II. ESTIMATED CAUSAL EFFECTS

IIa. Data Utilized

Gauteng Province

The data I collected for Gauteng Province consist of observations for 494 secondary schools between the years 2000 and 2007. The Gauteng Online project included 217 of those schools, and computer deployments began in the year 2002 with the largest deployments in 2003, 2004 and 2005.

The outcome variable used in Gauteng Province comes from school-level results on the national senior certificate examination, also known as the matric examination. Each Grade 12 student who wishes to earn a senior certificate must take and pass the exam, and a senior certificate is a prerequisite for most forms of post-secondary education in South Africa as well as for many types of employment. Its equivalent in the United States education system is the high school diploma.

Data are available from the Gauteng Department of Education (GDE) Examinations and Assessment Directorate for each school in the province offering the exam. Overall pass rates and pass rates disaggregated by gender at the school level are used as an outcome variable. Each set of pass rates is normalized to a mean of zero and standard deviation of one in each year. The total number of students sitting for the examination is used as a covariate in some specifications.

Additional data from the GDE provide information on important school demographics. Each year, the schools in the province are required by law to complete the Annual School Survey, which collects a variety of information about students, educators and grade cohorts within the school. In my analysis, data on student body racial composition and overall school enrollment for each school from this survey's results are used, provided by the department's Education Management Information Systems (EMIS) directorate.

Since the year 2000, funding in South Africa has been allocated within provinces on a per-capita basis according to school poverty levels. The method of allocating funds is prescribed by national policy. Provinces are required to assess the poverty level of each schools' respective community and allocate more funds per capita to the schools it determines to be serving the poorest students. This is done by ranking school poverty levels on an index from 1 to 100 and dividing schools into five quintiles based on this ranking. The schools in the quintile with the highest poverty ranking must be allocated more funding per capita than schools in the next quintile, and so forth. Also, provinces must make determinations on what level of autonomy schools are allowed in spending their allocated funding, based on the quality of school management and governance.

Data are available from the GDE's Funding, Subsidies and Auditing Directorate on each school's poverty ranking, assigned quintile, and per capita funding level in each year from 2000 through 2007. Additionally, data are available in each of these years for the level of autonomy each school was allowed in spending this allocated funding. This amounts to information on whether the school could directly spend its funds or needed to seek district approval for expenditures in each of three spending categories. This division of spending authority between school and district officials is known as the school's "functional status".

I also use data on several static characteristics of each school. Gauteng Province is divided into 12 educational districts based on schools' geographic location. These districts mostly serve to provide administrative oversight to schools, and serve such functions as certifying Annual School Survey results and managing funding. Additionally, an important characteristic of schools in South Africa today is their historical status, specifically the former educational authority responsible for the school and former "model" of the school. These refer to the apartheid-era structure of South Africa's education system, where schools were administered by various education authorities that broke down along racial lines and where the most well-resourced schools could only be attended by white students. Since aspects of the apartheid education system were in place as recently as 1994, major differences still exist between schools based on their apartheid-era status. Data on schools' district, former educational authority and former model status were provided by the GDE's EMIS Directorate.

Finally, data on the year each school in the Gauteng Online project received computers are constructed from two sources. The Gauteng Online project office provided information on which schools were included in each phase of the project. I included only schools from the first two phases of the project since subsequent phases included very small numbers of schools and in some cases focused only on schools with special needs. Additionally, schools in phases 1 and 2 represent the vast majority of schools in the project and received their deployments first.

While data on which schools were included in the project were provided by the project office, the office could not provide data on the specific years in which the schools received their deployments. The reason for this was that GDE officials were unable to extract the appropriate data from the project's archived files. As a substitute, data from the Annual School Survey were utilized, as the survey asked about the number of computers at the school used for educational purposes. Since Gauteng Online deployments were always a standard number of computers, and since the vast majority of project schools did not procure computers of their own, it was possible to identify the year of treatment based on the survey data.

