What do you mean by 'good'? Interrogating quality and choice preferences in South Africa's non-fee paying school system

Draft July 2017

Gabrielle Wills[1]

Abstract

Whether school choice, and implementing policies to promote this, could improve service delivery outcomes requires certain conditions for success. At the very least, accessible schools of choice for the poor must exist and system actors should be able to recognize and respond to learning quality differentials across schools. This paper attempts to provide more specificity on the validity of these two assumptions using a combination of both Universal Annual National Assessment data and alternate data sources of school recommendations collected for a DFID/ESRC funded project entitled "Succeeding Against the Odds: Understanding resilience and exceptionalism in high-functioning township and rural primary schools in South Africa". The first part of the paper describes the outcomes of a rigorous data collection process to find and verify the quality of what could *potentially* be accessible schools of choice for the poor in three of South Africa's nine provinces. Our search yields no examples of choice non-fee paying schools, although a continuum of functionality is evident as well as the presence of more resilient children even in underperforming schools. The second part analyses a school recommendations dataset identifying if perceptions of performance (in the non-fee paying system) align with higher levels of learning. Overall the purposively selected sample of respondents who are likely to be more aware about school performance than the average citizen, can identify slightly better non-fee paying schools than the average performing school. For certain groups, however, specifically education district officials, enrolment growth appears to be a better indicator of their preferred school than measures of student performance.

JEL codes: I20, I21

Preliminary findings from the project entitled "Succeeding Against the Odds: Understanding resilience and exceptionalism in high-functioning township and rural primary schools in South Africa". Funded by DFID/ESRC.

[1] Gabrielle Wills, Research on Socio-Economic Policy, University of Stellenbosch. Email: gabriellewills@gmail.com

I. Introduction

School choice has long been considered an important consideration in policies to improve educational outcomes, particularly for the poor. As an extension of social accountability discussions in the 2004 World Development Report "Making Services Work for Poor People", Bruns, Filmer and Patrinos (2011) argue that "school choice" is an important mechanism through which citizens impact on standards of education service delivery. Benefits of school choice can be direct through accessing higher quality schools but there may be indirect benefits for system improvement as effective choice drives competition and induces providers to improve quality. Practically, the presence of more functional schools can provide "best practice" examples from which policy makers, districts, school managers and teachers may learn and emulate these in under-performing contexts.

As growing numbers of implemented social accountability interventions generate collective evidence on their effectiveness, a key issue being highlighted is that more attention needs to be given to the assumptions underlying social accountability mechanisms, such as school choice, and the assumed causal mechanisms by which they may realign relationships of power and alter parent/citizen behaviour (McGee and Gaventa, 2010; Read and Atinc, 2017). For school choice to operate as a channel for educational improvement it requires that there are at the very least some better school options that can be accessed by the poor and that quality is detectable and valued by system actors. In some countries, especially those with underdeveloped private school sectors, the availability of choice schools is not necessarily a given. There may be complete deserts of access to quality schooling, as evidenced for example in a search for quality schools in Guinea Bissau, West Africa (Boone, 2013). Even if better schools exist, if higher quality is not recognised by individuals working within education systems or citizens in general the full benefits of school choice are not realisable.

Using a combination of both Universal Annual National Assessment data (U-ANA), a system wide test of mathematics and literacy, and alternate data sources of school recommendations collected for a DFID/ESRC funded project entitled "Succeeding Against the Odds: Understanding resilience and exceptionalism in high-functioning township and rural primary schools in South Africa", this paper contributes by providing more specificity on the validity of these two assumptions in a context of limited data availability.

A handful of studies contribute to school choice debates in South Africa with respect to policy, access and patterns of choice (De Kadt et al., 2014; Hunter, 2015; Kotze, 2017; Msila, 2005; Woolman and Fleisch, 2006; Zoch, 2016). However, the local literature is still in its infancy. Limited understanding exists of contextual realities that shape this domain especially in relation to the prevalence of better primary schools within the majority non-fee paying school sector. Most of the work exploring perceptions of quality and preferences for schools is informed through qualitative studies, not quantitative work. This is in part due to availability of less system-wide data on primary school performance than secondary school performance and the lack of publicly available datasets linking student households to the schools they attend.

However, disadvantaged "schools that work" at the *secondary* level continue to receive policy priority focus largely due to the presence of matriculation data to identify these schools (Christie et al., 2007; Department of Basic Education (DBE), 2017; Jansen and Blank, 2014). But these stories of so called best practice may often just be a function of selection effects, student background factors playing a larger role than schools themselves. Significant dropout which occurs at higher grades, and weeding out of underperformers in schools (Hunter, 2015), makes these apparent outlier secondary schools much easier to identify than at the primary level. If poorer schools are indeed serving the poor, they need to be found at the lower grade levels and after accounting for student background effects. In a

seminal work by Klitgaard and Hall identifying unusually effective schools in the United States, as early as 1973, they identify that the important question to ask is

> *"Do some schools consistently produce outstanding students even after allowance is made for the different initial endowments of their students and for chance variation?" This is an important policy question. Even if unusually effective schools are rare, so long as some exist and can be identified there is hope that their superior performance might be replicated throughout the educational system. But if no exceptional schools exist, we may have to consider alternatives radically different from present attempts to discover and diffuse "best practice". We may need to make substantial changes in educational expenditures, or even overhaul the entire educational system. Thus, investigating the existence of unusually effective schools is not merely a matter of scientific curiosity, but a necessary foundation for a rational public policy toward educational improvement."* (Klitgaard and Hall, 1973, pp. 3–4)

At the outset, it is important to clarify that this paper does not consider private sector provision of schooling in the discussion of the existence of better schools for the poor in South Africa. This stands in contrast to the increased proliferation of studies in developing countries on school choice as a policy relevant lever for learning improvements, particularly with respect to burgeoning private sector school provision, how access to private institutions impacts on learning outcomes and in modelling demand for schools (Carneiro et al., 2013; Day Ashley et al., 2014; Glick and Sahn, 2006; Schneider et al., 2006). While there is considerable anecdotal mention of how the private schooling sector is growing in South Africa, providing alternate schooling options for the poor, this sector in fact remains very small serving at most 5% of all registered school-going students in 2016 (Department of Basic Education (DBE), 2016a).[2] For a dedicated discussion on low-fee private schooling in South Africa the reader is referred to van der Berg *et al* (2017). What is prioritised in this paper is answering the following overarching questions:

1) *Are there available primary schools of choice in the non-fee paying public school sector*?

The largest proportion of South *Africans* are schooled in non-fee paying institutions, and continued system improvements requires a ratcheting up of quality here. The first part of this paper considers whether there are any schools of choice for the poor in three selected provinces through a dedicated search. A fuller discussion on existing studies of school quality and choice in South Africa is given its own attention in the next section.

2) *To what extent are perceptions of quality in this sector aligned with actual learning outcomes?*

Even if better school options do exist and learner performance data on schools is disseminated within the system, it is not necessarily the case that the public will respond in the right direction or that better quality schools are even recognised by authorities. The null-impacts on student literacy and numeracy of the Kenyan UWEZO campaign to disseminate information to parents about their children's performance (combined with materials about how to be more involved in their children's learning) paints a stark reality in this regard (Lieberman et al., 2014). Lepine (2015) highlights that school performance is not always the main determinant of choice in Brazil, and preferences regarding schools are heterogenous across socio-economic groups (Hastings et al., 2009). For example, distance and school demographics are valued more highly than school performance in countries such as Pakistan and Chile (Carneiro et al., 2013; Schneider et al., 2006) while South African qualitative studies suggest that demand for English rather than African medium instruction may be the most critical determinant of choice (Hunter, 2015; Msila, 2005). These preferences can become a constraint to parents and system actors responding to information on school performance as social accountability theories

[2] Until recent developments with the piloting of 'collaboration' schools in the Western Cape, charter-like schools have not existed in South Africa.

would predict. Data limitations prevent us from exploring parent school choice decisions, and their responses to information, but an alternative data source collected for the ESRC/DFID project lend some specificity to how perceptions of school quality across various system actors align with actual school performance.

After a background discussion on school choice in South Africa, the first part of the paper describes the outcomes of a rigorous data collection process to find and verify the quality of what could *potentially* be accessible schools of choice for the poor in three of nine provinces. Annual National Assessment data is corroborated with a dataset collected on potential good schools recommended by various system actors to establish a potential sample of 31 better performing schools. The method and data section maps out in detail the selection process, data used and collected to identify these schools.

Section IV discusses the outcomes of the quality verification process, namely testing literacy proficiencies among grade 3 and grade 6 students in the 31 schools. In benchmarking these schools' performance, our grade 6 sample is juxtaposed against students from other lower to middle income countries that wrote the same released items in previous rounds of the Progress in International Literacy Study (PIRLS). In summary, we fail to find quality non-fee paying primary schools that are performing at levels comparable to other lower to middle income countries. However, within the student sample (augmented through the inclusion of 30 additional typical or underperforming schools to support a matched pair analysis), there are a handful of more functional schools with regards to developing English language proficiencies. A residual analysis, similar to that used by Klitgaard and Hall (1975), indicates that higher learning in these more resilient schools is not merely the result of differences in student background but a school effect is most likely present.

Having discussed the fieldwork verification results in the search for quality schools, Section V analyses the school recommendations dataset collected in the process of identifying quality schools. A purposeful selection process was used to obtain recommendations on potentially 'good' schools in rural and township areas from groups and peoples likely to be more 'in-the-know' about school performance. Using a combination of descriptive methods and multivariate analysis, I identify whether school recommendations reflects a preference for actual higher levels of learning by linking the recommendations dataset to U-ANA. These results interpreted in relation to part A, the search for better schools, leads to the final discussion sections.

II. Background: Existing research on schools of choice in South Africa

It is important to qualify that South Africa certainly does have quality, well-resourced public schools that are performing at or above low international benchmarks of learning. But this is a fee-paying system of schools that only reaches about 10-15% of school-goers and has historically served a privileged white (and to a lesser extent Indian) population who were favoured under Apartheid. School choice discussions in South Africa are typically dominated from the perspective of poorer, and typically black students accessing these historically fee-paying schools. Evidence of increased racial integration in former white and Indian schools along with small-scale quantitative and qualitative studies on school choice patterns provide evidence that some less advantaged South African's are making substantial investments in pursuit of these higher-quality schools (De Kadt et al., 2014; Hunter, 2015). Attending this small proportion of high functioning schools has significant implications for accessing higher levels of learning as discussed below, as well as acquiring describes as 'symbolic (especially English language development) and social capital (networks)' (Hunter, 2015).

New evidence adds impetus to the notion that accessing a quality school, particularly in the primary phase, matters for life trajectories and labour market equalities in South Africa. The critical decision point affecting children's future educational outcomes is the choice made about what school they attend. This is first implied in the work of Taylor and Yu (2009) who investigate the relationship between on the one hand individual *students'* socio-economic status and on the other hand *school* socio-economic status (SES) on students educational outcomes. They note that it is largely school SES rather than student SES that accounts for the largest explained component of student performance, implying that once a school choice decision is made, school effects become more important than background in determining learning trajectories. More recently work by Shepherd (2015), Von Fintel (2015), Von Fintel and Van der Berg (2017) and Kotze (2017) using different sets of nationally representative data, identify effects of school quality equivalent to about one additional year of learning even after controlling for various selection issues that may drive this result. Von Fintel (2015), for example, using National Systemic Evaluation Survey data (2007-2009) finds that black students who attended former white schools were ahead of their peers in non-white schools by almost a year but school quality effects are greater in lower grades. More recent research based on tracking children across schools finds similarly large effects of attending a better performing school (Von Fintel and Van Der Berg, 2017). But school quality doesn't just matter for success on tests, school completion rates and university acceptance. Higher levels of skills obtained through access to better quality schools are important for social mobility and are rewarded in the labour market (Burger and Teal, 2016; Moses et al., 2017; Zoch, 2016).

There are however severe limitations to the extent to which the functioning fee-paying school sector can absorb poorer students. On the supply-side there are too few of these schools, and enrolment intake is very small relative to the size of the non-fee paying sector. At the grade 1 level (the first year of formal schooling) about 70% of students are attending non-fee paying schools (Van der Berg, 2015, p. 35), typically highly under-resourced and characterized by histories of dysfunction - a complexity of institutional Apartheid design and resulting political contestation in these school environments (Fiske and Ladd, 2004). On the demand side, affordability of fee paying schools (in the public system these schools are also referred to as Quintile 4 and 5 schools) is a major constraint due to a combination of direct school fees and at times even higher indirect costs, especially transport given extensive geographical distances between poorer communities to wealthier suburbs with better schools.

