88495

Expanding Opportunities for South African Youth through Math & Science:

The impact of the Dinaledi program

Jurgen Blum, Nandini Krishnan, and Arianna Legovini[1]

World Bank 2010

Abstract

In 2001, the Government of South Africa launched the Dinaledi Schools program, aimed at increasing the number of African and Indian secondary students entering and succeeding in mathematics and physical sciences Senior Certificate exams. The Dinaledi program provides selected high schools with a combination of supplementary inputs, including teachers, training, textbooks and calculators as well as close monitoring by the National Department of Education. This paper estimates the impact of the Dinaledi program for about 350 schools assigned to the program in the course of its expansion in 2005. It exploits administrative panel data for the years 2005-07 using a difference-in-difference estimation strategy. This is combined with a propensity score matching approach that compares Dinaledi schools with schools that are observably similar prior to the implementation of the program.

It finds the program to be particularly effective in improving Senior Certificate results in Higher Grade (HG) physical sciences and mathematics. Estimates suggest that in 2007, the Dinaledi program nearly doubled the number of students who entered and wrote HG physical sciences. In 2007, an average of about 35 students entered HG physical sciences in Dinaledi schools, compared to an average of 18 students in non-Dinaledi schools that had similar observed characteristics in 2004. On average, about 6 or 60% more students passed in HG physical sciences. Preliminary findings indicate that the program had strongly heterogeneous effects across different school types. It was particularly effective in the provinces of KwaZulu Natal and Limpopo, which comprise a large share of former Bantustan schools. By contrast, it had no statistically significant effect on the number of students passing HG physical sciences in Free State, Gauteng, Mpumalanga, North West and Western Cape. These findings call for further differential analysis of Dinaledi program effects.

[1] Jurgen Blum is Public Sector Management Specialist, Nandini Krishnan is Senior Economist, and Arianna Legovini is Manager, Development Impact Evaluation, Research Group, at the World Bank.

TABLE OF CONTENTS

TABLES AND FIGURES

TABLES

FIGURES

'Special attention will need to be given to the compelling evidence that the country has a critical shortage of mathematics, science and language teachers, and to the demands of the new information and communication technologies.'

President Thabo Mbeki, State of the Nation Address, 2000

I. INTRODUCTION

This paper presents the results of a retrospective impact evaluation of the South African Government's Dinaledi Schools Program. It estimates the effect of a supply-side intervention package in secondary schools including teacher training and provision of learning materials on mathematics and physical sciences learning outcomes. It aims at informing the Government of South Africa's policy in extending, targeting and redesigning the Dinaledi program by assessing:

(1) to which extent the Dinaledi program has increased the number and performance of secondary students in mathematics and physical sciences Senior Certificate examinations;
(2) to which extent the impact of the Dinaledi program differed by province and between different types of schools.

The Government of South Africa adopted the Dinaledi program in 2001 as part of its National Strategy for Mathematics, Science and Technology Education (MSTE) in order to address a severe shortage of and ethnic imbalance among secondary students passing the Senior Certificate in mathematics and physical sciences. As a heritage of the Apartheid system, the number of high school graduates in these subjects is particularly low among African and Indian students. This situation has far-reaching consequences. South Africa's share of university graduates with natural science and engineering degrees compares poorly to countries with similar GDP levels. According to the National Department of Education (NdoE), in 2004 it was less than half has high as in Chile, Cuba, and Brazil[2]. At the beginning of the millennium, only 1.6% of South Africa's chartered accountants, 2.4% of engineers, 1.6% of dentists, 7.5% of medical doctors and 2.4% of actuaries were African.[3]

The Dinaledi program seeks to improve these mathematics and physical sciences education outcomes and to reduce ethnic and gender imbalances through focused supply-side interventions in selected schools. It provides eligible schools with a needs-based combination of supplementary inputs, including additional teachers, in-service teacher training, textbooks, scientific calculators, computers and other learning support materials. Dinaledi schools are monitored through school surveys, on-site school visits and phone interviews. The program also creates school-level incentives, such as performance awards, for improving mathematics and physical sciences learning outcomes. In order to enhance its cost-effectiveness, the program targets only selected public high schools, based on a combination of needs and minimum capacity criteria.

[2] Department of Education (2004), "National Strategy for Mathematics, Science and Technology Education (NSMSTE)", presented to Portfolio Committee on Education on 3 February 2004, http://www.pmg.org.za/docs/2004/appendices/040203strategymst.ppt.
[3] Ibidem.

This evaluation exploits administrative panel data in order to estimate the effect of enrolment in the Dinaledi program for about 350 schools newly added to the program in 2004. The data originates from two sources. Information on school characteristics is drawn from South Africa's Education Management Information Systems (EMIS) for the baseline year 2004. The National Department of Education provided information on student test scores, performance and enrolment in science and mathematics for the school leaving exam from for the period 2004-2007.

The identification strategy employed for estimating the Dinaledi program effect on physical sciences learning outcomes combines a difference-in-difference (DID) method with propensity score matching as suggested by Heckman et al. (1997). This combined strategy serves to distinguish the program's effect from effects due to considerable pre-existing differences between Dinaledi and non-Dinaledi schools. Such differences could bias estimation results. First, propensity score matching minimizes *observable* differences by matching each Dinaledi school with a control school that had similar observed characteristics in 2004, based on an aggregate indicator of similarity (the propensity score). Second, the DID component seeks to eliminate bias from *unobserved* pre-existing differences (such as teaching capacity) by comparing Dinaledi and non-Dinaledi schools (the first DID-"difference") with respect to the *change* (the second DID-"difference") in learning outcomes over time. The DID approach eliminates bias from unobserved differences if the assumption holds that outcomes in Dinaledi and non-Dinaledi schools would have pursued parallel trends, had Dinaledi schools not received the program. For mathematics learning outcomes, due to lack of reliable baseline data, a simple propensity score matching method is employed, which requires stronger identification assumptions.

The study's main finding is that the Dinaledi program has substantially improved Senior Certificate results in Higher Grade (HG) physical sciences and mathematics. Preliminary estimates suggest that in 2007, on average about 17 more students entered and wrote HG physical sciences and about 6 more students passed in Dinaledi schools than in non-Dinaledi schools that had similar observed characteristics in 2004. Equivalently, in 2007, on average 12 more students entered and 5 more students passed HG mathematics in Dinaledi than in non-Dinaledi schools. These results are associated with major HG enrolment rate increases. The Dinaledi program on average doubled enrolment rates of grade 12 students in HG physical sciences from 6% in matched comparison schools to 12%. In HG mathematics, it nearly doubled enrolment rates from on average about 5.5% to 11% in 2007. By contrast, the Dinaledi program had, on average, no statistically significant effect on passing rates[4] in HG mathematics and physical sciences in 2007. For HG mathematics, however, estimates suggest that the Dinaledi program did increase passing rates on average by 14 percentage points in 2005 and by 8 percentage points in 2006.

In contrast to HG physical sciences, the Dinaledi program has on average not affected the absolute

[4] This study employs two distinct definitions of (mean) passing rates. In both definitions, the term passing rate refers to the number of students passing the Senior Certificate examination (p_n) in a given subject divided by the number of the students entering (e_n) (not: writing) the respective subject in school n. While the first definition comprises only a restricted sample of schools with at least one student entering the respective subject, the second definitions also comprises schools with no students entering the respective subject, setting the passing rate to 0 for these schools. The first definition is referred to simply as passing rate, the second as entry-corrected passing rate (EC PR). Analogously, a writing-corrected passing rate definition (WC PR) is used for the share of students who pass out of the number of students who write the SC examination, setting the passing rate to zero in schools where no student *writes* the SC examination in the respective subject.

number of students passing Standard Grade (SG) physical sciences at statistically significant levels in 2006 and 2007. However, the Dinaledi program apparently *did* increase the number of students passing SG mathematics. In 2007, on average about 56 students passed SG mathematics in Dinaledi schools, compared to 47 students in comparison schools (see Table 19). For SG physical sciences, two countervailing factors influence this result. The Dinaledi program slowed the growth of enrolment rates in comparison schools, but also slowed a rapid decline in passing rates. In SG mathematics, the program did on average not (or only temporarily) reduce enrolment rates at statistically significant levels and increased passing rates.

The Dinaledi program effects vary over time. For example, its effect on the number of students entering HG physical sciences increased between from 2005 to 2007. In 2005, on average about 11 more students entered in Dinaledi than in matched comparison schools, compared to 14 students in 2006 and 17 students in 2007. However, its effect on the number of students passing HG physical sciences peaked in 2006 with about 8 students, compared to about 6 students in 2005 and 2007. Multiple factors may explain these trends. The program effect may vary *within individual Dinaledi schools*, due to punctual interventions (e.g. teacher training, textbook provision) whose effect may change over time. Also, the average program effect may vary, because not all schools that were formally part of the Dinaledi expansion in 2005 were actually treated at the same time. Finally, the program may interact with other important time trends, such as a remarkable growth in grade 12 enrolments between 2004 and 2007. Given that Dinaledi schools received different treatments at different times, attributing these trends to any individual factors is not possible.

The Dinaledi program has had strongly heterogeneous effects across different types of schools and provinces. Estimates suggest that it had, on average, no statistically significant effect on the number of students passing HG physical sciences in five out of nine provinces – Free State, Gauteng, Mpumalanga, North West and Western Cape. Yet, it showed strong effects in particular in Limpopo and KwaZulu Natal, but also in Eastern Cape[5]. In the two former provinces, it increased the average number of students passing HG physical sciences by about factor five or by 18 students, relative to comparison schools (22 students versus 4 students).

The Dinaledi program has been more effective in provinces with a high share of former Bantustan schools. A breakdown by former education authority under the Apartheid system suggests that assignment to the Dinaledi program showed particularly strong effects in former Bantustan schools, while no statistically significant effects can be observed in former House of Assembly schools (formerly for white students). KwaZulu Natal and Limpopo comprise a large share of former Bantustan schools, while in Free State, Gauteng, North West and Western Cape there are few if any Bantustan schools. Regarding HG physical sciences enrolment rates, the program has on average been most effective in schools with low pre-treatment enrolment rates, while having no statistically significant effect in schools whose enrolment rates were greater than 20% in 2004. Further analysis in particular by pre-treatment school and parental household characteristics will provide useful insight into context-dependent programs effects that may inform its future targeting.

[5] Estimating the Dinaledi program effect for Northern Cape is impossible by lack of sufficient observations.

There are several important threats to the validity of estimated Dinaledi program effects (see Technical Appendix I for details). They imply that estimates may be biased for Dinaledi schools which benefitted from the program in 2005 to 2007 (internal validity) and/or that these estimates cannot be extrapolated to non-Dinaledi schools, which might receive the program in the future (external validity):

First and foremost, the credibility of the provided estimates hinges on the assumption that Dinaledi schools are matched to non-Dinaledi schools such that both groups would have pursued parallel trends[6] in the absence of the Dinaledi program. However, due to lack of baseline data on key school performance determinants – such as the quality of teaching inputs and students' socioeconomic background – this study's ability to identify a plausible counterfactual of matched schools remains constrained.

Second, without further information, this study does not permit comparing the program's (cost-) effectiveness between different types of target schools. Such comparisons are bound to be lopsided because different Dinaledi schools received different input packages at different costs, based on needs. This study merely estimates the average effect of *assignment to the Dinaledi program* on the sample of current Dinaledi schools, ignoring the specific input package received by each school. Its findings thus may be poor predictors for the effect of a specific input package on any individual school.

Third, caution is warranted in expecting the Dinaledi program to show similar effects if expanded. The study shows that schools selected into the Dinaledi program by far outperformed other schools *before* benefitting from the program. It is very likely that these superior pre-treatment school qualities may interact with the Dinaledi treatment. If this is the case, a future expansion of the program would yield different average returns in schools with lower pre-treatment performance levels.

Given limited external validity of the Dinaledi program effects, qualitative research on the channels through which the Dinaledi program enhanced learning outcomes will be important for complementing this study's results. Such information will be particularly important for designing future policies aimed at enhancing mathematics and physical sciences learning outcomes.

The remainder of this paper is organized as follows. Section II describes the Dinaledi program and Section III describes the data. Section IV motivates the choice of the DID-matching identification strategy for estimating the Dinaledi program effect and discusses the plausibility of underlying assumptions. Section V presents the estimated program effects. Section VI concludes. Four technical appendices discuss methodological questions in further detail. Technical Appendix I discusses major threats to the internal and external validity of the estimated program effects. Technical Appendix II discusses the plausibility of the identification strategy set out in Section IV. Technical Appendix III conducts sensitivity checks for different sets of matching criteria and algorithms as well as for OLS regression as the standard alternative method for selection on observables. Technical Appendix IV summarizes different methodologies employed for estimating heterogeneous effects.

[6] The difference-in-difference method employed in this study serves to eliminate bias due to differences in outcomes (e.g. passing rates) that existed *prior* to the Dinaledi intervention between Dinaledi and comparison schools. Eliminating such bias is only possible under the (identification) assumption that these differences would have *remained constant over time* in the absence of the Dinaledi program, i.e. that Dinaledi and comparison schools would have pursued parallel trends.

II. BACKGROUND

The Government of South Africa (GoSA) adopted the Dinaledi program with the goal of improving secondary students' mathematics and science performance in Senior Certificate examinations, especially of African and female learners. Specifically, it set the target of doubling the number of mathematics HG passes from about 24,000 in 2004 to 50,000 by 2009.

A. MATHEMATICS AND SCIENCE EDUCATION IN SOUTH AFRICA

Policies aimed at improving mathematics and science education in South Africa serve two overarching objectives – overcoming the Apartheid-heritage of severe ethnic inequalities in access to quality education and developing critical human capital for the country's growth.

Under Apartheid, unequal access to education for different ethnic groups was institutionalized. The National Party government established parallel, separately administered education systems for white, colored, Indian and African students, disposing of very unequal teaching resources (for details see e.g. Kallaway, 2002). For example, according to Case and Deaton (1999), per student funding in "House of Assembly" (HoA) schools for white students was about 1.85 times higher than in "Department of Education and Training" (DoET) and about 2.5 times higher than in Bantustan schools for African students. Training institutions preparing teachers for African schools often did not offer mathematics as a specialization. As a consequence, non-white and particularly African students were systematically underexposed to mathematics and science education (OECD, 2008).

Since 1994, the GoSA has replaced these racially duplicated institutions by a single national education system. It has undertaken persistent efforts to promote equity in access to education (see e.g. Fiske et al., 2004). Despite these efforts, African students still have severely underprivileged access to mathematics and physical sciences education. As Khan (2004) shows, while the average passing rate for students of all ethnicities enrolled in HG mathematics in 2001 was 55.9%, only 20.0% of African students passed, according to the "language proxy method"[7]. While narrowing, the gap in enrolment and passing rates in HG mathematics and physical sciences between African and white/non-African students is far from closed.

This ethnic imbalance in access to education in mathematics and sciences has far-reaching consequences. It limits access to tertiary education in engineering and actuarial sciences, as most universities require a HG pass for these subjects (OECD, 2008). It is also seen as a major contributing factor to the underrepresentation of Africans in professions requiring quantitative skills. According to labor statistics for 2000, only about 9% of employed South Africans aged 15 to 65 were in occupations that require some mathematical competence, such as technicians and accountants (OECD, 2008).

International benchmarks suggest that South Africa's education system performs poorly in mathematics and science education, compared to other African countries at similar economic and human development levels. As shown in Table 1, in 2003, South African eighth-graders performed poorly compared to their peers from Tunisia, Egypt, Morocco, Botswana and Ghana, according to the Trends in

[7] The Language proxy method served to identify the ethnicity of exam candidates, which was not explicitly recorded from 1991 to 2002. It identifies candidates as "African" if they take an African language as an examination subject (Kahn, 2004).

International Mathematics and Science Survey (TIMSS) 2003[8]. Investing in mathematics and physical sciences education is crucial because a shortage of human capital is diagnosed as a binding constraint for enhancing South Africa's growth and reducing unemployment (Hausmann et al., 2008).

B. THE DINALEDI PROGRAM

The Dinaledi program is part of a broader set of policies aimed at improving mathematics and physical sciences learning outcomes. In secondary education, it is complemented by increased schooling funding, teacher recruitment and schooling infrastructure investments on the supply-side[9]. On the demand-side, the mathematics Olympiads by the African Institute for Mathematical Sciences, for instance, seek to strengthen students' incentives for learning in mathematics. In 2008, the National Senior Certificate examination system was redesigned. It no longer differentiates between Higher and Standard Grade exams but only offers a single examination levels, largely corresponding to HG. In primary education, it is complemented by programs such as the Minister of Education's Foundation for Learning Initiative, which seeks to strengthen fundamental numeracy skills.

The GoSA designed the Dinaledi program primarily as a supply-side intervention. It diagnosed the lack of qualified mathematics and physical sciences teachers as well as other crucial teaching inputs such as textbooks, computers, calculators and laboratory equipment as the binding constraint on the number of students successfully passing the Senior Certificate in these subjects. In 1997, a report found that only about 85% of mathematics and science teachers were professionally qualified as educators, and only 50% and 42% were qualified in their respective subjects (see Arnott and Kubeka, 1997). The difficulty of attracting qualified mathematics and science graduates to the teaching profession is seen as a major bottleneck aggravating this situation. Given their scarcity on the labor market, qualified graduates find more attractive opportunities, for instance, in private sector jobs. As a result of the lack of qualified teachers, reportedly many South African secondary schools do not offer classes in mathematics and physical sciences at all.

The Dinaledi program also comprised a demand-side component. In particular, it provided students with career guides in Grade 9, potentially reducing systematic undervaluations of the expected returns to passing mathematics and physical sciences matriculation exams.

Schools included in the Dinaledi program received, on the basis of need, a combination of the following treatments:

1. Teacher training program focused on improving teachers content knowledge in mathematics and sciences;
2. Recruitment of additional teachers;
3. Provision of textbooks in Mathematics and Physical sciences;
4. Provision of scientific calculators and/or computer resources;

[8] The TIMSS 2003 provides the latest available international benchmark of South African learners performance in mathematics ad sciences. South Africa did not participate in the TIMSS 2007.

[9] *Source*: National Department of Education, Technical Report National Senior Certificate Results, 2008.

5. Provision of Learning Support Material (LSM), including Grade 9 career guidance guides, exemplar question papers to support assessment of the curriculum (past High Grade mathematics and science question papers with marking memorandum), audio-visual cassettes;
6. Monitoring: school surveys, on-site school visits, phone interviews;

In addition, a number of Dinaledi schools received financial and in-kind support from private companies. Data on the allocation of Dinaledi treatments indicates that different Dinaledi schools received very different input packages. Some inputs, such as calculators (86% of Dinaledi-schools), higher education guides (67%), sample exam papers (41%) and textbooks (18% to 54% of Dinaledi-schools) were allocated to a substantial share of schools – but in very different quantities. Other program components only benefitted a smaller number of schools, in particular “adoptions” (24%) or financial support (9%) by private companies. Data on the amount of teacher training and additional teaching staff provided to schools is lacking. Table 2 provides an overview of the distribution of the above inputs across schools.

The Dinaledi program was rolled out in several phases. At its start in 2001, 102 schools were selected. In 2005, the program saw a major expansion, with the number of schools increasing to 400. In 2007, 29 schools of the 400 schools were de-listed and 117 schools added resulting in 488 Dinaledi Schools. This study estimates the effect of the Dinaledi program’s expansion on these 488 schools[10] between 2005 and 2007. The number has further grown to 500 schools in 2008. Table 3 shows a breakdown of the number of Dinaledi schools by province. Its last column provides the number of schools includes in this study.

During the program’s major expansion in 2005, the GoSA defined guiding principles for Dinaledi eligibility, in order to target resources cost-effectively to selected schools. Measurable eligibility criteria comprise a minimum number of 35 African Senior Certificate mathematics learners, and a minimum 50% Senior Certificate passing rate in mathematics and physical sciences. In addition, schools were required to “display basic levels of functionality” and to have “potential to improve both participation and performance in SC mathematics and physical science”. In view of these criteria, each provincial government preselected schools for final approval by the central government. However, an ex-post comparison between Dinaledi and non-Dinaledi schools shows that these criteria were only loosely applied (see Section III). One explanation is that different provincial governments weighted needs and capacity criteria differently.

