the father) is more educated. Two of these triple interactions are significant at the 5% level, and a joint test decisively rejects the hypothesis of no impact of the three triple interactions (p-value of 0.002). Thus students who are less likely to have VAD (those with higher hemoglobin levels and higher height-for-age Z-scores) and who have less educated parents benefit more from the provision of eyeglasses.

The last column in Table 7 considers how much of the difference across the two counties is due to these differential effects by hemoglobin levels, height-for-age and household head's education. It adds the first column County 1 interaction terms to the second column regression. The results suggest that about half of the difference reflects differences in these three variables. In particular, the statistically significant triple interaction effect in column 1 falls from 0.205 to 0.102 and loses significance, while the triple interactions in column 2 retain their significance.

It is also possible that the program and/or data collection were not properly implemented in County 2.[21] First, in County 2 data were collected in a decentralized way, with electronic Excel files sent to schools to be filled in by teachers or school officials and returned to the CDC. This may have reduced data quality, as teachers received little training and lacked strong incentives to collect the data carefully. In contrast, in County 1 all data were collected by a small group (6 to 8) of CDC professional staff trained by the authors, and their work was monitored by county and provincial CDC staff. Indeed, the initial data files received had many more problems for County 2. Second, program implementation may also have been superior in County 1; County 2 is more populous, which made monitoring by the CDC more difficult in that county. [22]

## 8. If the Benefits Are So Large, Why Do Some Children not Wear Eyeglasses?

The glasses provided by the GVIP cost about 120 yuan (about $15 U.S.). Our estimates of their impact on learning after one year is equivalent to one third to one half of a year of schooling, which should lead to higher wages when these students are adults. De Brauw and Rozelle (2007) estimate that a year of schooling in rural China raises wages by 9.3% for those under 35 years old. Our estimate based on a Mincerian wage function using data on individuals aged 15 to 35 from Wave 2 (in 2004) of the GSCF is lower: 4.6%. The GSCF data also indicate that a wage earner age 15-25 who completes lower secondary (grade 9) earned about 710 yuan per month, that is 8,520 yuan per year. Using our lower estimate of the impact of schooling on wages, and assuming that the program effect is equivalent to only one third of a year in school, the program should

[21]It is difficult to imagine plausible reasons why such problems would overestimate program impacts. Manipulation of data to find stronger positive impacts in County 1 is implausible, as pre-treatment test scores were collected from school records before eyesight was tested or eyeglasses provided, and the post-treatment test scores were also collected from school records without reference to the treatment status of the students (all test scores were collected from administrative records). Moreover, there was no incentive for the data collectors in County 1 to exaggerate program benefits, and the intervention rewarded neither schools nor teachers for student performance.

[22] Checks of the data used in the analysis revealed no large differences in obvious errors between the two counties. Some data anomalies, such as 2005 test scores identical to 2004 test scores, were more common in County 2. But others, such as data indicating that children with good vision received eyeglasses, were more frequent in County 1.

increase such a wage earner's annual income by 128 yuan (8520×0.33×0.0456). Assuming that this person works for 40 years after finishing grade 9, the present discounted value (PDV) of this wage increase easily exceeds the cost of glasses; using a 10% discount rate, the PDV is 830 yuan, and using a more plausible 5% rate yields a PDV of 1,834 yuan.

These large benefits from eyeglasses, relative to their cost, and the many refusals of free glasses[23] and almost no wearing of glasses absent the intervention, suggest a failure to make a high-return investment. Understanding this failure can have important policy implications. Is the cost of glasses too high, especially for the poor, who may be credit-constrained? Even if offered at no cost, parents may hesitate since accepting the offer may create an obligation to purchase glasses in later years if the original pair is lost or broken, or if the prescription needs updating.

Alternatively, parents may not know of their children's vision problems, or may believe (incorrectly) that glasses will weaken their children's eyes or that poor vision has little effect on learning at a young age. Even if told that their child needs glasses, parents may doubt this, or think that the vision problems are minor and so need no correction. Other parents may view glasses as useful only for schooling, and may have low educational aspirations for their child. Community norms may also influence parents' value of eyeglasses and education. To explore these hypotheses, we estimate which children accept the GVIP's eyeglasses, and use the 2004 GSCF to estimate the determinants of wearing glasses absent an intervention.

Table 8 presents probit estimates of the factors associated with accepting the GVIP's offer of free eyeglasses. First, as one would expect better visual acuity (average over both eyes) has a significantly negative impact on accepting glasses. The standard deviation of the visual

[23] As explained above, only about 70% of the students with poor vision in the program schools accepted the eyeglasses, even though they entailed no cost. The stated reasons for not accepting them are not very informative, the two most common being "child refused" and "household head refused" (see Table 2).

acuity variable is 0.234, so raising visual acuity by a standard deviation reduces the probability of accepting glasses by 11.2 percentage points (0.234×0.479).

One unusual result is that girls accepted eyeglasses less often than boys: 66.0% vs. 73.6%. In particular, they have a significant 8.1 percentage point lower probability of accepting glasses. The reasons for this are unclear; boys' and girls' stated reasons for refusing glasses are similar. Anecdotal evidence suggests that girls may worry more that glasses make them less attractive.[24]

Four other factors have significant impacts. First, the few children with poor vision who already had glasses were 17.5 percentage points more likely to accept new ones, as expected since they likely valued glasses and may have needed new prescriptions. Surprisingly, children in households headed by a schoolteacher or a village cadre were less likely to accept glasses. These effects are large: 22 percentage points less for schoolteachers and 33.6 percentage points less for cadres. Perhaps these authority figures doubt the value of glasses. Lastly, students in wealthier townships were more likely to accept eyeglasses.

Further insights are obtained from the 2004 GSCF data. We examine 925 children in primary school (and between age 8 and 15) in that year; 413 were GSCF "index" children and 512 were younger siblings of index children. These data contain much more detail, including vision-related information, than do the GVIP students' school records. The GSCF data indicate that parents are often unaware of their children's vision problems. Mothers were asked to assess their child's vision using five categories, from very good to very bad. As seen in Table 9, nearly all (86%) mothers of children with good vision (as measured by optometrists) correctly report that their child had good or very good vision. Yet 82% of those whose children had fair vision, and 62% whose children had poor vision also felt that their child had good or very good vision.

[24] Li et al. (2010) find that worries about appearance are not a major reason why children in rural China with vision problems do not wear glasses. The most common worry is that wearing glasses causes further deterioration in vision.

Findings are similar when children are asked about vision problems. In Table 10, children with good or fair vision rarely report vision problems (difficulty seeing the blackboard in school, trouble doing homework due to poor vision, and eye pain when studying in dim light at home). Children with poor vision report problems more often – 30.4% cite difficulty seeing the blackboard in school, 26.1% report trouble doing homework due to poor vision, and 29.0% cite eye pain when studying in dim light at home – yet for each about 70% report not having the problem.

Regression analysis of the GSCF data sheds light on almost all of the above hypotheses. Of the 925 8-15 year old primary school children in the data, 23 (2.5%) report wearing glasses. The following variables from the GSCF data can be used to assess these hypotheses: mothers' and fathers' assessments of their child's vision; mothers' estimates of the cost of glasses and the distance to the nearest locality that sells glasses; parents' wearing of glasses; community literacy rates; and parental aspirations for their child's education.

Table 11 reports results for five regressions. Since few children wear glasses, the marginal effects of changes in the explanatory variables are small in percentage point terms. Yet the results are highly suggestive of factors that affect the wearing of eyeglasses. We focus on the statistically significant results, and report marginal effects of the fullest specification in the last column.[25]

The first regression, the most parsimonious, shows that children's visual acuity has a significantly negative impact on the probability of having glasses, as expected. Unlike Table 8, child sex has no effect. Older children are more likely to report having glasses; this may seem to reflect that such children have worse vision, but the regression controls for visual acuity so this may reflect more parental acceptance of eyeglasses for older children. Mothers', but not fathers', education has a strong positive impact on having glasses. Finally, children in better off households are more likely to have glasses. Column 2 in Table 11 considers whether lack of awareness

[25] Adding parental aspirations (not shown) reduces the sample size, and these variables are generally insignificant.

of vision problems reduces children's probability of having glasses. Mothers who think their child has poor vision are more likely to provide glasses, but fathers' assessments have no effect. Note that the impact of the child's visual acuity falls from -1.33 to -0.86, suggesting that mothers' perceptions matter as much as actual acuity, and that mothers may not know their child's acuity.[26]

Column 3 examines whether perceived price and distance dissuade parents from obtaining glasses for their child. Price has the expected negative effect (significant at 10% level), yet distance has no significant effect. An interaction between price and per capita expenditure was insignificant. Column 4 in Table 11 shows that, as expected, having a parent who wears glasses has a strong positive effect on the probability that a child has glasses. Finally, Column 5 in Table 11 examines whether community characteristics have any effects. Only the community literacy rate, an indicator of the value placed on education by the community, has a significant impact, increasing a child's probability of having glasses.