Western Cape Province

The data I collected for Western Cape Province consist of observations for 1,121 schools with a Grade 3 between the years 2002 and 2006. The Khanya project included 437 of those schools, and computer deployments began in the year 2001 with over 40 schools deployed in each year since 2002.

The outcome variable used in the evaluation of Western Cape Education Department's (WCED) Khanya project comes from school-level results on the province's systemic evaluation. Systemic evaluations began in Western Cape province in 2002 and serve as the only province-wide objective assessment of student achievement prior to the senior certificate exam. They are conducted at the Grade 3 and Grade 6 levels in alternating years and include a standardized examination on literacy and numeracy. Each set of scores is normalized to a mean of zero and standard deviation of one in each year. Literacy and numeracy are the two learning areas on which the Khanya project focuses its efforts in primary schools.

Data were provided on the systemic evaluation results from the WCED's Education Research Directorate, and include overall school-level results for the literacy and numeracy portions of the evaluation as well as results disaggregated by gender. In my analysis I use the Grade 3 results, which are available for the years 2002, 2004 and 2006. In 2002, the evaluation did not include all of the province's schools—844 of roughly 1100 schools with a Grade 3 (77%) were included. However, data are available in 2002 for a substantially higher proportion of schools in the Khanya project—389 of 437 (89%). In 2004 and 2006 observations for all of the province's schools are included. All of the results presented below are robust to the exclusion of schools for which 2002 exam data are not available.

Much like Gauteng Province, data are also available on schools' poverty ranking, poverty quintile, per capita funding and functional status from the WCED's

Financial Management Directorate. Racial composition, enrollment and class size data collected through the Annual School Survey as well as static data on each school's district and former educational authority were received from the department's Education Research Directorate.

Data on the receipt of computers were received from the Khanya project office. The office was able to provide a specific account of which schools received computers from the project and in which year the computers were received.

Namibia

The data I collected for Namibia include observations for 464 secondary and combined[2] schools between the years 1998 and 2006. The SchoolNet project included 133 of those schools, and computer deployments began in the year 2001 with over 20 schools deployed in each year from 2001 through 2005, with the exception of 2003 when only eight schools were deployed.

The outcome measure used to evaluate Namibia's SchoolNet project is the set of school-level results on the national junior secondary certificate (JSC) examination. The exam is administered at the end of Grade 10 and used to determine which students will earn the certificate. A JSC is necessary for Namibian students to proceed to the senior secondary phase of their education, which consists of grades 11 and12. Unlike previous grade levels, students who fail the JSC exam may not repeat Grade 10 in a conventional public school, and as a result this has become a major school exit point for many Namibians. Many employers also use the JSC as a prerequisite for entry-level positions of employment.

[2] Combined schools in Namibia typically include primary and the early secondary grades, most commonly Grade 1 through Grade 10.

The JSC exam results for all subjects were made available by the Namibian Ministry of Education's (MoE) Directorate of National Examinations and Assessment (DNEA). The final outcome variable used consists of composite results for six subjects—mathematics, physical science, life science, history, geography and English—representing each core subject area in which all students must take an examination. Overall results and results disaggregated by gender are available at the school level annually from 1998 to 2006. Each set of scores is normalized to a mean of zero and standard deviation of one in each year. Data are also available on the number of students taking the examination.

In each year from 1998 to 2006 data are also available for each school on average teacher qualification levels, average teacher years of experience and average class size. These data are collected on a survey comparable to the Annual School Survey in South Africa and conducted by the MoE's own EMIS division. Each school in the country is required to complete and return the questionnaire which collects mostly demographic information on schools' teachers and students. Average teacher qualification levels are calculated using the MoE's own index, which corresponds roughly to a one unit increase for every 2 years of training.