III. DATA

A. DATA SOURCES AND MEASUREMENT QUALITY

The administrative panel data for this study come from two sources: First, information on school characteristics in the year 2004 is drawn from South Africa’s Education Management Information Systems (EMIS). This data serves to match Dinaledi schools with control schools that had similar observed characteristics at baseline, prior to the expansion of the Dinaledi program in 2005. Second, the National Department of Education provided information on the number and performance of students

[10] These 488 schools are referred to as “Dinaledi 2005 schools”.

taking Secondary Certificate exams for the pre-intervention year 2004 and three subsequent years 2005-2007. Both datasets were merged based on national school identification numbers and exam centre numbers.

The raw data covers a comprehensive sample of schools. With 3347 schools, it comprises nearly the entire population of South Africa's secondary schools. It includes data for 447 out of a total of 488 Dinaledi schools in 2007. In order to ensure basic comparability between Dinaledi and other schools, the sample underlying the remainder this study was restricted to public secondary, combined and intermediate schools with certain specializations[11]. These restrictions reduce the sample size to 3135 schools, comprising 440 Dinaledi and 2695 other schools. The resulting sample comprises 72 Dinaledi schools admitted to the initial program in 2001.

Shortcomings in data completeness and quality regarding *school characteristics* are a major constraint for this study. The EMIS provides detailed data on the number of students per school by grade, ethnicity and gender – but has less information on key determinants of mathematics and physical sciences learning outcomes, such as the number of mathematics and physical sciences teachers per school, class sizes or the availability of computers and other relevant teaching inputs. Data on student or household characteristics is lacking and proxy indicators such as school fees or whether the school is in an urban or rural setting are not available. School-level data varies significantly in completeness and quality between provinces.

Outcome data on student test participation and performance from the NDoE is detailed and comprehensive for physical sciences. The mathematics outcome data has the major caveat that pre-treatment data (for 2004) is apparently unreliable[12]. Without this data, estimates of the Dinaledi program effects on mathematics learning outcomes could be severely biased (see Section IV.D). Therefore, this study primarily focuses on estimating the Dinaledi effect on physical sciences learning outcomes.

B. DESCRIPTIVE STATISTICS: OBSERVED DIFFERENCES BETWEEN DINALEDI AND NON-DINALEDI SCHOOLS

Schools admitted to the Dinaledi program differed significantly from non-selected schools before the start of the intervention in 2004. Dinaledi schools comprised most schools with a specialization in mathematics and sciences (see Table 4). On average, they were about 1½ times larger than other schools (see Table 6). Dinaledi schools outperformed others in terms of their students' success in mathematics and physical sciences Secondary Certificate examinations (see Table 8). Both in HG and SG physical sciences, the average number of students from Dinaledi schools participating in, writing and passing matriculation exams was about twice as high as in other schools. While this does to some extent

[11]The sample was restricted to schools with the following specializations: "COMPREHENSIVE", "COMPUTER; MATHS SCIENCE AND TECHNOLOGY", "DANCE; MATHS SCIENCE AND TECHNOLOGY", MATHS SCIENCE AND TECHNOLOGY", "ORDINARY", "ORDINARY & TECHNICAL" and "TECHNICAL".

[12]Mathematics outcome data for 2004 are apparently frequently miscoded as zero when in fact missing. In 2004, over 93% of all schools in the raw dataset report that no student were enrolled in SG mathematics, whereas for the subsequent years this share declines to less than 2%. The respective values for HG mathematics, are 96% for 2004 and less than 40% in all subsequent years. Thus both in absolute terms and compared to subsequent years, the mathematics data for 2004 thus is clearly implausible.

simply reflect the fact that Dinaledi schools have a larger number of students on average, their enrolment and passing rates are also significantly higher, especially for HG physical sciences (see Table 8). For example, with a mean of 50%, the entry-corrected passing rate for HG physical sciences is nearly twice as high as in other schools (28%).

These differences reflect the targeting of the Dinaledi program towards schools that were already performing comparatively well in mathematics and physical sciences. In addition, they may partly be due to the fact that a majority of schools that benefitted from the first Dinaledi intervention in 2001 were readmitted to the 2004 program. Out of the total of 72 Dinaledi in the sample for this study, 57 were reselected as Dinaledi schools in 2004. The implications of this overrepresentation of Dinaledi 2001 schools for the identification strategy are discussed in Section IV.B.1.

The Dinaledi targeting towards African and female students does not translate into major differences in the ethnic or gender composition of the student body between Dinaledi and non-Dinaledi schools. In Dinaledi schools, on average 86% of students are African, in non-Dinaledi-schools 84%. In Dinaledi schools, there were on average 100 female African for 90 male African students, in contrast to a nearly balanced gender-ratio in other schools (100 female to 97 male students) (see Table 7).

C. DISPARITIES IN MATHEMATICS AND PHYSICAL SCIENCES LEARNING OUTCOMES

South Africa's Apartheid legacy as well as its geographic and socio-economic diversity implies great variance in mathematics and physical sciences learning outcomes and in access to quality schooling. This section briefly highlights these disparities along two dimensions – by provinces, which bear part of the administrative responsibility for the post-apartheid education system, and by schools' former attribution to the different, racially segregated education administrations under Apartheid.

1. By province

As shown in Table 9, the distribution of Dinaledi schools across provinces in the sample is equitable based on population. While the absolute number of Dinaledi schools per province ranges from 17 schools in Northern Cape to 101 schools in Gauteng, the number of Dinaledi schools per capita only varies slightly and favors less populous provinces, e.g. Free State, Mpumalanga, Northern Cape and North West Province.

Pre-treatment outcomes, in particular the shares of grade 12 students that enter HG and SG physical sciences Senior Certificate exams varied starkly across provinces at baseline. In Free State, Gauteng, Mpumalanga, Northern Cape, North West Province and Western Cape in 2004 on average at least one in ten students entered SG and about one in twenty entered HG physical sciences (see Table 10 and Table 11)[13]. Eastern Cape had the highest average enrolment rate for SG physical sciences (20%), but a relatively low enrolment rate for HG physical sciences (2%). By contrast, in Kwazulu Natal and Limpopo, average entry rates for HG and SG physical sciences were zero. Inter-provincial discrepancies regarding the passing rate in physical sciences existed in 2004, but were less pronounced. They were highest in

[13] All Figures reported comprise both Dinaledi and non-Dinaledi schools in the respective province.

Western Cape with 89% on average for SG and 78% for HG and lowest in Mpumalanga with 67% and 41% respectively.

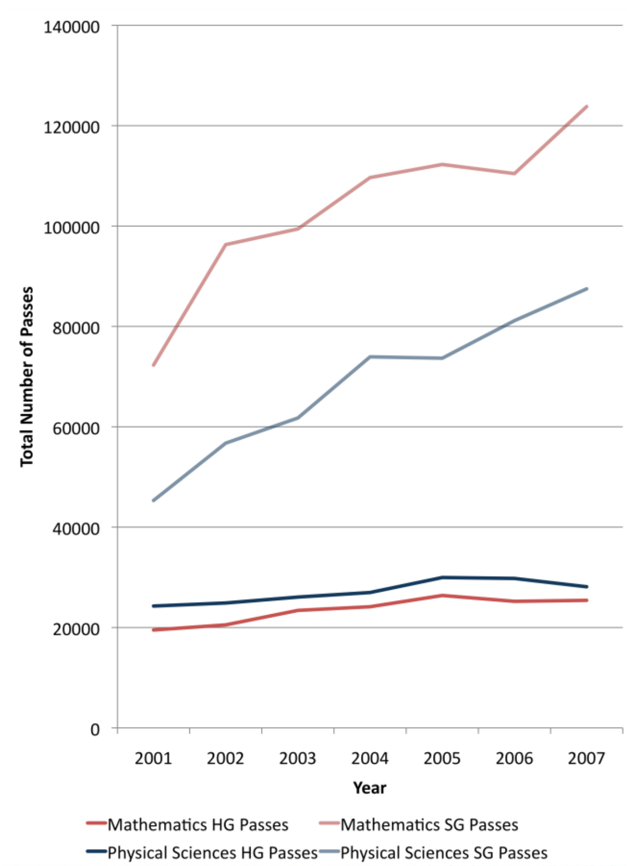
2. By Apartheid administration

Categorizing schools according to their administration under Apartheid shows that strong discrepancies prevail in particular between former House of Assembly schools for white students and Bantustan schools for Africans living in the homelands. While Bantustan schools have practically exclusively African students, Africans remain a minority in former HoA schools (see Table 12). As shown in Table 13 enrolment (14 % versus 2%) and passing rates (86% vs. 43%) in HG physical sciences were still multiple times higher in 2004 in former HoA schools than in former Bantustan schools included in this study's sample.

D. TRENDS

How has the number and passing rate of mathematics and physical sciences Senior Certificate takers evolved before and after the baseline year 2004? Aggregate data for the entire population of schools of the Ministry of Education (NDoE) shows that the number of Standard Grade (SG) passes has grown continuously and at a relatively rapid pace from 2001 to 2007 – by 71% for SG mathematics and by 93% for SG physical sciences. By contrast, the number of HG passes has only grown by 30% and 16 % respectively (see Figure 1) over the same period. These trends mostly reflect a general growth in the number of graduates, not a major increase in the shares of grade 12 students who take and pass the Senior Certificate in mathematics and physical sciences. As shown in Figure 2 and Figure 3, the share of mathematics writers and passers (including HG and SG) in total has nearly stagnated since 2001 – it has only grown from 58.2% to 61.5% in 2007 and from 38.1% to 40.2% respectively.

Figure 1. Growth of the absolute number of mathematics and science passes 2001-2007



Source: NDoE report, 2007

What has been the contribution of Dinaledi

2005 schools to this trend? A breakdown (see Figure 4) based on a sample of 1,571 schools[14] suggests different trends for SG and HG physical sciences. Both in SG and HG physical sciences, Dinaledi schools outperformed other schools between 2004 and 2006 regarding the growth of enrolment rates. But whereas this growth differential was modest for SG physical sciences, it was strikingly high for HG physical sciences. From very similar rates of 4 to 5% in 2004, Dinaledi schools nearly doubled their average HG enrolment rate to 8% in 2006, while it slightly declined in non-Dinaledi schools. This growth differential is consistent with this study's finding that the Dinaledi program considerably affected HG physical science enrolment rates (see Section V.A.1).

Trends in aggregate passing rates also differed considerably. In SG physical sciences, Dinaledi and non-Dinaledi schools pursued similar trends. Average SG physical science passing rates more or less stagnated above the 70% mark between 2004 and 2007 and remained about 7-10 percentage points higher in Dinaledi schools than in other schools. By contrast, HG passing rates fell continuously from on average 72% in 2004 to 58% in 2007. While Dinaledi and other schools both started from the same 72%-mark in 2004, in 2005 passing rates in Dinaledi schools dropped 13 percentage points *below* passing rates in other schools (63% versus 76%). This differential persisted till 2007. This study's findings suggest that this differential is not attributable to the Dinaledi program, but to pre-existing differences between Dinaledi and other schools (see Section V.A.1).

IV. IDENTIFICATION STRATEGY

A. OVERVIEW

Estimating the Dinaledi program effect is difficult because major differences between Dinaledi and other schools clearly *existed before* the program, but are – at least in part – *unobserved*. Dinaledi schools may outperform their peers simply because they had a considerable head start. To estimate program effects accurately, Dinaledi recipients should be compared to schools that started from a similar mark but did not benefit from the program. For example, comparable schools should on average have as many and as well qualified mathematics and physical sciences teachers per student, have students from similar socioeconomic backgrounds and dispose of similar quality textbooks and other teaching materials before the program. By comparing schools with similar baseline characteristics, claiming that learning outcome differences between Dinaledi and non-Dinaledi schools in 2007 were due to the Dinaledi program – and not to a head start or other pre-existing differences – would become more plausible. But data on such key determinants of mathematics and physical sciences learning outcomes is very limited. In order to estimate the Dinaledi program effect (on physical sciences learning outcomes) despite this lack of data, this study employs a twin identification strategy:

First, it seeks to *minimize differences between Dinaledi and comparison schools*, by approximating *unobserved* through *observed* characteristics. Data *is* available on pre-treatment *outcomes*. Dinaledi and non-Dinaledi schools can thus be matched based on the share of their students entering and passing the Secondary Certificate in physical sciences. In addition, *observed* school characteristics that do *not*

[14] The sample comprises schools for which complete data were available for all years between 2004 and 2007. It is not necessarily representative of South Africa's secondary school population.

directly determine school performance are sometimes *correlated* with unobserved performance determinants and thus serve as *indirect proxies*. For example, the size of a school may be correlated with its student-teacher ratio. Section IV.B provides further detail on the method employed in order to minimize observed differences (propensity score matching), and Technical Appendix II contains a technical discussion.

Second, this study seeks to eliminate bias from *unobserved differences* – provided they remain constant over time – by comparing Dinaledi and non-Dinaledi schools based on the *change* in mathematics and physical sciences learning outcomes over time – not based on *absolute* outcomes. This is possible thanks to the availability of time series data on learning outcomes. For example, rather than comparing Dinaledi physical sciences passing rates in 2007 to passing rates in comparison schools in 2007, the study compares the *growth in passing rates* from 2004 to 2007 between treated and comparison schools. Even if Dinaledi schools had a head start, i.e. higher passing rates than non-Dinaledi schools in 2004, this approach would correctly identify the Dinaledi program effect, provided that passing rates would have evolved in parallel in both groups of schools had none of them benefited from the Dinaledi program. Section IV.C discusses the plausibility of this parallel trends assumption, which is crucial for the difference-in-difference (DID) estimation. Appendix II adds technical details.

For mathematics learning outcomes, it was not possible to use the second DID-component of this identification strategy, due to the lack of reliable time-series outcome data, which is not available for mathematics in 2004 (see Section III.A). The estimation of Dinaledi program effects on mathematics learning outcomes thus needs to rely on stronger identification assumptions – and is more likely to provide biased estimates than for physical sciences. Section IV.D discusses the implications of this lack of data for the reliability of mathematics estimates.

B. PROPENSITY SCORE MATCHING

This study seeks to reduce *bias due to observable differences* by matching Dinaledi with similar non-Dinaledi schools. In principle, for each Dinaledi-school, a comparison school is identified that has similar baseline characteristics. This method requires choosing both matching characteristics (How to measure school similarity?) and a matching algorithm (Given a measure of similarity, how to identify matches?) for selecting or weighting comparison schools. This section briefly describes and motivates these two choices.

1. Matching criteria

Prior to matching comparison schools to Dinaledi schools, the sample of non-Dinaledi schools is restricted to plausible matches by excluding schools that do not belong to the same categories as Dinaledi schools. Nearly all Dinaledi schools are public secondary, combined and intermediate schools with certain specializations. Therefore only schools that belong to these categories are used as comparison schools. For example, private, elementary or high schools specializing in arts and music are omitted, as they are likely to be poor comparisons. Second, all schools which already benefitted from the Dinaledi program in 2001 are excluded from the sample. If these schools were included, estimates of the Dinaledi expansion's effects in 2005 may risk partly capturing the effects of the program's first phase. It seems unlikely that 2001 Dinaledi schools would have evolved in parallel to non-Dinaledi schools, had the program not been expanded (see Section IV.C). Finally, schools that lacked data on relevant baseline characteristics were also excluded. Imposing these restrictions reduces the sample of Dinaledi schools from a total of 447 schools to about 350 schools[15] (see Technical Appendix II for further details).

The guiding criterion for the choice of matching characteristics (covariates) is whether they may cause bias if omitted because they *both* influence learning outcomes *and* are correlated with Dinaledi school status. Ideally, matching covariates should, for example, include indicators of the quantity and quality of (mathematics & physical sciences) teachers and of other school inputs, as well as child and household characteristics. However, as mentioned above, these key determinants of students' learning outcomes are *not observed* in existing data sources. Failing to control for these characteristics is particularly apt to cause bias, because the Dinaledi program was explicitly targeted towards schools that differed significantly from other schools in their ability to teach mathematics and physical sciences. In order to mitigate the risk of bias to the extent possible in the absence of such data, the covariates of this study's main specification have been selected according to three criteria (see Technical Appendix II for details):

1. *Their suitability as proxy indicators of school quality and student socio-economic backgrounds*, based on theory and evidence from the empirical education literature. In particular, schools are matched based on their students' performance in physical sciences before the Dinaledi program in 2004 – assuming that schools of similar size that produced similar numbers of students entering and passing physical sciences, have – on average – comparable teaching capacity (see Technical Appendix II for details).

2. *Their suitability for predicting Dinaledi eligibility.* Second, the choice of school characteristics controlled for in this study's main specification is driven by information about the Dinaledi program's design and targeting. Including the number of African and female students reflects its targeting towards these groups. School size measures matter particularly because the Dinaledi effect on learning outcomes – e.g. the number of students entering and writing physical sciences matriculation exams – is measured in absolute numbers, rather than in shares of students enrolled in the final grade 12. To make outcomes comparable, Dinaledi and non-Dinaledi schools thus need to have similar size. School size also matters, because large schools are likely to differ from small schools in unobserved ways. For example, large schools may be found more often in urban areas with students from different socio-economic backgrounds, have more specialized teachers due to scale effects and may be more or less likely to be eligible to the Dinaledi program. Finally, the number of students with English as language of instruction is an important control, as it is a proxy indicator of the student bodies' ethnic composition. These qualitative considerations are backed by testable criteria as suggested by Heckman et al. (1997) (see Technical Appendix II for details).

[15] The sample slightly varies in size for each post-treatment year (between 346 schools in 2006 and 353 schools in 2005), as outcome data is not available for all post-treatment years (2005-2007) for all schools. For the estimation of the Dinaledi program effect in 2007, 88 Dinaledi schools were excluded from the original sample of 440 Dinaledi schools, (a) because they already were Dinaledi schools in 2001 (57 observations dropped), or (b) because they lacked data on characteristics used for estimating the PS (18 observations dropped), or (c) because they lacked outcome data for 2007 (6 observations dropped) or (d) by lack of comparison schools to meet the common support condition (7 observations dropped). For the 2006 and 2005 estimations, similar numbers apply.

3. *Their suitability for balancing baseline characteristics among Dinaledi and control schools*. Finally, the choice of matching characteristics should ideally lead to balanced baseline characteristics between treated and matched control schools after matching, lending credibility to the assumption that – at least for observables characteristics – Dinaledi schools and selected matched schools are on average comparable. Tables 6 to 8 in the Technical Appendices show that the main specification successfully balances most observed baseline characteristics for the DID matching results for years 2006 and 2007 (see Technical Appendix II for details).

BOX 1:
MAIN SPECIFICATION MATCHING COVARIATES

A. School characteristics: size, ethnic and gender composition, language of instruction

- Total male enrolments
- Total female enrolments
- Number of African learners in the school
- Enrolments in grade 10
- Enrolments in grade 11
- Male Enrolments in grade 12
- Female Enrolments in grade 12
- Number of African learners in grades 10 to 12
- Number of learners with English as medium of instruction

B. Learning outcome measures

- Total number of students writing SG physical science
- Total number of students writing HG physical science
- Total number of students passing SG physical science
- Total number of students passing HG physical science
- Entry-corrected passing rate out of those writing SG physical science
- Writing-corrected passing rate out of those writing HG physical science

C. Province dummies

Box 1 lists the pre-treatment covariates that have been selected based on these criteria for this study's main specification. They comprise selected baseline characteristics, pre-treatment Secondary Certificate outcome measures and province dummies. Baseline characteristics include size, ethnic and gender characteristics as well as the language of instruction. Outcome measures include the number of students writing and passing HG and SG physical sciences and the respective passing rates[16] (see Technical Appendix II for details).

2. Matching algorithm

Choosing the matching algorithm essentially implies choosing the number of comparison schools used for constructing the counterfactual as well as the weights attributed to these schools. At one extreme, single nearest neighbor matching identifies only a single match for each Dinaledi school. At the other extreme, algorithms include all non-Dinaledi schools in the counterfactual, but give more weight to good matches than to poor matches (e.g. Kernel matching). Between these extremes, the choice of a matching algorithm needs to be guided by two competing objectives: On the upside, using only few, precise matches reduces bias by ensuring that control schools are very similar to Dinaledi schools. On the downside, using few matches limits the sample size – and thus increases the smallest program effect that can be detected at standard significance levels. Inversely, using many matches allows detecting smaller program effects, but risks introducing bias by including poorer matches.

This study's main specification one-sidedly seeks to minimize bias by using a single nearest neighbor matching algorithm, accepting that small program effects cannot be detected at statistically significant levels. One-sidedly opting for bias-minimization is possible thanks to the large sample of schools

[16] Pre-treatment mathematics learning outcomes are not used as matching covariates due to the lack of reliable data.

available for this study, comprising the majority of South Africa's high school population. With about 350 Dinaledi and 350 control schools, the resulting sample is still large enough to detect sufficiently small program effects.