## 9. Summary and Conclusion

Vision problems affect about 10% of primary school age children in both developed and developing countries. Fortunately, most vision problems can be corrected by properly fitted eyeglasses. In almost all developed countries, public programs pay for children's eye exams and provide free eyeglasses to poor children. Yet in developing countries very few children with vision problems have glasses, and they are rarely assisted by public or private organizations.

This paper examines the impact of providing glasses to rural children with poor vision in Gansu, one of China's poorest provinces. A randomized controlled trial was implemented in two counties in Gansu, where 13% of students in grades 3-5 had poor vision. Our estimates indicate that offering glasses to students with poor vision raises their test scores by 0.11 to 0.16 standard

[26] A related issue is whether parents think that glasses deteriorate their child's vision. The 2004 GSCF has no data on this, but the 2007 GSCF asked a new sample of mothers; about 25% opined that glasses worsen children's vision.

deviations of the distribution of those scores. This is an average over both counties; the impact was larger in County 1 (0.26 to 0.39 standard deviations) but small and insignificant in County 2.

About one third of the children, or their parents, did not accept the offer of free eyeglasses. Thus the impact of actually wearing glasses is about 50% higher, which is confirmed by instrumental variable estimates. The average estimates over both counties are large, equivalent to one third to one half of a year of schooling. Simple calculations suggest that the benefits in terms of higher wages greatly exceed the costs. Thus provision of eyeglasses is a low cost, easily implementable intervention that could raise learning among a substantial proportion of primary (and secondary) school students in developing countries.

Our finding that providing glasses to children with poor vision increases their test scores is not surprising, although the impact likely varies across contexts. The more important questions are: 1. What determines heterogeneity in the impact of providing eyeglasses? and 2. Why do some parents not obtain eyeglasses for their children who need them?

Our results suggest heterogeneity in impacts by students' nutritional status and parental education. In addition, among students with poor vision, weaker students, as measured by pre-program test scores, appear to benefit more from glasses than stronger students, which suggests that providing eyeglasses may equalize education outcomes. Regarding the second question, our estimates suggest that parental misperceptions, especially their belief that their child's eyesight is adequate, play a major role. There is also evidence that low income and perceived high prices reduce the wearing of eyeglasses. Yet more research is needed to understand both heterogeneity in the impact of eyeglasses and parents' choices regarding eyeglasses for their children. Answers to these questions will improve the design of policies to ensure that all school age children in developing countries who can benefit from eyeglasses will have them.

## References


Afridi, Farzana. 2011. "The Impact of School Meals on School Participation: Evidence from Rural India", *Journal of Development Studies* 41(11): 1636-1656.

Angle, John, and David Wissmann. 1980. "The Epidemiology of Myopia." *American Journal of Epidemiology* 111(2):220-228.

Barro, Robert. 1991. "Economic Growth in a Cross-Section of Countries," *Quarterly Journal of Economics.* 106(2):407-443.

Bruhn, Miriam, and David McKenzie. 2009. "In Pursuit of Balance: Randomization in Practice in Development Field Experiments." *American Economic Journal: Applied Economics*. 1(4):200-232.

Bundy, Donald, Arun Joshi, Megan Rowlands and Yung-Ting Kung. 2003. "EnVISIONing Education in Low Income Countries." The World Bank. Washington, D.C.

Cameron, Colin, Jonah Gelbach and Douglas Miller. 2008. "Bootstrap-Based Improvements for Inference with Clustered Errors." *Review of Economics and Statistics* 90(3):414-427.

de Brauw, Alan, and Scott Rozelle. 2007. "Returns to Education in Rural China" in Emily Hannum and Albert Park, eds., *Education and Reform in China* (London: Routledge).

*Gansu Statistical Yearbook*. 2005. Beijing: China Statistics Press.

Glewwe, Paul, Eric Hanushek, Sarah Humpage and Renato Ravina. 2013. "School Resources and Educational Outcomes in Developing Countries", in P. Glewwe, ed., *Education Policy in Developing Countries*. University of Chicago Press.

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.

Greiner, Ted, Yunguo Liu, Jingxiong Jiang and Diying Yan. 2001. "Vitamin A Deficiency and Iron Deficiency in China: Problems, Programs and Policies". Consultant Report. World Bank. http://www.global-breastfeeding.org/pdf/Vitamin%20A%20and%20iron%20in%20China.doc

Hannum, Emily, Jere Behrman, Meiyan Wang, and Jihong Liu. 2008. "Education in the Reform Era", in L. Brandt and T. Rawski, eds., *China's Great Economic Transformation*. New York: Cambridge University Press.

Hannum, Emily, Albert Park, and Kai-ming Cheng. 2007. "Introduction: Market Reforms and Educational Opportunity in China", in Emily Hannum and Albert Park, eds. *Education and Reform in China* (Oxford: Routledge).

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, Eric, and Ludger Woessmann. 2008. "The Role of Cognitive Skills in Economic Development." *Journal of Economic Literature* 46(3):607-668.

He, Mingguang,.et al. 2007. "Refractive Error and Visual Impairment in School Children in Rural Southern China." *Ophthalmology* 114(2):374-382.

Krueger, Alan, and Mikael Lindahl. 2001. "Education for Growth: Why and for Whom?" *Journal of Economic Literature* 39(4):1101-1136.

Li, Liping, et al. 2010. "Attitudes of Students, Parents, and Teachers Towards Glasses Use in Rural China". *Archives of Ophthalmology* 128(6):759-765.

Li, Wen, Albert Park, and Sangui Wang. 2009. "School Equity in Rural China", in E. Hannum and A. Park, eds., *Education and Reform in China*. New York: Routledge.

Lu, Bei, and others. 2009. "Associations between Near Work, Outdoor Activity and Myopia among Adolescent Students in Rural China: The Xichang Pediatric Refractive Error Study: Report Number 2." *Archives of Ophthalmology* 127(6):769-775.

Luo, Renfu and others. 2012. "Nutrition and Educational Performance in Rural China's Elementary Schools: Results from a Randomized Control Trial In Shaanxi Province". *Economic Development and Cultural Change* 41(4):1012-1025.

Ma, Xiaochen, and others. 2014. "Effect of providing free eyeglasses on children's outcomes in rural China: cluster randomized controlled trial." *British Medical Journal* 349: 5740-5752.

Mankiw, N. Gregory, David Romer and David Weil. 1992. "A Contribution to the Empirics of Economic Growth," *Quarterly Journal of Economics* 107(2):407-437.

Miguel, Edward, and Michael Kremer. 2004. "Worms: Identifying Impact on Education and Health in the Presence of Treatment Externalities", *Econometrica* 1(1):159-217.

Mutti, Donald, et al. 2002. "Parental Myopia, Near Work, School Achievement, and Children's Refractive Error." *Investigative Ophthalmology and Visual Science* 43(12):3633-3640.

National Bureau of Statistics. 2005. *Statistical Yearbook of China.* Beijing: China Statistical Press.

Shi, Xinzheng. 2012. "Does an intra-household flypaper effect exist? Evidence from an educational fee reduction reform in rural China". *Journal of Development Economics* 99(2):459-73.

Vermeersch, Christel and Michael Kremer. 2004. "School Meals, Educational Achievement, and School Competition: Evidence from a Randomized Evaluation". *World Bank Policy Research Working Paper No. 3523.* The World Bank. Washington, D.C.

World Bank. 2000. *China: Overcoming Rural Poverty.* The World Bank. Washington, D.C.

Wooldridge, Jeffrey. 2010. *Econometric Analysis of Cross Section and Panel Data*, second edition. Cambridge, MA: MIT Press.

Zhang, J.G., B. Zhang B, W. W. Du, and others. 2011. "Analysis of Vitamin A Deficiency Status among Children under Five Years Old in Poor Rural Areas of Six Western Chinese Provinces in 2009. *Chinese Journal of Epidemiology* 32(12): 1224-1226 (in Chinese).

Zhao, JiaLiang, and others. 2000. "Refractive Error Study in Children: Results from Shunyi District, China." *American Journal of Ophthalmology.* 129(4):427-435.