Data on schools' static characteristics were also made available from the EMIS division. These data include the name of each school's region, circuit and cluster. In Namibia, schools are divided into 13 regions, managed by regional education offices, which correspond to the country's political regions. In many cases the regions also correspond roughly to areas historically populated predominantly by one of the country's several ethnic groups. Within regions, schools are divided into organizational units called circuits which are overseen by a circuit office and circuit inspectors. Within circuits schools are divided into clusters, small groups of schools which are located in close proximity to each other that serve the same general community and often collaborate on tasks such as writing year-end exams. Additionally, data are available from the EMIS division on whether or not the school had fixed-line telephone access in 2004. I use telephone access a proxy for rural or non-rural location of Namibian schools.

Data on receipt of computers were retrieved from the SchoolNet Namibia website. The website lists the name of each school which received computers through the organization as well as the year in which computers were received.

IIb. Descriptive Statistics

In all three of the projects I am examining, there exist substantial differences between the schools chosen to receive computers and those which were not chosen. These differences are driven by projects' varying selection criteria, which differed substantially between projects. In Gauteng Province, GDE officials reported that selection focused primarily on schools serving poorer communities. This is evidenced in the summary statistics for Gauteng Province schools shown in Table IIb-1. Treated schools in Gauteng score substantially higher on the GDE's school poverty index than those schools which were not selected into the Gauteng Online project. Not unexpectedly, senior certificate examination pass rates are therefore also substantially lower in treated schools than other schools in the province.

Racial composition of the student body may also have played a role in selection of schools into the Gauteng Online project. In an interview, one individual familiar with the project selection process reported that some less-poor schools may have been chosen to participate in the project if they had racially diverse student bodies. The explanation offered was that the department did not want to disadvantage students from poorer backgrounds attending such schools by disqualifying them from inclusion. Despite this selection factor, there are still major differences between treated and untreated schools in racial composition, as shown in Table IIB-1.

Table IIb-1: Differences Between Treated and Untreated Schools in Gauteng

	Pass Rate	Pass Rate (Male)	Pass Rate (Female)	Poverty Score	Percent White Students	Number of Schools
Treated	-.384	-.338	-.376	41.4	5.0	212 (44%)
Untreated	.306	.273	.309	30.1	32.7	266 (56%)

Notes: Summary statistics of 2002 values for schools which were chosen to participate in the Gauteng Online project versus schools which were not chosen. Senior certificate examination pass rates normalized at the school level to a mean of 0 and standard deviation of 1. Poverty score ranked on a scale from 1-100, with 100 being most poor.

Unlike Gauteng Province, South Africa's Western Cape province has a substantial rural population, and the majority of its schools are located in rural areas. While in Gauteng schools were selected at the provincial level, the WCED selects new Khanya schools at the district level. Four of the province's educational districts are considered urban areas, covering the city of Cape Town and the surrounding metropolitan areas. The remaining three districts cover the province's rural areas.

Table IIb-2 shows differences between treated and untreated schools in the province's urban and rural areas. In the urban districts, schools selected for the Khanya project have lower test scores, higher poverty levels, and a substantially different racial composition than other urban schools. The situation is reversed in the rural areas, however, with treated schools having higher exam scores and lower poverty levels than untreated schools. It is difficult to pinpoint the underlying reason for this differential selection between urban and rural districts.

While selection of schools varied from district to district, both district officials and Khanya staff reported that selection of schools in all districts is largely based on two factors—school poverty and management quality. The differences in treated groups between district types is evidence either that balancing the emphasis on each of these two selection criteria produced very different results across districts or that there are other important selection factors which are unknown. In the analysis to follow, I attempt to control for these differences with district-level fixed effects and by only matching treated and untreated schools within districts.