Accordingly, a single match is identified for each Dinaledi school. Given the large number of selected matching covariates, matches are identified based on an aggregate indicator of their observed characteristics, the propensity score. The propensity score estimates the likelihood of being selected for the Dinaledi program, given a school's covariates. Rosenbaum and Rubin (1983) show that in principle controlling for the propensity score provides the same estimates of program effects as controlling for the covariates used for estimating the propensity score. Matches are allowed to be used multiple times, if they are the closest match for several Dinaledi schools (see Technical Appendix II for further details on the matching algorithm).

C. DIFFERENCE-IN-DIFFERENCE

Besides reducing bias due to observable differences via matching, this study seeks to *make unobserved differences between Dinaledi and matched non-Dinaledi schools negligible* – provided that these differences would have remained constant over time in the absence of the Dinaledi program (difference-in-difference (DID) component).

How plausible is this parallel trends assumption? For example, in 2004, the number of students entering and writing HG physical sciences was on average by six students *lower* in Dinaledi schools than in matched control schools[17] (20 vs. 26 students). If this difference had remained constant until 2007 in the absence of the Dinaledi program, the DID approach would provide an unbiased estimate of its effect on the number of students entering and writing HG physical sciences. By contrast, this assumption would fail, if for example overall grade 12 enrolment had grown between 2004 and 2007 (as it did, see Section III.D) *and* if Dinaledi schools on average had students with stronger preferences for taking HG physical sciences (an unobserved characteristic). In this case, the enrolment growth would translate into a higher number of students taking HG physical sciences in Dinaledi than in non-Dinaledi schools. Consequently, the initial gap in HG physical sciences enrolment would have narrowed *without the Dinaledi program* – and the DID strategy would overestimate the Dinaledi program effect.

While the parallel trends assumption is not directly testable for the treatment period (2005 to 2007), it *is* testable for the pre-treatment period (2004 and prior). If it can be shown that physical sciences learning outcomes in Dinaledi and matched control schools pursued parallel trends prior to 2004, this would lend strong support to the assumption that they would have continued to evolve in parallel after 2004 in the absence of the Dinaledi program (see Technical Appendix II for details).

D. IDENTIFICATION STRATEGY FOR MATHEMATICS LEARNING OUTCOMES

The DID-component of the above identification strategy cannot be employed for estimating the Dinaledi program effect on mathematics learning outcomes. It requires reliable pre-treatment (2004 or prior) learning outcome data, which is available for physical sciences, but lacking for mathematics (see Section

[17] For the sample of schools with outcomes data availability from 2004-2007 (see Appendix Table 2, row 4).

III.A). Accordingly, stronger identification assumptions underlie the reported preliminary estimates of the Dinaledi program effect on mathematics outcomes. These estimates would only hold if, in 2004, matched control schools had not differed from Dinaledi schools in observable and unobservable characteristics that influence mathematics learning outcomes. Given the lack of data on key determinants of mathematics learning outcomes, this is a strong and highly implausible assumption. For example, if HG mathematics enrolment had differed between Dinaledi and control schools prior to the Dinaledi program (as is the case for HG physical sciences), this difference would erroneously be attributed to the Dinaledi program, if constant over time. Obtaining reliable pre-treatment mathematics outcome data is therefore necessary to provide plausible estimates of the Dinaledi program effects on mathematics learning outcomes.

V. RESULTS

A. AVERAGE EFFECTS ON DINALEDI SCHOOLS (ATET)

1. <u>Physical Sciences</u>

Under the parallel trends identification assumption, this study's main finding is that assignment to the Dinaledi program on average increased the number of students entering, writing and passing *HG physical sciences* substantially. The Dinaledi program approximately doubled the enrolment rate of grade 12 students in HG physical sciences from six to twelve percent, at the 1%-significance level (see Table 14). However, it had no positive effect on HG physical sciences passing rate (see Table 14). Estimates suggest that the Dinaledi program entailed a slight reduction of HG physical sciences passing rates but mostly at statistically insignificant levels. In absolute terms, on average about 35 students entered HG physical sciences in Dinaledi schools in 2007, compared to only about 18 students in matched control schools (see Table 15). Analogously, on average 16 students passed HG physical sciences in Dinaledi schools, compared to only 10 students in control schools (see Table 13). These results hold at the 1%-significance level. As Table 16 and Table 17 show, these estimates are robust to a bias-corrected estimation algorithm as suggested by Abadie and Imbens (2002) (see Technical Appendix II for details).

In contrast to HG physical sciences, the Dinaledi program has on average not affected the absolute number of students *passing SG physical sciences* at statistically significant levels. In 2007, on average 35 students passed SG physical sciences both in Dinaledi and comparison schools (see Table 15). Two countervailing factors influence this result. On the one hand, the Dinaledi program reduced the growth rate of students entering and writing SG physical science. While the SG physical sciences enrolment rate grew by on average 5 percentage points from 2004 to 2007 (peaking in 2006) in comparison schools, it only grew by 3 percentage points in Dinaledi schools (see Table 14). On the other hand, estimates suggest that the Dinaledi program slowed a rapid decline in SG physical sciences passing rates – but only temporarily. From 2004 to 2006, SG physical sciences passing rates declined by about 10 percentage points in non-Dinaledi schools, compared to only 3 percentage points in Dinaledi schools (see Table 14). However, this differential does not persist in 2007.

Estimates of the Dinaledi program effects on physical sciences outcomes vary considerably for the three post-treatment years 2005 to 2007. For interpreting these trends, further inquiry is needed into their correlation with the sequencing of Dinaledi implementation and with overall enrolment trends. Including the 2008 matriculation results in this study would be helpful for corroborating the findings for previous years.

2. Mathematics

As argued in Section IV.D, estimates of the Dinaledi program effect on mathematics learning outcomes are likely to be severely biased, as they are based on implausible assumptions. Hence, the following preliminary results should be interpreted with care.

The Dinaledi program had similarly substantial effects on the number of students entering, writing and passing the Senior Certificate in HG mathematics as for physical sciences. It nearly doubled the HG mathematics enrolment rate from on average about 5.5 to 11 percent (see Table 18). This result holds at the 1%-significance level. By contrast to physical sciences, it temporarily increased passing rates[18] by about 8 percentage points in 2006 compared to control schools, but this effect does not persist in 2007 (see Table 18). As a result, in 2007, on average about 25 students entered and wrote HG mathematics in Dinaledi schools, compared to only about 13 students in matched control schools (see Table 19). In the same year, on average about 14 students passed HG mathematics in Dinaledi schools, compared to only 9 students in matched control schools. These results hold at the 1%-significance level and are robust to bias-correction (see Table 20 and Table 21).

By contrast to physical sciences, the Dinaledi program apparently *did* increase the number of students passing SG mathematics. In 2007, on average about 56 students passed SG physical sciences in Dinaledi schools, compared to 47 students in control schools (see Table 19). This difference is significant at the 1% level. This is partly due to higher passing rates: In 2007, passing rates (of those entering SG mathematics) were 8 percentage points higher in Dinaledi schools than in control schools, at the 1%-significance level (see Table 18). Second, by contrast to physical sciences, the Dinaledi program did not (or only temporarily) reduce SG mathematics enrolment rates at statistically significant levels.

B. HETEROGENEOUS TREATMENT EFFECTS

1. By former department under Apartheid[19]

Persistent inequalities in access to quality education in today's South Africa can partly be traced back to segregated, parallel education systems for different ethnic groups under the Apartheid system. This section presents estimates of heterogeneous Dinaledi program effects on three groups[20] of schools that belonged to separate Apartheid education systems: House of Assembly (HoA) schools (formerly for white students), Department of Education and Training schools (DoET) (formerly for black students) and

[18] Passing rates of students entering HG mathematics.

[19] The discussion of heterogeneous treatment effects does not cover mathematics learning outcomes for the time being, constrained by lack of reliable pre-treatment data.

[20] Heterogeneous effects for House of Representative (for … students) and House of Delegates schools (for … students) are not estimated, by lack of observations. Regrouping all Bantustan schools may hide heterogeneous effects between Bantustans, which merit further inquiry.

schools in nine out of ten[21] of the semi-autonomous black homelands or Bantustans. Heterogeneous effects of the Dinaledi program on these groups of schools are likely given the major differences that persisted between them in 2004 (see Section III.C.2).

The Dinaledi program had strikingly different effects on these three groups of schools (see Table 22 to Table 25). On average, it had practically no statistically significant effect on former House of Assembly schools. By contrast, in Bantustan schools in 2007, it more than tripled the number of students enrolled in and septupled the number of students passing HG physical sciences compared to control schools. In 2007, on average, 43 students entered and 15 students passed HG physical sciences in Dinaledi Bantustan schools, compared to only 11 and 2 students respectively in non-Dinaledi schools. These results hold at the 1%-significance level. They are largely explained by a drastic increase in HG physical sciences enrolment rates due to the Dinaledi program by 11% to 16% compared to an average of 5% in control schools in 2007. HG physical sciences passing rates were not affected at statistically significant levels. As shown in Table 33, these heterogeneous effects are robust to controlling for pre-treatment (propensity score) matching criteria and in particular for pre-treatment HG physical sciences enrolment rates.

In former Department of Education and Training schools, the Dinaledi program effects are similar, but much less pronounced. In 2007, on average about 10 more students entered and 4 more students passed HG physical sciences in Dinaledi schools than in control schools, at the 1% significance level. Both values are nearly twice as high as in control schools. In 2007, HG physical sciences enrolment rates were 4% higher in Dinaledi DoET schools than in control schools, at the 1%-significance level. By contrast, HG physical sciences passing rates were not affected at statistically significant levels.

It is important to emphasize that – without additional information – these findings do not necessarily imply a rationale for targeting the Dinaledi program to Bantustan schools in the future. They *do* imply that assignment to the Dinaledi program led to stronger improvements in mathematics and physical sciences learning outcomes in former Bantustan schools than in former HoA schools. But, besides general constraints on external validity (see Technical Appendix I for details), concluding that the Dinaledi program has been more (cost-)effective in Bantustan schools may be wrong for a number of reasons:

First, the Dinaledi program may have meant very different things in HoA and Bantustan schools. As the program components were tailored to individual school needs, former HoA schools may on average have received less – and less costly – support than Bantustan schools. Second, decreasing marginal returns may explain that the Dinaledi program was less effective in HoA schools. For example, the impact of providing training in physical sciences education to teachers on HG physical sciences enrolment in a school that has no qualified teachers may be higher than in a school that already disposes of a number of qualified teachers in this area. Finally, the classification as HoA and Bantustan schools may only be a rough proxy for other (observed and unobserved) school characteristics that are the root cause for the strikingly heterogeneous treatment effects. In this case, the root cause of heterogeneous effects should serve as targeting criterion for a potential extension of the Dinaledi

[21] Former Qwaqwa schools are not included, because no observations are available.

program. Further inquiry into heterogeneous treatment effects in breakdowns by other school characteristics is likely to shed more light on the factors determining heterogeneous treatment effects (see Section V.B.3). Proxy measures for students' socio-economic background – e.g. household survey data – would be helpful for conducting further differential analysis.

2. By province

Estimating heterogeneous treatment effects by province is constrained by the number and quality of observations available for each province. For South Africa's four northern provinces – Northern Cape, North West, Limpopo and Mpumalanga – the sample of schools with sufficient baseline data is too small for at all providing reliable within-province matching estimates. For the five southern provinces – Western Cape, Eastern Cape, Kwazulu Natal, Free State and Gauteng – there are sufficient observations for estimating the Dinaledi program effect, but matching results are mostly poor because there are few valid comparison schools for Dinaledi schools within each province. Consequently, significant baseline differences remain between Dinaledi schools and matched controls from the same province and risk causing bias.

In order to enlarge the pool of available matches and to improve the quality of matching results, this study reports heterogeneous treatment effects for three sub-sets of provinces, rather than for individual provinces (see Table 26 to Table 29). The first sub-set comprises Free State, Gauteng, Mpumalanga, North West and Western Cape, the second Kwazulu Natal and Limpopo. Eastern Cape shapes the third sub-set as a single province. The Dinaledi effect for Northern Cape is not estimated, due to lack of observations[22]. Provinces were attributed to the same sub-set if the Dinaledi program effects were expected to be in a similar order of magnitude for all provinces within a set, based on preliminary province-by province estimations.

Regarding the absolute number of students passing HG physical sciences, the Dinaledi program had very different effects in the three sets of provinces in 2007 (see Table 26 to Table 29). In Gauteng, Free State, Mpumalanga, North West and Western Cape, it had on average no statistically significant effect on the number of students passing HG physical sciences. By contrast, in KwaZulu Natal and Limpopo, the program on average increased the number of students passing HG physical sciences by about factor five or by 18 students, compared to control schools (about 22 vs. 4 students), at the 1%-significance level. In Eastern Cape, the Dinaledi program more than doubled the number of students passing HG (about 6.1 vs. 2.4 students) compared to control schools, at the 1%-significance level.

These heterogeneous results are related to different Dinaledi effects on HG physical sciences *enrolment rates* in different province sets. In Free State, Gauteng, Mpumalanga, North West and Western Cape, the program increased enrolment rates only moderately from on average about 9% in control schools to 12% in Dinaledi schools, at the 5%-significance level. In KwaZulu Natal and Limpopo, the program more than doubled HG enrolment rates from about 9% in control schools to 22% in Dinaledi schools, at the 1%-significance level. In Eastern Cape, it about doubled HG enrolment rates from 5% to 11%, at the 5%-

[22] There are only three Dinaledi schools in Northern Cape with sufficient baseline information that meet selection criteria.

significance level. Across all three sub-sets of provinces, the Dinaledi program did not have a statistically significant effect on *passing rates* of students entering HG physical sciences.

Contrary to HG physical sciences, the Dinaledi program had no statistically significant effect on the number of students passing SG physical sciences in any province sub-set in 2007. This is the case despite some statistically significant Dinaledi effects on SG enrolment or passing rates. In Gauteng, Free State, Mpumalanga, North West and Western Cape, the D program increased SG passing rates by 8 percentage points (to 79% compared to about 71% in control schools). This change may have been compensated by a slight, but insignificant decline in SG enrolment rates. In Kwazulu Natal and Limpopo, the Dinaledi program slowed the growth of SG enrolment rates in control schools by about 6 percentage points, at the 1%-significance level, but without significant impact on the number of students passing SG physical sciences.

The Dinaledi program's starkly heterogeneous effects across provinces and across former departments under Apartheid are closely correlated. The two provinces where the Dinaledi program was most effective in increasing the number of students passing HG physical sciences – Kwazulu Natal and Limpopo – also comprise a large population of former Bantustan schools. In the Kwazulu Natal province, 62% of Dinaledi schools of the analyzed sample formerly belonged to the Kwazulu Natal Bantustan (see Table 30). In Limpopo, 92% of Dinaledi schools are former Bantustan schools, with 14% formerly belonging to Gazankulu, 38 % to Lebowa and 40 % to Venda. By contrast, Free State, Gauteng and Western Cape, where the program showed no significant effect, also did not comprise a single Dinaledi school from a Bantustan. Mpumalanga and North West province do not fit this pattern. Here, the Dinaledi program had no significant effect, although about 41% of Dinaledi schools were former Bantustan schools in Mpumalanga (from Kangwane (28%), Kwandebele (6%) and Transvaal (6%)) and 32% in North West Province (from Bophuthatswana).

3. <u>By other baseline characteristics: pre-treatment HG physical sciences enrolment rates</u>

The Dinaledi program effect also varied depending on other observed baseline school characteristics. In particular, the program's considerable average effect on HG physical sciences enrolment rates raises the question whether this effect varies depending on the pre-treatment enrolment rate in HG physical sciences in 2004. This section seeks to shed light on this question. It illustrates heterogeneity in one important dimension, but should be complemented by systematic inquiry into heterogeneity in other observable baseline characteristics, in particular with respect to parental household characteristics, if relevant data can be obtained.

The Dinaledi effect on HG physical sciences enrolment rates rapidly decreases with increasing pre-treatment enrolment rates. While estimations (see Figure 5 to Figure 7) suggest that the Dinaledi program increased HG physical sciences enrolment rates on average by over 6 percentage points in schools with pre-treatment enrolment rates in the 0% to about 5%- range, it had no statistically significant effect in schools with initial enrolment rates greater than 20%, holding constant pre-treatment characteristics used for matching (see Box 1). This finding is robust to different techniques for estimating heterogeneous effects (see Technical Appendix IV for details).

VI. CONCLUSION

This study's main findings are the following:

1. Severe regional and ethnic imbalances in mathematics and physical sciences learning outcomes persisted in South Africa in 2004, prior to the Dinaledi program. Major discrepancies in enrolment and passing rates in HG physical sciences across provinces and between former House of Assembly schools and Bantustan schools are indications for these imbalances.

2. The Dinaledi program has substantially increased the number of students entering, writing and passing the Senior Certificate Examination in HG physical sciences and probably in HG mathematics in the considered sample of about 350 Dinaledi schools. It has substantially increased *enrolment rates* of grade 12 students in HG physical sciences, without having major effects on *passing rates*. These findings are valid under the assumption that physical sciences learning outcomes in Dinaledi and matched control schools would have pursued parallel trends from 2004 to 2007 in the absence of the program. Obtaining pre-treatment outcome data for years prior to 2004 will be crucial to corroborate this assumption.

Preliminary estimates for mathematics learning outcomes show similarly positive Dinaledi effects on the number of students entering, writing and passing HG mathematics. However, these estimates are likely to be severely biased. Reliable mathematics outcome data for 2004 or before is necessary for providing more reliable estimates.

3. An initial analysis of heterogeneous program effects shows that assignment to the Dinaledi program affected different types of schools very differently. For example, it had particularly strong effects on HG physical sciences passes in former Bantustan schools, and accordingly in the KwaZulu Natal and Limpopo province, where a large part of the school population consists of former Bantustan schools. By contrast, it had less pronounced effects on former Department of Education and Training schools and it had no statistically significant effects on former House of Assembly schools. It also had no statistically significant effect on HG physical sciences passes in the Gauteng, Free State, Mpumalanga, North West Western Cape province. Further inquiry into these heterogeneous treatment effects would be desirable to shed more light on their root cause.

What are the implications of these findings for the future extension, targeting and design of the Dinaledi program?

This study identifies strong reasons for caution in extrapolating from past to potential future Dinaledi program effects. The primary reason is that the Dinaledi program was explicitly targeted towards well-performing schools – and is likely to have different effects if extended to less-performing schools. Second, the strikingly heterogeneous program effects – e.g. between former Bantustan and House of Assembly schools – do, without further information, not provide an unambiguous rationale for the program's future targeting. Given the program's needs-based design, higher effects e.g. in Bantustan schools may be explained by the fact that they received more – but also more costly – inputs than other schools. Further differential analysis is required for identifying the likely root cause of heterogeneous program effects, for which the former Bantustan status may only be a rough proxy. Finally, this study does not permit conclusions as to which Dinaledi input packages were more effective than others. Detailed information on the packages received by each school will be required to shed light on this question.

In sum, this study's findings should be considered as only one among different necessary sources of information for designing future policies aimed at improving mathematics and physical sciences learning outcomes in South Africa. Its value–added lies in providing quantitative estimates of the (heterogeneous) Dinaledi program effects as well as of prevailing disparities in mathematics and physical sciences learning outcomes. However, this study's approach is not suited to shed light on the crucial question *through which channels* the Dinaledi program affected learning outcomes. Qualitative research on the determinants and constraints of mathematics and physical sciences learning outcomes is therefore a necessary complement to this study for future policy design.

VII. BIBLIOGRAPHY

Abadie, Alberto (2003), "Semiparametric Difference-in-Difference Estimators", Review of Economic Studies, Vol. 72, No. 1, pp. 1-19.

Abadie, Alberto, A. Diamond and J. Hainmueller (2007), "Synthetic Control Methods for Comparative Case studies. Estimating the effect of California's Tabacco Control Program", Working Paper 12831, NBER Working Paper Series, NBER, Cambridge, http://www.nber.org/papers/w12831.

Abadie, A. and G. Imbens (2002), "Simple and Bias-Corrected Matching Estimators". Tech. rep., Department of Economics, UC Berkeley, http://emlab.berkeley.edu/users/imbens/.