**Table 1: Descriptive Statistics by County**

| | 25 Compliant Townships | | | All 37 Townships | | |
|---|---|---|---|---|---|---|
| | County 1 | County 2 | Both Counties | County 1 | County 2 | Both Counties |
| Number of children in grades 4-6 in 2004-05 | 6,130 | 12,772 | 18,902 | 10,217 | 18,581 | 28,798 |
| Children with vision problems | 787 | 1,742 | 2,529 | 1,552 | 2,625 | 4,177 |
| | (12.8%) | (13.6%) | (13.4%) | (15.2%) | (14.1%) | (14.5%) |
| Of which: | | | | | | |
| Had glasses already | 23 | 36 | 59 | 68 | 67 | 135 |
| | (2.9%) | (2.1%) | (2.3%) | (4.4%) | (2.6%) | (3.2%) |
| Did not have glasses | 764 | 1,706 | 2,470 | 1,484 | 2,558 | 4,042 |
| | (97.1%) | (97.9%) | (97.7%) | (95.6%) | (97.4%) | (96.8%) |
| Test scores in spring 2004 (before the intervention): | | | | | | |
| Students without vision problems: | | | | | | |
| Chinese | 78.6 | 79.0 | 78.9 | 78.6 | 79.0 | 78.8 |
| Mathematics | 79.0 | 79.2 | 79.1 | 78.6 | 79.4 | 79.1 |
| Science | 80.6 | 80.8 | 80.7 | 80.2 | 80.7 | 80.6 |
| Students with vision problems: | | | | | | |
| Chinese | 77.1 | 78.7 | 78.2 | 77.9 | 79.2 | 78.7 |
| Mathematics | 76.8 | 79.2 | 78.5 | 77.5 | 79.8 | 79.0 |
| Science | 80.2 | 80.8 | 80.6 | 80.2 | 81.1 | 80.8 |

Notes:

1. Vision problem is defined as a visual acuity score < 4.9 in one or both eyes. As explained in the text, although the 298 children in the full sample for whom one or both eyes had a score of 4.9 were offered glasses, only 18 (6.0%) accepted the glasses, so the analysis focuses on children for whom one or both eyes had a score of less than 4.9.

**Table 2: Implementation of Gansu Vision Intervention Project**

| | 25 Compliant Townships | | | All 37 Townships | | |
|---|---|---|---|---|---|---|
| | County 1 | County 2 | Both Counties | County 1 | County 2 | Both Counties |
| Students in grades 4-6 in 2004-05 with vision problems | 787 | 1,742 | 2,529 | 1,552 | 2,625 | 4,177 |
| Of which: | | | | | | |
| In control schools | 112 | 889 | 1,001 | 703 | 1,496 | 2,199 |
| In program schools | 675 | 853 | 1,528 | 849 | 1,129 | 1,978 |
| Students in program schools who: | | | | | | |
| Accepted the offer to receive glasses | 417 | 649 | 1,066 | 521 | 863 | 1,384 |
| Did not accept the offer to receive glasses | 258 | 204 | 462 | 328 | 266 | 594 |
| Reasons given for not accepting glasses: | | | | | | |
| Household head refused | 54 | 91 | 145 | 65 | 122 | 187 |
| Child refused | 42 | 38 | 80 | 43 | 46 | 89 |
| Cannot adjust to glasses | 58 | 0 | 58 | 61 | 0 | 61 |
| Mixed astigmatism | 11 | 0 | 11 | 12 | 0 | 12 |
| Optometrist not available | 27 | 7 | 34 | 30 | 13 | 43 |
| Pathological change in Fundus | 33 | 30 | 63 | 36 | 43 | 79 |
| Eye problem cannot be corrected by glasses | 5 | 0 | 5 | 5 | 0 | 5 |
| Astigmatism | 1 | 0 | 1 | 1 | 0 | 1 |
| Vision not correctable | 0 | 19 | 19 | 0 | 22 | 22 |
| Child is handicapped | 0 | 2 | 2 | 0 | 2 | 2 |
| Missing | 27 | 17 | 44 | 75 | 18 | 93 |

Notes:

1. Vision problem is defined as a visual acuity score < 4.9 in one or both eyes.

**Table 3: Pre-Program Differences between Treatment and Control Groups**
**(25 townships where randomization was correctly implemented, separately by county)**

| *Variable* | *Treatment Mean* | *Control Mean* | *Difference* | *p-values for Differences based on Wild Bootstrap* |
|---|---|---|---|---|
| | | ***All Children, County 1*** | | |
| Chinese test | -0.133 | -0.336 | 0.203 | 0.543 |
| Math test | -0.112 | -0.308 | 0.196 | 0.608 |
| Science test | -0.032 | -0.386 | 0.354 | 0.472 |
| Average test | -0.112 | -0.416 | 0.304 | 0.868 |
| Ethnic minority | 0.388 | 0.495 | -0.106 | 0.016** |
| Visual acuity | 5.04 | 5.09 | -0.05 | 0.025** |
| Poor vision | 0.130 | 0.122 | 0.008 | 0.163 |
| Male | 0.542 | 0.561 | -0.019 | 0.864 |
| Head yrs educ | 8.62 | 7.45 | 1.17 | 0.461 |
| Age | 9.78 | 10.35 | -0.57 | 0.286 |
| Joint test (F-test [p-value]) | | | 1.06 [0.452] | |
| | | ***All Children, County 2*** | | |
| Chinese test | -0.058 | -0.158 | 0.100 | 0.462 |
| Math test | -0.079 | -0.085 | 0.005 | 0.703 |
| Science test | 0.001 | -0.107 | 0.108 | 0.555 |
| Average test | -0.055 | -0.141 | 0.086 | 0.561 |
| Ethnic minority | 0.004 | 0.010 | -0.006 | 0.795 |
| Visual acuity | 5.01 | 5.04 | -0.03 | 0.290 |
| Poor vision | 0.121 | 0.155 | -0.035 | 0.784 |
| Male | 0.530 | 0.531 | -0.001 | 0.586 |
| Head yrs educ | 9.23 | 8.73 | 0.50 | 0.301 |
| Age | 10.97 | 11.18 | -0.21 | 0.388 |
| Joint test (F-test [p-value]) | | | 3.88** [0.019] | |

Statistical significance of mean differences is based on regressions that include school random effects and account for clustering at the township level.
Asterisks denote statistical significance: * 10% level, ** 5% level, *** 1% level.

**Table 4: Estimated Program Effect After 1 Year: ITT Results, Compliant Sample**

| | *Dependent Variables* | | | | | |
|---|---|---|---|---|---|---|
| | Average Test Scores | | | Subject Scores (County 1 only) | | |
| *Explanatory Variables* | Both Counties | County 1 | County 2 | Chinese | Math | Science |
| **Equation (1): Compliant Sample, Only Students with Poor Vision** | | | | | | |
| Treatment Township (β) | 0.158** | 0.393*** | 0.079 | 0.413*** | 0.269** | 0.259** |
| | (0.078) | (0.125) | (0.094) | (0.124) | (0.124) | (0.114) |
| | [0.188] | [0.026] | [0.624] | [0.080] | [0.164] | [0.148] |
| Sample Size | 2,474 | 732 | 1,742 | 745 | 733 | 732 |
| **Equation (2): Compliant Sample, All Students** | | | | | | |
| Poor Vision (π) | -0.022 | -0.121** | -0.016 | -0.162** | -0.116 | -0.038 |
| | (0.030) | (0.059) | (0.034) | (0.064) | (0.119) | (0.042) |
| Treatment Township (τ) | -0.013 | -0.022 | -0.028 | 0.005 | -0.047 | -0.017 |
| | (0.064) | (0.130) | (0.077) | (0.119) | (0.074) | (0.127) |
| Poor Vision×Treatment | 0.109** | 0.257*** | 0.073 | 0.289*** | 0.212* | 0.146*** |
| Township (β) | (0.049) | (0.065) | (0.068) | (0.075) | (0.128) | (0.051) |
| | [0.046] | [0.026] | [0.374] | [0.060] | [0.106] | [0.020] |
| Combined Effect (τ + β) | 0.096 | 0.235* | 0.045 | 0.294** | 0.165 | 0.129 |
| {p-value} | {0.174} | {0.096} | {0.616} | {0.018} | {0.238} | {0.321} |
| | [0.240] | [0.242] | [0.700] | [0.138] | [0.396] | [0.434] |
| Sample Size | 18,504 | 5,736 | 12,768 | 5,788 | 5,744 | 5,742 |

Notes: 1. Coefficients on lagged test scores, constant terms, and strata dummy terms are not shown (to reduce clutter).

2. Standard errors in parentheses; wild bootstrap p-values in brackets; p-values for combined effects in curly brackets. All models include school random effects and allow for heteroscedasticity, including correlation (clustering) among observations within the same township, of unknown form for the individual level error. Asterisks denote statistical significance: * 10% level, ** 5% level, *** 1% level.