With a limited number of schools in the wealthy functional school system and underdeveloped private sector, addressing the service delivery challenge in basic education is currently impossible without the development of quality schools *within* the majority non-fee paying public education system. These schools are synonymous with the Department of Education's Quintile 1 to 3 classification (which receive higher per child allocations than fee-charging Quintile 4 and 5 schools) but are technically not allowed to charge fees. One possible lever for improving service delivery in this system is studying poorer schools that are working and transferring systems, procedures or pedagogical approaches from these functional poor schools to others. This could present a more viable transfer process of best practice than attempting to emulate success from contextually removed former Model C school environments which face different incentives, are resourced completely differently and cater to a much wealthier student composition. But this proposition is premised on an *a priori* assumption that quality or 'outlier' non-fee paying schools exist. As Klitgaard and Hall (1975) reflect, and apt for the South African context as well, we expect that effective schools exist because "parents and children, administrators and teachers, journalists and taxpayers seem to *act* as if some schools were unusually effective". But our *a priori* assumption that these schools exist supporting the DFID/ESRC study was based on more than just anecdotal behaviour. The possibility that exceptional township and rural schools exist was premised on the following:

- The increased success of black South African students in the National Senior Certificate[3] where a strong link between early grade and later grade performance implies that this success was established through access to a functional primary school environment.
- Acknowledged existence of outlier township and rural *secondary* schools as evidenced in national Senior Certificate results and popularized in research such as Christie et al's (2007) "Schools that Work", a more recent update of such institutions (Department of Basic Education (DBE), 2017) and anecdotal media coverage of schools or students that beat the odds.
- Case studies that have highlighted the higher levels of functionality of certain poorer schools over others (Hoadley and Galant, 2015; Levy and Shumane, 2016).
- Growing evidence on national improvements in the South African schooling system (albeit off a very low base) as measured in international tests of numeracy and literacy which are not just limited to wealthier student groups (Department of Basic Education (DBE), 2016b; Reddy et al., 2016, 2015).

But possibly the dominant motivator for this *a priori* assumption is that studies highlight the existence of a 'second pattern' of school choice in South Africa, involving choice between more local schools but less travel or financial investment than decisions to access high-quality fee-paying schools (De Kadt et al., 2014). In a study of children's daily travel to school in Soweto (a township in Johannesburg) using Birth to 20 cohort data, de Kadt *et al* (2014) identify that *only* 15% of primary school-goers in their sample attend their nearest school (on average about 400m from their home). If a one-way distance of 3km or more from home to school requires making use of transport and thus financial investment, they argue this provides a crude proxy for distinguishing between the first and second patterns of choice. Their findings imply that the majority (roughly two thirds) of school choice decisions that are being made in Soweto can be classified by the second pattern. They posit that

> "even those families who do not have the resources to travel long distances and pay high school fees still engage in school choice in a more local context, and appear to use this as a tool to improve the educational opportunities available to their children. This stands in sharp contrast to findings from other contexts that less-advantaged parents are often less engaged in school choice, raising additional questions about the implications of school choice for equality of access to educational opportunities." (De Kadt et al., 2014, p. 184)

Although a second pattern of school choice exists, empirical evidence is necessary to verify whether these school choice patterns do in fact lead to improved access to quality education. This depends first on whether adequate non-fee paying schools exist which in turn requires system-wide testing data to determine this. Unlike many African states which have a primary school leaving certificate, South Africa neither has this nor any other consistent metric for measuring learning outcomes across *all* primary schools. However, the introduction of the short-lived Universal Annual National Assessments (U-ANA)– a universal testing system for grades 1-6 and grade 9 implemented between 2011-2014 - provided important data with which to better understand the performance of specifically the primary schooling system.[4] Kotze (2017) drawing on U-ANA for 2012-2014, investigates the prevalence and performance of poor schools which manage to perform above a demographic expectation. She notes that only 5% of all Quintile 1 – 3 schools (more generally referred to as non-fee paying schools), serving only 4% of the learner population, manage to perform at this level. Despite less advantaged citizens engaging in school choice decisions, the extremely low prevalence of quality non-fee schools presents a major barrier to social mobility of most South African families. This paints

[3] The NSC also known as 'matric' is a school leaving examination at grade 12.

[4] In 2015, the ANA's were boycotted by teachers' unions. They have been indefinitely discontinued in the face of union resistance and methodological criticisms. It is not clear that a universal ANA will be implemented again although sample based ANA-like system could be introduced.

an unfortunate reality that at least with regards to primary schooling, for the majority who engage in a 'second pattern' of school choice this is unlikely to provide children with access to even *adequate* levels of learning.

Kotze's quantification established a very useful foundation to proceed in identifying how many choice non-fee paying primary schools exist in the non-fee public sector. However, more grass-roots level research is necessary to verify the quality of the schools identified among her '5%' and add supportive evidence to her sobering findings. The notion of deserts of access to quality education is a hard pill to swallow for public and policy makers, but there has also been widespread concern about the validity and reliability of the ANAs as a testing system, with numerous accounts of cheating through leakage of scripts, teaching to the test, and inconsistencies across school level results and those achieved in national datasets (Ross, 2016; Taylor and Draper, 2014). One of the primary points of inefficiency is that teachers within schools mark these tests. For the most part these concerns would lead one to think that Kotze's estimates are *overestimated*, but more generally this produces some uncertainty about the validity of results. The school search process described in this paper for the ESRC/DFID funded project attempts to address some of these validity concerns through supplementing U-ANA with word-of-mouth recommendations.

Other than the ANAs, and systemic testing conducted in the Western Cape, there is scant tracking of primary school performance using anything that closely resembles standardised testing methods in eight other provinces. Much mention is made of the use of common tests in some provinces, but there is no systematic strategy for standardised assessment. There has also not been a systematic approach taken to disseminate ANA or Western Cape Systemic data into the public to assess how it may affect school choice patterns, although officials and educators may be exposed to results. The lack of systemic data prevents the implementation of existing legislation that holds school leaders accountable for underperformance (Republic of South Africa, 2000). In the context of high levels of information asymmetries in the primary education sector, I investigate whether system actors can identify better quality schools. Studies in other countries indicate that even in the absence of information, system actors can make quality judgements about schools, however quality can become a more important determinant of choice when data are available (Gomez et al., 2012).

III. Method and Data

a) About the ESRC/DFID funded project entitled "Identifying exceptionalism and resilience among township and rural primary schools in South Africa"

"Identifying exceptionalism and resilience among township and rural primary schools in South Africa" more affectionately known as "Leadership for literacy" is an education research project lead by a multi-disciplinary team of researchers across Stellenbosch University, UCT, JET, UNISA and the Department of Basic Education in South Africa. The project is funded by the Economic and Social Research Council (ESRC) and the Department for International Development (DFID) in the United Kingdom. It runs over the period September 2016 to September 2018. The project was initiated in reaction to a deficit discourse where much research has focussed on the real realities of a highly underperforming schooling system in South Africa particularly among the non-fee paying public schools. Where solutions are desperately needed, less consideration has been given to exploring pockets of excellence that may exist. There has been little rigorous research on higher-performing schools in challenging contexts at least at the primary level and what may be learned from these schools, particularly with respect to how the organisation is led and managed. The project was premised on the assumptions that these schools exist, where reasons for this *a priori* assumption are presented earlier.

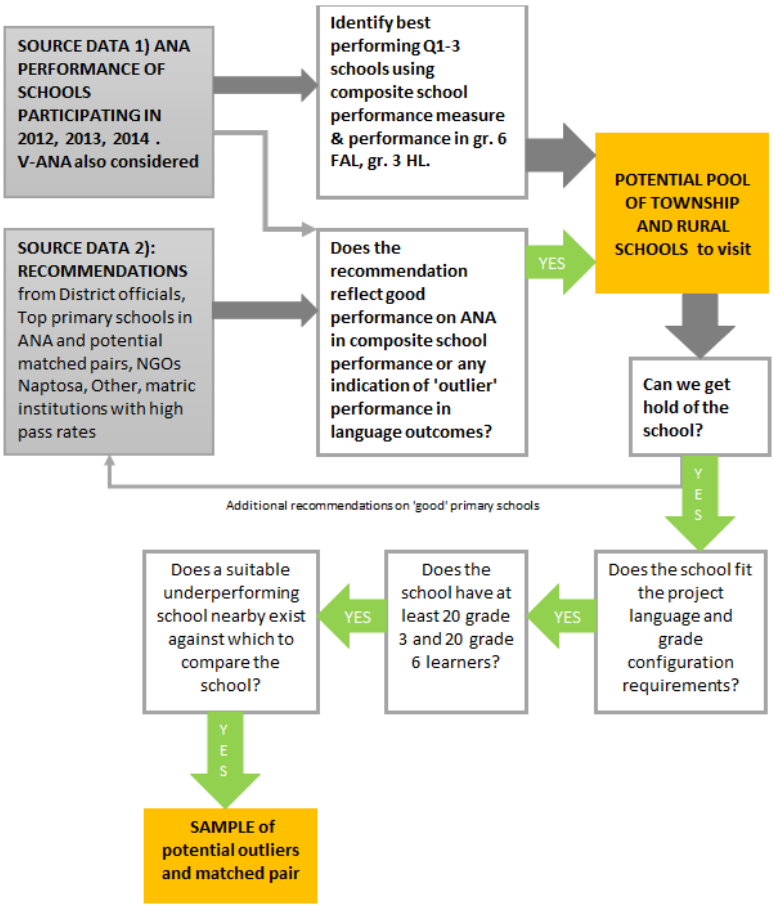
In South Africa schools are administered under nine provincial departments with their own set of characteristics and unique bureaucratic and political dynamics. We decided to locate the study within three departments of distinct administrative functionality: Gauteng (a typically highly functional administration), KwaZulu-Natal (medium functionality) and Limpopo (low functionality).

At project onset, a mixed methods approach with a matched pairs design was envisaged. Each outlier school is paired with a nearby typical or underperforming school. The matched pairs approach assumes that given a similar geographical position each school should display roughly equal socioeconomic characteristics and share similar cultural/political/local dynamics. This largely supports the qualitative component of the project - by making comparisons across a high-performing and low-performing schools one can factor out the influence of some unobserved characteristic on the findings. However, the first challenge to establishing a set of schools to visit was to *identify* the outlier school pairs.

b) Method of establishing for potential pool of 'outlier' schools in three provinces

Figure 1 depicts the decision process we used in establishing a sample of potential better performing township and rural school across our three provinces visited for round 1 fieldwork, February 2017. (A second round of fieldwork is envisaged for later in the same year). The process starts with using Annual National Assessment data, collecting our own dataset on school recommendations and then a decision process determining whether potential schools meet the language and grade configuration dimensions of the project.

Figure 1: Mapping out the process used in establishing our final sample of schools to visit



Data source 1, Annual National Assessment data:

The cornerstone of our school identification process was a three-year school level dataset constructed by Kotze (2017) using Annual National Assessment data (2012, 2013 and 2014). Additionally, we looked for best performing Quintile 1-3 schools in the sample-based verification ANA dataset. It is important to point out, however, that it is often the case that schools do not perform consistently well across all grades, subjects and years. To limit the complication this presents, we focused largely on composite measures of school performance (average performance across grades, subjects and even testing years) as well as performance in our testing areas of interest for baseline fieldwork, namely grade 3 home language and grade 6 first additional language.

In defining better performing schools in ANA we loosely set our benchmark as the average performance of South Africa's Quintile 5 schools (public schools receiving the least allocation per child but charging the highest fees). Using a composite school performance ANA measure, standardized to have a mean of 0 across all schools participating in ANA, Quintile 5 schools are on average scoring 1 standard deviation above the national average. The rationale for using this benchmark is that students in Quintile 5 schools typically perform at around the low international benchmark in international tests of numeracy and literacy such as TIMSS and PIRLS (Van der Berg, 2015). This is consistent with Kotze's (2017) approach, however ours is slightly more conservative as we benchmark *at* the average rather a standard deviation about the average. We argue that this is a conservative benchmark for "good" school performance. Our project research team identified specific Quintile 5 schools that we would personally be comfortable sending our children to given our knowledge of schools we have attended, our friends' children attend and have known to be historically great schools. With respect to ANA performance, these institutions are performing around 2 standard deviations above the national mean (not 1 standard deviation) in the composite performance measure.

Table 1: The number of non-fee paying schools performing at or above the average of fee paying Quintile 5 schools, U-ANA 2012-2014.

	Q1-3 schools performing at or above Q5 average				**Q1-3 schools performing below Q5 average on composite school performance**
	School size >=250		**School size < 250**		
Province	**Number**	**% of Q1-3 schools in sample**	**Number**	**% of Q1-3 schools in sample**	
EC	15	0.4%	74	1.8%	4081
FS	5	0.7%	39	5.7%	642
GT	26	3.7%	2	0.3%	702
KZ	85	2.6%	122	3.6%	3164
LP	11	0.5%	15	0.6%	2310
MP	1	0.1%	4	0.4%	933
NC	6	2.3%	6	2.2%	255
NW	1	0.1%	1	0.1%	927
WC	0	0.0%	11	2.0%	538
Total	**150**	**1.1%**	**274**	**2.0%**	**13552**

Source: Universal Annual National Assessment, panel of 2012-2014 data. **Notes:** Averages are calculated using a composite measure of performance across grades1-6, all subjects and three years of data. "Q" is an abbreviation for quintile. Quintile 1-3 schools are non-fee paying schools while Quintile 4, and 5 schools are typically fee-paying schools receiving lower per child allocations from the state but charging fees.

Unfortunately, we had to loosely apply this mere 1 standard deviation benchmark, simply because so few schools are performing at or above this level. Table 1 indicates that even with the possibilities of cheating, teaching to the test and other irregularities, only 3% of all Q1-3 schools are performing at or above the average of Quintile 5 schools. Using this ANA metric, in Gauteng there are at most 26 "outlier" schools with enrolment of at least 250 students that we could consider, 85 in KwaZulu-Natal and 11 in Limpopo.