Acemoglu, Daron and J.D. Angrist (2001), "Consequences of Employment Protection? The Case of the Americans with Disabilities Act", Journal of Political Economy, Vol. 109, pp. 915-957.

Angrist, Joshua D. and J.-S. Pischke (2009), Mostly Harmless Econometrics. An Empiricist's Companion, Princeton University Press, Princeton.

Ardington, Cally, D. Lam and M. Leibbrandt (2008), "Explaining the Persistence of Racial Gaps in Schooling in South Africa", Paper prepared for submission to the XXVI IUSSP International Conference Marrakech, Morocco, September 2009.

Arnott, A. and Z. Kubeka (1997), "Mathematics and Science Educators in South Africa. Demand, Utilization, Supply and Training", EduSource, Vol. 97/01.

Bonetti, M. and R. D. Gelber (2000), "A graphical method to assess treatment–covariate interactions using the Cox model on subsets of the data", *Statistics in Medicine, Vol.* 19, pp. 2595–2609.

Caliendo, Marco and S. Kopeinig (2005), "Some practical Guidance for the implementation of propensity score matching", Discussion Paper No. 1588, Institute for the Study of Labor, Bonn, Germany.

Case, Anne and A. Deaton (1999), "School Inputs and Educational Outcomes in South Africa", The Quarterly Journal of Economics, Vol. 114, No. 3, pp. 1047-1084.

Cochran, W. and D.B. Rubin (1973), "Controlling Bias in Observational Studies: A Review", Sankhyā: The Indian Journal of Statistics, Series A, Vol. 35, No. 4, pp. 417-446.

Corak, Miles (2001), "Death and Divorce: The Long Term Consequences of Parental Loss on Adolescents", Journal of Labor Economics, 2001, Vol. 19, No. 3, pp. 682-715.

Dehejia, R.H and S. Wahba (1999), "Causal Effects in Non-Experimental Studies: Re-Evaluating the Evaluation of Training Programmes", Journal of the American Statistical Association, Vol. 94, pp. 1053-1062.

Dehejia, R.H. and S. Wahba (2002), "Propensity Score Matching Methods for Non-Experimental Causal Studies", Review of Economics and Statistics, Vol. 84, No. 1, pp. 151-161.

Fiske, Edward, H. F. Ladd (2004), Elusive, Equity. Education Reform in post-Apartheid South Africa, The Brookings Institution Press, Washington D.C.

Hausmann, Ricardo (2007), "Final recommendations of the International Panel on Growth", http://www.treasury.gov.za/publications/other/growth/

Heckman, J.J., H. Ichimura, H. and P.E. Todd (1997), "Matching As An Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme", Review of Economic Studies, Vol. 64, pp. 605-654.

Heckman, J.J., H. Ichimura, H. and P.E. Todd (1998), "Matching as an Econometric Evaluation Estimator", Review of Economic Studies, Vol. 65, pp. 261-294.

Heckman, J.J., H. Ichimura, H. and P.E. Todd (1998), "Characterising Selection Bias Using Experimental Data", Econometrica, Vol. 66, No. 5, pp. 1017-98.

Heckman, J.J., R.J. LaLonde and J.A. Smith (1999), "The Economics and Econometrics of Active Labour Market Programmes", in Ashenfelter, O. and D. Card (eds.) (1999), The Handbook of Labour Economics, Volume III.

Howie, Sarah J. (2003), "Language and other background factors affecting secondary pupils' performance in Mathematics in South Africa", African Journal of Research in Mathematics, Science and Technology Education, Vol. 7, pp. 1-20.

James, Angela, J. Naidoo and H. Benson (2008), "CASME'S Approach to the Sustainability of Science Education in South Africa", paper presented at the XIII.IOSTE Symposium, The Use of Science and Technology Education for Peace and Sustainable Development, September 21-26, 2008, Kuşadası / Turkey, http://www.casme.org.za/

Kahn, Michael (2004), "For Whom the School Bell Tolls: Disparities in Performance in Senior Certificate Mathematics and Physical Science"**,** in: Perspectives in Education, Vol. 22, No. 1, pp. 149-156.

Kallaway, Peter (2002), The History of Education under Apartheid. 1948-1994. The Doors of Learning and Culture Shall be Opened, Pearson Education, Cape Town, South Africa.

OECD (2008), Reviews of National Policies for Education. South Africa, OECD, Paris.

Reddy Vijay (ed.) (2007), Mathematics and Science Achievement at South African Schools in TIMSS 2003, Human Sciences Research Council Press, Cape Town, South Africa.

Reddy Vijay (ed.) (2006), Marking Matric. Colloquium Proceedings, Human Sciences Research Council Press, Cape Town, South Africa.

Rosenbaum, P.R. and D.B. Rubin (1983), "The Central Role of the Propensity Score in Observational Studies for Causal Effects", Biometrika, Vol. 70, No. 1, pp. 41-55.

Rosenbaum, P.R. and D.B. Rubin (1984) "Reducing Bias in Observational Studies using Subclassification on the Propensity Score", Journal of the American Statistical Association, Vol. 79, pp. 516-524.

Royston, Patrick, D. Altman, and W. Sauerbrei (2006), "Dichotomizing continuous predictors in multiple regression: a bad idea", StatMed 2006, Vol. 25, pp. 127-141.

Sauerbrei, W., P. Royston, and K. Zapien (2007), "Detecting an interaction between treatment and a continuous covariate: A comparison of two approaches", Computational Statistics and Data Analysis, Vol. 51, pp. 4054–4063.

Smith, Jeffrey, and P. E. Todd (2005), "Does Matching overcome Lalonde's Critique of Nonexperimental Estimators?", Journal of Econometrics, Vol. 125, pp. 305–353.

Taylor, Nick, J. Muller and P. Vinjevold (eds.) (2003), Getting Schools Working. Research and Systemic School Reform in South Africa, Pearson Education South Africa, Cape Town, South Africa.

UNDP (2003), Human Development Report 2003. Millenium Development Goals, UNDP, Geneva.

Table 1. Scale scores and key indicators of African country participants in TIMSS 2003

County	Average mathematics score (SE)	Average science score (SE)	Population	Life expectancy	Net enrolment (primary)	Net enrolment (secondary)	GNI per capita in US$
Tunisia	410 (2.2)	404 (2.1)	9.8	73	97	68	1990
Egypt	406 (3.5)	421 (3.9)	66.4	69	90	78	1470
Morocco	387 (2.5)	396 (2.5)	29.6	68	88	31	1170
Botswana	366 (2.6)	365 (2.8)	1.7	38	81	55	3010
Ghana	276 (4.7)	255 (5.9)	20.3	55	60	30	270
South Africa	264 (5.5)	244 6.7)	45.3	46	90	62	2500

Source: UNDP (2003)
See Reddy (2006) for a differential analysis of South Africa's TIMSS performance by province and school types.

Table 2. Distribution of Dinaledi program treatment components across schools

Input category	Input	No. of schools that received input	No. of schools with no record of reception[1]	Percentage of schools that received input	Minimum input (>0)	Maximum input	Input unit
1. / 2.	Teacher training and additional teachers	No data available					
3.	Received Grade 11 mathematics textbooks	95	325	22.6%	5	539	textbooks
	Received Grade 11 physical sciences textbooks	65	355	15.5%	3	435	textbooks
	Received Grade 12 mathematic textbooks	227	193	54.0%	1	279	textbooks
	Received Grade 12 physical sciences textbooks	76	344	18.1%	5	181	textbooks
4.	Received calculators in 2007	359	61	85.5%	50	390	calculators
5.	Received higher education guides in 2006-2007	280	140	66.7%	60	180	guides
	Received exam papers math	171	249	40.7%	1	360	papers
	Received exam papers physical science	171	249	40.7%	1	360	papers
6.	Received at least one monitoring visit in 2006-2007	169	251	40.2%	1	2	visits
7.	Were adopted by private companies	99	321	23.6%	-	-	-
	Received financial support from private companies	36	384	8.6%	100000	150000	Rand (ZAR)
Other	Received Department of Education rewards	37	383	8.8%	10000	40000	Rand (ZAR)
	Received Tutor Finance	6	414	1.4%	8000	12000	Rand (ZAR)

Source: Department of Education (2007)
[1]This Table reports data for 420 Dinaledi schools from the sample underlying this study for which detailed input data are available. School with missing values or reportedly no inputs are grouped in this column.

Table 3. Phases of Dinaledi expansion

Province	2001	2005	2007	Number of Dinaledi Schools in the sample
Eastern Cape	15	60	60	55
Free State	6	30	35	34
Gauteng	11	70	101	93
KZN	23	70	84	76
Limpopo	23	50	51	46
Mpumalanga	7	30	44	39
Northern Cape	4	10	17	11
North West	7	40	51	42
Western Cape	6	40	45	44
Total	102	400	488	440

Source: NDoE, Government of South Africa, 2007.

Table 4. Descriptive statistics. Categorical school characteristics

Category	**Dinaledi 2004**	**Non-Dinaledi 2004**	**Percentage of Dinaledi 2004 in category**
Dinaledi 2001 schools	57	15	79.2%
Combined schools	13	237	5.2%
Intermediate schools	1	10	9.1%
Secondary schools	426	2,448	14.8%
Public schools	440	2,695	14.0%
Schools by specialization			
Comprehensive	59	346	14.6%
Computer, mathematics science and technology	2	0	100.0%
Dance, math, science and technology	1	0	100.0%
Math, science and technology	40	8	83.3%
Ordinary	318	2,283	12.2%
Ordinary & technical	14	31	31.1%
Technical	6	27	18.2%
N	440	2,695	14.0%

Table 5. Descriptive statistics. Dinaledi schools by former department under Apartheid

Department under Apartheid	Dinaledi 2004	Non-Dinaledi 2004	Percentage of Dinaledi 2004 in Category
BOPHUTATSWANA (BOP)	18	110	16.4%
CAPE EDUCATION DEPARTMENT (CED)	12	153	7.8%
CISKEI	4	160	2.5%
DEPARTMENT OF EDUCATION AND TRAINING (DoET)	148	693	21.4%
FREE STATE EDUCATION DEPARTMENT	1	3	33.3%
GAZANKULU	7	87	8.0%
HOUSE OF ASSEMBLY (HOA)	56	184	30.4%
HOUSE OF DELEGATES (HOD)	6	43	14.0%
HOUSE OF REPRESENTATIVES (HOR)	34	168	20.2%
KANGWANE	12	65	18.5%
KWANDEBELE	2	54	3.7%
KWAZULU NATAL EDUCATION DEPARTMENT	49	134	36.6%
LEBOWA	19	196	9.7%
NEW EDUCATION DEPARTMENT	21	242	8.7%
TRANSKEI	26	269	9.7%
TRANSVAAL EDUCATION DEPARTMENT (TED)	2	5	40.0%
VENDA	15	65	23.1%
WESTERN CAPE EDUCATION DEPARTMENT	2	18	11.1%
INDEPENDENT	1	0	-
TO BE UPDATED	5	46	10.9%
N	440	2,695	14.0%

Table 6. Descriptive statistics. Pre-treatment learner, enrolment and language characteristics

Baseline characteristics (2004)	Dinaledi 2004		Non-Dinaledi 2004			Difference-in-means	SE
	N	Mean (SD)	N	Mean	SD	Dummy for 488 Dinaledi schools (2005)	
Total number of learners	440[4]	5475.63 (3037.40)	2695[4]	3747.46	(2950.07)	1,728.169***	(155.446)
Number of African learners in the school	420	4928.82 (3201.53)	2580	3127.37	(2893.51)	1,801.453***	(166.158)
Number of male African learners	420	1341.71 (890.97)	2580	859.48	(736.81	482.227***	(45.796)
Number of female African learners	420	1518.40 (918.08)	2580	937.51	(1299.50)	580.885***	(51.557)
Learners in grades 10, 11 and 12	440	921.93 (451.59)	2695	568.66	(367.07)	353.274***	(22.644)
Number of African learners in grades 10 though 12	440	807.39 (519.75)	2695	480.51	(387.47)	326.876***	(25.859)
Total enrolment	440	3142.05 (1580.78)	2695	2088.75	(1486.39)	1,053.308***	(80.56)
Total male enrolments	427	1503.70 (810.26)	2610	1030.80	(735.65)	472.900***	(41.741)
Total female enrolments	427	1734.01 (855.43)	2610	1125.97	(785.53)	608.043***	(44.129)
Enrolments in grade 10	427	443.46 (208.26)	2610	279.07	(180.87)	164.390***	(10.674)
Enrolments in grade 11	427	346.37 (151.77)	2610	213.42	(132.20)	132.950***	(7.782)
Enrolments in grade 12	427	221.78 (101.06)	2610	128.09	(91.18)	93.690***	(5.202)
Male Enrolments in grade 10	427	141.00 (71.83)	2610	89.02	(59.53)	51.984***	(3.664)
Male Enrolments in grade 11	427	105.49 (52.01)	2610	65.15	(42.66)	40.338***	(2.65)
Male Enrolments in grade 12	427	66.87 (34.58)	2610	39.06	(28.68)	27.816***	(1.764)
Female Enrolments in grade 10	427	157.69 (75.46)	2610	97.41	(64.52)	60.277***	(3.861)
Female Enrolments in grade 11	427	128.82 (59.20)	2610	77.52	(49.16)	51.306***	(3.02)
Female Enrolments in grade 12	427	82.89 (42.74)	2610	47.05	(37.75)	35.845***	(2.195)
Number of learners with English as medium of instruction	428	537.60 (271.28)	2615	299.63	(223.95)	237.972***	(13.814)

Table 7. Descriptive statistics. Ratio measures

Ratio measure	Dinaledi 2004					Non-Dinaledi 2004					Difference-in-means (SE)
	N1	Mean	SD	Min	Max	N0	Mean	SD	Min	Max	
Proportion of African learners in total learners	416	0.86	0.30	0.00	1.00	2561	0.84	0.33	0.00	1.00	0.016 (0.016)
Proportion of African learners in learners in grades 10 to 12	416	0.86	0.30	0.00	1.00	2542	0.84	0.34	0.00	1.00	0.014 (0.016)
*Male to female ratio among African learners	406	0.90	0.49	0.00	6.50	2489	0.97	0.70	0.00	18.00	-0.075*** (0.028)
*Male to female enrolments ratio grades 10 through 12	420	0.88	0.42	0.00	5.60	2590	2.18	43.48	0.00	1576.00	-1.299 (0.855)
*Male to female enrolment ratio in grades 10 to 12	420	0.86	0.40	0.00	5.27	2565	1.01	4.76	0.00	237.00	-0.145 (0.096)
*Share of learners with English as medium of instruction out of total number of l	416	0.15	0.92	0.00	18.92	2550	0.09	0.13	0.00	5.66	0.053 (0.045)

*Values >1 are data to data errors.

Table 8. Descriptive statistics. Outcomes measures

Outcomes measure	Dinaledi 2004			Non-Dinaledi 2004			Difference-in-means (SE)
	N1	Mean	SD	N0	Mean	SD	
Total number of students entering SG physical science	440	29.69	(30.50)	2695	16.66	(20.73)	13.033*** (1.507)
Total number of students entering HG physical science	440	13.95	(20.31)	2695	5.57	(13.44)	8.380*** (1.002)
Total number of students writing SG physical science	440	29.61	(30.39)	2695	16.57	(20.60)	13.041*** (1.501)
Total number of students writing HG physical science	440	13.93	(20.27)	2695	5.56	(13.41)	8.367*** (1.000)
Total number of students passing SG physical science	440	24.88	(24.91)	2695	11.90	(15.41)	12.973*** (1.223)
Total number of students passing HG physical science	440	9.66	(16.20)	2695	3.73	(11.57)	5.927*** (0.803)
Passing rate out of those entering SG physical science[1]	440	0.64	(0.41)	2695	0.57	(0.39)	0.072*** (0.021)
Passing rate out of those entering HG physical science[1]	440	0.50	(0.40)	2695	0.28	(0.39)	0.223*** (0.02)
Passing rate out of those writing SG physical science[1]	440	0.64	(0.41)	2695	0.57	(0.39)	0.071*** (0.021)
Passing rate out of those writing HG physical science[1]	440	0.50	(0.40)	2695	0.28	(0.39)	0.224*** (0.02)
Proportion entering SG Physical Sciences of students enrolled in grade 12	426	0.14	(0.13)	2573	0.13	(0.13)	0.069 (0.07)
Proportion entering HG Physical of students enrolled in grade 12	426	0.06	(0.09)	2573	0.04	(0.09)	0.027*** (0.005)

[1]The passing rate is set as zero for schools with no students entering the respective subject (entry-corrected passing rate).

Table 9. Descriptive Statistics. Observations by province

Province	Number of Dinaledi schools in the population (2007 – 2008)[3]	No. of Dinaledi 2004 Schools in sample[1]	No. of Non-Dinaledi 2004 in sample	Total no. of schools per province in sample	Share of Dinaledi 2004 schools per province	Population in m (mid-2007 estimate)[2]	Dinaledi schools per m population[4]
Eastern Cape (EC)	60	55	735	790	12.50%	6.9	8.7
Free State (FS)	35	34	249	283	7.70%	2.9	12.1
Gauteng (GT)	101	93	305	398	21.10%	9.6	10.5
Kwazulu Natal (KZ)	84	76	189	265	17.30%	10	8.4
Limpopo (LP)	51	46	417	463	10.50%	5.4	9.4
Mpumalanga (MP)	44	39	261	300	8.90%	3.5	12.6
Northern Cape (NC)	17	11	48	59	2.50%	1.1	15.5
North West Province (NW)	51	42	235	277	9.60%	3.4	15.0
Western Cape (WC)	45	44	256	300	10.00%	4.8	9.4
Total	**488**	**440**	**2695**	**3135**	**100%**	**47.6**	**10.3**

[1]These figures are based on the restricted sample of 3135 schools underlying this study.
[2]Source: http://www.southafrica.info/about/people/popprov.htm
[3]Source: MoE Report, 2007
[4]Based on the population number of Dinaledi schools as in column one.