**Table 5: Interaction Effects Between Program, Visual Acuity and 2004 Test Scores**

| *Explanatory Variables* | *Dependent Variable: Average Test Score* | |
|---|---|---|
| **Equation (1): School Random Effects, Only Students with Poor Vision** | | ***N* = 2,474** |
| Treatment Township (β) | 0.173** | 0.138* |
| | (0.078) [0.130] | (0.080) [0.236] |
| Very Poor Vision | 0.053 | -- |
| | (0.040) | |
| Very Poor Vision × Treatment Township | -0.083 | -- |
| | (0.070) [0.282] | |
| Avg. Test Score 2004 × Treatment Township | -- | -0.104 |
| | | (0.090) [0.284] |
| **Equation (2): School Random Effects, All Students *N* = 18,478** | | |
| Poor Vision (π) | -0.026 | -0.013 |
| | (0.033) | (0.030) |
| Treatment Township (τ) | -0.011 | -0.007 |
| | (0.064) | (0.062) |
| Poor Vision × Treatment Township (β) | 0.120** | 0.083* |
| | (0.052) [0.054] | (0.044) [0.092] |
| Very Poor Vision | 0.030 | -- |
| | (0.053) | |
| Very Poor Vision × Treatment Township | -0.065 | -- |
| | (0.078) [0.410] | |
| 2004 Avg. Test Score × Treatment Township | -- | 0.033 |
| | | (0.078) |
| 2004 Avg. Test Score × Poor Vision | -- | 0.067** |
| | | (0.031) |
| 2004 Avg. Test Score × Poor Vision | | -0.163*** |
| × Treatment Township | | (0.054) [0.014] |

Notes: 1. Coefficients on lagged test scores, constant terms, and strata dummy terms are not shown (to reduce clutter).

2. Standard errors in parentheses; wild bootstrap p-values in brackets. All models allow for heteroscedasticity, including correlation (clustering) among observations within the same township, of unknown form for the individual level error term. Asterisks denote statistical significance: * 10% level, ** 5% level, *** 1% level.

**Table 6: Means of Hemoglobin, Height-for-Age Z-scores and Head Schooling, by County (Compliant Sample Only, Includes Only Students with Poor Vision)**

| *Variable* | *County 1* | *County 2* | *Difference (robust std. error)* |
|---|---|---|---|
| Hemoglobin level (g/l) | 138.7 | 120.0 | 18.7*** (1.76) |
| Height-for-age Z-score | 0.021 | -0.908 | 0.928*** (0.255) |
| Household head years of schooling | 7.75 | 8.98 | -1.24*** (0.312) |

Statistical significance of mean differences is based on regressions that account for clustering at the township level. Asterisks denote statistical significance: * 10% level, ** 5% level, *** 1% level.

**Table 7: Heterogeneity of Treatment Effects by County, Hemoglobin Levels, Height-for-age Z-scores, and Head of Household Schooling (Compliant Sample Only)**

| *Variable* | *Heterogeneity by county* | *Heterogeneity by student characteristics* | *Heterogeneity by county &student characteristics* |
|---|---|---|---|
| Poor Vision | -0.008 (0.031) | 0.053 (0.312) | -0.149 (0.311) |
| County 1 × Poor Vision | -0.140** (0.068) | -- | -0.122 (0.096) |
| Hemoglobin | -- | -0.002 (0.002) | -0.002 (0.002) |
| Height-for-age Z-score | -- | -0.004 (0.014) | -0.004 (0.014) |
| Head Years of Schooling | -- | 0.004 (0.008) | 0.005 (0.008) |
| Poor Vision × Hemoglobin | -- | -0.001 (0.002) | 0.001 (0.002) |
| Poor Vision × Height-for-age Z-score | -- | -0.029* (0.017) | -0.028 (0.018) |
| Poor Vision × Head years of schooling | -- | 0.006 (0.011) | 0.004 (0.010) |
| Treatment Township | -0.019 (0.080) | -0.089 (0.266) | -0.114 (0.263) |
| County 1 × Treatment Township | 0.015 (0.130) | -- | 0.008 (0.134) |
| Poor Vision × Treatment Township | 0.068 (0.068) | -0.023 (0.362) | 0.138 (0.353) |
| County 1 × Poor Vision × Treatment Township | 0.205** (0.094) | -- | 0.102 (0.131) |
| Hemoglobin × Treatment Township | -- | 0.0004 (0.0020) | 0.0006 (0.0020) |
| Height-for-Age Z-score × Treatment Township | -- | -0.005 (0.014) | -0.005 (0.014) |
| Head Years of Schooling × Treatment Township | -- | 0.002 (0.009) | 0.002 (0.009) |
| Poor Vision × Treatment Township × Hemoglobin | -- | 0.0033 (0.0024) | 0.0018 (0.0028) |
| Poor Vision × Treatment Township × Height-for-age Z-score | -- | 0.037** (0.018) | 0.036* (0.019) |
| Poor Vision × Treatment Township × Head Years of Education | -- | -0.029** (0.013) | -0.027** (0.013) |
| Sample Size | 18,504 | 18,472 | 18,472 |
| Joint test 3 treatment township triple interactions: $\chi^2(3)$ [p-value] | -- | 0.002*** | 0.073* |

Statistical significance of mean differences is based on regressions with school random effects that account for township level clustering. Asterisks denote stat. significance: * 10% level, ** 5% level, *** 1% level.

**Table 8: Probit Estimates of Factors Associated with Accepting Eyeglasses**

| Variable | Mean | Standard Deviation | Coefficient | Marginal Effects |
|---|---|---|---|---|
| Average visual acuity | 4.550 | 0.234 | -1.424** | -0.479** |
| | | | (0.563) | (0.203) |
| Female | 0.498 | 0.500 | -0.242*** | -0.081*** |
| | | | (0.059) | (0.019) |
| Had glasses before program began | 0.032 | 0.176 | 0.653* | 0.175* |
| | | | (0.382) | (0.078) |
| Household head is a teacher | 0.017 | 0.128 | -0.585** | -0.220** |
| | | | (0.236) | (0.095) |
| Household head is a village leader | 0.017 | 0.128 | -0.880* | -0.336* |
| (cadre) | | | (0.484) | (0.183) |
| Township per capita income, 2003 | 1519.1 | 467.6 | 0.00040** | 0.00013** |
| (yuan/yr) | | | (0.00020) | (0.00006) |
| Head years of schooling | 8.59 | 2.60 | -0.015 | -0.005 |
| | | | (0.020) | (0.007) |
| Test score, spring 2004 | -0.190 | 1.047 | -0.012 | -0.004 |
| (avg. for 3 subjects) | | | (0.069) | (0.023) |
| County 1 | 0.432 | 0.496 | -0.119 | -0.040 |
| | | | (0.216) | (0.073) |
| Tibetan | 0.144 | 0.351 | 0.022 | 0.007 |
| | | | (0.156) | (0.052) |
| Grade in 2003-2004 (3, 4 or 5) | 4.26 | 0.801 | -0.076 | -0.026 |
| | | | (0.125) | (0.043) |
| Observations | | | 1497 | |

Notes: 1. Constant term is not shown (to reduce clutter).

2. Standard errors are in parentheses. The specification allows for both heteroscedasticity and clustering at the township level of unknown form. Asterisks denote statistical significance: * 10% level, ** 5% level, *** 1% level.

3. The sample consists of all children in the program schools in grades 4-6 in 2004-05 who were deemed to have poor vision (one or both eyes with visual acuity below 4.9).

**Table 9: Mother's Assessment of Vision and Actual Visual Acuity (children age 8-15 who were enrolled in primary school in 2004)**

| *Measured* | *Mother's Assessment* | | | | | |
|---|---|---|---|---|---|---|
| *Acuity* | Very Bad | Bad | Fair | Good | Very good | Don't Know |
| Good (≥ 5.0) | 1 | 4 | 92 | 251 | 367 | 4 |
| Fair (4.8-4.99) | 0 | 0 | 18 | 29 | 52 | 0 |
| Poor (< 4.8) | 1 | 7 | 17 | 14 | 29 | 1 |

Source: Gansu Survey of Children and Families.

**Table 10: Children's Reports of Vision Problems, by Actual Visual Acuity (children age 8-15 who were enrolled in primary school in 2004)**

| *Measured Visual Acuity* | *Child Reports of Vision Problems* | | |
|---|---|---|---|
| | Difficulty seeing blackboard (%) | Trouble doing homework due to poor vision (%) | Felt pain in eyes when studying at home in dim light (%) |
| Good (≥ 5.0) | 8.9 | 7.0 | 19.4 |
| Fair (4.8-4.99) | 7.8 | 7.8 | 20.6 |
| Poor (< 4.8) | 32.0 | 25.3 | 30.7 |

Source: Gansu Survey of Children and Families.