Data source 2, recommendations of township and rural school:
Given concerns about the reliability of ANA, we also initiated a process of collecting a dataset on school recommendations eliciting stakeholder opinions on "good" township and rural schools with respect to academic performance, particularly in language. This dataset of school recommendations is not derived from a random selection of South Africans. We purposefully identified individuals and organisations that we assumed would be more aware about school performance including provincial education offices, district officials (specifically subject advisors), non-profit organisations working in the education sector, non-fee paying secondary schools with high NSC pass rates, and national and provincial branches of a teachers' union. We also asked a large education blog readership to identify what they believe are outlier schools. Importantly, school recommendations from principals or administrative clerks from non-fee paying primary schools also comprised a large portion of our responses. This was a direct result of contacting well-performing ANA schools (and potential matched pairs) to scope out the possibility of i) actually being able to contact them (given terribly outdated EMIS data on school telephone numbers), determining whether their language of learning and teaching in the foundation phase matched our project needs and ii) whether they would be willing for us to visit their school.

Table 2: Breakdown of provincial school recommendations received and institutions or individuals from whom suggestions were provided.

	Gauteng		**KwaZulu-Natal**		**Limpopo**		**All recommended schools**	
	Q1-3	**Q4-5**	**Q1-3**	**Q4-5**	**Q1-3**	**Q4-5**	**Q1-3**	**Q4-5**
District official suggestions	25		55	23	50	1	130	24
Q1-3 Secondary schools with high matric pass rates	37	5	25	20	15		77	25
Worse performing ANA primary schools	45	9	14	5	14		73	14
Better performing ANA primary schools	31	4	14	3	5		50	7
NGO informants	16	2	14	1	16		46	3
Other	8	2	15	8	2		25	10
Total	**162**	**22**	**137**	**60**	**102**	**1**	**401**	**83**
Combined totals	**184**		**197**		**103**		**484**	

Notes: Better performing ANA primary schools are defined here as performing at 0.5 standard deviations or more above the national mean while worse performing ANA primary schools are performing below 0.5 standard deviations on a composite measure of school performance. "Q" is an abbreviation for quintile.

Over the period July 2016 to January 2017, we collected 519 recommendations of township and rural schools across Gauteng, Limpopo and KwaZulu-Natal.[5] 486 of these recommendations coming from about 245 unique sources could be matched to the three-year school level ANA dataset as well as national schools' data to identify characteristics of the recommended schools. A breakdown of the numbers of recommendations received per province, with a broad category for respondent type are provided in Table 2. About 83 recommendations were for fee paying schools (typically low fees of less than R2000 a year) which is not surprising given that low fee-paying schools exist in township and rural areas.

Given the intentionally non-random nature of the dataset it cannot be used to generalize about average perceptions of school quality. Nevertheless, this data source is useful in achieving two sets of research objectives. First, it aids the identification of a potential pool of township and rural schools to consider for sample selection for the project. Second, it can be used to explore how system actors' perceptions of quality align with actual quality in the poorer part of the schooling system, even after accounting for observed school resourcing differences or other observed differences across schools that may inform perceptions. The empirical strategy used to evaluate perceptions versus quality is discussed in section V.

c) Method of narrowing down the potential pool of 'outlier' schools to verify performance

Having established a potential pool of good schools to visit from data sources described above, the next step was to narrow it down to the actual schools in which we could conduct fieldwork, verifying performance. The filtering process was largely operationalized by corroborating information on the recommendations received with performance on ANA. However, it was further complicated by language and grade configuration requirements for language testing, national schools' data inaccuracies as well as the matched pair design of the project. These challenges are described in more detail in the appendix. However, attention is drawn here to the language dimensions and matched pair design of the project that placed the greatest squeeze on our potential sample.

Language of testing

At project initiation, our plan was to use 2015, 2016 and 2017 universal ANA results which tests in maths and language. This plan ended abruptly with union resistance in late 2014 to the continuation of ANA. In response to this setback and to growing concerns about abysmally low levels of basic literacy skills being acquired in both African languages and English across primary schools (Spaull et al., 2016) a decision was taken to focus on just testing *literacy* outcomes at two critical grade transition points: grade 3 and grade 6.[6]

For the sake of reduced complexity and cost, a decision was taken to develop reading tests at the grade 3 level in only three of the 11 official South African languages. The predominance of isiZulu in KwaZulu-Natal and its proliferation in many other parts of the country, including Gauteng, made this

[5] It must be noted that while we collected a few hundred recommendations obtaining these involved *thousands* of phone calls. Telephone numbers of officials or schools as per national school lists (EMIS) or even obtained directly from districts often do not exist, just ring, or go to voicemail resulting in low ratios of recommendations to phone-calls made. This is a major barrier to fieldwork projects in schooling more generally.

[6] South African basic education at the primary level is split into three phases; the foundation phase (grades 1-3), the intermediate phase (grades 4-6) and senior phase (grade 7). Students typically learn in their home language in grades 1 to 3, then a language switch to English takes place in grade 4. The curriculum assumes that children have acquired basic reading skills in both their home language and English by grade 3. Testing both African language and English proficiency at this point is an important indicator of the readiness of the child to proceed to further grades and keep pace with the demands of the national curriculum.

an obvious test language of choice. Anticipating a shortage of acceptable schools in Limpopo with its accepted lower levels of learning at the primary school level relative to other provinces, we expanded the language of testing to two areas: Sepedi and Xitsonga.[7]

Initially we wanted to limit our sample to only schools teaching in African home language in the foundation phase. But the options became so limited that we had no choice but to lift this criteria and include a few English LOLT schools in our sample, provided the dominant home language of the class matched our language testing area.

Table 3 summarises our reasons for selecting the final 31 schools as potential 'outliers' other than the fact that they all meet language and grade configuration requirements of the project. In particular, it shows on which ANA measures schools are performing at or above the 90th percentile in the national distribution of *all* schools participating in ANAs from 2012 to 2014. It also identifies if and how many respondents recommended each school. Three important points are noted from the table:

- In Gauteng and KwaZulu-Natal our final selected 'good' pairs were typically recommended at least once by a word of mouth source, and were performing at or above the 90th percentile in at least one of the ANA measures considered. In the exceptions where a word of mouth recommendation had not been received for the school, but we included it in the sample anyway, this was because it really appeared to be an outlier in terms of the composite school performance measure in ANA. (There is one school in KZN which we included in spite of its average ANA performance because as many as 5 different 'word of mouth' sources indicated this was a good school – the maximum number of recommendations we had received for one school).
- In Limpopo, the selection criteria had to differ. For only 3 of the 10 potential outlier schools selected did we receive a word of mouth recommendation. Despite collecting roughly 100 suggestions for primary schools to visit, there were so few cases where recommendations aligned with good performance on ANA and project language dimensions– in most cases recommended schools had dismally low ANA performance in the national context although they may be better performing relatively within the province. This is observed in Figure 2, the distribution of average school performance in ANA in quintile 1-3 schools across Gauteng, KwaZulu-Natal and Limpopo. The Limpopo distribution lies substantially to the left of the other two provinces.
- Not all the final schools selected were non-fee paying schools. With few potential non-fee paying outliers emerging we included five *low* fee-paying (under R2800 per annum) public schools in the sample.

The matched underperforming pairs were chosen on the basis of being the closest school to the better pair that i) had the same language of learning and teaching and ii) was underperforming relative to the better potential pair on key ANA metrics. Due to these criteria, the matched pair was not always a nearby school. Note, only schools with at least 20 learners enrolled in grade 3 and 20 in grade 6 were included.

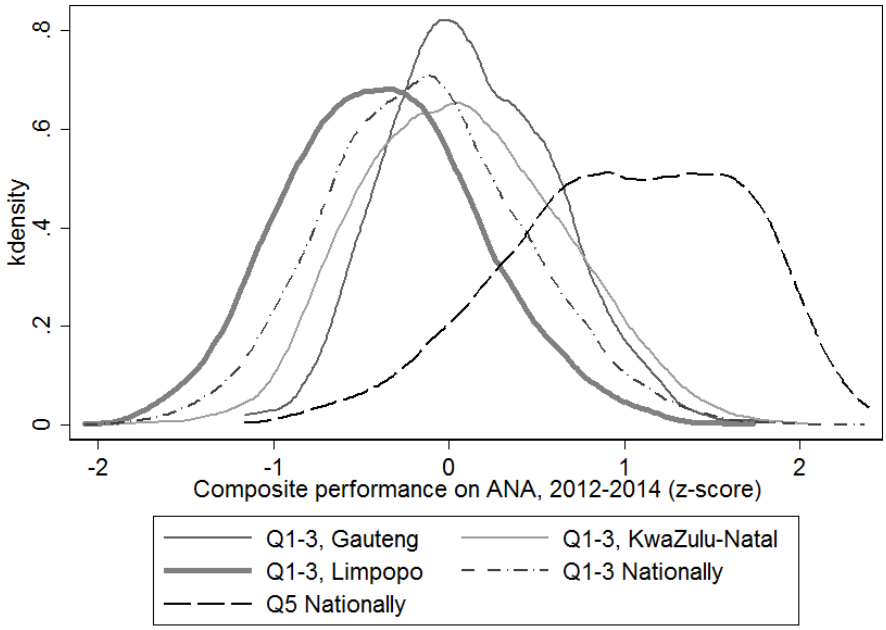
[7] It is important to qualify that despite testing in the most frequently occurring languages in these provinces, it became evident after further investigation that best performing quintile 1-3 schools in ANA in each province are not dominated by a language group. In this respect, we have potentially missed some quality township and rural schools across the three provinces in consideration, in addition to 6 other provinces which were not considered.

Table 3: Reasons for selection of potential "outlier" schools

	Data source 2		Data source 1: Annual National Assessment Performance					
	Low-fee paying	# of 'word of mouth' recommendations	Composite performance 3 year (z-score) 90th pct.	Gr 3 home language 2014, 90th pct.	Gr 6 home language 2014, 90th pct.	Gr 4 first additional language 2014, 90th pct.	Gr 5 first additional language 2014, 90th pct.	Gr 6 first additional language 2014, 90th pct.
Gauteng potential "good" pairs								
GP10		0	X			X	X	X
GP6		0	X			X	X	X
GP0 (no pair)		1	X	X		X	X	X
GP8		1	X	X		X	X	X
GP9	X	1	X	X				X
GP3		1	X		X			
GP5		2	X			X	X	X
GP7		2	X			X	X	X
GP1		3	X			X	X	X
GP2		5	X		X			
KwaZulu-Natal potential "good" pairs								
KZN20		0	X					
KZN14		0	X	X		X	X	X
KZN15		1	X					
KZN12		1	X	X		X		X
KZN13		2						X
KZN0 (no pair)	X	2	X		X			
KZN18		2	X	X			X	X
KZN19		2	X			X	X	
KZN11		2	X	X		X	X	
KZN17		5						
KZN16	X	5	X	X		X	X	X
Limpopo potential "good" pairs								
LP24		0			X	X	X	X
LP30	X	0			X	X		X
LP28		0	X	X		X		
LP29		0	X	X			X	X
LP26		0	X	X				X
LP21	X	0			X		X	X
LP27		0	X			X	X	X
LP22		1	X				X	X
LP25		1			X	X	X	X
LP23		3		X		X	X	

Notes: Percentiles (pct.) of performance in ANA are calculated including all (Q1-5) schools that are in the 2012-2014 ANA panel dataset.

Figure 2: Distribution of schools' average composite performance on ANA, 2012-2014



Source: Annual National Assessment performance data. Q1-3 stands for quintile 1-3 schools that are non-fee paying. Q5 are Quintile 5 schools that receive the lowest per child state allocations but typically charge the highest fees.

d) Fieldwork: Verifying the performance of potential 'outlier' schools

Eventually we visited and assessed students in 61 schools in three provinces of which 31 were potential high performing schools.[8] Quantitative fieldwork was conducted for one day in each school between early February into the beginning of March 2017. National and provincial education department approval was obtained for the study as well as ethical clearance from Stellenbosch University.

At the grade 6 level, we administered a written test to one full class per school visited. A total of 2652 grade 6 students were tested and background information was collected on each student. For much of the analysis that follows I give most attention to these grade 6 outcomes as it involved testing an entire class – a larger sample than for the grade 3s where time constraints of one-on-one Oral Reading Fluency testing limits the potential for larger samples. Additionally, a number of instruments were administered to capture school characteristics, school climate, school functionality indicators, teacher perceptions and leadership and management practices in the school.

[8] Due to school access challenges during fieldwork we surveyed only 29 pairs rather than 30 as initially intended (10 in Gauteng, 10 in KwaZulu-Natal, 9 in Limpopo). One Limpopo school which seemed to be high performing in ANA was visited but due to bad weather and strike action we could not access its intended underperforming pair. Instead we visited 2 additional schools for which we received recommendations, bringing the total sample to 61.