Table 10. Descriptive statistics by province. Baseline and physical sciences outcomes measures (ratio and absolute measures)

	Eastern Cape			Free State			Gauteng			Kwazulu Natal			Limpopo		
Variable	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
Baseline characteristics (2004)															
Learners	790	2656.94	2421.88	283	4426.46	2571.38	398	6130.87	2884.11	265	4337.56	2539.7	463	3351.29	1901.94
African Learners	759	2350.83	2189.46	280	3841.01	2794.83	384	4776.81	3341.07	255	4367.31	2462.07	450	3417.11	1837.71
Male African Learners	759	637.83	581.86	280	1067.64	759.35	384	1346.81	947.83	255	1204.35	696.4	450	940.34	506.04
Female African Learners	759	739.09	626.25	280	1113.87	794.62	384	1394.24	961.2	255	1276.37	671.16	450	1008.32	566.08
Learners 10 to 12	790	503.11	349.94	283	642.07	374.21	398	904.87	427.9	265	644.49	362.58	463	528.19	302.31
African Learners 10 to 12	790	457.75	355.62	283	581.61	412.67	398	688.48	536.28	265	623.6	368.88	463	523.58	302.68
Enrolment	790	1486.99	1304.53	283	2434.53	1355.93	398	3428.75	1573.34	265	2570.25	1398.37	463	1848.62	1055
Male Enrolment	772	706.15	661.2	281	1189.96	686.34	386	1706.89	788.08	263	1234.34	733.06	453	898.96	495.57
Female Enrolment	772	815.51	727.98	281	1261.89	721.54	386	1828.45	832.99	263	1355.45	782.98	453	990.47	558.55
Enrolment 10	772	254.31	182.04	281	331.66	200.61	386	436.87	197.04	263	302.79	158.85	453	211.56	121.42
Enrolment 11	772	179.66	123.68	281	213.72	134.27	386	325.61	146.25	263	259.22	129.01	453	234.26	140.33
Enrolment 12	772	112.59	82.56	281	123.86	78.94	386	214.03	105.52	263	152.81	81.48	453	109.16	78.15
Male Enrolment 10	772	76.2	55.41	281	109.74	68.59	386	145.12	68.06	263	99.77	59.07	453	67.3	39.63
Male Enrolment 11	772	52.09	37.22	281	66.96	45.05	386	102.53	50.2	263	81.07	46.59	453	72.25	43.66
Male Enrolment 12	772	33.18	25.49	281	39.25	26.75	386	66	37.59	263	47.68	29.21	453	33.48	24.57
Female Enrolment 10	772	93.31	68.93	281	112.46	69.57	386	148.58	71.27	263	102.68	53.68	453	74.35	43.3
Female Enrolment 11	772	68.16	49.21	281	76.01	48.38	386	116.83	56.65	263	92.98	47.27	453	84.72	52.64
Female Enrolment 12	772	42.84	34.31	281	43.58	30.94	386	78.18	41.14	263	56.22	36.24	453	39.6	29.24
English Instruction	782	224.34	173.8	275	352.08	223.88	387	473.6	316.92	263	447.4	213.37	452	330.24	186.52
Outcomes measure															
SG EC PR of entering[1]	790	0.66	0.31	283	0.76	0.32	398	0.75	0.25	265	0.18	0.38	463	0.15	0.35
HG EC PR of entering[1]	790	0.18	0.35	283	0.48	0.41	398	0.59	0.35	265	0.08	0.26	463	0.13	0.32
SG EC PR of writing[1]	790	0.67	0.32	283	0.76	0.32	398	0.76	0.25	265	0.18	0.38	463	0.15	0.35
HG EC PR of writing[1]	790	0.18	0.35	283	0.48	0.41	398	0.59	0.35	265	0.08	0.26	463	0.13	0.32
SG passing rate of entering	711	0.74	0.24	255	0.84	0.2	391	0.77	0.23	54	0.87	0.34	84	0.81	0.38
HG passing rate of entering	259	0.55	0.4	211	0.64	0.34	367	0.64	0.32	43	0.47	0.48	148	0.4	0.46
SG passing rate of writing	711	0.74	0.24	255	0.84	0.2	391	0.77	0.23	52	0.9	0.3	84	0.81	0.38
HG passing rate of writing	258	0.55	0.4	211	0.64	0.34	367	0.64	0.32	42	0.49	0.49	148	0.4	0.47
SG enrolment rates	748	0.2	0.14	276	0.14	0.12	383	0.17	0.12	263	0	0.01	449	0	0.01
HG enrolment rates	750	0.02	0.07	276	0.06	0.08	383	0.09	0.1	263	0	0.01	449	0	0.01
Students entering SG	790	24.17	23.81	283	16.61	16.27	398	34.41	25.14	265	0.32	0.79	463	0.25	0.59
Students entering HG	790	2.65	7.26	283	8.98	14.84	398	20.8	26.43	265	0.21	0.53	463	0.45	0.8
Students writing SG	790	24.04	23.7	283	16.57	16.24	398	34.27	25.03	265	0.31	0.78	463	0.25	0.59
Students writing HG	790	2.64	7.24	283	8.96	14.83	398	20.76	26.38	265	0.2	0.52	463	0.45	0.79
Students passing SG	790	17.68	18.38	283	13.45	13.15	398	25.71	20.12	265	0.29	0.78	463	0.2	0.54
Students passing HG	790	1.86	6.46	283	5.9	12.56	398	15.15	23.68	265	0.09	0.33	463	0.19	0.52

[1]The passing rate is set as zero for schools with no students entering the respective subject (entry-corrected).

Table 11. Descriptive statistics by province. Baseline and physical sciences outcomes measures (ratio and absolute measures)

	Mpumalanga			Northern Cape			North West			Western Cape		
Variable	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
						Baseline characteristics (2004)						
Learners	300	4882.33	3454.75	59	2232.27	3042.36	277	3239.57	4126.01	300	5073.52	2725.01
African Learners	291	4932.58	3390.45	29	1661.79	2172.63	255	3299.41	4126.91	297	1576.99	2910.35
Male African Learners	291	1351.33	904.75	29	459.76	599.57	255	864.96	611.51	297	415.33	781.35
Female African Learners	291	1418.2	924.91	29	465.62	624.75	255	1125.65	3419.97	297	524.6	920.84
Learners 10 to 12	300	686.8	393.34	59	245.86	396.29	277	612.35	432.42	300	644.59	408.7
African Learners 10 to 12	300	680.18	398.09	59	135.03	344.7	277	593.94	440.49	300	219.28	417.66
Enrolment	300	2663.75	1830.18	59	1302.03	1680.24	277	1714.32	1283.81	300	2985	1298.89
Male Enrolment	292	1310.33	888.45	30	1248.67	770.81	263	865.99	612.17	297	1395.25	656.6
Female Enrolment	292	1426.4	930.46	30	1312	769.22	263	939.59	678.79	297	1619.91	747.31
Enrolment 10	292	319.46	187.97	30	242.3	204.22	263	320.28	217.08	297	334.45	197.44
Enrolment 11	292	256.33	146.74	30	173.13	144.66	263	223.7	149.69	297	226.6	126.27
Enrolment 12	292	152.22	85.59	30	100.97	84.91	263	144.68	95.63	297	166.65	132.46
Male Enrolment 10	292	101.98	60.22	30	80.4	67.73	263	104.66	70.53	297	102.4	62.05
Male Enrolment 11	292	78.83	47.44	30	57.53	55.37	263	69.46	46.51	297	65.35	37.84
Male Enrolment 12	292	46.47	26.95	30	33.4	29.86	263	44.63	29.44	297	48.32	36.22
Female Enrolment 10	292	111.24	67.01	30	80.93	69.5	263	108.68	76.72	297	122.61	74.69
Female Enrolment 11	292	91.62	53.24	30	61.83	48.52	263	79.2	54.82	297	87.44	51.96
Female Enrolment 12	292	55.12	31.23	30	34.63	29.62	263	52.14	36.28	297	64.39	65.63
English Instruction	292	440.68	227.54	29	133.98	206.21	266	300.1	224.7	297	265.2	281.1
						Outcomes measure						
SG EC PR of entering[1]	300	0.64	0.27	59	0.78	0.33	277	0.66	0.29	300	0.81	0.3
HG EC PR of entering[1]	300	0.33	0.36	59	0.51	0.47	277	0.31	0.38	300	0.57	0.43
SG EC PR of writing[1]	300	0.64	0.27	59	0.78	0.33	277	0.66	0.29	300	0.81	0.3
HG EC PR of writing[1]	300	0.33	0.36	59	0.51	0.47	277	0.31	0.38	300	0.57	0.43
SG passing rate of entering	287	0.67	0.24	52	0.88	0.17	257	0.71	0.23	272	0.89	0.15
HG passing rate of entering	241	0.41	0.36	37	0.82	0.32	167	0.51	0.37	220	0.78	0.29
SG passing rate of writing	285	0.67	0.24	52	0.88	0.17	256	0.72	0.23	272	0.89	0.15
HG passing rate of writing	239	0.41	0.36	37	0.82	0.32	167	0.51	0.37	220	0.78	0.29
SG enrolment rates	291	0.16	0.1	30	0.11	0.09	257	0.18	0.13	297	0.11	0.08
HG enrolment rates	292	0.05	0.07	30	0.05	0.06	258	0.04	0.07	297	0.07	0.09
Students entering SG	300	22.22	16.21	59	12.85	13.62	277	25.07	24.15	300	19.68	27.08
Students entering HG	300	8.05	11.1	59	5.53	9.19	277	6.39	11.69	300	11.56	18.58
Students writing SG	300	22.05	16.13	59	12.81	13.53	277	24.99	24.08	300	19.58	26.73
Students writing HG	300	8.03	11.08	59	5.53	9.19	277	6.38	11.68	300	11.52	18.49
Students passing SG	300	14.2	10.44	59	10.86	11.3	277	17.6	18.99	300	16.9	22.42
Students passing HG	300	2.83	4.69	59	4.58	7.87	277	3.31	6.71	300	9.96	17.5

[1]The passing rate is set as zero for schools with no students entering the respective subject (entry-corrected).

Table 12. Differences in baseline characteristics between former House of Assembly and former Bantustan schools. Pre-treatment learner, enrolment and language characteristics

Baseline characteristics (2004)	former House of Assembly schools			former Bantustan schools			Difference in means	SE
	N	Mean	SD	N	Mean	SD		
Total number of learners	240	4634.54	2498.06	1292	3038.58	2161.07	1596***	(171.9)
Number of African learners in the school	226	1721.79	2126.28	1249	3127.93	2104.66	-1406***	(153.3)
Number of male African learners	226	467.14	574.79	1249	862.99	576.09	-395.8***	(41.51)
Number of female African learners	226	554.81	676.59	1249	935.27	603.98	-380.5***	(48.08)
Learners in grades 10, 11 and 12	240	603.21	367.1	1292	578.01	359.07	25.20	(25.69)
Number of African learners in grades 10 though 12	240	197.53	247.09	1292	577.21	359.3	-379.7***	(18.81)
Total enrolment	240	2740.07	1371.64	1292	1686.31	1164.37	1054***	(94.16)
Total male enrolments	226	1389	761.53	1268	812.41	564.26	576.6***	(53.00)
Total female enrolments	226	1520.8	804.72	1268	905.82	616.88	615.0***	(56.19)
Enrolments in grade 10	226	261.52	135.84	1268	263.35	172.55	-1.828	(10.24)
Enrolments in grade 11	226	226.03	121.58	1268	223.45	139.57	2.579	(8.976)
Enrolments in grade 12	226	193.82	115.33	1268	126.82	84.25	67.00***	(8.017)
Male Enrolments in grade 10	226	87.16	54.2	1268	82.88	55.45	4.287	(3.923)
Male Enrolments in grade 11	226	72.34	48.65	1268	67.81	43.83	4.532	(3.458)
Male Enrolments in grade 12	226	61.42	45.57	1268	38.38	26.09	23.04***	(3.114)
Female Enrolments in grade 10	226	92.62	56.51	1268	92.43	62.28	0.189	(4.141)
Female Enrolments in grade 11	226	83.23	53.22	1268	81.5	52.48	1.726	(3.829)
Female Enrolments in grade 12	226	71.89	48.84	1268	46.73	33.44	25.17***	(3.377)
Number of Learners with English as language of instruction	224	246.94	263.8	1275	312.29	199.95	-65.34***	(18.47)

Table 13. Differences in baseline characteristics between former House of Assembly and former Bantustan schools. Pre-treatment measures of learning outcomes

Pre-Treatment outcomes measures (2004)	former House of Assembly schools			former Bantustan schools			Difference in means	SE
	N	Mean	SD	N	Mean	SD		
Passing rate out of those entering SG physical science	217	0.93	0.13	762	0.71	0.26	0.219***	(0.0130)
Passing rate out of those entering HG physical science	205	0.86	0.19	485	0.43	0.4	0.429***	(0.0228)
Passing rate out of those writing SG physical science	217	0.93	0.13	759	0.72	0.26	0.216***	(0.0128)
Passing rate out of those writing HG physical science	205	0.86	0.19	482	0.43	0.4	0.429***	(0.0228)
Proportion entering SG Physical Sciences of students enrolled in grade 12	226	0.14	0.15	1247	0.11	0.14	0.0313***	(0.0105)
Proportion entering HG Physical of students enrolled in grade 12	226	0.14	0.12	1250	0.02	0.04	0.122***	(0.00814)
Total number of students entering SG physical science	240	23.05	25.18	1292	14.61	22.55	8.436***	(1.740)
Total number of students entering HG physical science	240	28.1	31.54	1292	2.21	6.45	25.89***	(2.041)
Total number of students writing SG physical science	240	23	25.11	1292	14.53	22.45	8.468***	(1.735)
Total number of students writing HG physical science	240	28.06	31.48	1292	2.21	6.44	25.86***	(2.037)
Total number of students passing SG physical science	240	21.23	22.67	1292	10.18	16.68	11.05***	(1.533)
Total number of students passing HG physical science	240	24.58	29.1	1292	0.85	2.49	23.73***	(1.877)

Figure 2. Evolution of the share of mathematics candidates writing the Senior Certificate

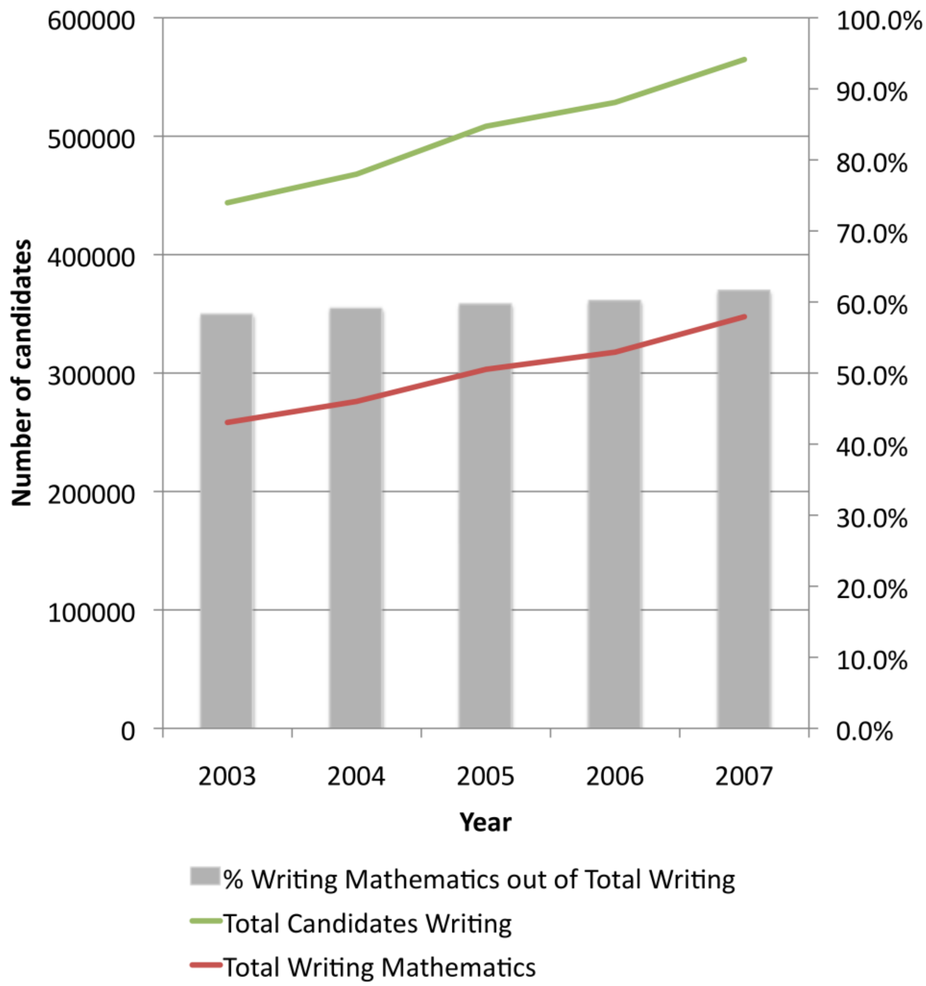


Figure 3. Evolution of the share of mathematics candidates passing the Senior Certificate

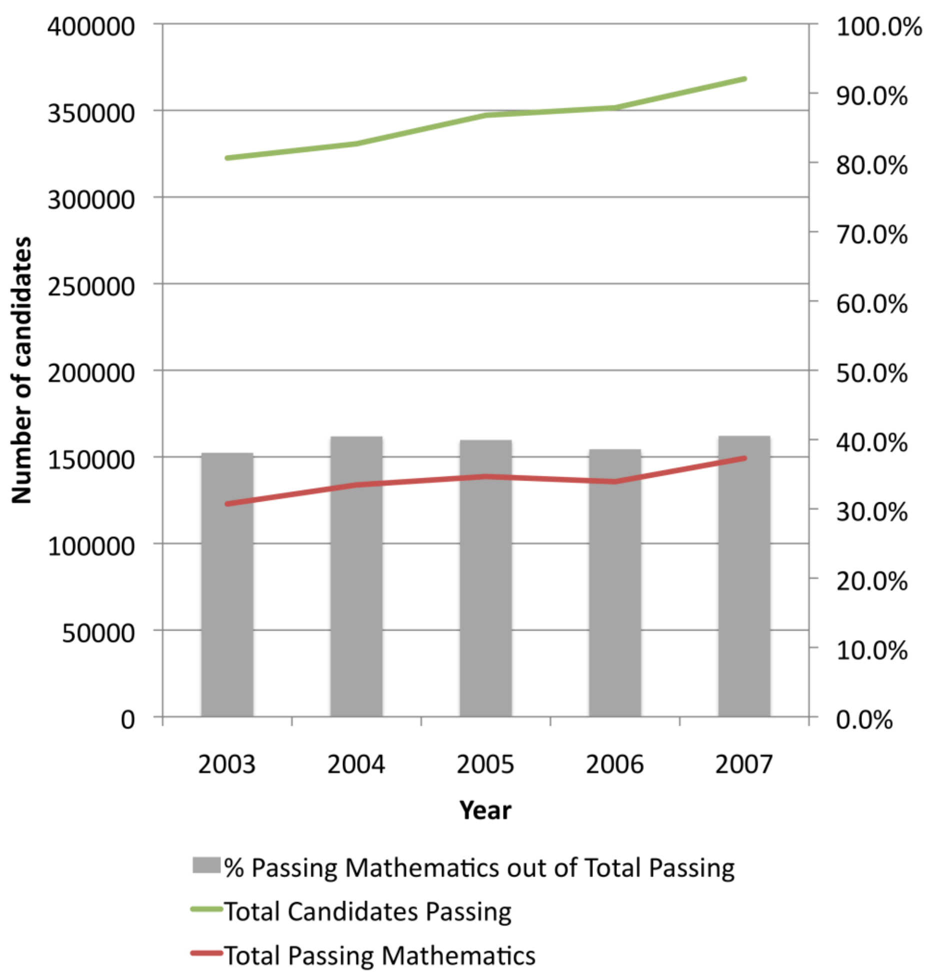
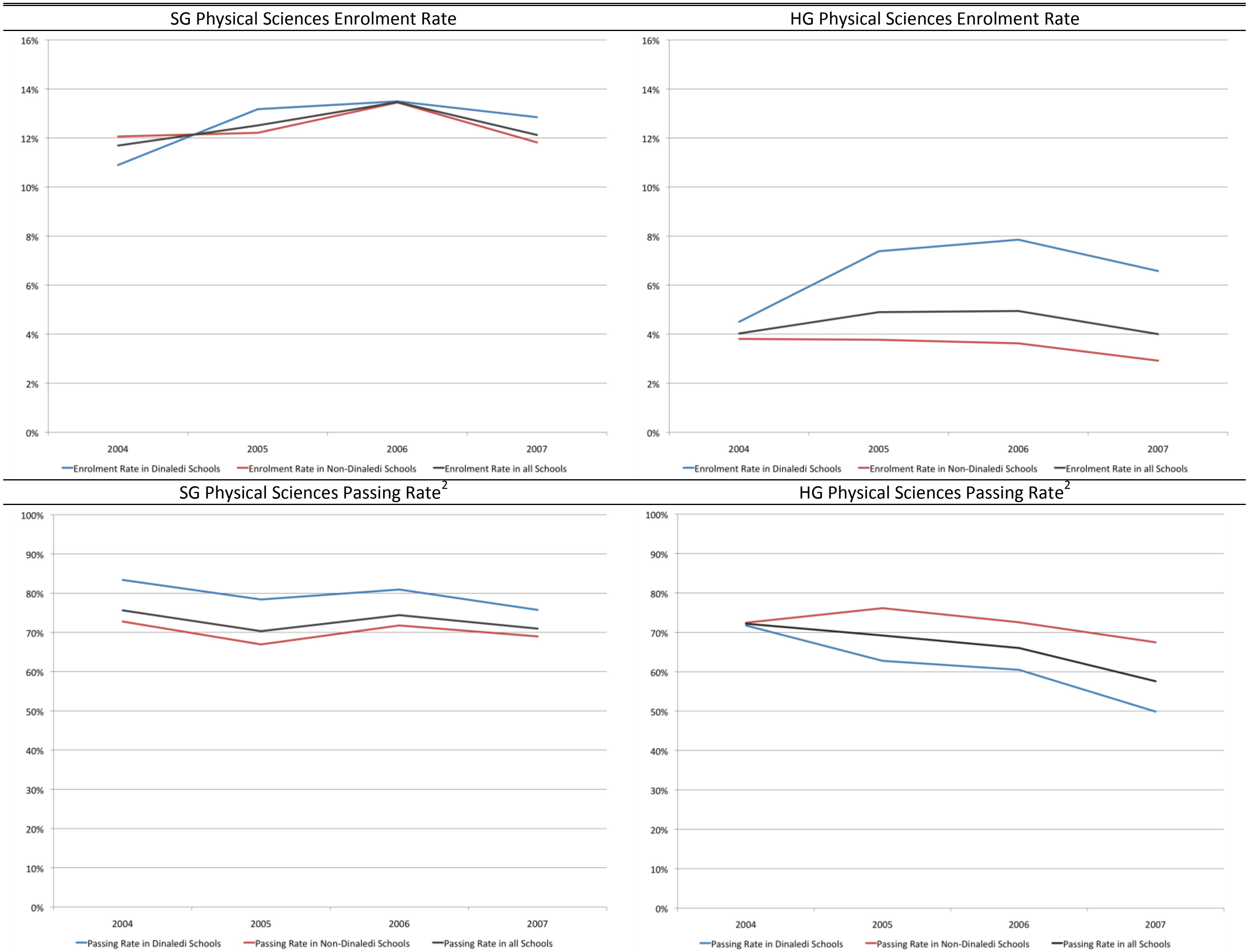


Figure 4. Breakdown of enrolment rates and passing rates for Dinaledi and non-Dinaledi schools 2004-2007[1]



[1]The enrolment and passing rates in this chart have been calculated for a sub-set of 1,571 schools for which data on grade 12 enrolment and on the number of students entering and passing SG and HG Physical Sciences is available for each year from 2004 to 2007. The sample comprises 1,230 Non-Dinaledi schools and 341 Dinaledi schools. It is not representative of the entire school population, such that population enrolment and passing rates may differ from the values reported here.