Table IIb-2: Differences Between Treated and Untreated Schools in Western Cape

		Lit./Num. Scores		Lit./Num. Scores (M)		Lit./Num. Scores (F)		Pov. Score	% White	No. of Schools
Urban	Treated	-.014	.157	-.004	.131	-.015	.184	49.4	3.6	198 (39%)
	Untreated	.191	.375	.164	.361	.177	.350	44.4	11.7	307 (61%)
Rural	Treated	.092	-.048	.072	-.068	.106	-.037	57.0	12.3	234 (40%)
	Untreated	-.218	-.379	-.186	-.338	-.216	-.382	63.1	9.0	354 (60%)

Notes: Summary statistics of 2004 values for schools which were chosen to participate in the Khanya project versus schools which were not chosen, disaggregated into groups of urban and rural schools. Systemic evaluation exam scores for literacy and numeracy normalized at the school level to a mean of 0 and standard deviation of 1. Poverty score ranked on a scale from 1-100, with 100 being most poor.

In Namibia, there was no firm set of selection criteria consistently applied from the beginning of the SchoolNet project. Rather, schools were either chosen for the project as part of a sub-project funded by an outside source or showed interest in receiving a computer lab from SchoolNet by applying directly to the organization and were later chosen to receive computers. SchoolNet did, however, generally focus on schools in historically disadvantaged parts of the country, especially in four of the country's political regions known collectively as the north-central region where a large proportion of Namibia's population resides.

Table IIb-3 illustrates the differences between treated and untreated groups of schools in Namibia's north-central region as well as other regions. The proportion of schools selected into the SchoolNet project was much higher in the north central region. Additionally, treated schools in the north-central region had much higher exam score levels than untreated schools, while in other regions the treated group had only slightly higher scores.

Table IIb-3: Differences Between Treated and Untreated Schools in Namibia

		JSC Scores	JSC Scores (Male)	JSC Scores (Female)	Avg. Class Size	Number of Schools
North-Central Region	Treated	-.111	-.096	-.114	37.2	91 (33.6%)
	Untreated	-.281	-.237	-.271	38.3	180 (66.4%)
Other Regions	Treated	.263	.223	.256	31.0	42 (21.8%)
	Untreated	.239	.205	.237	29.2	151 (78.2%)

Notes: Summary statistics of 2000 values for schools which were chosen to participate in the SchoolNet Namibia project versus schools which were not chosen, disaggregated by region. Junior Secondary Certificate exam scores normalized at the school level to a mean of 0 and standard deviation of 1.

Another important characteristic to note in the Namibia data is that of significant differential trends related to the selection of schools. While schools in the north-central region were more likely to be selected into the SchoolNet project, those schools also improved exam scores over time at a much faster rate than schools in other regions. Table IIb-4 shows the results of a simple regression of the following specification:

$$Y_{it} = \beta_0 + \beta_1 R_i + \beta_2 t + \beta_3 R_i * t \qquad (1)$$

where Y is the normalized exam score, R is a binary variable indicating whether a school is in the north-central region, t indicates the year of the observation in the panel, and R_i*t is an interaction term showing the linear time trend for north-central region schools relative to other schools.

Table IIb-4: Differential Trends in the Namibia Data

	Differential Trend	Region Main Effect	Year Main Effect
Dependent Variable: Normalized Exam Scores	.104*** (.013)	-.776*** (.080)	-.055*** (.010)

* Significant at the p<.1 level, **p<.05 level, ***p<.01 level

Notes: Regression results are specified according to equation (1). Standard errors are adjusted for clusters by school.

The difference in exam score trends between north-central region schools and schools in other parts of Namibia is highly significant both in a statistical and practical sense. North-central region schools improved at an average rate of .104 standard deviations per year over the course of the panel. This example highlights the importance of accounting for differential trends in my attempt at evaluating the Namibia project.

IIc. Identification Strategy

In order to estimate the causal effect of computer availability on student performance in the three projects of interest, my analysis must overcome the challenges induced by the nonrandom selection of schools for treatment. As outlined in the previous section, there are stark observable differences between treated and untreated schools both on outcome variables and other key covariates. Additionally, differential trends in the data related to school selection are also an important source of potential bias which must be considered in any analysis.