Table 4: PIRLS analysis of the two comprehension tests used, proportional mark allocations by type of item

	Text 2	Text 1	
Type of questions	**Purposes for reading: Acquire and Use Information**	**Purposes for reading: Literary Experience**	**Both texts**
Examine and Evaluate Content, Language, and Textual Elements	11.1%	12.5%	11.8%
Focus on and Retrieve Explicitly Stated Information	22.2%	31.3%	26.5%
Interpret and Integrate Ideas and Information	38.9%	43.8%	41.2%
Make straightforward Inferences	27.8%	12.5%	20.6%
Grand Total	**100.0%**	**100.0%**	**100.0%**
Maximum possible marks	**18**	**16**	**34**
Grade appropriate word analysis*	**Gr 3.3**	**Gr 5.3**	

Source: Own analysis of PIRLS 2011, and PIRLS 2011 item information files (TIMSS 2011 Assessment and/or PIRLS 2011 Assessment. Copyright © 2013 International Association for the Evaluation of Educational Achievement (IEA). Publisher: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.) *Internal word analysis.

The grade 6 test focused on English literacy where permission was received from the IEA to administer two comprehension tests - released items from previous rounds of the 4th grade Progress in International Reading and Literacy Study (PIRLS). Table 4 provides an analysis of the type of literacy skills tested in the two comprehensions as identified by the IEA. An internal word analysis conducted by our research team on the two texts, indicates that first comprehension text is appropriate for a grade 3 level while the second is appropriate for a grade 5 level. In addition to the PIRLS texts, which allow us to internationally benchmark student proficiencies in our sample, vocabulary testing was also included with four sections testing most common to less common words in English. Total possible marks obtainable in the entire test was 106.

The following section discusses the test results from February data collection, positioning performance within an international comparison while also considering variation across the 61 school sample. The analysis relies on both descriptive and multivariate methods.

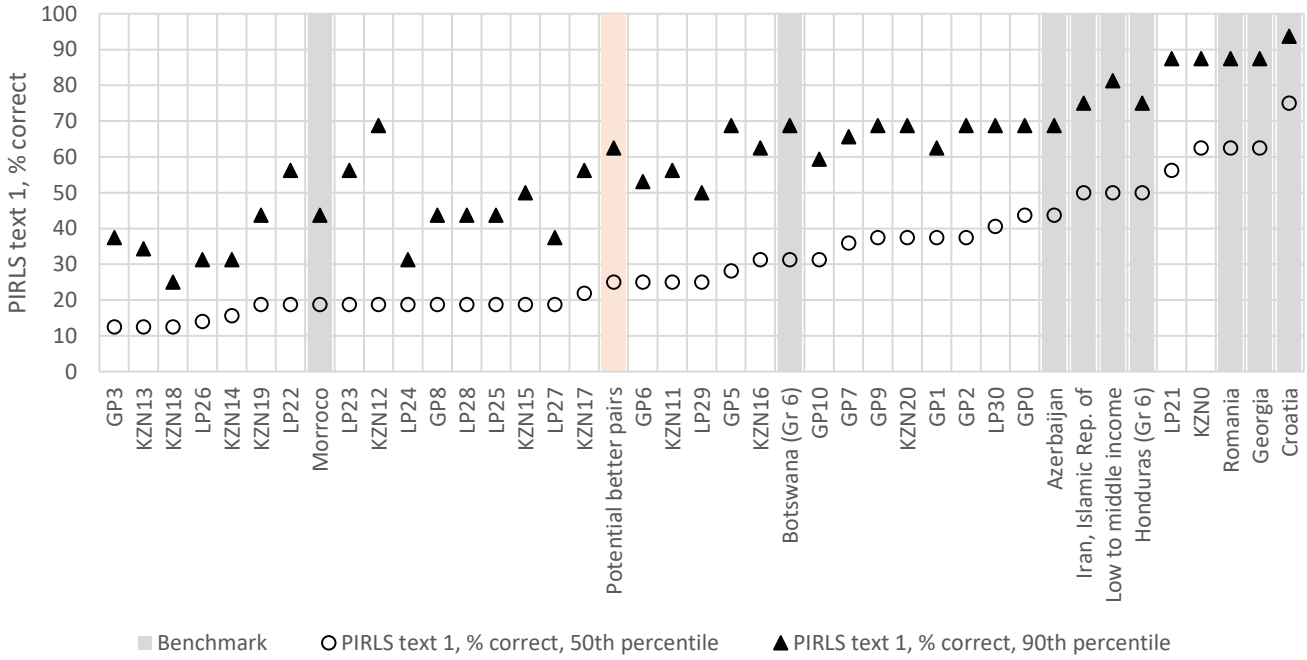
IV. Findings part A: Did we identify any 'outlier' schools?

a) Benchmarking performance against other lower to upper middle income countries

This section benchmarks the performance of the schools we visited to other country samples. Figures 2 and 3 juxtapose the performance of students in each of the 31 potential better performing schools (equivalent to 31 different grade 6 classrooms) to students in lower to upper middle income country samples[9] that wrote the two PIRLS texts described in section III. The figures reflect performance on each text separately, reflecting the percentage correct across each test at the 50th and 90th percentiles. All countries reflected administered these tests at *the grade 4 level*, with the exception of Botswana and Honduras which administered the tests at the grade 6 level. The samples in the figures are sorted by median performance.

Evidently, the overall performance of the 31 school sample does not fare well relative to the lower to upper-middle income countries in question. Across both texts there are only two schools in our sample, LP21 and KZN0 (both low-fee paying schools), with performance that exceeds the median and 90th percentile performance of the combined comparator countries. But even these two best schools do not stand out as much better than country samples of students in Romania, Georgia and Croatia that are at a two-grade level disadvantage. What is most sobering is the contrast with students in neighbouring Botswana. As many as 21 of the 31 purposefully selected and recommended grade 6 school classes are performing worse at the 50th percentile in the first text than a random sample of Botswanan grade 6 students (28 of 31 in the second text).

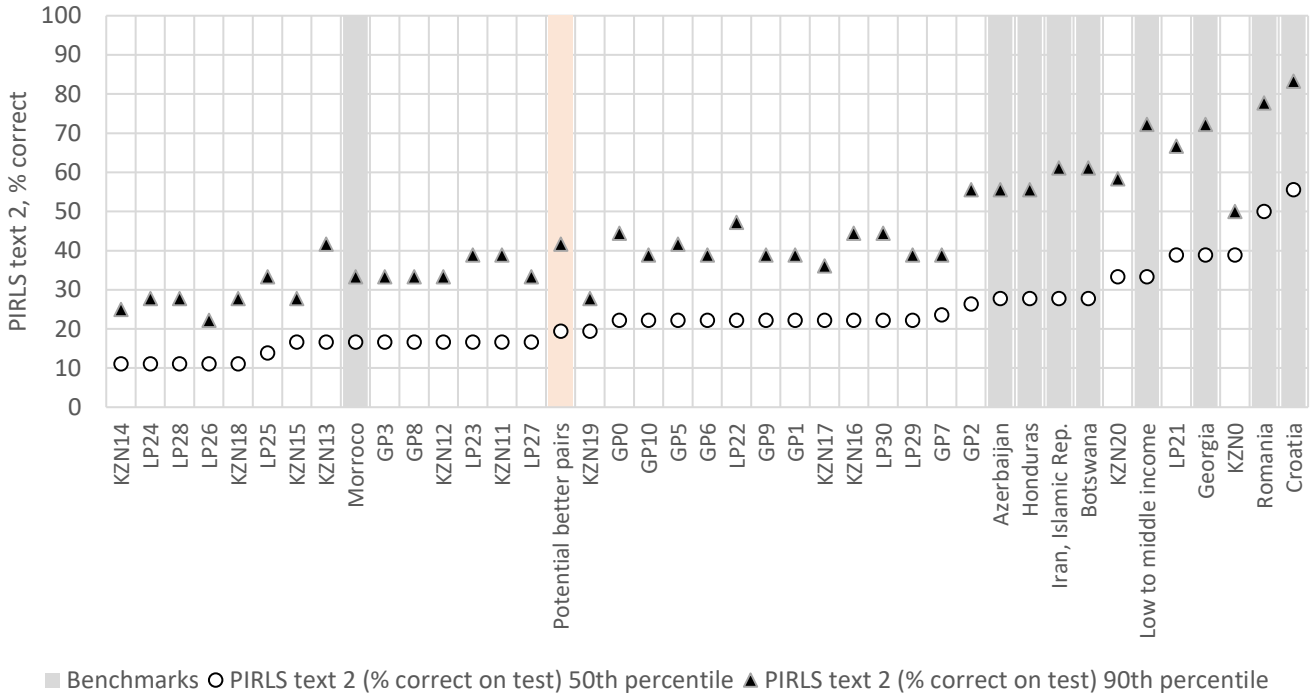
Figure 3: International comparison of potential better performing pairs on PIRLS text 1 (literacy experience), % correct on entire comprehension



Source: Own collected data, analysis of PIRLS 2011 international database. Except for Honduras and Botswana, all benchmarking countries write the test at the grade 4 level.

[9] Azerbaijan, Botswana, Croatia, Iran and Romania are classified as upper-middle income economies along with South Africa. Honduras, Georgia and Morocco are classified as lower-middle income economics using World Bank classifications.

Figure 4: An international comparison of potential better pairs on PIRLS text 2 (acquire and use information), % correct on entire comprehension



Source: Own collected data, analysis of PIRLS 2011 international database. Except for Honduras and Botswana, all benchmarking countries write the test at the grade 4 level.

b) 61 school sample residual analysis

While we may not have found any township and rural schools with unusual performance in English literacy by select middle-income country standards, it is evident from Figure 3 and Figure 4 that there exists some variation across the school sample, especially on the PIRLS easier text. Two low-fee paying schools outperform 29 other potential better performing pairs; and performance at the 90th percentile reflects that even in schools with weak median performance there may be some more resilient students. However, these schools and students may be performing better simply due to their background differences, such as exposure to English, Early Childhood development, parental involvement and access to a literacy rich environment. This section considers whether there any schools that are statistically unique in the 61 school sample, even after controlling for students' background factors.

Estimation strategy

Following the work of Klitgaard and Hall (1975), I identify more effective schools in our sample using a residual analysis. Rather than focusing on the properties of the regression that may explain the performance of schools, the aim here is to avoid overcontrolling and explaining away school effects. Using a standard production function regression framework, I control only for socioeconomic and other non-school background factors and assume that remaining variation could be attributable to the school. Of course, one can't rule out that the variation remaining could also be the result of factors such as measurement errors, omitted variables or other random noise. Nevertheless, if at least some unusual residual performance remains after controlling for background factors then it provides some possibility that there may be some schools in the sample that are having a more positive impact on

students. To estimate the residuals I use the following regression framework where Y reflects the performance of the individual student in school s, measured in total marks on the combined PIRLS texts and a vocabulary test. X is a vector of a student's background characteristics but school characteristics are intentionally excluded from the model with effects rather captured in e, the residual.

$$Y_{is} = \beta_1 + \beta_2 X_{is} + e$$

To test the sensitivity of the residuals to the inclusion of controls, I estimate six different models which sequentially include more background controls. The most parsimonious model only includes a composite index of student's socio-economic status (SES) derived through a principal components analysis of asset ownership indicators and a square of this index to allow for non-linearity in the relationship between learning outcomes and SES. However, the asset index appears to be truncated at the upper end and possibly not effective enough at capturing higher levels of student wealth. For this reason, additional wealth controls are necessary. Model 2 includes controls for whether the mother's child speaks the language of the test at home (in this case English), and whether the child attended grade R or crèche before grade 1. Model 3 then includes indicators for whether the child lives with their mother and father and access to literacy material at home including indicators for number of books in the home and whether the child has any story books. The age and gender of the student, as well as an indicator to capture the student's attitudes towards reading are included in model 4. At the start of the testing we presented students with some simple questions to ascertain how they felt about school using emoticons. Our indicator for reading attitude takes on a value of 1 if the student circled a 'very happy' emoticon face when asked "how do you feel about reading?" (reference category includes very unhappy, unhappy, happy). Model 5 adds an index of the wealth of the small place area in which the school is situated and its squared derived from Census 2011[10] as well as an indicator for the rural status of the school which also likely proxies for the rural/urban status of the student's home. Finally, a class average index of student's individual asset indices is included and its square in model 6.

Results

Moving across each model's results in Table 5 it is evident that an asset index of student SES is not sufficient to account for student background. Although residency with parents and access to literacy material at home does not appear to contribute much to the models, indicators for grade R or crèche attendance and attitude towards reading are significant and positive. Age[11] and gender of the student enter significantly with an advantage to girls evident in the estimations. Rural status is negative but not consistently significant while the small place area in which the school is located (and likely where students live) reflects a strong convex relationship with literacy. Despite the inclusion of a student asset index, the small place area Census index contributes significantly to the explained variation. It is important to note that student residuals in schools with the wealthiest student compositions are most sensitive to the inclusion of additional wealth controls (suggesting that the asset index is not good at capturing enough wealth variation at the top end).

[10] The author is grateful to Chris van Wyk and Asmus Zoch for the construction of this index.

[11] The inclusion of the age of the child is somewhat problematic because while this may be influenced by background factors (e.g decisions about when to send a child to school), grade for age could be influenced by school factors if poor teaching or assessments contribute to failure. Progression rules however limit the number of times children can fail in a school and therefore school effects are unlikely to strongly drive the presence of over age children.

Although Model 6 shows that mean student SES is significant and adds explanatory power to the estimation, in the spirit of avoiding overcontrolling and giving better schools every opportunity to exhibit by regarding the residual as a school effect, I use the residuals in model 5.