[2]In the aggregate, standard and entry-corrected passing rates are equal.

Table 14. Physical sciences results. Ratio measures. DID estimates of Dinaledi program effects for years 2005-2007

	(1) Enrolment rate of grade 12 students in SG physical science	(2) Enrolment rate of grade 12 students in HG physical science	(3) EC PR SG physical science	(4) EC PR HG physical science	(5) WC PR SG physical science	(6) WC PR HG physical science	(7) Passing rate out of those entering SG physical science	(8) Passing rate out of those entering HG physical science	(9) Passing rate out of those writing SG physical science	(10) Passing rate out of those writing HG physical science
2004-2005										
Dummy for 488 Dinaledi schools (2005)	-0.0193**	-0.0153**	-0.109***	-0.0365	-0.111***	-0.0367	0.0339**	0.0424	0.0324**	0.0424
	(0.00977)	(0.00720)	(0.0264)	(0.0303)	(0.0264)	(0.0304)	(0.0142)	(0.0274)	(0.0142)	(0.0274)
Year 2005	0.0116	0.0122	-0.0666***	-0.0185	-0.0581***	-0.0159	-0.0979***	-0.0723**	-0.0890***	-0.0677**
	(0.00948)	(0.00763)	(0.0218)	(0.0302)	(0.0217)	(0.0303)	(0.0167)	(0.0294)	(0.0165)	(0.0294)
Interaction Dinaledi Dummy * Year2005	0.0247*	0.0499***	0.180***	0.143***	0.180***	0.142***	0.0559**	0.0544	0.0557**	0.0509
	(0.0133)	(0.0117)	(0.0338)	(0.0407)	(0.0338)	(0.0408)	(0.0218)	(0.0378)	(0.0216)	(0.0378)
Constant	0.156***	0.0789***	0.764***	0.551***	0.767***	0.553***	0.835***	0.661***	0.838***	0.664***
	(0.00683)	(0.00523)	(0.0155)	(0.0217)	(0.0156)	(0.0218)	(0.0102)	(0.0207)	(0.0102)	(0.0208)
Observations	1396	1397	1412	1412	1412	1412	1251	1200	1251	1199
R-squared	0.012	0.042	0.022	0.016	0.023	0.016	0.058	0.018	0.052	0.016
2004-2006										
Dummy for 488 Dinaledi schools (2005)	-0.00636	0.000491	-0.0343	0.0367	-0.0360	0.0377	0.0347**	0.0825***	0.0297*	0.0792***
	(0.0101)	(0.00685)	(0.0294)	(0.0309)	(0.0295)	(0.0309)	(0.0157)	(0.0291)	(0.0154)	(0.0290)
Year 2006	0.0712***	0.00653	0.0446*	0.0264	0.0539**	0.0274	-0.101***	-0.0277	-0.0861***	-0.0253
	(0.00984)	(0.00639)	(0.0241)	(0.0300)	(0.0241)	(0.0300)	(0.0178)	(0.0298)	(0.0170)	(0.0297)
Interaction Dinaledi Dummy * Year2006	-0.0342**	0.0714***	0.0721**	0.107***	0.0695*	0.108***	0.0692***	-0.00721	0.0618***	-0.00667
	(0.0136)	(0.0110)	(0.0360)	(0.0404)	(0.0360)	(0.0405)	(0.0228)	(0.0385)	(0.0221)	(0.0383)
Constant	0.140***	0.0617***	0.681***	0.470***	0.684***	0.471***	0.833***	0.616***	0.839***	0.622***
	(0.00723)	(0.00471)	(0.0199)	(0.0222)	(0.0199)	(0.0222)	(0.0119)	(0.0224)	(0.0116)	(0.0223)
Observations	1370	1369	1384	1384	1384	1384	1200	1141	1195	1135
R-squared	0.056	0.095	0.017	0.030	0.020	0.031	0.063	0.017	0.051	0.015
2004-2007										
Dummy for 488 Dinaledi schools (2005)	-0.00790	-0.000699	-0.0387	0.00915	-0.0402	0.0102	0.0221	0.0622**	0.0175	0.0590**
	(0.0105)	(0.00687)	(0.0291)	(0.0307)	(0.0292)	(0.0308)	(0.0149)	(0.0285)	(0.0146)	(0.0284)
Year 2007	0.0520***	0.0133**	0.0314	0.00257	0.0457**	0.00581	-0.112***	-0.0632**	-0.0991***	-0.0642**
	(0.00993)	(0.00660)	(0.0233)	(0.0298)	(0.0233)	(0.0298)	(0.0160)	(0.0294)	(0.0154)	(0.0293)
Interaction Dinaledi Dummy * Year2007	-0.0235*	0.0625***	0.0572	0.0458	0.0510	0.0443	0.0308	-0.0611	0.0283	-0.0555
	(0.0136)	(0.0107)	(0.0349)	(0.0402)	(0.0350)	(0.0403)	(0.0221)	(0.0382)	(0.0216)	(0.0381)
Constant	0.143***	0.0644***	0.690***	0.502***	0.693***	0.503***	0.846***	0.638***	0.852***	0.644***
	(0.00779)	(0.00476)	(0.0197)	(0.0222)	(0.0198)	(0.0222)	(0.0111)	(0.0219)	(0.0107)	(0.0219)
Observations	1378	1377	1408	1408	1408	1408	1228	1185	1225	1181
R-squared	0.032	0.090	0.010	0.004	0.014	0.004	0.067	0.023	0.055	0.023

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 15. Physical sciences results. Absolute measures. DID estimates of Dinaledi program effects for years 2005-2007

	(11) Total number of students entering SG physical science	(12) Total number of students entering HG physical science	(13) Total number of students writing SG physical science	(14) Total number of students writing HG physical science	(15) Total number of students passing SG physical science	(16) Total number of students passing HG physical science
2005						
Dummy for 488 Dinaledi schools (2005)	-8.388***	-5.150***	-8.249***	-5.108***	-6.538***	-2.184
	(2.800)	(1.854)	(2.771)	(1.846)	(2.275)	(1.437)
Year 2005	6.329*	1.742	5.244	1.564	-0.924	1.028
	(3.574)	(2.048)	(3.431)	(2.035)	(2.658)	(1.573)
Interaction Dinaledi Dummy * Year2005	3.629	11.33***	4.193	11.32***	7.329**	5.926***
	(4.209)	(2.841)	(4.075)	(2.824)	(3.162)	(2.147)
Constant	38.02***	19.18***	37.80***	19.12***	31.35***	12.28***
	(2.363)	(1.496)	(2.332)	(1.488)	(1.936)	(1.101)
Observations	1412	1412	1412	1412	1412	1412
R-squared	0.018	0.030	0.016	0.029	0.008	0.015
2006						
Dummy for 488 Dinaledi schools (2005)	-3.425	-0.150	-3.275	-0.116	-2.266	1.483
	(2.852)	(1.693)	(2.821)	(1.686)	(2.304)	(1.254)
Year 2006	20.35***	2.483	18.70***	2.408	7.017***	0.757
	(4.132)	(1.680)	(3.871)	(1.672)	(2.717)	(1.123)
Interaction Dinaledi Dummy * Year2006	-10.38**	14.08***	-9.208**	13.97***	0.315	7.630***
	(4.703)	(2.617)	(4.460)	(2.603)	(3.246)	(1.840)
Constant	32.67***	13.97***	32.45***	13.90***	26.70***	8.431***
	(2.411)	(1.279)	(2.378)	(1.272)	(1.962)	(0.834)
Observations	1384	1384	1384	1384	1384	1384
R-squared	0.042	0.074	0.039	0.073	0.015	0.052
2007						
Dummy for 488 Dinaledi schools (2005)	-3.801	-0.756	-3.682	-0.727	-3.165	0.352
	(2.845)	(1.721)	(2.815)	(1.714)	(2.323)	(1.309)
Year 2007	21.54***	3.739**	19.62***	3.628**	7.687***	0.849
	(4.146)	(1.757)	(3.880)	(1.744)	(2.847)	(1.259)
Interaction Dinaledi Dummy * Year2007	-5.864	17***	-4.636	16.43***	1.810	6.378***
	(4.900)	(2.884)	(4.661)	(2.857)	(3.456)	(1.906)
Constant	33.25***	14.82***	33.05***	14.76***	27.78***	9.759***
	(2.413)	(1.325)	(2.382)	(1.318)	(1.991)	(0.925)
Observations	1408	1408	1408	1408	1408	1408
R-squared	0.045	0.088	0.043	0.084	0.019	0.030

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 16. Physical sciences results. Ratio measures. Bias-corrected DID estimates of Dinaledi program effects for years 2005-2007

	(1) Enrolment rate of grade 12 students in SG physical science	(2) Enrolment rate of grade 12 students in HG physical science	(3) EC PR SG physical science	(4) EC PR HG physical science	(5) WC PR SG physical science	(6) WC PR HG physical science	(7) Passing rate out of those entering SG physical science	(8) Passing rate out of those entering HG physical science	(9) Passing rate out of those writing SG physical science	(10) Passing rate out of those writing HG physical science
					2004-2005					
SATT	0.0264***	0.0498***	0.180***	0.143***	0.180***	0.142***	0.0435*	0.0172	0.0441*	0.0159
	(0.0101)	(0.00940)	(0.0309)	(0.0394)	(0.0305)	(0.0396)	(0.0257)	(0.0356)	(0.0249)	(0.0356)
Observations	2241	2243	2287	2287	2287	2287	1742	1126	1740	1124
					2004-2006					
SATT	-0.0345***	0.0695***	0.0721*	0.107***	0.0696*	0.108***	0.0516	-0.0717	0.0215	-0.0646
	(0.0132)	(0.0115)	(0.0400)	(0.0416)	(0.0402)	(0.0416)	(0.0368)	(0.0438)	(0.0297)	(0.0437)
Observations	1782	1782	1830	1830	1830	1830	1476	855	1471	853
					2004-2007					
SATT	-0.0233*	0.0593***	0.0572	0.0457	0.0510	0.0442	0.0120	-0.115**	0.0155	-0.110**
	(0.0134)	(0.0106)	(0.0360)	(0.0425)	(0.0358)	(0.0427)	(0.0308)	(0.0462)	(0.0292)	(0.0459)
Observations	1817	1818	1886	1886	1886	1886	1543	921	1529	906

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 17. Physical sciences results. Absolute measures. Bias-corrected DID estimates of Dinaledi program effects for years 2005-2007

	(11) Total number of students entering SG physical science	(12) Total number of students entering HG physical science	(13) Total number of students writing SG physical science	(14) Total number of students writing HG physical science	(15) Total number of students passing SG physical science	(16) Total number of students passing HG physical science
			2004-2005			
SATT	3.629	11.33***	4.193*	11.33***	7.333***	5.927***
	(2.632)	(2.131)	(2.498)	(2.121)	(1.771)	(1.160)
Observations	2287	2287	2287	2287	2287	2287
			2004-2006			
SATT	-10.41***	14.08***	-9.233**	13.97***	0.314	7.632***
	(4.035)	(3.074)	(3.891)	(3.058)	(2.173)	(1.495)
Observations	1830	1830	1830	1830	1830	1830
			2004-2007			
SATT	-5.893	17.00***	-4.661	16.43***	1.813	6.378***
	(4.174)	(3.229)	(3.965)	(3.168)	(2.795)	(1.298)
Observations	1886	1886	1886	1886	1886	1886

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 18. Math results. Ratio measures. Propensity score matching estimates of Dinaledi program effects for years 2005-2007

	(1) Enrolment rate of grade 12 students in SG mathematics	(2) Enrolment rate of grade 12 students in HG mathematics	(3) EC PR SG mathematics	(4) EC PR HG mathematics	(5) WC PR SG mathematics	(6) WC PR HG mathematics	(7) Passing rate out of those entering SG mathematics	(8) Passing rate out of those entering HG mathematics	(9) Passing rate out of those writing SG mathematics	(10) Passing rate out of those writing HG mathematics
2005										
Dummy for 488 Dinaledi schools (2005)	0.0450***	0.0200**	0.116***	0.178***	0.115***	0.177***	0.121***	0.143***	0.121***	0.141***
	(0.0115)	(0.00849)	(0.0206)	(0.0288)	(0.0205)	(0.0289)	(0.0200)	(0.0264)	(0.0199)	(0.0264)
Constant	0.328***	0.0724***	0.578***	0.519***	0.587***	0.523***	0.584***	0.638***	0.594***	0.644***
	(0.00807)	(0.00549)	(0.0157)	(0.0222)	(0.0157)	(0.0223)	(0.0156)	(0.0218)	(0.0155)	(0.0219)
Observations	693	697	706	706	706	706	696	602	696	602
R-squared	0.022	0.008	0.043	0.051	0.043	0.050	0.050	0.048	0.050	0.047
2006										
Dummy for 488 Dinaledi schools (2005)	-0.0216*	0.0533***	0.120***	0.178***	0.117***	0.176***	0.129***	0.0835***	0.126***	0.0659**
	(0.0114)	(0.00844)	(0.0202)	(0.0270)	(0.0202)	(0.0271)	(0.0199)	(0.0257)	(0.0199)	(0.0255)
Constant	0.385***	0.0532***	0.521***	0.493***	0.532***	0.498***	0.521***	0.620***	0.532***	0.641***
	(0.00839)	(0.00494)	(0.0151)	(0.0211)	(0.0151)	(0.0213)	(0.0151)	(0.0204)	(0.0151)	(0.0203)
Observations	682	683	692	692	692	692	687	605	687	599
R-squared	0.005	0.055	0.049	0.059	0.046	0.057	0.058	0.018	0.056	0.011
2007										
Dummy for 488 Dinaledi schools (2005)	-0.0178	0.0486***	0.0771***	0.153***	0.0740***	0.152***	0.0807***	0.0413	0.0776***	0.0321
	(0.0112)	(0.00721)	(0.0192)	(0.0273)	(0.0192)	(0.0273)	(0.0191)	(0.0273)	(0.0191)	(0.0271)
Constant	0.378***	0.0552***	0.552***	0.466***	0.561***	0.472***	0.552***	0.591***	0.561***	0.607***
	(0.00842)	(0.00406)	(0.0145)	(0.0214)	(0.0145)	(0.0215)	(0.0145)	(0.0217)	(0.0145)	(0.0216)
Observations	679	680	704	704	704	704	702	623	702	618
R-squared	0.004	0.062	0.022	0.043	0.021	0.042	0.025	0.004	0.023	0.002

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 19. Math results. Absolute measures. Propensity score matching estimates of Dinaledi program effects for years 2005-2007

	(11) Total number of students entering SG mathematics	(12) Total number of students entering HG mathematics	(13) Total number of students writing SG mathematics	(14) Total number of students writing HG mathematics	(15) Total number of students passing SG mathematics	(16) Total number of students passing HG mathematics
			2005			
Dummy for 488 Dinaledi schools (2005)	-0.813	4.079**	0.326	4.096**	9.649***	2.210
	(4.668)	(1.993)	(4.473)	(1.979)	(2.726)	(1.584)
Constant	86.09***	15.90***	83.61***	15.78***	45.17***	11.64***
	(3.843)	(1.260)	(3.643)	(1.250)	(2.250)	(1.184)
Observations	706	706	706	706	706	706
R-squared	0.000	0.006	0.000	0.006	0.017	0.003
			2006			
Dummy for 488 Dinaledi schools (2005)	-11.99**	12.04***	-10.23**	12.01***	9.708***	6.957***
	(4.952)	(1.704)	(4.634)	(1.695)	(2.350)	(1.250)
Constant	95.85***	10.69***	92.65***	10.58***	40.61***	7.604***
	(4.270)	(0.821)	(3.948)	(0.819)	(1.808)	(0.737)
Observations	692	692	692	692	692	692
R-squared	0.008	0.068	0.007	0.068	0.024	0.043
			2007			
Dummy for 488 Dinaledi schools (2005)	-5.026	12.43***	-3.264	12.00***	8.920***	5.381***
	(5.592)	(2.068)	(5.289)	(2.028)	(2.807)	(1.373)
Constant	102.7***	13.11***	99.60***	13.01***	47.09***	9.054***
	(4.431)	(0.998)	(4.101)	(0.997)	(2.185)	(0.902)
Observations	704	704	704	704	704	704
R-squared	0.001	0.049	0.001	0.047	0.014	0.021

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 20. Math results. Ratio measures. Bias-corrected Propensity Score matching estimates of Dinaledi program effects for years 2005-2007

	(1) Enrolment rate of grade 12 students in SG mathematics	(2) Enrolment rate of grade 12 students in HG mathematics	(3) EC PR SG mathematics	(4) EC PR HG mathematics	(5) WC PR SG mathematics	(6) WC PR HG mathematics	(7) Passing rate out of those entering SG mathematics	(8) Passing rate out of those entering HG mathematics	(9) Passing rate out of those writing SG mathematics	(10) Passing rate out of those writing HG mathematics
					2005					
SATT	0.0443***	0.0193*	0.116***	0.178***	0.115***	0.177***	0.118***	0.127***	0.118***	0.127***
	(0.0161)	(0.0110)	(0.0314)	(0.0433)	(0.0311)	(0.0434)	(0.0313)	(0.0397)	(0.0310)	(0.0398)
Observations	2249	2256	2287	2287	2287	2287	2256	1414	2255	1404
					2006					
SATT	-0.0215	0.0537***	0.120***	0.178***	0.117***	0.176***	0.127***	0.0784**	0.124***	0.0566*
	(0.0176)	(0.0118)	(0.0303)	(0.0391)	(0.0302)	(0.0393)	(0.0303)	(0.0355)	(0.0302)	(0.0343)
Observations	1793	1797	1830	1830	1830	1830	1824	1113	1823	1106
					2007					
SATT	-0.0170	0.0493***	0.0773**	0.153***	0.0741**	0.152***	0.0792**	0.0156	0.0760**	0.00729
	(0.0165)	(0.0104)	(0.0313)	(0.0415)	(0.0313)	(0.0416)	(0.0313)	(0.0395)	(0.0313)	(0.0392)
Observations	1824	1830	1886	1886	1886	1886	1882	1177	1876	1146

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 21. Math results. Absolute measures. Bias-corrected Propensity Score matching estimates of Dinaledi program effects for years 2005-2007

	(11) Total number of students entering SG mathematics	(12) Total number of students entering HG mathematics	(13) Total number of students writing SG mathematics	(14) Total number of students writing HG mathematics	(15) Total number of students passing SG mathematics	(16) Total number of students passing HG mathematics
			2005			
SATT	-0.852	4.078	0.288	4.095	9.632**	2.210
	(6.963)	(3.351)	(6.591)	(3.340)	(4.031)	(3.119)
Observations	2287	2287	2287	2287	2287	2287
			2006			
SATT	-12.08	12.04***	-10.31	12.01***	9.684***	6.955***
	(8.196)	(2.267)	(7.778)	(2.261)	(3.562)	(1.856)
Observations	1830	1830	1830	1830	1830	1830
			2007			
SATT	-5.148	12.43***	-3.380	11.99***	8.882**	5.380*
	(8.122)	(3.417)	(7.619)	(3.392)	(4.275)	(2.909)
Observations	1886	1886	1886	1886	1886	1886

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 22. Heterogeneous treatment effects by former department. Ratio measures. DID estimates of Dinaledi program effects for years 2005-2007