**Table 11: Determinants of Student Wearing of Eyeglasses (from 2004 GSCF data)**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| VARIABLES | *Coefficient Estimates from Probit Specification* | | | | | *Marginal Effects* |
| Visual acuity | -1.329*** | -0.858** | -0.852** | -0.670 | -0.783* | -0.0196* |
| (best eye) | (0.374) | (0.412) | (0.420) | (0.423) | (0.417) | (0.0120) |
| Female | -0.228 | -0.210 | -0.228 | -0.145 | -0.155 | -0.00396 |
| | (0.169) | (0.177) | (0.173) | (0.173) | (0.173) | (0.00417) |
| Age (years) | 0.0849** | 0.0902*** | 0.0861** | 0.101*** | 0.101*** | 0.00253*** |
| | (0.0339) | (0.0341) | (0.0346) | (0.0375) | (0.0386) | (0.00106) |
| Father's education | -0.0210 | -0.0114 | -0.0104 | -0.00014 | 0.0023 | 0.00006 |
| | (0.0325) | (0.0341) | (0.0334) | (0.0365) | (0.0375) | (0.00093) |
| Mother's education | 0.0919*** | 0.0895*** | 0.0936*** | 0.0851** | 0.0722* | 0.00180* |
| | (0.0316) | (0.0326) | (0.0331) | (0.0364) | (0.0372) | (0.00105) |
| Log per capita | 0.553*** | 0.515*** | 0.530*** | 0.497** | 0.463** | 0.0116** |
| expenditure | (0.184) | (0.190) | (0.189) | (0.204) | (0.201) | (0.0056) |
| Mother assessment | | 0.898*** | 0.887*** | 0.858*** | 0.809*** | 0.0508*** |
| of child's vision | | (0.326) | (0.329) | (0.307) | (0.298) | (0.0364) |
| Father assessment | | 0.603 | 0.605 | 0.558 | 0.542 | 0.0247 |
| of child's vision | | (0.380) | (0.385) | (0.357) | (0.364) | (0.0251) |
| Estimated cost of | | | -0.00366* | -0.00375* | -0.00451** | -0.00011** |
| glasses | | | (0.00215) | (0.00218) | (0.00225) | (0.00007) |
| Estimated distance | | | 0.00164 | 0.00070 | 0.00087 | 0.00002 |
| to buy glasses | | | (0.00422) | (0.00392) | (0.00388) | (0.00010) |
| Parent wears | | | | 0.955*** | 0.962*** | 0.0678*** |
| glasses | | | | (0.322) | (0.318) | (0.0439) |
| Village literacy rate | | | | | 1.300** | 0.0325** |
| | | | | | (0.645) | (0.0171) |
| Constant | -0.392 | -2.710 | -2.668 | -3.679 | -3.890 | |
| | (2.360) | (2.732) | (2.751) | (2.731) | (2.898) | |
| Observations | 921 | 921 | 921 | 921 | 921 | 921 |

Notes: 1. Standard errors in parentheses. The specification allows for both heteroscedasticity and clustering at the village level of unknown form. * 10% level, ** 5% level,*** 1% level.

2. Estimated cost of glasses and distance to buy glasses are medians of parental reports.

**Online Appendix Table A.1: Pre-Program Differences between Treatment and Control Groups (25 townships where randomization was correctly implemented, separately by county)**

| *Variable* | *Treatment Mean* | *Control Mean* | *Difference* | *p-values for Differences based on Wild Bootstrap* |
|---|---|---|---|---|
| | | ***All Children, County 1*** | | |
| Chinese test | -0.133 | -0.336 | 0.203 | 0.543 |
| Math test | -0.112 | -0.308 | 0.196 | 0.608 |
| Science test | -0.032 | -0.386 | 0.354 | 0.472 |
| Average test | -0.112 | -0.416 | 0.304 | 0.868 |
| Ethnic minority | 0.388 | 0.495 | -0.106 | 0.016** |
| Visual acuity | 5.04 | 5.09 | -0.05 | 0.025** |
| Poor vision | 0.130 | 0.122 | 0.008 | 0.163 |
| Male | 0.542 | 0.561 | -0.019 | 0.864 |
| Head yrs educ | 8.62 | 7.45 | 1.17 | 0.461 |
| Age | 9.78 | 10.35 | -0.57 | 0.286 |
| Joint test (F-test [p-value]) | | | 1.06 [0.452] | |
| | | ***Children with Poor Vision Only, County 1*** | | |
| Chinese test | -0.314 | -0.223 | -0.091 | 0.994 |
| Math test | -0.296 | -0.115 | -0.181 | 0.551 |
| Science test | -0.111 | -0.279 | 0.168 | 0.522 |
| Average test | -0.292 | -0.250 | -0.042 | 0.906 |
| Ethnic minority | 0.359 | 0.339 | 0.019 | 0.071* |
| Visual acuity | 4.61 | 4.70 | -0.09*** | 0.006*** |
| Male | 0.496 | 0.514 | -0.017 | 0.774 |
| Head yrs educ | 7.82 | 7.34 | 0.48 | 0.183 |
| Age | 10.16 | 10.54 | -0.37* | 0.232 |
| Joint test (F-test [p-value]) | | | 0.91 [0.539] | |
| | | ***All Children, County 2*** | | |
| Chinese test | -0.058 | -0.158 | 0.100 | 0.462 |
| Math test | -0.079 | -0.085 | 0.005 | 0.703 |
| Science test | 0.001 | -0.107 | 0.108 | 0.555 |
| Average test | -0.055 | -0.141 | 0.086 | 0.561 |
| Ethnic minority | 0.004 | 0.010 | -0.006 | 0.795 |
| Visual acuity | 5.01 | 5.04 | -0.03 | 0.290 |
| Poor vision | 0.121 | 0.155 | -0.035 | 0.784 |
| Male | 0.530 | 0.531 | -0.001 | 0.586 |
| Head yrs educ | 9.23 | 8.73 | 0.50 | 0.301 |
| Age | 10.97 | 11.18 | -0.21 | 0.388 |
| Joint test (F-test [p-value]) | | | 3.88** [0.019] | |
| | | ***Children with Poor Vision Only, County 2*** | | |
| Chinese test | -0.105 | -0.184 | 0.079 | 0.714 |
| Math test | -0.101 | -0.027 | -0.075 | 0.784 |
| Science test | -0.082 | -0.044 | -0.038 | 0.698 |
| Average test | -0.117 | -0.103 | -0.014 | 0.894 |
| Ethnic minority | 0.001 | 0.009 | -0.008 | 0.794 |
| Visual acuity | 4.50 | 4.65 | -0.15*** | 0.006*** |
| Male | 0.503 | 0.457 | 0.046 | 0.418 |
| Head yrs educ | 9.18 | 8.79 | 0.39 | 0.372 |
| Age | 11.25 | 11.27 | -0.03 | 0.120 |
| Joint test (F-test [p-value]) | | | 5.80*** [0.005] | |

Statistical significance of mean differences is based on regressions that include school random effects and account for clustering at the township level.
Asterisks denote statistical significance: * 10% level, ** 5% level, *** 1% level.

**Online Appendix Table A.2: Pre-Program Differences between Treatment and Control Groups (All 37 Townships, separately by county)**

| *Variable* | *Treatment Mean* | *Control Mean* | *Difference* | *p-values for Differences based on Wild Bootstrap* |
|---|---|---|---|---|
| | | ***All Children, County 1*** | | |
| Chinese test | -0.103 | -0.244 | 0.141 | 0.816 |
| Math test | -0.097 | -0.257 | 0.160 | 0.814 |
| Science test | -0.027 | -0.261 | 0.234 | 0.438 |
| Average test | -0.092 | -0.308 | 0.216 | 0.666 |
| Ethnic minority | 0.418 | 0.508 | -0.090 | 0.432 |
| Visual acuity | 5.05 | 5.04 | 0.01 | 0.890 |
| Poor vision | 0.132 | 0.185 | -0.053 | 0.858 |
| Male | 0.533 | 0.553 | -0.020 | 0.332 |
| Head yrs educ | 8.53 | 7.46 | 1.07 | 0.270 |
| Age | 9.82 | 10.12 | -0.29 | 0.294 |
| Joint test (F-test [p-value]) | | | 0.81 [0.615] | |
| | | ***Children with Poor Vision Only, County 1*** | | |
| Chinese test | -0.228 | -0.221 | -0.006 | 0.558 |
| Math test | -0.245 | -0.180 | -0.065 | 0.884 |
| Science test | -0.053 | -0.211 | 0.158 | 0.322 |
| Average test | -0.212 | -0.248 | 0.035 | 0.632 |
| Ethnic minority | 0.379 | 0.445 | -0.066 | 0.462 |
| Visual acuity | 4.60 | 4.64 | -0.04** | 0.014** |
| Male | 0.485 | 0.486 | -0.0005 | 0.690 |
| Head yrs educ | 7.89 | 7.22 | 0.67 | 0.008*** |
| Age | 10.21 | 10.22 | -0.002 | 0.372 |
| Joint test (F-test [p-value]) | | | 1.05 [0.437] | |
| | | ***All Children, County 2*** | | |
| Chinese test | -0.143 | -0.045 | -0.098 | 0.496 |
| Math test | -0.110 | 0.019 | -0.129 | 0.536 |
| Science test | -0.087 | -0.004 | -0.083 | 0.362 |
| Average test | -0.137 | -0.012 | -0.125 | 0.452 |
| Ethnic minority | 0.020 | 0.009 | 0.011 | 0.763 |
| Visual acuity | 5.03 | 5.01 | 0.02 | 0.692 |
| Poor vision | 0.108 | 0.184 | -0.076** | 0.124 |
| Male | 0.527 | 0.534 | -0.007 | 0.664 |
| Head yrs educ | 8.88 | 8.67 | 0.20 | 0.484 |
| Age | 11.13 | 11.19 | -0.06 | 0.836 |
| Joint test (F-test [p-value]) | | | 3.61** [0.011] | |
| | | ***Children with Poor Vision Only, County 2*** | | |
| Chinese test | -0.161 | -0.047 | -0.113 | 0.500 |
| Math test | -0.126 | 0.081 | -0.208 | 0.442 |
| Science test | -0.136 | 0.051 | -0.187 | 0.104 |
| Average test | -0.171 | 0.034 | -0.206 | 0.318 |
| Ethnic minority | 0.017 | 0.008 | 0.009 | 0.912 |
| Visual acuity | 4.49 | 4.62 | -0.13*** | 0.002*** |
| Male | 0.496 | 0.470 | 0.026 | 0.684 |
| Head yrs educ | 8.94 | 8.66 | 0.28 | 0.376 |
| Age | 11.34 | 11.32 | 0.02 | 0.368 |
| Joint test (F-test [p-value]) | | | 5.76*** [0.001] | |