To assess whether students in a school are performing better than expected given their background factors, I plot standardized residuals (e) against unadjusted (standardized) total marks for each school's grade 6 performance, focusing on the 50th and 90th percentile. As shown in Figure X, literacy performance in two low-fee paying schools reflected in blue (LP21 and KZN0) continue to exceed the demographic expectations of the sample suggesting particularly positive school effects relative to the sample. At the 90th percentile students in these schools are performing 3-4 standard deviations above sample expectations after accounting for their background characteristics. Promisingly, there is also one non-fee paying school in KwaZulu-Natal which exhibits unusual performance at the 90th percentile after the background adjustments. Two non-fee paying schools in Gauteng, GP1 and GP2, also appear to be exceeding sample expectations at the 50th percentile with residual performance on a par with the low-fee paying Limpopo school (LP21).

The analysis also highlights how there are some seemingly underperforming schools which are achieving much better results than expected given the relatively poorer backgrounds of their students. Rural school LP22 is such an example. Without adjusting for student background, this school would be labelled as underperforming, potentially targeted with interventions and scrutinized by its district when in fact it is potentially more efficient than other schools given its student composition. Conversely some schools that may be praised for good performance are actually very inefficient as reflect in the residual analysis; for example low-fee paying schools GP9 and KZN16 in blue. In future national attempts to identify underperformance or exceptionalism, whether at the primary or secondary level (especially at matric level), appropriate adjustments must be made to account for student background differentials. Even in the presence of information on school performance one can make erroneous conclusions about school and teacher efficacy if performance is not juxtaposed against student background realities (which also span a continuum within the non-fee paying school sector).

Another significant finding from the residual analysis is the presence of more resilient students regardless of general school effectiveness. Even in very weak schools, there are some students performing at 3 or more standard deviations above sample expectations after adjusting for their background characteristics. This is highlighted in a boxplot of residuals from model 5 for all 61 schools. The boxplot of class performance shows the median, 25th and 75th percentiles, but what is of interest here are upper ends of the whiskers and the outlier values indicated by the scatted dots. While the focus of this project is exploring the characteristics and workings of more effective township and rural *schools*, there are students that appear to exceed expectations despite their home *and* their school environments. This is not just an anomalous feature of our data. A look at PrePIRLS data 2011 reveals that about 5-6% of South African grade 4 children writing African language tests reach high benchmarks in reading (550 points or above). Despite comprising a minority, they are found in over half of all classrooms or schools tested in African languages. There are literally one two high achievers in these classrooms.

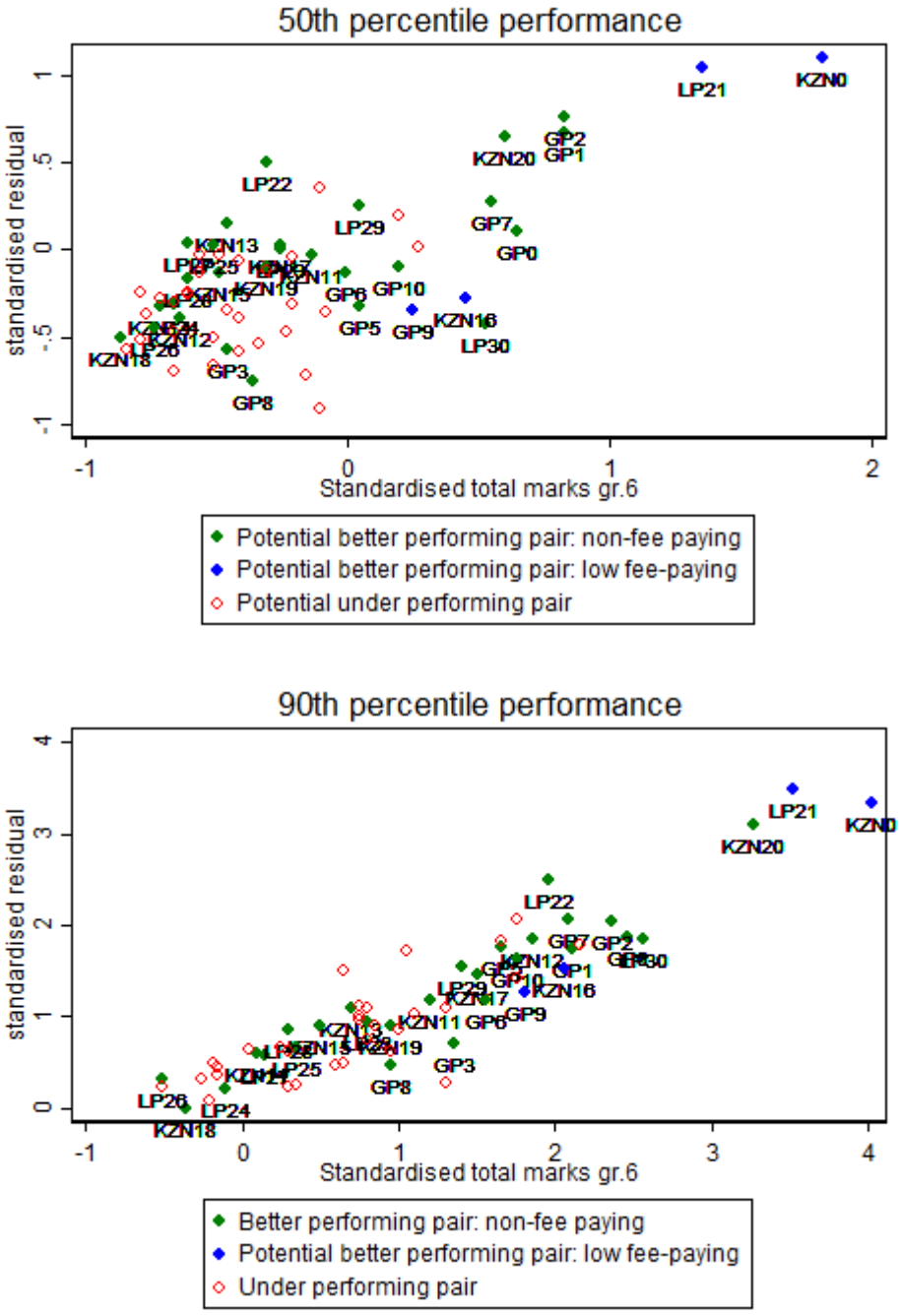
Future research would do well to better understand the processes and characteristics underlying their resilience in the South African context and how the potential of these students can be nurtured. As evidenced in recent case studies observing the teaching of reading and writing in foundation phase classrooms in the North West province, top performers are not being extended enough as teachers fail to differentiate instruction to suit the varied ability level of students in their classrooms (Reeves, 2017, pp. 56–58).

Table 5: Estimations of grade 6 total marks controlling for student background

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Student SES index	1.833***	1.797***	1.759***	1.714***	1.122***	0.494***
	(16.87)	(16.55)	(15.70)	(15.49)	(8.46)	(3.52)
Student SES index squared	0.164***	0.157***	0.153**	0.167***	0.0877	0.0506
	(3.50)	(3.33)	(3.26)	(3.64)	(1.87)	(1.03)
Mother speaks English		2.190	2.436	2.483	3.173*	1.855
		(1.52)	(1.72)	(1.88)	(2.30)	(1.41)
Mother's language is missing		-6.447***	-5.669***	-4.839***	-4.643***	-4.976***
		(-10.30)	(-8.57)	(-7.40)	(-7.08)	(-7.96)
Attended grade R or crèche		1.302***	1.377***	1.158***	1.123***	0.979**
		(4.05)	(4.27)	(3.64)	(3.61)	(3.12)
Lives with mother			0.585	0.472	-0.0389	-0.0975
			(1.06)	(0.87)	(-0.07)	(-0.19)
Lives with father			-1.463	-0.961	-1.201	-1.063
			(-1.59)	(-1.07)	(-1.37)	(-1.27)
Lives with mother X Lives with father			0.513	0.161	0.559	0.578
			(0.51)	(0.16)	(0.58)	(0.63)
Has a few books (10)			0.874*	0.878*	0.847*	0.397
			(1.99)	(2.04)	(2.04)	(0.97)
Has enough books to fill one shelf (20)			-0.0331	-0.00797	-0.0431	-0.430
			(-0.05)	(-0.01)	(-0.07)	(-0.69)
Has enough books to fill one bookcase (50)			2.405	2.416	2.353	2.495
			(1.63)	(1.69)	(1.66)	(1.79)
Has enough books to fill two or more bookcases (100)			-4.931**	-4.954**	-4.956**	-4.857*
			(-2.74)	(-2.70)	(-2.70)	(-2.16)
Books missing			-2.632***	-2.168***	-2.029***	-2.408***
			(-4.14)	(-3.43)	(-3.33)	(-3.99)
Number of own story books at home			0.0132	0.00197	0.00793	0.00377
			(0.47)	(0.07)	(0.29)	(0.14)
Age				-1.487***	-1.476***	-1.302***
				(-9.13)	(-9.31)	(-8.51)
Is a girl				2.146***	1.982***	1.926***
				(5.80)	(5.42)	(5.36)
Child feels very happy about reading				0.797*	1.000**	1.142**
				(2.19)	(2.82)	(3.28)
Census small area wealth index					-22.30***	-19.19**
					(-3.77)	(-3.04)
Census small area wealth index squared					34.88***	20.60**
					(4.96)	(2.60)
Rural					-1.505*	-0.426
					(-2.44)	(-0.68)
Average class SES						2.598***
						(8.98)
Average class SES squared						0.434***
						(3.55)
Constant	12.53***	11.63***	11.25***	27.16***	30.73***	29.55***
	(55.48)	(34.27)	(19.18)	(12.79)	(12.02)	(11.68)
Observations	2652	2652	2652	2652	2652	2652
Adjusted R-squared	0.114	0.135	0.147	0.190	0.223	0.254

Source: February 2017 data collection, Leadership for Literacy. **Notes:** Statistically significant * p<0.05, **p<0.01, ***p<0.001. Indicators for missing information are also included for mother's language, books at home, lives with mother and lives with father.

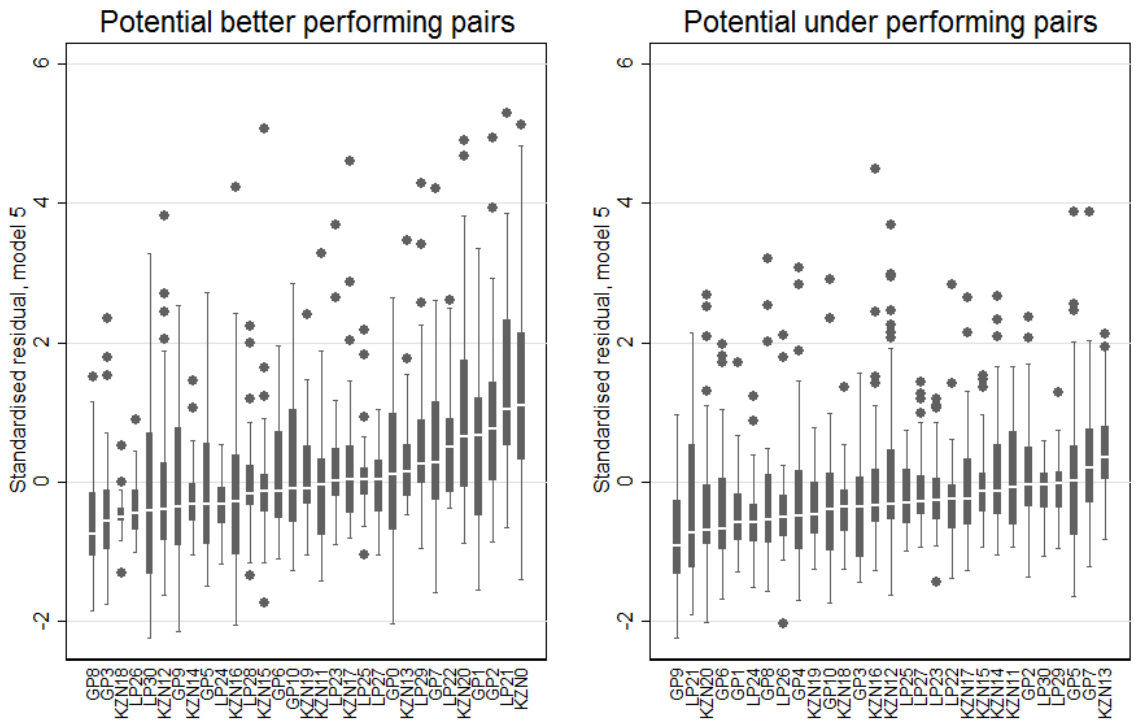
Figure 5: Performance of median grade 6 students in the class – adjusted versus unadjusted performance on grade 6 English literacy test



Notes: The standardized residual from a linear regression of total marks on student's background characteristics, model 5. This is the unexplained performance of the median student after accounting for their age and gender, socio-economic status (proxied by an index of asset ownership, small census place area wealth index and rural indicators), residential status with parents (whether child lives with their mother, lives with their father), grade R or crèche attendance and language support at home (mother speaks English at home, anyone reads to the child at home, number of books at home).