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Enrolment rate of grade 12 students in SG physical science	Enrolment rate of grade 12 students in HG physical science	EC PR SG mathematics	EC PR HG mathematics	WC PR SG mathematics	WC PR HG mathematics	Passing rate out of those entering SG physical science	Passing rate out of those entering HG physical science	Passing rate out of those writing SG physical science	Passing rate out of those writing HG physical science
former House of Assembly schools										
Dummy for 488 Dinaledi schools (2005)	-0.0540**	0.00246	-0.0831*	-0.0143	-0.0820	-0.00405	0.0404	0.0627*	0.0417	0.0741**
	(0.0230)	(0.0260)	(0.0501)	(0.0484)	(0.0502)	(0.0486)	(0.0270)	(0.0375)	(0.0270)	(0.0372)
Year 2007	-0.0424**	-0.0131	0.0816***	-0.0412	0.0863***	-0.0404	0.0816***	-0.0412	0.0863***	-0.0404
	(0.0208)	(0.0214)	(0.0163)	(0.0394)	(0.0161)	(0.0395)	(0.0163)	(0.0394)	(0.0161)	(0.0395)
Interaction Dinaledi Dummy * Year2007	0.0464	0.0286	0.0600	0.130**	0.0607	0.120*	-0.0635**	0.0528	-0.0630**	0.0415
	(0.0294)	(0.0352)	(0.0521)	(0.0637)	(0.0520)	(0.0638)	(0.0306)	(0.0558)	(0.0303)	(0.0556)
Constant	0.200***	0.174***	0.877***	0.738***	0.877***	0.739***	0.877***	0.738***	0.877***	0.739***
	(0.0157)	(0.0163)	(0.0132)	(0.0236)	(0.0132)	(0.0236)	(0.0132)	(0.0236)	(0.0132)	(0.0236)
Observations	206	206	208	208	208	208	201	203	201	203
R-squared	0.041	0.008	0.105	0.034	0.111	0.033	0.081	0.055	0.094	0.060
former Department of Education and Training schools										
Dummy for 488 Dinaledi schools (2005)	-0.0181	0.00938	0.0530	-6.03e-05	0.0503	-0.000643	0.0360	0.0542	0.0329	0.0536
	(0.0152)	(0.00768)	(0.0385)	(0.0444)	(0.0386)	(0.0444)	(0.0225)	(0.0403)	(0.0224)	(0.0402)
Year 2007	0.0491***	0.00176	-0.0888**	-0.0712	-0.0801**	-0.0675	-0.184***	-0.0615	-0.165***	-0.0574
	(0.0158)	(0.00756)	(0.0352)	(0.0443)	(0.0354)	(0.0444)	(0.0270)	(0.0427)	(0.0263)	(0.0428)
Interaction Dinaledi Dummy * Year2007	-0.0160	0.0392***	0.0898*	0.0250	0.0932**	0.0245	0.107***	-0.0759	0.0997***	-0.0769
	(0.0214)	(0.0111)	(0.0470)	(0.0603)	(0.0471)	(0.0604)	(0.0351)	(0.0568)	(0.0344)	(0.0568)
Constant	0.176***	0.0436***	0.706***	0.568***	0.709***	0.569***	0.801***	0.609***	0.805***	0.610***
	(0.0117)	(0.00570)	(0.0281)	(0.0315)	(0.0281)	(0.0315)	(0.0166)	(0.0302)	(0.0166)	(0.0302)
Observations	456	456	472	472	472	472	447	435	445	435
R-squared	0.045	0.107	0.050	0.009	0.047	0.008	0.170	0.032	0.147	0.030
former Bantustan schools										
Dummy for 488 Dinaledi schools (2005)	-0.0112	0.0120***	-0.00494	0.108**	-0.00492	0.109**	0.0310	0.130*	0.0312	0.131*
	(0.0198)	(0.00457)	(0.0564)	(0.0517)	(0.0565)	(0.0517)	(0.0350)	(0.0775)	(0.0349)	(0.0776)
Year 2007	0.0796***	0.0365***	0.205***	0.0894*	0.225***	0.0916*	-0.164***	-0.111	-0.138***	-0.0876
	(0.0178)	(0.00742)	(0.0460)	(0.0500)	(0.0464)	(0.0500)	(0.0331)	(0.0754)	(0.0329)	(0.0758)
Interaction Dinaledi Dummy * Year2007	-0.0417*	0.116***	-0.00145	0.0333	-0.0158	0.0351	0.0368	-0.0839	0.0236	-0.0950
	(0.0251)	(0.0165)	(0.0688)	(0.0682)	(0.0692)	(0.0683)	(0.0478)	(0.0924)	(0.0473)	(0.0927)
Constant	0.114***	0.00894***	0.421***	0.223***	0.422***	0.223***	0.822***	0.537***	0.825***	0.537***
	(0.0139)	(0.00165)	(0.0395)	(0.0352)	(0.0396)	(0.0352)	(0.0257)	(0.0623)	(0.0257)	(0.0623)
Observations	488	487	492	492	492	492	345	320	343	314
R-squared	0.060	0.346	0.067	0.046	0.075	0.047	0.103	0.053	0.082	0.045

Table 23. Heterogeneous treatment effects by former department. Absolute measures. DID estimates of Dinaledi program effects for years 2005-2007

	(11) Total number of students entering SG physical science	(12) Total number of students entering HG physical science	(13) Total number of students writing SG physical science	(14) Total number of students writing HG physical science	(15) Total number of students passing SG physical science	(16) Total number of students passing HG physical science
former House of Assembly Schools						
Dummy for 488 Dinaledi schools (2005)	-13.90***	-4.308	-13.96***	-4.365	-9.346**	-2.058
	(4.795)	(6.662)	(4.787)	(6.650)	(3.952)	(6.409)
Year 2007	-1.981	5.442	-2.231	5.385	3.769	0.942
	(4.419)	(6.462)	(4.410)	(6.447)	(3.837)	(6.716)
Interaction Dinaledi Dummy * Year2007	7.135	2.423	7.212	2.462	1.635	4.808
	(6.427)	(9.405)	(6.414)	(9.384)	(5.701)	(9.148)
Constant	47.27***	45.29***	47.27***	45.25***	39.90***	36.37***
	(3.221)	(4.756)	(3.221)	(4.746)	(2.469)	(4.784)
Observations	208	208	208	208	208	208
R-squared	0.055	0.012	0.055	0.012	0.054	0.004
former Department of Education and Training schools						
Dummy for 488 Dinaledi schools (2005)	-3.958	0.288	-3.669	0.305	-3.424	1.424
	(5.426)	(2.447)	(5.344)	(2.435)	(4.273)	(1.046)
Year 2007	38.47***	1.814	34.91***	1.720	6.805	-1.347
	(9.573)	(2.655)	(8.916)	(2.633)	(5.758)	(0.917)
Interaction Dinaledi Dummy * Year2007	-22.71**	10.26***	-20.03**	10.17***	1.703	4.636***
	(10.55)	(3.442)	(9.933)	(3.415)	(6.506)	(1.356)
Constant	43.13***	12.70***	42.75***	12.67***	34.34***	6.136***
	(4.773)	(2.145)	(4.684)	(2.133)	(3.860)	(0.827)
Observations	472	472	472	472	472	472
R-squared	0.078	0.070	0.073	0.069	0.013	0.086
former Bantustan schools						
Dummy for 488 Dinaledi schools (2005)	-12.48**	1.244*	-12.43**	1.228*	-11.51**	0.593
	(5.337)	(0.690)	(5.309)	(0.688)	(4.647)	(0.546)
Year 2007	29.78***	8.561***	27.24***	8.211***	11.31*	0.602
	(7.559)	(1.705)	(7.491)	(1.651)	(5.946)	(0.475)
Interaction Dinaledi Dummy * Year2007	-8.228	31.86***	-6.610	30.54***	2.171	13.14***
	(9.041)	(4.854)	(8.945)	(4.782)	(7.037)	(1.786)
Constant	31.76***	2.211***	31.63***	2.211***	27.41***	1.691***
	(4.643)	(0.404)	(4.618)	(0.404)	(4.126)	(0.353)
Observations	492	492	492	492	492	492
R-squared	0.087	0.286	0.078	0.275	0.042	0.273

Table 24. Heterogeneous treatment effects by former department. Ratio measures. Bias-corrected DID estimates of Dinaledi program effects for years 2005-2007

	(1) Enrolment rate of grade 12 students in SG physical science	(2) Enrolment rate of grade 12 students in HG physical science	(3) EC PR SG physical science	(4) EC PR HG physical science	(5) WC PR SG physical science	(6) WC PR HG physical science	(7) Passing rate out of those entering SG physical science	(8) Passing rate out of those entering HG physical science	(9) Passing rate out of those writing SG physical science	(10) Passing rate out of those writing HG physical science
House of Assembly schools										
SATT	0.0635**	0.0327	0.0419	0.150**	0.0425	0.140*	-0.0675	0.0661	-0.0691*	0.0542
	(0.0321)	(0.0294)	(0.0700)	(0.0732)	(0.0691)	(0.0718)	(0.0424)	(0.0572)	(0.0408)	(0.0540)
Observations	184	184	190	190	190	190	175	172	175	171
former Department of Education and Training schools										
SATT	-0.0153	0.0332**	0.0900	0.0248	0.0934	0.0243	0.109*	-0.0708	0.0970*	-0.0705
	(0.0256)	(0.0162)	(0.0869)	(0.123)	(0.0871)	(0.123)	(0.0578)	(0.123)	(0.0579)	(0.123)
Observations	555	555	584	584	584	584	536	366	533	362
former Bantustan schools										
SATT	-0.0416	0.113***	-0.00237	0.0333	-0.0167	0.0351	0.0461	-0.150	0.0425	-0.130
	(0.0294)	(0.0209)	(0.0800)	(0.0772)	(0.0796)	(0.0776)	(0.0972)	(0.154)	(0.0923)	(0.152)
Observations	572	573	598	598	598	598	394	118	388	114

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 25. Heterogeneous treatment effects by former department. Absolute measures. Bias-corrected DID estimates of Dinaledi program effects for years 2005-2007

	(11) Total number of students entering SG physical science	(12) Total number of students entering HG physical science	(13) Total number of students writing SG physical science	(14) Total number of students writing HG physical science	(15) Total number of students passing SG physical science	(16) Total number of students passing HG physical science
House of Assembly schools						
SATT	8.363	1.329	8.431	1.380	1.703	5.100
	(6.854)	(10.00)	(6.843)	(9.956)	(7.049)	(6.641)
Observations	190	190	190	190	190	190
former Department of Education and Training schools						
SATT	-22.78*	10.23*	-20.09*	10.13*	1.754	4.648***
	(13.14)	(6.031)	(11.59)	(6.021)	(5.486)	(1.689)
Observations	584	584	584	584	584	584
former Bantustan schools						
SATT	-8.136	31.85***	-6.512	30.52***	2.167	13.12***
	(8.369)	(5.843)	(8.245)	(5.791)	(6.358)	(2.093)
Observations	598	598	598	598	598	598

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 26. Heterogeneous treatment effects by province. Ratio measures. DID estimates of Dinaledi program effects for years 2005-2007

	(1) Enrolment rate of grade 12 students in SG physical science	(2) Enrolment rate of grade 12 students in HG physical science	(3) EC PR SG physical science	(4) EC PR HG physical science	(5) WC PR SG physical science	(6) WC PR HG physical science	(7) Passing rate out of those entering SG physical science	(8) Passing rate out of those entering HG physical science	(9) Passing rate out of those writing SG physical science	(10) Passing rate out of those writing HG physical science
Gauteng, Free State, Mpumalanga, North West and Western Cape										
Dummy for 488 Dinaledi schools (2005)	-0.00393	0.00381	0.00522	-0.0127	0.00324	-0.0133	0.0134	-0.0241	0.0114	-0.0247
	(0.0123)	(0.0105)	(0.0163)	(0.0290)	(0.0163)	(0.0290)	(0.0153)	(0.0260)	(0.0153)	(0.0260)
Year 2007	0.0326**	-0.00768	-0.144***	-0.0947***	-0.129***	-0.0880***	-0.133***	-0.0908***	-0.104***	-0.0803***
	(0.0128)	(0.00942)	(0.0207)	(0.0320)	(0.0208)	(0.0321)	(0.0201)	(0.0296)	(0.0191)	(0.0295)
Interaction Dinaledi Dummy * Year2007	-0.0271	0.0271**	0.0905***	0.0195	0.0846***	0.0156	0.0799***	-0.00265	0.0593**	-0.0104
	(0.0168)	(0.0136)	(0.0272)	(0.0431)	(0.0272)	(0.0432)	(0.0256)	(0.0404)	(0.0247)	(0.0403)
Constant	0.178***	0.0939***	0.849***	0.683***	0.852***	0.684***	0.849***	0.724***	0.852***	0.725***
	(0.00963)	(0.00774)	(0.0114)	(0.0208)	(0.0114)	(0.0208)	(0.0114)	(0.0183)	(0.0114)	(0.0183)
Observations	829	829	844	844	844	844	837	804	833	803
R-squared	0.014	0.013	0.084	0.018	0.069	0.016	0.088	0.027	0.060	0.025
Kwazulu Natal and Limpopo										
Dummy for 488 Dinaledi schools (2005)	-0.00290	-0.000577	-0.0102	0.0877**	-0.0102	0.0928**	0.254	0.338***	0.254	0.357***
	(0.00205)	(0.000901)	(0.0444)	(0.0399)	(0.0444)	(0.0407)	(0.158)	(0.109)	(0.158)	(0.110)
Year 2007	0.178***	0.0848***	0.441***	0.168***	0.457***	0.170***	0.0164	0.0910	0.0338	0.0934
	(0.0118)	(0.00923)	(0.0433)	(0.0347)	(0.0437)	(0.0348)	(0.117)	(0.0747)	(0.118)	(0.0747)
Interaction Dinaledi Dummy * Year2007	-0.0641***	0.134***	0.0218	0.162***	0.0131	0.156***	-0.150	-0.125	-0.149	-0.140
	(0.0164)	(0.0183)	(0.0655)	(0.0562)	(0.0661)	(0.0568)	(0.163)	(0.116)	(0.163)	(0.118)
Constant	0.00379*	0.00283***	0.112***	0.0544**	0.112***	0.0544**	0.579***	0.178**	0.579***	0.178**
	(0.00202)	(0.000756)	(0.0321)	(0.0225)	(0.0321)	(0.0225)	(0.114)	(0.0686)	(0.114)	(0.0686)
Observations	391	390	392	392	392	392	201	234	200	233
R-squared	0.476	0.499	0.329	0.241	0.336	0.237	0.050	0.141	0.049	0.148
Eastern Cape										
Dummy for 488 Dinaledi schools (2005)	0.0232	-0.0205	-0.0195	0.0311	-0.0197	0.0319	-0.0195	0.0169	-0.0197	0.0180
	(0.0257)	(0.0151)	(0.0224)	(0.0973)	(0.0220)	(0.0973)	(0.0224)	(0.0516)	(0.0220)	(0.0512)
Year 2007	-0.0586***	-0.0283*	-0.146***	-0.124	-0.112***	-0.123	-0.127***	-0.114	-0.0920***	-0.0440
	(0.0216)	(0.0162)	(0.0362)	(0.0982)	(0.0331)	(0.0982)	(0.0316)	(0.0809)	(0.0273)	(0.0694)
Interaction Dinaledi Dummy * Year2007	0.00736	0.0556**	0.0774	0.137	0.0512	0.146	0.0588	-0.121	0.0318	-0.164*
	(0.0341)	(0.0214)	(0.0472)	(0.133)	(0.0445)	(0.133)	(0.0438)	(0.104)	(0.0404)	(0.0940)
Constant	0.228***	0.0575***	0.927***	0.544***	0.930***	0.544***	0.927***	0.866***	0.930***	0.866***
	(0.0156)	(0.0126)	(0.0132)	(0.0690)	(0.0128)	(0.0690)	(0.0132)	(0.0387)	(0.0128)	(0.0387)
Observations	170	170	172	172	172	172	171	118	171	115
R-squared	0.073	0.042	0.125	0.023	0.088	0.025	0.117	0.114	0.082	0.108

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 27. Heterogeneous treatment effects by province. Absolute measures. DID estimates of Dinaledi program effects for years 2005-2007

	(11) Total number of students entering SG physical science	(12) Total number of students entering HG physical science	(13) Total number of students writing SG physical science	(14) Total number of students writing HG physical science	(15) Total number of students passing SG physical science	(16) Total number of students passing HG physical science
Gauteng, Free State, Mpumalanga, North West and Western Cape						
Dummy for 488 Dinaledi schools (2005)	-13.67***	-1.408	-13.29***	-1.412	-11.11***	-1.896
	(4.922)	(2.496)	(4.835)	(2.493)	(3.874)	(2.165)
Year 2007	25.31***	2.621	20.65***	2.403	3.896	-1.147
	(8.549)	(2.649)	(7.772)	(2.639)	(5.553)	(2.312)
Interaction Dinaledi Dummy * Year2007	-18.84**	4.024	-14.72*	4.100	-0.768	2.161
	(9.026)	(3.580)	(8.284)	(3.566)	(5.950)	(3.048)
Constant	54.07***	23.58***	53.60***	23.55***	44.52***	17.80***
	(4.600)	(1.926)	(4.509)	(1.924)	(3.644)	(1.655)
Observations	844	844	844	844	844	844
R-squared	0.049	0.010	0.044	0.009	0.019	0.001
Kwazulu Natal and Limpopo						
Dummy for 488 Dinaledi schools (2005)	-0.204*	0.133	-0.204*	0.122	-0.153	0.163***
	(0.117)	(0.122)	(0.117)	(0.121)	(0.114)	(0.0604)
Year 2007	42.55***	20.89***	41.18***	20.42***	21.61***	4.398***
	(3.925)	(2.257)	(3.808)	(2.205)	(2.095)	(0.571)
Interaction Dinaledi Dummy * Year2007	-8.235	35.15***	-7.857	34.05***	1.857	17.97***
	(6.063)	(5.767)	(5.922)	(5.688)	(3.934)	(2.357)
Constant	0.378***	0.357***	0.378***	0.357***	0.296***	0.0612**
	(0.104)	(0.0602)	(0.104)	(0.0602)	(0.103)	(0.0243)
Observations	392	392	392	392	392	392
R-squared	0.298	0.395	0.295	0.388	0.253	0.389
Eastern Cape						
Dummy for 488 Dinaledi schools (2005)	0.302	0.395	0.326	0.372	0.674	0.326
	(8.136)	(1.351)	(8.087)	(1.347)	(6.508)	(1.155)
Year 2007	4.605	-1.419	1.837	-1.488	-3.070	-1.512
	(10.52)	(1.217)	(10.31)	(1.215)	(7.818)	(0.990)
Interaction Dinaledi Dummy * Year2007	13.74	10.86***	15.74	9.651***	11.98	4.721***
	(13.52)	(3.094)	(13.28)	(3.010)	(9.886)	(1.742)
Constant	42.09***	4.767***	41.91***	4.767***	36.60***	3.907***
	(7.127)	(0.861)	(7.086)	(0.861)	(5.685)	(0.738)
Observations	172	172	172	172	172	172
R-squared	0.029	0.165	0.029	0.139	0.021	0.096

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 28. Heterogeneous treatment effects by provinces. Ratio measures. Bias-corrected DID estimates of Dinaledi program effects for years 2005-2007

	(1) Enrolment rate of grade 12 students in SG physical science	(2) Enrolment rate of grade 12 students in HG physical science	(3) EC PR SG physical science	(4) EC PR HG physical science	(5) WC PR SG physical science	(6) WC PR HG physical science	(7) Passing rate out of those entering SG physical science	(8) Passing rate out of those entering HG physical science	(9) Passing rate out of those writing SG physical science	(10) Passing rate out of those writing HG physical science
Gauteng, Free State, Mpumalanga, North West and Western Cape										
SATT	-0.0228	0.0209*	0.0906**	0.0199	0.0845**	0.0158	0.0771*	0.00875	0.0514	0.00281
	(0.0229)	(0.0125)	(0.0407)	(0.0605)	(0.0409)	(0.0605)	(0.0407)	(0.0589)	(0.0405)	(0.0589)
Observations	942	942	973	973	973	973	897	732	893	726
Kwazulu Natal and Limpopo										
SATT	-0.0667***	0.135***	0.0200	0.161**	0.0113	0.156**	-0.122	-0.00947	0.417*	0.0737
	(0.0216)	(0.0222)	(0.0799)	(0.0777)	(0.0811)	(0.0777)	(0.268)	(0.197)	(0.216)	(0.196)
Observations	242	241	243	243	243	243	47	53	44	51
Eastern Cape										
SATT	0.0212	0.0534***	0.0775	0.137	0.0513	0.146	0.0541	-0.178	0.0280	-0.159
	(0.0284)	(0.0160)	(0.0514)	(0.111)	(0.0497)	(0.113)	(0.0506)	(0.113)	(0.0480)	(0.107)
Observations	633	635	670	670	670	670	599	136	592	129