Statistical significance of mean differences is based on regressions that include school random effects and account for clustering at the township level. Joint tests based on regressions of treatment variable on all baseline variables. Asterisks denote statistical significance: * 10% level, ** 5% level, *** 1% level.

**Online Appendix Table A.3: Pre-Program Differences between Compliant and Non-Compliant Townships (separately by county)**

| *Variable* | *Compliant Mean* | *Non-Compliant Mean* | *Difference* | *p-values of differences based on Wild Bootstrap* |
|---|---|---|---|---|
| | | ***All Children, County 1*** | | |
| Chinese | -0.161 | -0.147 | -0.014 | 0.224 |
| Math | -0.139 | -0.180 | 0.041 | 0.348 |
| Science | -0.082 | -0.159 | 0.077 | 0.460 |
| Average | -0.155 | -0.197 | 0.042 | 0.292 |
| Ethnic minority | 0.404 | 0.522 | -0.118** | 0.054* |
| Poor vision | 0.128 | 0.187 | -0.059 | 0.362 |
| Visual acuity | 5.05 | 5.03 | 0.01 | 0.662 |
| Male | 0.545 | 0.534 | 0.011 | 0.796 |
| Head yrs educ | 8.44 | 7.67 | 0.77 | 0.708 |
| Age | 9.87 | 10.02 | -0.16 | 0.620 |
| | | ***Children with Poor Vision Only, County 1*** | | |
| Chinese | -0.302 | -0.148 | -0.154 | 0.470 |
| Math | -0.271 | -0.160 | -0.111 | 0.584 |
| Science | -0.134 | -0.116 | -0.018 | 0.740 |
| Average | -0.286 | -0.171 | -0.115 | 0.530 |
| Ethnic minority | 0.356 | 0.464 | -0.108 | 0.148 |
| Visual acuity | 4.62 | 4.62 | 0.00 | 0.994 |
| Male | 0.499 | 0.472 | 0.027 | 0.420 |
| Head yrs educ | 7.75 | 7.42 | 0.32 | 0.636 |
| Age | 10.22 | 10.21 | 0.01 | 0.932 |
| | | ***All Children, County 2*** | | |
| Chinese | -0.102 | -0.096 | -0.006 | 0.572 |
| Math | -0.082 | 0.009 | -0.090 | 0.274 |
| Science | -0.047 | -0.058 | 0.011 | 0.730 |
| Average | -0.094 | -0.059 | -0.035 | 0.496 |
| Ethnic minority | 0.007 | 0.034 | -0.027 | 0.694 |
| Poor vision | 0.136 | 0.152 | -0.016 | 0.598 |
| Visual acuity | 5.02 | 5.02 | 0.00 | 0.968 |
| Male | 0.531 | 0.530 | 0.001 | 0.962 |
| Head yrs educ | 9.01 | 8.31 | 0.70** | 0.092* |
| Age | 11.06 | 11.36 | -0.30** | 0.072* |
| | | ***Children with Poor Vision Only, County 2*** | | |
| Chinese | -0.145 | 0.001 | -0.146 | 0.566 |
| Math | -0.063 | 0.101 | -0.164 | 0.506 |
| Science | -0.062 | 0.036 | -0.098 | 0.730 |
| Average | -0.110 | 0.056 | -0.165 | 0.648 |
| Ethnic minority | 0.005 | 0.025 | -0.020 | 0.524 |
| Visual acuity | 4.58 | 4.54 | 0.04 | 0.362 |
| Male | 0.479 | 0.485 | -0.005 | 0.818 |
| Head yrs educ | 8.98 | 8.39 | 0.60 | 0.200 |
| Age | 11.26 | 11.46 | -0.20 | 0.134 |

Statistical significance is based on regressions that include school random effects and account for township level clustering. Strata dummy variables are excluded; they are perfectly collinear with the compliant dummy variable.
Asterisks denote statistical significance: * 10% level, ** 5% level, *** 1% level.

**Online Appendix Table A.4: Estimated Program Effect: ITT Results without 2004 Test Scores**

| | *Dependent Variables* | | | | | |
|---|---|---|---|---|---|---|
| | Average Test Scores | | | Subject Scores (County 1 only) | | |
| *Explanatory Variables* | Both Counties | County 1 | County 2 | Chinese | Math | Science |
| **Equation (1): Compliant Sample, Only Students with Poor Vision** | | | | | | |
| Treatment Township (β) | 0.163 | 0.372** | 0.097 | 0.407*** | 0.250* | 0.245* |
| | (0.100) | (0.155) | (0.122) | (0.157) | (0.137) | (0.142) |
| | [0.280] | [0.108] | [0.648] | [0.116] | [0.295] | [0.248] |
| Sample Size | 2,474 | 732 | 1,742 | 745 | 733 | 732 |
| **Equation (2): Compliant Sample, All Students** | | | | | | |
| Poor Vision (π) | 0.018 | -0.090 | 0.030 | -0.120** | -0.092 | -0.019 |
| | (0.023) | (0.059) | (0.024) | (0.059) | (0.121) | (0.042) |
| Treatment Township (τ) | 0.008 | -0.029 | 0.022 | 0.003 | -0.058 | -0.013 |
| | (0.089) | (0.158) | (0.110) | (0.146) | (0.097) | (0.144) |
| Poor Vision×Treatment | 0.082 | 0.234*** | 0.038 | 0.252*** | 0.194 | 0.130** |
| Township (β) | (0.056) | (0.065) | (0.087) | (0.073) | (0.129) | (0.051) |
| | [0.162] | [0.012] | [0.696] | [0.052] | [0.152] | [0.034] |
| Combined Effect (τ + β) | 0.091 | 0.205 | 0.060 | 0.254 | 0.135 | 0.117 |
| | {0.324} | {0.259} | {0.592) | {0.116} | {0.403} | {0.467} |
| | [0.456] | [0.470] | [0.644] | [0.306] | [0.670] | [0.574] |
| Sample Size | 18,505 | 5,737 | 12,768 | 5,789 | 5,747 | 5,745 |
| **Equation (1): Full Sample, Only Students with Poor Vision** | | | | | | |
| Treatment Township (β) | -0.092 | 0.174* | -0.206 | 0.229** | 0.171** | 0.023 |
| | (0.111) | (0.101) | (0.151) | (0.105) | (0.070) | (0.107) |
| | [0.648] | [0.230] | [0.434] | [0.140] | [0.052] | [0.916] |
| Sample Size | 4,093 | 1,468 | 2,625 | 1,491 | 1,469 | 1,468 |
| **Equation (2): Full Sample, All Students** | | | | | | |
| Poor Vision (π) | 0.041 | 0.044 | 0.040 | 0.038 | -0.014 | 0.077* |
| | (0.022) | (0.044) | (0.025) | (0.048) | (0.027) | (0.042) |
| Treatment Township (τ) | -0.156* | -0.046 | -0.212* | 0.033 | -0.071 | -0.072 |
| | (0.091) | (0.106) | (0.127) | (0.102) | (0.064) | (0.094) |
| Poor Vision×Treatment | 0.032 | 0.077 | 0.0001 | 0.061 | 0.104** | 0.025 |
| Township (β) | (0.047) | (0.056) | (0.069) | (0.066) | (0.047) | (0.054) |
| | [0.560] | [0.306] | [0.988] | [0.386] | [0.086] | [0.634] |
| Combined Effect (τ + β) | -0.124 | 0.031 | -0.211 | 0.094 | 0.034 | -0.047 |
| | {0.210} | {0.756} | {0.138} | {0.311} | {0.619} | {0.623} |
| | [0.350] | [0.764] | [0.314] | [0.408] | [0.674] | [0.670] |
| Sample Size | 28,272 | 9,695 | 18,577 | 9,786 | 9,714 | 9,712 |

Notes: 1. Coefficients on constant terms, and strata dummy terms are not shown (to reduce clutter).