In summary, our international comparison of sample performance presents a sobering reality about English proficiency levels in what may be some of our best non-fee paying (and a few low-fee paying) schools in three provinces. However, a continuum of relative effectiveness still exists *within* our school sample. This leads to the next discussion. Even in a context of little to no flagship schools in the non-fee paying sector, are system actors still able to distinguish higher levels of quality?

Figure 6: Box plot of standardized residuals (from model 5, estimation of total grade 6 English literacy test marks) for potential better and under-performing school pairs.



V. Findings, part B: Were respondents (data source 2) able to identify better quality schools?

Drawing on data source 2, school recommendations from various system actors linked to ANA performance and national schools data, this section identifies whether respondents were able to identify better performing schools. The discussion draws on a combination of descriptive and multivariate analyses as explained below.

Descriptive results

A difference in means test is used to compare the average academic performance of quintile 1-3 schools in Gauteng, KwaZulu-Natal and Limpopo that were not mentioned by respondents (first column of Table 8) against those that were recommended (second column). For benchmarking reasons, average performance of quintile 5 schools is also shown (last column).

Estimates for multiple measures are provided as respondent's awareness of school performance may vary depending on the subject, grade or year in question.[12] The first four of 12 are composite ANA

[12] For example, it may be easier to determine how well a school is doing on language than mathematics if performance is deduced informally through conversations with students. But schools may actually perform

averages across grades 1-6 and all subjects expressed as z-scores but differing by year of testing. The rest are subject specific measures, expressed in percentage terms, for home language (HL), first additional language (FAL) and mathematics for years 2012 and 2014, focusing on grades 3 and 6.

In 9 of the 12 measures recommended quintile 1-3 schools are performing slightly better than no-mentioned schools across the three provinces. Recommended schools are on average better than schools that were not mentioned by 10-19% of a standard deviation across the composite z-score measures. But in absolute terms their average performance is still very low at 0.04 standard deviations in the first school composite measure. By comparison the national mean across *all* ANA participating schools (regardless of quintile) is set to 0. The 0.04 average is far from our benchmark for 'good' (Quintile 5 schools which are performing on average at 1 standard deviation above the national mean). Only the top 5% of the recommended non-fee paying schools in the sample are performing at this level using a three-year composite ANA measure.

The 2014 grade 6 first additional language (FAL) results indicate that Quintile 1-3 schools that were *not* mentioned in recommendations achieved an average of 41%. In recommended schools, this average was slightly higher at 45% but considerably far off from Quintile 5 average performance of 60%.

Estimation strategy

Although recommended schools are achieving slightly higher levels of learning than those not mentioned, it is possible that learning outcomes do not directly inform quality perceptions. Rather respondents perceptions of quality may be a function of other observable characteristics such as language differences, student compositional or resourcing differences, that are positively correlated with learning outcomes. Holding these factors constant, are recommended schools still better than non-mentioned schools? This is operationalized by running a multivariate linear regression as reflected in the following framework,

$$Y_s = \beta_1 + \beta_2 D_s + \beta_3 X_s + e$$

where Y_s is one of 12 measures of performance and additionally 3 enrolment growth measures which are often considered as indicators school functionality. D_s is an indicator that takes on a value of 1 if the school was recommended and 0 otherwise.

I capture resourcing differences across schools, X_s through the following controls:

- Indicators for official DBE quintile status.
- Pupil to teacher ratios.
- A wealth index (and its squared term) for the small place area in which the school is situated as identified from the most recent Census 2011.
- Indicators for the school's languages of learning and teaching. South Africa has 11 official languages. English is set as the reference category.

Respondents may also choose schools based on differences in bureaucratic administrative processes at the provincial and district level that may raise overall school performance. To account for these differences, I include provincial indicators in all regressions, results are shown in Table 9. District level

better in certain subjects, grades or years (evident in ANA patterns) - the result of teacher quality differences, compositional differences in student bodies or changes in other school factors. ANA data dissemination within districts or provinces may have also differed depending on subjects, phases or years in consideration. Little is known about how and to whom school performance on ANA was disseminated, nevertheless it could have also had differential signalling effects across years.

fixed effects are also used, running estimations separately with results reported in Table 10. Each table shows the estimated coefficient on the indicator of interest, D_s. The independent or outcome performance measure being estimated is defined in the first column of each row. R-squared values and the sample of all schools in the estimations are also shown.

Multivariate results

Taken together the coefficients for each of the estimations in part A of the table confirm the earlier descriptive findings that on average respondents are able to recommend slightly better performing schools. Coefficients are positive and significant for 11 of the 12 performance measures and the average difference in favour of recommended schools is only slightly attenuated when controlling for observable school characteristics. (Coefficients in the estimation are not too different from the difference in means calculated in Table 8 despite controls for resourcing differences and provincial administration.) For the first four composite school measures, recommended schools are about 8 to 10 % of a standard deviation better than non-mentioned schools. Of course, it is possible that recommendations may be informed by other unobserved characteristics that are positively correlated with learning outcomes such as leadership and management proficiencies. The estimation is limited in this regard.

I also find, but results are not reported here, that average school performance is increasing in the number of times the school was recommended, adding legitimacy to our search strategy for choice schools. The most revealing findings, however, occur when we disaggregate the variable of interest by the type of respondent who made the recommendation. In part B of the table, the single variable of interest D_s is replaced with six indicators,

- R_1= recommended by a district official,
- R_2 = recommended by a respondent in a good secondary school (high matric pass rate),
- R_3= recommended by a respondent in a bad to ok performing primary school,
- R_4 = recommended by a respondent in a good primary school,
- R_5 = recommended by an NGO working in education,
- R_6= recommended by other source

The reference category remains that the school is not mentioned by any respondent. Results are identified in Table 9 and Table 10 which differ in the inclusion of district fixed effects. For brevity, Table 7 summarises the number of measures for which significant results are identified across the 6 groups of respondents. At face value the results are surprising. Of the six groups in our sample NGO respondents, the group of 'other' respondents, secondary school respondents and district officials are the *least* effective at identifying better performing schools. Significant positive effects on district official recommendations are only identified in at most 3 of the 12 estimations. Respondents from primary schools (typically school principals or administrative clerks) and specifically from the worse performing primary schools in ANA were the best at identifying better performing schools. For 11 of 12 school performance estimations, schools recommended by respondents from poorer performing primary schools are statistically significantly better performers. The size of the coefficients is also notable on recommendations from poorer performing primary schools: 30-40% of a standard deviation on composite measures of performance (reflected in the estimations with and without district fixed effects) and 7-8% better on grade 6 First Additional Language.

This is somewhat puzzling. Why is it that the very group one would least expect to know about school performance provide better recommendations than district officials or NGOs, initially assumed to be the experts with more information on a set of schools' performance? This may be a peculiarity of the non-randomness of the respondent sample but this may also be accounted for by informational asymmetries. Some groups may have access to different types of information than others on school performance, especially when there has been little systematic approach taken to publicly disseminate

available ANA data. Respondents from primary schools may be making more informed recommendations because of their geographical proximity to other schools and localised knowledge about neighbouring schools.

By linking the geographic coordinates of respondents located in specific primary schools to the primary school they recommended, it was possible to calculate distances (as the crow flies) between 141 pairs. The median distance between respondents from primary schools and the school they recommended was as little as 1.53 kilometres and the distance at the 90th percentile was less than 9km. A cumulative distribution frequency from a nearest neighbour analysis provided in Figure 7 indicates that the median recommended school was about the 6th nearest school to a respondent's primary school position. It would be instructive to explore further on what basis school quality judgements are being made at these localised levels. This has implications for the design of future informational and transparency interventions including dissemination of any future national testing data such as ANA. As noted by Read and Atinc (2017, p. 12), one can't just impose interventions without a fuller consideration of contextual dynamics and better understanding "*which* local actors use *what* types of information to *what end*…"

Figure 7: Nearest neighbour analysis of recommended schools from a respondent in a primary school location

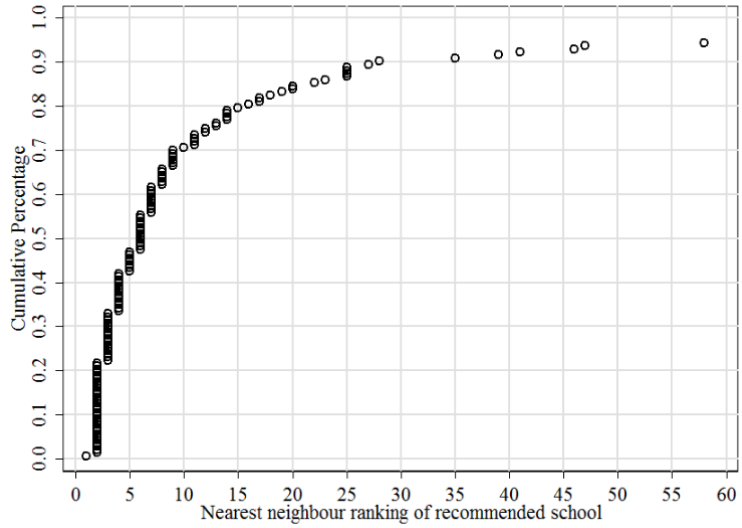


Table 6: Distance of respondent in a primary school location to the primary school they recommended.

	Distance of recommended school from recommending school (Kilometres)
P10	0.41
p50	1.53
p75	3.94
p90	8.67
N	**141**

On the one hand, it is surprising that the very respondents who should have more information about school performance, namely district officials, were the least effective at identifying better performing

non-fee paying primary schools. On the other hand, this is not unexpected if bureaucratic indicators of success redirect their attention to the wrong metrics. If too much value is placed on school compliance metrics with respect to fulfilling administrative duties as a gauge of school effectiveness, the wrong metrics for quality education service delivery may dominate district official perceptions of primary schools. Another possibility is that they have conflicting data on school performance within the same school which can be variable depending on the subject, grade or year in question. This would make it difficult to ascertain generally whether a school is doing well. What is interesting is that district officials are particularly good at recommending schools with higher enrolment growth (recommended schools had nearly 7% higher growth between 2012 and 2014 compared with non-mentioned schools). This is not surprising given they would be exposed to school enrolment data and the growth needs of schools. The opposite holds for our primary school respondents – they recommend schools with better performance but not higher enrolment.

For our sample of district officials' perceptions of success may also be based on parent demand for schools. This would only be an effective strategy of identifying better performing schools if enrolment growth patterns are in fact associated with higher learner performance. I simplistically test this by plotting school ANA performance measures against school enrolment growth percentages in Figure 8 and fitting a non-parametric local polynomial line. It reflects little systematic relationship between enrolment growth and school performance, assuming that ANA is a relatively reliable metric of learner performance. Enrolment growth is likely a very faulty metric for quality education service delivery. The picture as a whole implies that parents are *not* voting with their feet on the basis of ANA data, although one can't rule out that a spatial analysis which considers more localised decisions between competing schools may indicate otherwise. This would be an interesting and useful future research endeavour.

Table 7: Who is best among the set of respondents at identifying better performing schools? Number of measures in which recommended schools perform better than non-mentioned schools, by group of respondents

	Significant, positive coefficient in X out of 12 school performance estimates		Significant, positive coefficient in X out of 3 school enrolment estimates		
	Without district FE	**With district FE**	**Without district FE**	**With district FE**	**N recommend-ations**
District officials	1	3	2	3	130
A good secondary school	0	1	1	0	77
A bad to ok primary school	11	11	0	0	73
A good primary school	8	8	0	1	50
NGO respondents	0	2	2	2	46
Other	0	0	1	1	25

Notes: Summary of significant results for Table 8 and Table 9.

Table 8: Average performance of Quintile 1-3 schools that are recommended, not mentioned by respondents and Quintile 5 schools.

	Average performance of ANA schools in Gauteng, Limpopo and KwaZulu-Natal															
	Quintile 1-3 schools											All Quintile 5 schools				
	Schools not mentioned by any respondent					Recommended schools										
	Mean	Std. dev.	N	P90	P95	Mean	Std. dev.	N	P90	P95	diff in mean	Mean	Std. dev.	N	P90	P95
Ave. school perf. (z score)	-0.09	0.60	6094	0.72	0.94	0.04	0.59	401	0.82	1.04	0.13**	1.00	0.61	728	1.78	1.90
Ave. school perf. 2012 (z score)	-0.14	0.69	6094	0.79	1.08	-0.03	0.64	401	0.83	1.06	0.11**	1.02	0.68	728	1.88	2.06
Ave. school perf. 2013 (z score)	-0.05	0.69	6094	0.84	1.11	0.04	0.67	401	0.83	1.07	0.10*	1.01	0.64	728	1.81	1.94
Ave. school perf. 2014 (z score)	-0.07	0.70	6091	0.82	1.06	0.12	0.68	401	1.00	1.20	0.19**	0.94	0.74	728	1.76	1.87
Home Language Gr. 3 2012 (%)	52.32	13.41	5286	69.51	74.00	52.53	11.90	357	68.84	72.00	0.21	61.79	11.39	677	75.82	78.29
Home Language Gr. 3 2014 (%)	55.46	14.45	5733	72.88	77.39	56.94	13.10	385	71.85	76.51	1.48*	67.24	13.79	707	82.19	85.00
FAL Gr.6 2012 (%)	34.64	12.29	4375	51.55	57.21	39.62	12.12	276	54.76	59.31	4.98**	62.05	14.34	142	79.60	83.22
FAL Gr.6 2014 (%)	40.97	12.36	5305	56.72	61.36	45.38	13.12	320	60.93	67.28	4.41**	59.68	14.51	106	76.40	78.02
Math Gr 3. 2012 (%)	40.05	15.16	5360	60.00	66.73	41.11	14.57	367	61.88	67.69	1.06	53.19	13.44	676	69.85	74.11
Math Gr.3 2014 (%)	53.58	15.01	5758	72.17	78.06	54.65	13.22	385	70.74	75.38	1.07	67.93	12.79	708	81.63	84.12
Math Gr 6. 2012 (%)	25.53	12.24	5044	43.22	49.45	26.68	11.39	352	43.35	48.26	1.15*	40.20	12.50	632	57.03	59.97
Math Gr 6. 2014 (%)	39.26	13.73	5571	57.47	62.37	42.28	13.02	366	60.55	62.70	3.02**	58.74	13.50	702	74.43	77.23

Notes: FAL stands for First Additional Language. * means are significantly different at the 90% level, **means are significantly different at the 95% level. Performance over time is not directly comparable in ANA.