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 29. Heterogeneous treatment effects by provinces. Absolute measures. Bias-corrected DID estimates of Dinaledi program effects for years 2005-2007

	(11) Total number of students entering SG physical science	(12) Total number of students entering HG physical science	(13) Total number of students writing SG physical science	(14) Total number of students writing HG physical science	(15) Total number of students passing SG physical science	(16) Total number of students passing HG physical science
Gauteng, Free State, Mpumalanga, North West and Western Cape						
SATT	-19.05**	3.989	-14.88*	4.066	-0.769	2.165
	(8.709)	(4.892)	(7.628)	(4.851)	(4.139)	(1.570)
Observations	973	973	973	973	973	973
Kwazulu Natal and Limpopo						
SATT	-8.773	35.01***	-8.383	33.91***	1.566	17.94***
	(7.491)	(6.637)	(7.306)	(6.655)	(4.626)	(2.737)
Observations	243	243	243	243	243	243
Eastern Cape						
SATT	13.73	10.85**	15.73*	9.639***	11.93*	4.710***
	(9.906)	(4.601)	(9.546)	(2.174)	(6.556)	(0.941)
Observations	670	670	670	670	670	670

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 30. Characteristics of province sub-sets. Dinaledi schools by former department and province

	Province sub-set 1										Province sub-set 2				Province sub-set 3		Excluded		
	Free State		Gauteng		Mpumalanga		North West		Western Cape		Kwazulu Natal		Limpopo		Eastern Cape		Northern Cape		Total
Non-Bantustan	Nr	%	Nr	%	Nr	%	Nr	%	Nr	%	Nr	%	Nr	%	Nr	%	Nr	%	
CAPE EDUCATION DEPART	0	0.0%	0	0.0%	0	0.0%	0	0.0%	5	13.9%	0	0.0%	0	0.0%	4	8.9%	0	0.0%	9
DEPARTMENT OF EDUCATION AND TRAINING	24	80.0%	38	48.1%	13	40.6%	15	44.1%	10	27.8%	10	15.9%	2	5.4%	6	13.3%	1	33.3%	119
FREE STATE EDUCATION	1	3.3%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	1
HOUSE OF ASSEMBLY	4	13.3%	40	50.6%	0	0.0%	0	0.0%	0	0.0%	9	14.3%	0	0.0%	0	0.0%	1	33.3%	54
HOUSE OF DELEGATES	0	0.0%	1	1.3%	0	0.0%	0	0.0%	1	2.8%	1	1.6%	0	0.0%	1	2.2%	0	0.0%	4
HOUSE OF REPRESENTATIVES	0	0.0%	0	0.0%	0	0.0%	0	0.0%	18	50.0%	2	3.2%	0	0.0%	4	8.9%	1	33.3%	25
NEW EDUCATION DEPT.	0	0.0%	0	0.0%	2	6.3%	7	20.6%	0	0.0%	1	1.6%	0	0.0%	4	8.9%	0	0.0%	14
WESTERN CAPE EDUCATION	0	0.0%	0	0.0%	0	0.0%	0	0.0%	2	5.6%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	2
INDEPENDENT	1	3.3%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	1
TO BE UPDATED	0	0.0%	0	0.0%	1	3.1%	1	2.9%	0	0.0%	1	1.6%	1	2.7%	0	0.0%	0	0.0%	4
Bantustan																			
BOPHUTATSWANA (BOP)	0	0.0%	0	0.0%	3	9.4%	11	32.4%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	14
CISKEI	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	3	6.7%	0	0.0%	3
GAZANKULU	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	5	13.5%	0	0.0%	0	0.0%	5
KANGWANE	0	0.0%	0	0.0%	9	28.1%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	9
KWANDEBELE	0	0.0%	0	0.0%	2	6.3%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	2
KWAZULU NATAL EDUCATION DEPT.	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	39	61.9%	0	0.0%	0	0.0%	0	0.0%	39
LEBOWA	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	14	37.8%	0	0.0%	0	0.0%	14
TRANSKEI	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	23	51.1%	0	0.0%	23
TRANSVAAL EDUCATION DEPT.	0	0.0%	0	0.0%	2	6.3%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	2
VENDA	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	0	0.0%	15	40.5%	0	0.0%	0	0.0%	15
Bantustan total	**0**	**0.0%**	**0**	**0.0%**	**13**	**40.6%**	**0**	**32.4%**	**0**	**0.0%**	**39**	**61.9%**	**34**	**91.9%**	**26**	**57.8%**	**0**	**0.0%**	**112**
Total	**30**	**100.0%**	**79**	**100.0%**	**32**	**100.0%**	**34**	**100.0%**	**36**	**100.0%**	**63**	**100.0%**	**37**	**100.0%**	**45**	**100.0%**	**3**	**100.0%**	**359**

Table 31. Difference-in-means test between retained and non-retained Dinaledi schools. Pre-treatment outcome measures, robust standard errors

Baseline characteristics (2004)	Retained Dinaledi			Non-retained Dinaledi			Difference-in-means Dummy for retained Dinaledi schools	SE
	N	Mean	SD	N	Mean	SD		
Total number of students entering SG physical science	352	29.45	28.27	88	30.67	38.32	-1.224	-4.341
Total number of students entering HG physical science	352	14.06	20.61	88	13.52	19.16	0.537	-2.314
Total number of students writing SG physical science	352	29.37	28.16	88	30.58	38.22	-1.21	-4.329
Total number of students writing HG physical science	352	14.03	20.58	88	13.49	19.12	0.545	-2.309
Total number of students passing SG physical science	352	24.62	22.44	88	25.91	33.15	-1.293	-3.72
Total number of students passing HG physical science	352	10.11	17.36	88	7.85	10.19	2.259	-1.424
Passing rate out of those entering SG physical science	352	0.65	0.4	88	0.6	0.42	0.053	-0.05
Passing rate out of those entering HG physical science	352	0.51	0.4	88	0.47	0.39	0.044	-0.047
Passing rate out of those writing SG physical science	352	0.65	0.4	88	0.6	0.42	0.053	-0.05
Passing rate out of those writing HG physical science	352	0.51	0.4	88	0.47	0.39	0.045	-0.047
Proportion entering SG Physical Sciences of students enrolled in grade 12	352	0.14	0.13	74	0.13	0.15	0.001	-0.019
Proportion entering HG Physical of students enrolled in grade 12	352	0.06	0.09	74	0.07	0.1	-0.007	-0.013

Table 32. Difference-in-means test between retained and non-retained Dinaledi schools. Pre-treatment learner, enrolment and language characteristics, robust standard errors

Baseline characteristics (2004)	Retained Dinaledi			Non-retained Dinaledi			Difference-in-means Dummy for 352 retained Dinaledi schools	SE
	N	Mean	SD	N	Mean	SD		
Total number of learners	352	5736.14	2913.97	88	4433.59	3304.94	1,302.545***	-383.972
Number of African learners in the school	352	4938.07	3264.73	68	4880.96	2873.82	57.109	-388.044
Number of male African learners	352	1344.05	900.5	68	1329.57	846.22	14.48	-112.844
Number of female African learners	352	1519.04	935.93	68	1515.06	825.88	3.984	-111.461
Learners in grades 10, 11 and 12	352	966.31	410.96	88	744.43	554.82	221.878***	-62.886
Number of African learners in grades 10 though 12	352	847.44	500.63	88	647.17	565.22	200.273***	-65.717
Total enrolment	352	3249.7	1521.13	88	2711.49	1743.44	538.207***	-202.21
Total male enrolments	352	1512.93	814.84	75	1460.37	792.3	52.558	-100.93
Total female enrolments	352	1736.76	880.75	75	1721.11	730.02	15.658	-96.186
Enrolments in grade 10	352	443.35	210.22	75	444	200.14	-0.651	-25.597
Enrolments in grade 11	352	345.11	149.71	75	352.24	162.02	-7.126	-20.267
Enrolments in grade 12	352	223.43	96.18	75	214.03	121.82	9.402	-14.916
Male Enrolments in grade 10	352	141.12	71.92	75	140.41	71.9	0.712	-9.113
Male Enrolments in grade 11	352	105.05	50.9	75	107.56	57.25	-2.515	-7.12
Male Enrolments in grade 12	352	67.69	34.09	75	63.04	36.79	4.647	-4.604
Female Enrolments in grade 10	352	157.35	77.42	75	159.28	65.99	-1.928	-8.638
Female Enrolments in grade 11	352	128.4	60.11	75	130.8	55.02	-2.399	-7.092
Female Enrolments in grade 12	352	83.16	41.77	75	81.67	47.34	1.49	-5.881
Number of learners with English as medium of instruction	352	542.71	272.44	76	513.92	266.31	28.795	-33.711

Table 33. Heterogeneous treatment effects on HG physical sciences enrolment rates by former Bantustan status

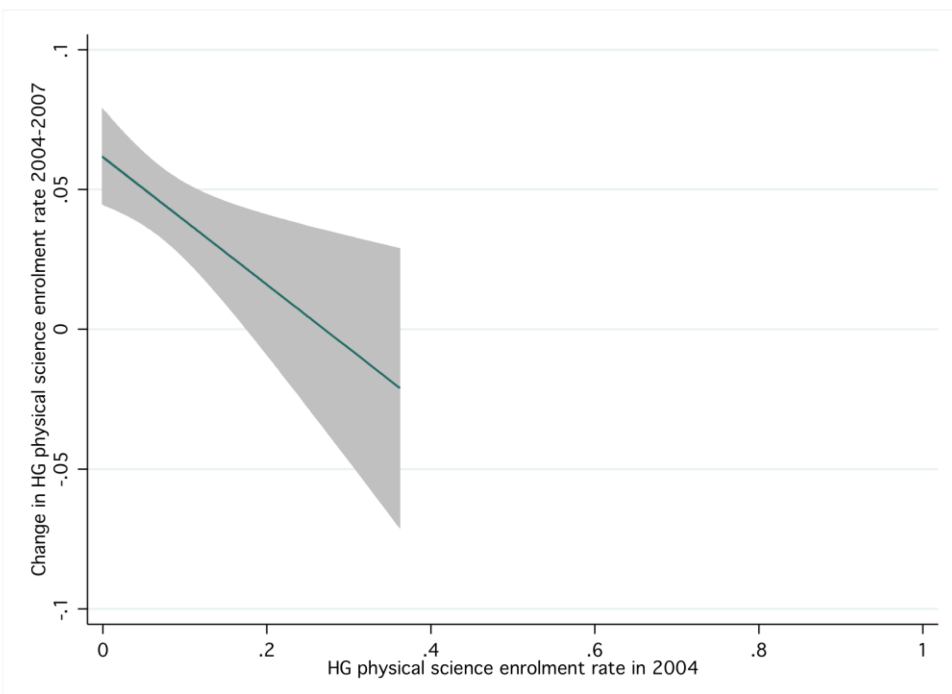
	(1) Change in enrolment rate of grade 12 students in HG physical sciences from 2004 to 2007	(2) Change in enrolment rate of grade 12 students in HG physical sciences from 2004 to 2007	(3) Change in enrolment rate of grade 12 students in HG physical sciences from 2004 to 2007
Department of Education and Training school (dummy)	-0.0514***	0.0108	0.0121
	(0.0105)	(0.0115)	(0.0112)
House of Assembly school (dummy)	-0.0661***	-0.000128	0.0125
	(0.0212)	(0.0229)	(0.0200)
House of Representatives school (dummy)	-0.0674***	-0.0411**	-0.0276*
	(0.0105)	(0.0166)	(0.0167)
Other school (dummy)	-0.0332***	-0.00161	0.00972
	(0.0118)	(0.0146)	(0.0143)
Dummy for 488 Dinaledi schools (2005)[1]	**0.103*** **	**0.107*** **	**0.107*** **
	(0.0171)	(0.0147)	(0.0147)
Interaction Department of Education and Training school (dummy) and Dinaledi	-0.0617***	-0.0746***	-0.0760***
	(0.0195)	(0.0177)	(0.0174)
Interaction House of Assembly school (dummy) and Dinaledi	-0.0679**	-0.0821***	-0.0828***
	(0.0326)	(0.0270)	(0.0250)
Interaction House of Representatives school (dummy) and Dinaledi	-0.0533**	-0.0652***	-0.0625***
	(0.0218)	(0.0204)	(0.0194)
Interaction Other school (dummy) and Dinaledi	-0.0818***	-0.0655***	-0.0664***
	(0.0232)	(0.0215)	(0.0212)
Proportion of students enrolled in grade 12 entering HG physical sciences			**-0.492*** **
			(0.128)
Constant	0.0500***	0.0614***	0.0804***
	(0.00820)	(0.0162)	(0.0160)
Controls used for propensity score estimation	No	Yes	Yes
Observations	678	678	678
R-squared	0.214	0.475	0.499

[1]The Dinaledi dummy estimates the Dinaledi program interaction effect with Bantustan schools, as the Bantustan school dummy is omitted from the regression for collinearity.
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 34. Heterogeneous treatment effects by pre-treatment HG physical sciences enrolment rates.
Subclassification in equal-size sub-sets of 60 observations with respect to enrolment rates

	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6
	HG physical science enrolment rate	HG physical science enrolment rate	HG physical science enrolment rate	HG physical science enrolment rate	HG physical science enrolment rate	HG physical science enrolment rate
Dummy for 488 Dinaledi schools (2005)	0***	0.000707*	-0.00344***	0.00321*	-0.00101	0.0292
	(0)	(0.000389)	(0.00127)	(0.00175)	(0.00297)	(0.0189)
Year 2007	0.0793***	0.0562***	0.0358***	0.0237**	-0.00900*	-0.0365**
	(0.0160)	(0.0103)	(0.00906)	(0.0117)	(0.00532)	(0.0172)
Interaction Dinaledi Dummy * Year2007	**0.0577****	**0.0875*****	**0.0590*****	**0.0192**	**0.0459*****	**0.00281**
	(0.0240)	(0.0213)	(0.0186)	(0.0188)	(0.0116)	(0.0277)
Constant	-0***	0.00103***	0.0206***	0.0426***	0.0902***	0.203***
	(0)	(0.000251)	(0.000825)	(0.00112)	(0.00210)	(0.0112)
Observations	235	234	225	216	240	219
R-squared	0.288	0.351	0.240	0.070	0.132	0.050
			Set characteristics			
Minimum enrollment in 2004	0	0	0.0067	0.032	0.064	0.12
Maximum enrollment in 2004	0	0.0067	0.032	0.064	0.12	1
N of Dinaledi schools in set (pre-matching)	60	60	60	60	60	59

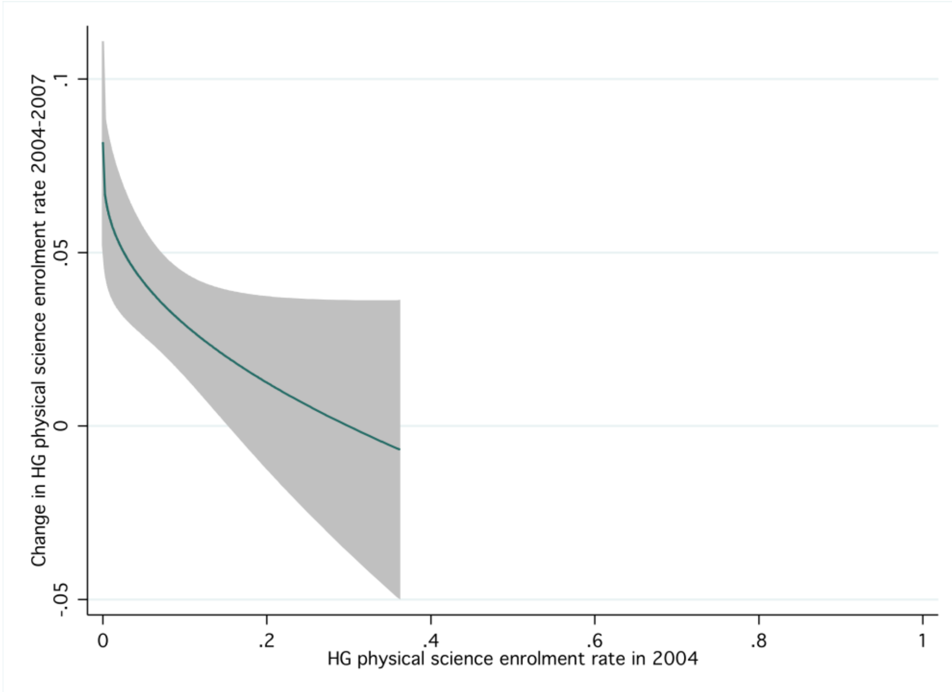
Figure 5. Heterogeneous treatment effects by pre-treatment HG physical sciences enrollment rates.
Multivariable fractional polynomials interactions with one degree of freedom for interaction terms (linear).



Notes: MFPI model with a single power term (1) and lower and upper 95%-confidence limits. The estimated model includes controls[23] for the same set of covariates employed for estimating the propensity score (see Box 1).

[23] The controls actually included in the confounder model are selected at the 5% significance level.

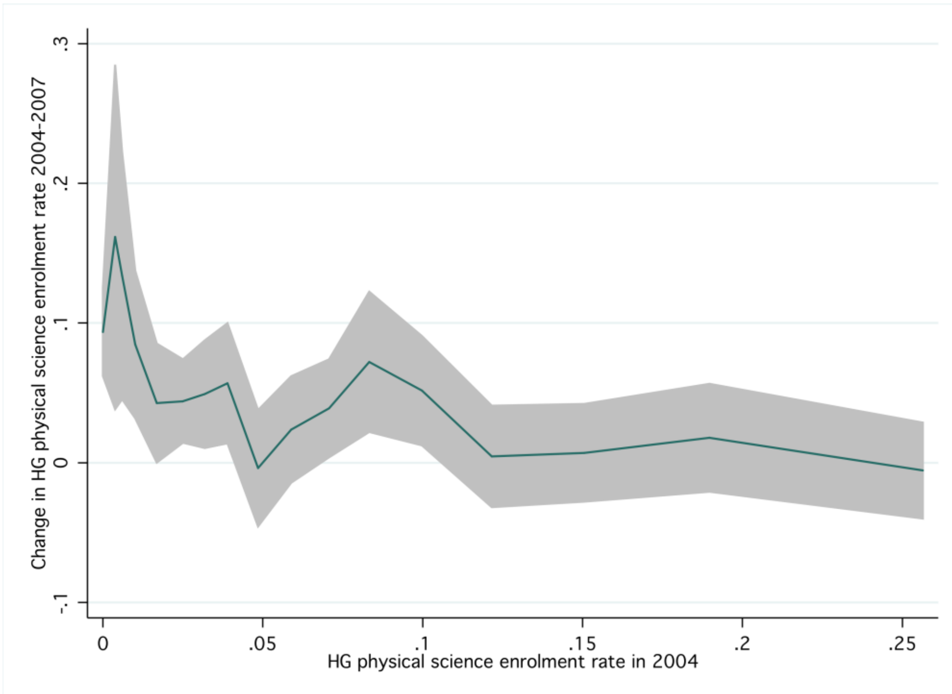
Figure 6. Heterogeneous treatment effects by pre-treatment HG physical sciences enrollment rates. Multivariable fractional polynomials interactions with two degrees of freedom for interaction terms.



Notes: MFPI model with two power terms (0 and -0.5) and lower and upper 95%-confidence limits. The estimated model includes controls[24] for the same set of covariates employed for estimating the propensity score (see Box 1).

[24] The controls actually included in the confounder model are selected at the 5% significance level.

Figure 7. Heterogeneous treatment effects by pre-treatment HG physical sciences enrollment rates. Subpopulation treatment-effect pattern plot (sliding window variant)



Notes: Sliding window STEPP variant with 60 observations in each subpopulation and an overlap of 30 observations between neighboring sub-populations and lower and upper 95%-confidence limits. The estimated model includes controls[25] for the same set of covariates employed for estimating the propensity score (see Box 1).

[25] The controls actually included in the confounder model are selected at the 5% significance level.