2. Standard errors in parentheses; wild bootstrap p-values in brackets; p-values for combined effect in curly brackets. All models include school random effects and allow for heteroscedasticity, including correlation (clustering) among observations within the same township, of unknown form for the individual level error term. Asterisks denote statistical significance: * 10% level, ** 5% level, *** 1% level.

**Online Appendix Table A.5: Estimated Program Effect After 1 Year: ITT Results, Full Sample**

| | *Dependent Variables* | | | | | |
|---|---|---|---|---|---|---|
| | Average Test Scores | | | Subject Scores (County 1 only) | | |
| *Explanatory Variables* | Both Counties | County 1 | County 2 | Chinese | Math | Science |
| **Equation (1): Full Sample, Only Students with Poor Vision** | | | | | | |
| Treatment Township (β) | -0.049 | 0.159 | -0.073 | 0.217** | 0.162** | 0.002 |
| | (0.077) | (0.095) | (0.086) | (0.093) | (0.065) | (0.110) |
| | [0.676] | [0.208] | [0.616] | [0.108] | [0.070] | [0.980] |
| Sample Size | 4,093 | 1,468 | 2,625 | 1,491 | 1,469 | 1,468 |
| **Equation (2): Full Sample, All Students** | | | | | | |
| Poor Vision (π) | 0.023 | 0.040 | 0.010 | 0.033 | -0.016 | 0.074* |
| | (0.022) | (0.044) | (0.027) | (0.047) | (0.027) | (0.043) |
| Treatment Township (τ) | -0.122** | -0.068 | -0.117* | 0.010 | -0.083* | -0.093 |
| | (0.054) | (0.084) | (0.065) | (0.080) | (0.049) | (0.079) |
| Poor Vision×Treatment | 0.041 | 0.075 | 0.022 | 0.064 | 0.103** | 0.024 |
| Township (β) | (0.040) | (0.055) | (0.054) | (0.064) | (0.047) | (0.054) |
| | [0.356] | [0.294] | [0.692] | [0.346] | [0.108] | [0.696] |
| Combined Effect (τ + β) | -0.081 | 0.007 | -0.095 | 0.074 | 0.020 | -0.069 |
| | {0.212} | {0.934} | {0.246} | {0.346} | {0.733} | {0.432} |
| | [0.312] | [0.864] | [0.382] | [0.416] | [0.764] | [0.498] |
| Sample Size | 28,271 | 9,694 | 18,577 | 9,785 | 9,709 | 9,707 |

Notes: 1. Coefficients on lagged test scores, constant terms, and strata dummy terms are not shown (to reduce clutter).

2. Standard errors in parentheses; wild bootstrap p-values in brackets. All models include school random effects and allow for heteroscedasticity, including correlation (clustering) among observations within the same township, of unknown form for the individual level error. Asterisks denote statistical significance: * 10% level, ** 5% level, *** 1% level.

**Online Appendix Table A.6: Effect of Eyeglasses After 1 Year: IV Results, Compliant and Full Samples**

| | *Dependent Variable* | | | | | |
|---|---|---|---|---|---|---|
| | Average Test Scores | | | Subject Scores (County 1 only) | | |
| *Explanatory Variables* | Both Counties | County 1 | County 2 | Chinese | Math | Science |
| **Equation (1): Compliant Sample, Only Students with Poor Vision** | | | | | | |
| Has Eyeglasses (β) | 0.224** | 0.677*** | 0.105 | 0.715*** | 0.465** | 0.447** |
| | (0.110) | (0.196) | (0.121) | (0.206) | (0.203) | (0.175) |
| Sample Size | 2,474 | 732 | 1,742 | 745 | 733 | 732 |
| **Equation (2): Compliant Sample, All Students** | | | | | | |
| Poor Vision (π) | -0.023 | -0.122** | -0.015 | -0.159** | -0.115 | -0.037 |
| | (0.030) | (0.059) | (0.033) | (0.063) | (0.116) | (0.041) |
| Has Eyeglasses (β) | 0.156** | 0.411*** | 0.093 | 0.456*** | 0.334* | 0.231*** |
| | (0.071) | (0.101) | (0.087) | (0.114) | (0.201) | (0.080) |
| Sample Size | 18,503 | 5,735 | 12,768 | 5,787 | 5,743 | 5,787 |
| **Equation (1): Full Sample, Only Students with Poor Vision** | | | | | | |
| Has Eyeglasses (β) | -0.099 | 0.542** | -0.126 | 0.752*** | 0.562*** | 0.009 |
| | (0.161) | (0.214) | (0.150) | (0.274) | (0.215) | (0.368) |
| Sample Size | 4,093 | 1,468 | 2,625 | 1,491 | 1,469 | 1,468 |
| **Equation (2): Full Sample, All Students** | | | | | | |
| Poor Vision (π) | 0.002 | -0.353 | 0.008 | -0.359 | -0.491 | 0.059 |
| | (0.043) | (0.537) | (0.038) | (0.455) | (0.589) | (0.362) |
| Has Eyeglasses (β) | 0.082 | 0.746 | 0.027 | 0.736 | 0.911 | 0.049 |
| | (0.096) | (0.896) | (0.090) | (0.778) | (0.994) | (0.610) |
| Sample Size | 28,270 | 9,693 | 18,577 | 9,784 | 9,708 | 9,706 |

Notes: 1. Coefficients on lagged test scores, constant terms, and strata dummy terms are not shown (to reduce clutter).

2. Standard errors in parentheses. All models include school random effects and allow for heteroscedasticity, including correlation (clustering) among observations within the same township, of unknown form for the individual level error term. Asterisks denote statistical significance: * 10% level, ** 5% level, *** 1% level.

3. The instrumental variable for having eyeglasses is a dummy variable that indicates that one has been selected into the program *and* has poor vision.

**Online Appendix Table A.7: Falsification Test: "Effect" of Program in 2003-04 School Year (Compliant Sample)**

| | *Dependent Variable (2004 test scores)* | | | | | |
|---|---|---|---|---|---|---|
| | Average Test Scores | | | Subject Scores (County 1 only) | | |
| *Explanatory Variables* | Both Counties | County 1 | County 2 | Chinese | Math | Science |
| **Equation (2): School Random Effects, All Students** | | | | | | |
| Poor Vision (π) | 0.019 | -0.015 | 0.023 | -0.045 | 0.031 | -0.026 |
| | (0.015) | (0.052) | (0.015 | (0.058) | (0.075) | (0.042) |
| Treatment Township (τ) | -0.010 | -0.080* | 0.012 | -0.124*** | -0.125** | 0.052 |
| | (0.029) | (0.047) | (0.035) | (0.029) | (0.053) | (0.046) |
| Poor Vision×Treatment | -0.018 | -0.011 | -0.002 | 0.034 | -0.068 | 0.013 |
| Township (β) | (0.031) | (0.062) | (0.041) | (0.076) | (0.087) | (0.050) |
| | [0.562] | [0.860] | [0.980] | [0.676] | [0.588] | [0.784] |
| Sample Size | 18,600 | 5,830 | 12,770 | 5,831 | 5,830 | 5,831 |

Notes: 1. Coefficients on lagged test scores, constant terms, and strata dummy terms are not shown (to reduce clutter).

2. Standard errors in parentheses; wild bootstrap p-values in brackets. All models include school random effects. All models allow for heteroscedasticity, including correlation (clustering) among observations within the same township, of unknown form for the individual level error term. Asterisks denote statistical significance: * 10% level, ** 5% level, *** 1% level.

**Online Appendix 1: Endogeneity and Measurement Error in the Poor Vision (PV) Variable**

This appendix shows that correlation between u and PV in equation (2) does not lead to bias in the estimate of program impact ($\beta$), and that the same is true if there is random measurement error in the poor vision (PV) variable.

To see the first point, assume that the correlation takes the form $u = \theta PV + \varepsilon$, where $\varepsilon$ is uncorrelated with both PV and P. Then equation (2) becomes $T = \alpha + \pi PV + \tau P + \beta PV*P + \theta PV + \varepsilon = \alpha + (\pi+\theta)PV + \tau P + \beta PV*P + \varepsilon$; this regression will yield biased estimates of $\pi$, but the estimate of $\beta$ is still unbiased. More generally, in equation (2) u is not correlated with PV*P after conditioning on PV.