Table 9: Multivariate estimations (excluding district fixed effects). Recommendations and performance of schools in Gauteng, KwaZulu-Natal and Limpopo.

	Key dependent variable								
	Coefficient on D_s: recommended (1), not mentioned (0)		**Coefficient on R_1 - R_6: Recommendation source. Reference category "not mentioned"**						
		R2 / *N*	**district officials**	**Secondary feeder**	**bad to ok primary**	**good primary**	**NGO respondents**	**Other**	**R2 / *N***
Performance outcomes:									
Composite performance 2012-2014 year [z-score]	0.10***	0.39	0.01	0.01	0.31***	0.20***	0.12	0.04	0.39
	(0.03)	*7347*	(0.04)	(0.06)	(0.06)	(0.07)	(0.08)	(0.09)	7347
Composite performance 2012 [z-score]	0.10***	0.31	0.04	0.01	0.34***	0.22**	0.05	0	0.31
	(0.03)	*7347*	(0.05)	(0.07)	(0.07)	(0.09)	(0.09)	(0.11)	7347
Composite performance 2014 [z-score]	0.12***	0.34	0.03	0.05	0.31***	0.23***	0.15	0.09	0.34
	(0.03)	*7346*	(0.05)	(0.07)	(0.07)	(0.09)	(0.09)	(0.11)	7346
Gr. 3 HL 2012 (%)	1.21*	0.13	0.78	0.02	4.51***	1.65	0.24	-1.02	0.13
	(0.65)	*6692*	(1.11)	(1.38)	(1.46)	(1.80)	(1.86)	(2.42)	6692
Gr. 3 math. 2012 (%)	0.67	0.15	-0.55	-0.61	5.67***	-0.08	0.9	-2.25	0.15
	(0.73)	*6774*	(1.25)	(1.55)	(1.62)	(2.10)	(2.14)	(2.62)	6774
Gr. 6 FAL 2012 (%)	3.02***	0.22	0.93	2.66	7.15***	4.79**	1.48	1.61	0.22
	(0.74)	*4698*	(1.24)	(1.70)	(1.55)	(2.01)	(2.05)	(2.98)	4698
Gr. 6 math. 2014 (%)	1.89***	0.21	-0.16	0.13	6.36***	4.78***	1.24	1.14	0.21
	(0.63)	*6082*	(1.08)	(1.29)	(1.38)	(1.76)	(1.82)	(2.45)	6082
Gr. 3 HL 2014 (%)	1.75***	0.21	1.91*	0.59	2.80*	2.44	0.11	2.98	0.21
	(0.65)	*7234*	(1.10)	(1.39)	(1.44)	(1.82)	(1.89)	(2.31)	7234
Gr. 3 math. 2014 (%)	0.67	0.28	0.19	-0.1	2.24	-0.32	1.08	2.03	0.28
	(0.64)	*7261*	(1.09)	(1.37)	(1.42)	(1.79)	(1.87)	(2.28)	7261
Gr. 6 FAL 2014 (%)	1.57**	0.18	0.18	-0.3	5.99***	3.48*	1.62	-0.81	0.18
	(0.68)	*5596*	(1.09)	(1.58)	(1.49)	(2.06)	(1.86)	(2.45)	5596
Gr. 6 math. 2014 (%)	1.59**	0.3	-0.94	1.18	4.18***	5.44***	1.71	1.42	0.3
	(0.65)	*6715*	(1.11)	(1.35)	(1.43)	(1.83)	(1.92)	(2.37)	6715

		R2 / *N*	district officials	Secondary feeder	bad to ok primary	good primary	NGO respondents	Other	R2 / *N*
Independent variable: Enrolment growth									
Enrolment growth 2012-2016 (%)	3.63***	0.1	4.28**	1.05	1.29	1.37	11.36***	6.27	0.1
	(1.13)	*7214*	(1.92)	(2.41)	(2.51)	(3.19)	(3.33)	(4.06)	7214
Enrolment growth 2012-2014 (%)	1.18	0.05	1.54	-0.95	0.35	-2.26	6.89***	5.24*	0.05
	(0.81)	*7319*	(1.36)	(1.71)	(1.79)	(2.27)	(2.42)	(2.89)	7319
Enrolment growth 2014-2016 (%)	3.04***	0.06	2.76*	6.00***	1.2	2.69	3.22	0.78	0.06
	(0.84)	*7189*	(1.41)	(1.77)	(1.85)	(2.37)	(2.50)	(2.99)	7189

Data source: Collected school recommendations dataset. **Notes:** Controls for quintile status, wealth index derived from 2011 Census small area places and its square, province, pupil to teacher ratios (2012), indicators for LOLT. Significant at *10% level, **5% level, ***1% level

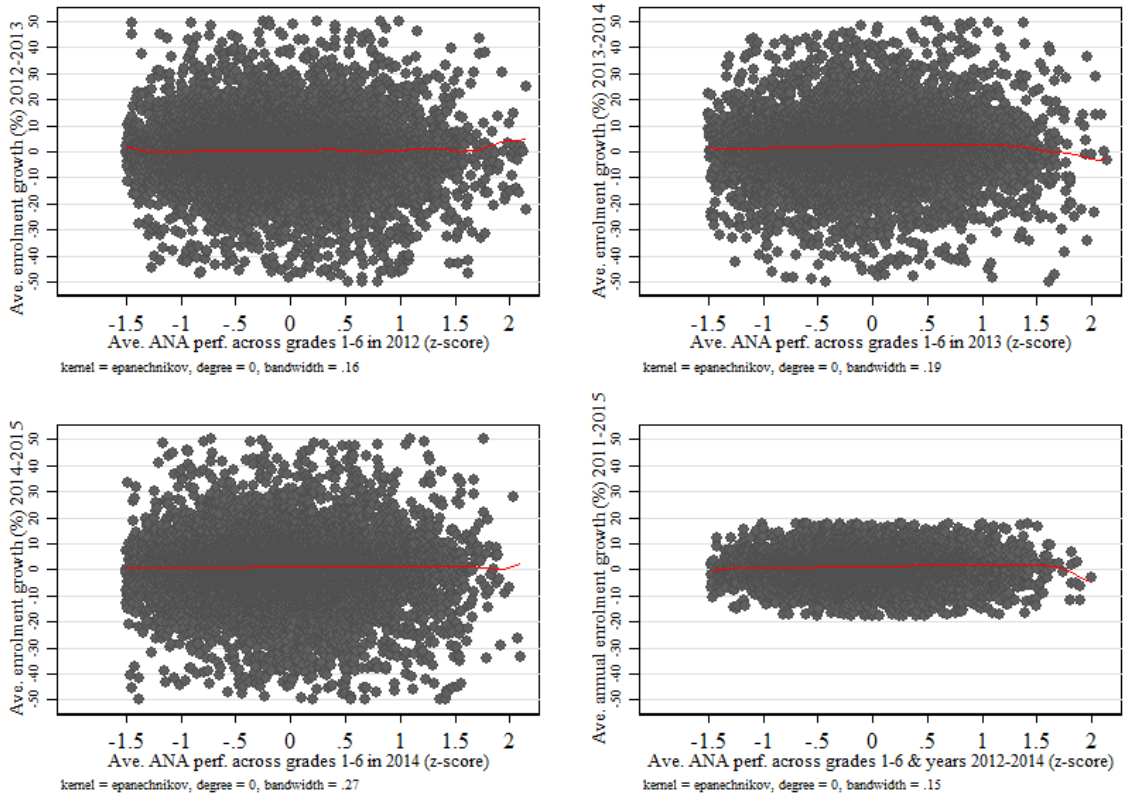
Table 10: Multivariate estimations (including district fixed effects). Recommendations and performance of schools in Gauteng, KwaZulu-Natal and Limpopo.

	Coefficient on D_S: Recommended (1), not mentioned (0)	**Coefficient on R_1 - R_6: Recommendation source. Reference category "not mentioned"**							
		District official	**Secondary Feeder**	**Bad to OK primary**	**Good Primary**	**NGO respondents**	**Other**	**R2-Within**	**N-groups / N**
Performance outcomes:									
Composite performance 2012-2014 (z-score)	0.08***	0.05	0.02	0.36***	0.20***	0.13**	0.04	0.05	37
	(0.02)	(0.04)	(0.05)	(0.06)	(0.08)	(0.07)	(0.11)		*7008*
Composite performance 2012 (z-score)	0.08***	0.09*	0.03	0.38***	0.20**	0.07	-0.01	0.06	37
	(0.03)	(0.05)	(0.06)	(0.07)	(0.08)	(0.07)	(0.12)		*7008*
Composite performance 2014 (z-score)	0.10***	0.06	0.06	0.35***	0.24***	0.17**	0.09	0.03	37
	(0.03)	(0.05)	(0.06)	(0.07)	(0.09)	(0.08)	(0.10)		*7007*
HL Gr. 3 2012 (%)	0.85	1.81*	0.55	5.63***	1.49	0.63	-0.58	0.03	37
	(0.72)	(1.04)	(1.39)	(1.29)	(1.42)	(1.82)	(2.50)		*6162*
Math Gr. 3 2012 (%)	0.2	0.28	0.5	6.73***	-1.07	1.24	-1.81	0.03	37
	(0.53)	(1.20)	(1.67)	(1.70)	(1.85)	(1.84)	(2.99)		*6254*
FAL Gr. 6 2012 (%)	2.66***	1.85	2.73*	7.90***	5.16***	1.97	1.74	0.05	37
	(0.73)	(1.23)	(1.49)	(1.51)	(1.90)	(1.24)	(2.97)		*4824*
Math Gr. 6 2012 (%)	1.50**	0.43	0.7	7.38***	4.33**	1.83	1.37	0.04	37
	(0.59)	(0.88)	(1.09)	(1.58)	(1.99)	(1.58)	(2.51)		*5777*
HL Gr. 3 2014 (%)	1.64**	2.70**	1.05	4.43***	3.01*	0.63	3.42	0.02	37
	(0.66)	(1.07)	(1.38)	(1.55)	(1.72)	(1.83)	(2.15)		*6696*
Math Gr. 3 2014 (%)	0.54	0.85	0.22	3.41**	0.13	1.53	2.68	0.02	37
	(0.66)	(1.06)	(1.38)	(1.49)	(1.87)	(1.64)	(2.34)		*6722*
FAL Gr. 6 2014 (%)	1.67**	1.12	0.68	6.96***	4.06**	2.37	0.31	0.03	37
	(0.80)	(1.03)	(1.47)	(1.57)	(1.62)	(1.94)	(2.34)		*5828*
Math Gr. 6 2014 (%)	0.7	-0.74	0.84	4.64***	5.20***	1.78	1.15	0.03	37
	(0.47)	(0.94)	(1.22)	(1.49)	(1.70)	(1.56)	(2.22)		*6383*

	Coefficient on D_s: Recommended (1), not mentioned (0)	District official	Secondary Feeder	Bad to OK primary	Good Primary	NGO respondents	Other	R2-Within	N-groups / N
Independent variable: Enrolment growth									
Enrol. Growth. 12-16 (%)	4.72***	6.60***	0.37	1.45	1.71	11.30*	3.69	0.02	37
	(1.19)	(2.17)	(2.30)	(1.56)	(1.83)	(5.97)	(3.34)		*6861*
Enrol. Growth 12-14 (%)	1.69*	1.81*	-0.11	0.62	-2.13	6.86	3.75*	0.02	37
	(0.84)	(0.98)	(1.33)	(1.05)	(2.13)	(4.69)	(2.25)		*6982*
Enrol. Growth 14-16 (%)	2.37**	4.56**	-0.18	0.9	2.43*	2.95**	-0.61	0.01	37
	(0.92)	(1.81)	(1.44)	(1.00)	(1.31)	(1.39)	(1.87)		*6834*

Data source: Collected school recommendations dataset. **Notes:** Controls for quintile status, wealth index derived from 2011 Census small area places and its square, province, pupil to teacher ratios (2012), indicators for LOLT. Significant at *10% level, **5% level, ***1% level

Figure 8: School enrolment growth (%) plotted against standardized measures of composite school performance in ANA



Notes: The red line is a non-parametric fitted local polynomial regression line. The scatter plots are truncated at extremely low and extremely high measures of ANA performance to limit the influence of outliers on the non-parametric line.