It should also be noted, however, that despite the nonrandom selection of schools for inclusion in the projects, the ways in which schools were selected and deployed do offer some advantages in evaluating the causal effect of computer availability on educational outcomes. Due to the rolling computer deployments in each project observed over several years, we can be more confident that differential effects unique to a specific time period are not substantially biasing the estimates. Additionally, evaluating projects where decisions regarding treatment were largely in the hands of education department or NGO officials eliminates many possible sources of endogeneity related to unobservable school characteristics, especially in cases where those responsible for selecting schools do not have much detailed information on individual schools.

For each project, I fit the panel data to a year and state fixed effects ordinary least squares (OLS) regression model. Most importantly, this model controls for school characteristics which do not vary over the length of the panel. It also controls for factors affecting all schools that are year-specific. Additionally, group-year effects are used to control for year-specific changes that act on significant groupings of schools, such as schools within the same circuit, district or poverty quintile. Allowing for such effects in the models controls for some possible differential time trends in the data. Beyond simple time trends, these controls absorb any factors affecting all schools in the given grouping in each year. Finally, important covariates for which data are available on an annual basis are also included in the panel regression models.

In order to demonstrate the robustness of the panel analyses and possibly bring to light potential biases in the estimates, a matching technique is also used to evaluate each project. In the Gauteng and Western Cape provinces, the level of decision-making in selecting schools into the projects is clear, and relatively consistent standards are assumed to be applied in each round of selection. Additionally, information is available from interviews with officials in both provinces on which major factors were considered in making school selections. Propensity score matching is therefore used as a secondary method of estimating the treatment effect in these provinces' respective projects. Propensity score matching estimates the pre-treatment probability of being selected into the treated group based on observable characteristics. The likelihood of treatment (propensity score) is then used to match treated schools to comparable untreated schools[3]. In Namibia, where significantly differing selection criteria may have been applied, a standard matching on observable variables method is used to identify untreated schools which are most similar to treated schools.

Finally, as an additional check for possible bias in the estimates, results are presented on the outcome variable of treated groups compared to control groups in the years prior to treatment. For the panel regressions, this involves adding one-, two- and three-year leads of the treatment variable to the original specification. In the cases of Gauteng Province and Namibia, where several pre-treatment

---

[3] Dehejia and Wahba (2002) provide a thorough discussion of the use of propensity score matching for estimating causal effects in nonexperimental settings.

observations are available per school, an interaction term is used to check for pre-treatment differences in trends between treated and untreated schools, conditional on the controls in the original specification. For the matching results, the treatment and constructed control group are compared on outcome variables in the years prior to treatment.

IId. Estimation Results

Gauteng Province

The regression equation used for the Gauteng Online project is:

$$Y_{it} = \beta_0 + C_{it} + C_{i(t+1)} + C_{i(t+2)} + C_{i(t+3)} + H_{it} + \gamma_{dt} + \delta_{qt} + \alpha_i + \lambda_t + W_{it} + \varepsilon_{it} \quad (2)$$

where Y_{it} is the normalized senior certificate pass rate, C_{it} represents a binary variable indicating receipt of computers from Gauteng Online, H_{it} represents a binary variable controlling for presence of Gauteng Online computers in later years, d and q represent districts and poverty quintiles, respectively, and W_{it} represents a set of variables of school characteristics varying annually—racial composition, provincial poverty index score, per-capita funding and number of students sitting for the senior certificate examination. The district and poverty quintile terms allow for district-year and quintile-year fixed effects. The *W* term includes important covariates that change annually for which data are available. This controls for any time trends related to these characteristics and adds parsimony to the specification by explaining some of the variation in test scores not captured by other controls.

The probit regression used to estimate the propensity score for Gauteng schools is:

$$\Pr(T = 1 \mid Q, R, D, M, F, \Delta E) = \Phi(\beta_0 + Q + R + D + M + F + \Delta E) \quad (3)$$