To see that random measurement error in the poor vision variable will not lead to under-estimation of program effects in estimation of equation (2), let PV* be the (unobserved) true indicator that a child has poor vision, and let PV denote the observed value of that variable. Thus $PV = PV^* + \varepsilon$, where $\varepsilon$ is the measurement error. Since both PV and PV* are dummy variables, $\varepsilon$ will clearly be correlated with PV*, so this is not classical measurement error.

Assume that both types of measurement error (PV = 0 when PV* = 1 and PV = 1 when PV* = 0) occur with the same frequency, denoted by $\theta$ where $\theta$ is assumed to be less than 0.5. Thus there are three possibilities:

| *Frequency* | *Value of PV** | *Value of PV* | *Value of ε* |
|---|---|---|---|
| 1-2θ (no error) | 1 or 0 | Same as PV* | 0 |
| θ | 0 | 1 | 1 |
| θ | 1 | 0 | -1 |

The assumption that both types of errors occur with the same frequency is plausible if the error in the underlying visual acuity variable is random and the density of the distribution of that variable is similar on both sides of the cutoff point (4.8), and the latter is approximately correct.

Measurement error alters equation (2) as follows:

$$T = \alpha + \pi PV^* + \tau P + \beta PV^*P + u$$

$$= \alpha + \pi(PV - \varepsilon) + \tau P + \beta(PV - \varepsilon)*P + u$$

$$= \alpha + \pi PV + \tau P + \beta PV*P + (u - \pi\varepsilon - \beta\varepsilon P)$$

Bias in OLS estimation of $\beta$ will be primarily determined by whether the interaction term PV*P is correlated with the composite error term $(u - \pi\varepsilon - \beta\varepsilon P)$. Focusing on bias due to measurement error, this will be determined by whether PV*P is correlated with $\pi\varepsilon + \beta\varepsilon P$. Covariance formulas imply $Cov(\pi\varepsilon + \beta\varepsilon P, PV*P) = \pi Cov(\varepsilon, PV*P) + \beta Cov(\varepsilon P, PV*P)$. The following derivations show that $Cov(\varepsilon, PV*P) = Cov(\varepsilon P, PV*P) = \theta E[P]$:

$$Cov(\varepsilon, PV*P) = E[(\varepsilon - E[\varepsilon])(PV*P - E[PV*P])]$$
$$= E[\varepsilon(PV*P - E[PV*P])]$$

$$= E[\varepsilon PV^*P] - E[\varepsilon]E[PV^*P]$$
$$= E[\varepsilon PV^*P]$$
$$= \text{Prob}[\varepsilon = 0]\times E[\varepsilon PV^*P|\ \varepsilon = 0] + \text{Prob}[\varepsilon = -1]\times E[\varepsilon PV^*P|\ \varepsilon = -1] + \text{Prob}[\varepsilon = 1]\times E[\varepsilon PV^*P|\ \varepsilon = 1]$$
$$= (1 - 2\theta)\times 0 + \theta\times(-1\times 0\times E[P]) + \theta\times(1\times 1\times E[P])$$
$$= 0 + 0 + \theta E[P] > 0$$

$$\text{Cov}(\varepsilon P, PV^*P) = E[(\varepsilon P - E[\varepsilon P])(PV^*P - E[PV^*P])]$$
$$= E[(\varepsilon P - E[\varepsilon]\times E[P])(PV^*P - E[PV^*P])]$$
$$= E[\varepsilon P(PV^*P - E[PV^*P])]$$
$$= E[\varepsilon PV^*P^2 - E[\varepsilon P]E[PV^*P])]$$
$$= E[\varepsilon PV^*P - E[\varepsilon]E[P]E[PV^*P])]$$
$$= E[\varepsilon PV^*P]$$
$$= \theta E[P] > 0$$

Thus we have:

$$\text{Cov}(\pi\varepsilon + \beta\varepsilon P, PV^*P) = \pi\text{Cov}(\varepsilon, PV^*P) + \beta\text{Cov}(\varepsilon P, PV^*P) = (\pi + \beta)\theta E[P]$$

This derivation has two implications. First, as measurement error decreases ($\theta \rightarrow 0$), this correlation goes to zero and so bias goes to zero. Second, it is reasonable to assume that $\pi = -\beta$. Quite simply, if we expect a child with poor vision to score lower on tests by a factor of $\pi$ (note that $\pi < 0$), then providing that child with glasses should remove the problem, which implies an impact of the same magnitude but in the opposite direction ($\beta > 0$). This thus implies that the two terms in the bias tend to cancel each other out. Note that the fact that $\pi$ may be estimated with bias does not matter for this derivation, which is based on the true underlying value of $\pi$, not an estimate of $\pi$.

## Online Appendix 2: Implications of Having Different Tests in Each School

Suppose that a "common" test (the same test) had been implemented in all the schools. For such a test, denote the mean and standard deviation of students' scores as:

$\mu^c$ = mean test score over all schools for common test
$\sigma^c$ = standard deviation of distribution over all schools for common test

Each school (denoted by s) would also have had its own mean and standard deviation if all of its students had taken this common test. They can be denoted as:

$\mu_s{}^c$ = mean test score of school s if it had used the common test
$\sigma_s{}^c$ = standard deviation of students in school s if they had taken the common test

Clearly, the $\mu_s{}^c$ for each school are related to $\mu^c$ as follows:

$$\mu^c = \sum_{s=1}^{S} w_s\mu_s{}^c$$

where S is the total number of schools and $w_s$ is the share of students in school s.

For the tests that were actually used, denote the means and standard deviations as:

$\mu_s^a$ = mean test score of school s for the actual test taken
$\sigma_s^a$ = standard deviation of students in school s for the actual test taken

For each school, normalize both the common test and the actual test at the school level:

$\frac{t_{is}^c - \mu_s^c}{\sigma_s^c}$ = normalized (at school level) score of student i in school s on common test: $n_{is}^c$

$\frac{t_{is}^a - \mu_s^a}{\sigma_s^a}$ = normalized (at school level) score of student i in school s on actual test: $n_{is}^a$

Assume that these two normalizations are equal to each other. That is, assume that $t_{is}^c$ and $t_{is}^a$ have the same "shape" but a different mean and variance. This assumption implies that $n_{is}^c = n_{is}^a$, which further implies that:

$$\frac{t_{is}^c - \mu_s^c}{\sigma_s^c} = \frac{t_{is}^a - \mu_s^a}{\sigma_s^a}$$

$$t_{is}^c = (\sigma_s^c/\sigma_s^a)\times t_{is}^a + \mu_s^c - (\sigma_s^c/\sigma_s^a)\mu_s^a$$

Consider a simple regression equation using the standardized common test:

$$t_{is}^c = \alpha + \beta P_s + u_{is}$$

The above derivations imply that we can express this relationship as:

$$(\sigma_s^c/\sigma_s^a)\times t_{is}^a + \mu_s^c - (\sigma_s^c/\sigma_s^a)\mu_s^a = \alpha + \beta P_s + u_{is}$$

$$(\sigma_s^c/\sigma_s^a)\times t_{is}^a = \alpha + \beta P_s + u_{is} + [(\sigma_s^c/\sigma_s^a)\mu_s^a - \mu_s^c]$$

$$t_{is}^a = \alpha(\sigma_s^a/\sigma_s^c) + \beta(\sigma_s^a/\sigma_s^c)P_s + (\sigma_s^a/\sigma_s^c)u_{is} + [\mu_s^a - (\sigma_s^a/\sigma_s^c)\mu_s^c]$$

The term in brackets is a standard school random effect. The main concern is that the program effect, β, is multiplied by school specific ratio, $(\sigma_s^a/\sigma_s^c)$. Because one can choose any scale for our hypothetical common test, one can choose it so that, on average (averaging across all S schools), $(\sigma_s^a/\sigma_s^c) = 1$. Thus the plim of $\beta(\sigma_s^a/\sigma_s^c)$ over all S schools will be β. In fact, for any given school the estimated impact could be different, even if the same (common) test had implemented over all schools; one could specify a random coefficients model even if all schools had used the same test. Yet it is standard practice to estimate a specification that assumes the same impact, and doing so estimates the average impact over all schools. Likewise, by regressing the actual test score (instead of a score on a hypothetical common test) on a constant and $P_s$, one estimates the average of $\beta(\sigma_s^a/\sigma_s^c)$, which is β. A final point is that regressions that compare good vision and poor vision students in the same school are comparing students on the same test, and so there is no problem of comparing students on different tests. Even so, it is still useful to have random effect to account for other unobserved school-level factors.