In summary, this analysis conducted among a sample of educators and other respondent groups who are likely to be more aware about school performance than the average citizen, indicates that some groups are able to identify slightly better non-fee paying schools than the average performing school. However for certain groups in our sample, specifically education district officials, enrolment growth appears to be a better indicator of their preferred school than measures of student performance. But at face value there appears to be little systematic evidence to indicate that school performance is associated with higher enrolment growth.

It is important to qualify, however, that average performance of schools recommended by our respondents is still relatively far from our benchmark for good, average performance of Quintile 5 schools. This is possibly due to two factors. First, there are simply not enough good schools as evidenced in our search in part A and perhaps we can't expect respondents' choice (at least among those groups whose recommendations are more closely aligned with school performance) to be any better in such an environment. In this respect, we can't disentangle how supply side constraints influence respondent's ability to detect quality service delivery. Quality predictions would possibly be better if there was more variation in performance amongst non-fee paying schools and there was a larger pool of quality non-fee paying schools from which to make a selection. Without these factors system actors and citizens in general have a limited frame of reference against which to evaluate what

a good school may be. This is exacerbated in a context where there is little reliable information disseminated on school performance.[13]

VI. Discussion

At this point it is useful to mention a key qualitative finding that emerged in collecting the recommendations data of good schools. When posing the question to respondents "Can you recommend a good township or rural school?" we were often met with this surprising response: "What do you mean by 'good'?" The extent to which we heard this phrase warranted quantification but we had not anticipated this would occur so frequently to anticipate this at the onset as a 'variable' worthy of collection.

The question "What do you mean by 'good'?" was followed with clarifications from the respondent such as "Do you mean resources, facilities, the principal or academics?" In these cases we then specified that we wanted recommendations of schools where children were learning and reading well. In wealthier contexts, if one were to ask a person to recommend a good school this request is unlikely to be met with this clarification question as one assumes that excellence is institutionalized, reflected in a gambit of factors which comprise a functioning learning environment such as good teaching, adequate resourcing, academic proficiency among students and management proficiency. It seems that when identifying good township and rural schools, it is not clear one can conceptualize excellence in the same way as in wealthier schooling contexts. These schools may contain aspects of excellence which may be tied to specific teachers, school leaders and a point in time. It was often the case that respondents would refer to a school as *once* being good, at a certain juncture, but they were no longer certain about the quality of the school as there had been a change in leadership. Where excellence is tied to individuals rather than embedded within the organization, achieving quality schooling becomes a fleeting notion and it becomes difficult to envision sustained gains in quality.

This question "what do you mean by good?" also highlights a fundamental constraint to developing excellence in South African schooling: inadequate reference frames for "good" schooling from which citizens or even those within the system can assess the quality of a school, compare the quality of one against the other, or establish whether a set of teaching practices are more effective than another. This is not surprising in the absence of best practice or even adequate non-fee paying schools. But inadequate reference frames are also likely attributable to apartheid legacies. Spatial controls, and racially desegregated education environments prevented exposure and sharing across lesser and more privileged school environments. These spatial legacies are also enduring, at least among teachers and school leaders which exhibit a low level of mobility (Gustafsson, 2016; Wills, 2015). Teacher moves that are made are typically to nearby schools, further limiting their experience of new teaching environments and seeing other examples.

In the absence of best practice schools in poorer contexts and the lack of teacher exposure to more functional environments, teacher education and training must start filling a gap in this regard – disrupting existing frameworks for 'good' with clearer examples and practices of what this may look like. But we also need more information on primary school performance to establish standards and challenge the public's frame of reference on quality education. It is possible that quality choices are

[13] There was no implemented strategy to make ANA performance results publicly available although performance information was fed back to districts and schools. Teachers also marked their own tests. That this information spilled into wider communities however is not clear. Simple correlations indicate that parents have not responded by voting with their feet, i.e. have not shifted school choice behaviour.

made even in the absence of actual test information – this is observed in the ability of some system actors to identify better quality schools. However as Gomez, Chumacero, Paredes (2012) identify in the context of Chile, while parents were making choices aligned with better quality schools even in the absence of publicly disseminated data on school performance quality became an increasingly important determinant of school choice when results were made public.

While it is not clear how and in what ways information could shift dynamics and perceptions within the system, we can expect that it will. Information and its media coverage can shift public perceptions of quality as suggested in recent changes in public satisfaction with education service delivery. Following a year that included a number of high-profile university based protests of tuition fees, education has become more of a major issue for the South African public. Pew Research conducting perceptions research among a representative sample of South African adults identify that the percentage naming poor-quality schools as a very big problem in 2016 increased 13 percentage points from 2015 (Pew Research Center, 2016). Information also shifts resourcing within the sector. For example, with high stakes matric testing there is some suggestions that more district human resources are devoted to supporting teachers in higher grades than lower grades (Wills, 2016) and there is evidence that matric results may inform teacher distributions as teachers tend to make moves to better performing matric institutions (Gustafsson, 2016).

VII. Conclusion

This paper set out to answer two key research questions:

1) *Are there available primary schools of choice in the non-fee paying public school sector*?
2) *To what extent are perceptions of quality in this sector aligned with actual learning outcomes?*

Despite a rigorous search for better non-fee paying schools in three provinces, relying largely on corroborating the shorted lived U-ANA system-wide testing of learner performance in primary schools with collected recommendations on better schools, we struggled to find schools that stand out as best practice examples after testing literacy performance in 31 purposively selected schools. Arguably the international benchmarking process used a very conservative approach, juxtaposing our student sample to students in lower to middle-income country samples.

It is evident that we should revise our *a priori* assumptions of excellence within the non-fee paying system. But how does one reconcile a lack of existing best practice non-fee primary schools with evidential improvements in the system, the media coverage of excellent matric students from poorer backgrounds and under-resourced schools that achieve excellent matric results? Dropout and merit based entrance into some secondary schools may explain some of this (Hunter, 2015). But answers may also lie in variation that exists in the performance within the system, however small it may be, as well as the presence of more resilient students despite their background or even the schools they attend.

Within our sample of potential better schools there is evidence of a continuum of school effectiveness. After discounting for student background factors which may underlie school performance differentials, two to three of our 31 potential better schools stand out from others. This suggests that striving for higher levels of quality is possible within the non-fee paying system. These levels may be far from good (or perhaps even adequate) but there is a middle ground, a rightward movement away from dysfunction that can be reached. However, these schools are rare as evidenced in the work by Kotze (2017) and confirmed here, implying geographic deserts of access to adequate primary education for the poor.

Encouragingly, however, even in highly underperforming school contexts there appear to be some more resilient students who significantly outperform their peers. There may be significant gains to be had from recognizing, supporting and protecting their potential. More research is required to understand what factors contribute to their success and how to further realise their potential, particularly in large class contexts where differentiated instruction suited to the ability levels of students may be challenging to implement.

Finally, in the absence of truly best practice non-fee paying schools but the possibility that a continuum of effectiveness exists, collecting and disseminating reliable system-wide information could go some way to fostering the following:

i) enabling citizens to make more informed choices in sending their children to incrementally better institutions. Even if the poor can't access a school that is adequate, there may be gains to choosing a school slightly better than the next. As identified by Kotze (2017), quality effects are not just limited to former white schools. There are learning gains to be realised from attending better schools even within the non-fee paying system. Of course, it is acknowledged that informational campaigns have had mixed results across the world in promoting better school choice (Read and Atinc, 2017), but in South Africa we currently have little understanding about how existing localised information and imposed testing information could shift school choices. This is a topic worth our attention. The country is grappling with the future of implementing standardised testing in the absence of contextual information about if and how this could shift public perceptions of quality and behaviour.

ii) Providing additional information for teachers, principals, education authorities and other education stakeholders such as unions to recognize something better, especially when the variation may be hard to detect and where effectiveness is probably better conceived of as existing in pockets within a school not something that is institution wide. Nevertheless, even in an underdeveloped data environment for school performance, the analysis in part B indicates that some system actors are able to detect better quality to some extent, even after controlling for observed differences across schools that may be driving perceptions of quality. The analysis of distance between schools recommended and locations from which the recommendation was made, suggests the use of localised information to make quality judgements. However, in other cases faulty metrics such as enrolment growth are more likely to guide perceptions of quality, at least among our sample of district officials.

Clear dissemination of primary performance data may go somewhere to rectifying underdeveloped frames of reference (and realigning perceptions with actual learning quality) even in the absence of choice or best practice schools. Without the right information, it is simply idealist to assume that we can promote public accountability for education improvements or realise any benefits of school choice.

VIII. References

Boone, P., 2013. The surprisingly dire situation of children's education in rural West Africa: Results from the Creo Study in Guinea-Bissau (No. 18971), NBER Working Paper Series. Cambridge, MA.

Bruns, B., Filmer, D., Patrinos, H.A., 2011. Making Schools Work: New Evidence on Accountability Reforms. The World Bank, Washington, D.C.

Burger, R.P., Teal, F.J., 2016. Convexity or Heterogeneity? Estimates of the Shape of the Earnings Profile, in: Unpublished Paper Presented at ReSEP Workshop, 21 November 2016.

Carneiro, P., Das, J., Reis, H., 2013. Parental Valuation of School Attributes in Developing Countries: Evidence from Pakistan, Working Paper.

Christie, P., Butler, D., Potterton, M., 2007. Ministerial Committee on Schools that work. Report to the Minister of Education. Pretoria.

Day Ashley, L., Mcloughlin, C., Aslam, M., Engel, J., Wales, J., Rawal, S., Batley, Kingdon, G., Nicolai, S., Rose, P., 2014. The role and impact of private schools in developing countries: a rigorous review of the evidence.

De Kadt, J., Norris, S.A., Fleisch, B., Richter, L., Alvanides, S., 2014. Children's daily travel to school in Johannesburg- Soweto, South Africa: geography and school choice in the Birth to Twenty cohort study. Child. Geogr. 12, 170–188. doi:10.1080/14733285.2013.812304

Department of Basic Education (DBE), 2017. Minister Angie Motshekga hosts round table discussions on the NEEDU National Report on Schools That Work, 10 July 2017 [WWW Document]. URL http://www.education.gov.za/Newsroom/MediaRelease/tabid/347/ctl/Details/mid/5986/ItemID/4418/Default.aspx (accessed 7.21.17).

Department of Basic Education (DBE), 2016a. 2016 School Realities. Pretoria.

Department of Basic Education (DBE), 2016b. Report on progress in the schooling sector against key learner performance and attainment indicators. Report prepared by Gustafsson, M. Pretoria.

Fiske, E.B., Ladd, H.F., 2004. Elusive equity: education reform in post apartheid South Africa. Brookings Institution Press, Washington, D.C.

Glick, P., Sahn, D.E., 2006. The demand for primary schooling in Madagascar: Price , quality , and the choice between public and private providers. J. Dev. Econ. 79, 118–145. doi:10.1016/j.jdeveco.2005.01.001

Gomez, D., Chumacero, Romulo, A., Paredes, R.D., 2012. School choice and information. Estud. Econ. 39, 143–157.

Gustafsson, M., 2016. Teacher supply and the quality of schooling in South Africa Patterns over space and time (No. 03/16), Stellenbosch Economic Working Papers.

Hastings, J.S., Kane, T.J., Staiger, D.O., 2009. Heterogeneous Preferences and the Efficacy of Public School Choice, National Bureau of Economic Research Working Papers. Cambridge, MA.

Hoadley, U., Galant, J., 2015. The organisation of schools succeeding against the odds. South African Rev. Educ. 12, 29–52.

Hunter, M., 2015. International Journal of Educational Development Schooling choice in South Africa : The limits of qualifications and the politics of race , class and symbolic power. Int. J. Educ. Dev. 43, 41–50. doi:10.1016/j.ijedudev.2015.04.004

Jansen, J., Blank, M., 2014. How to fix South Africa's schools: Lessons from schools that work. Bookstorm.

Klitgaard, R.E., Hall, G.R., 1975. Are There Unusually Effective Schools? J. Hum. Resour. 10, 90–106. doi:10.2307/145121

Klitgaard, R.E., Hall, G.R., 1973. A Statistical Search for Unusually Effective Schools. Prepared for the Carnegie Corporation of New York. Santa Monica, California.

Kotze, J., 2017. Social Gradients, Early Childhood Education and Schools Performing Above the Demographic Expectation: Empirical Insights into Educational Issues. Stellenbosch University.

Lepine, A., 2015. School Reputation and School Choice in Brazil: a Regression Discontinuity Design (No. 2015–38), Working Paper Series.

Levy, B., Shumane, L., 2016. School governance is a fragmented political and bureaucratic environment: Case studies from South Africa's Eastern Cape Province (No. 84), ESID. Manchester, UK.

Lieberman, E.S., Posner, D.N., Tsai, L.L., 2014. Does information lead to more active citizenship? Evidence from an education intervention in rural Kenya. World Dev. 60, 69–83. doi:10.1016/j.worlddev.2014.03.014

McGee, R., Gaventa, J., 2010. Synthesis Synthesis report: Review of impact and effectiveness of transparency and accountability initiatives. London.

Moses, E., Van Der Berg, S., Rich, K., 2017. A society divided: How unequal education quality limits social mobility in South Africa. Synthesis report for the Programme to Support Pro-Poor Policy Development (PSPPD). Stellenbosch.

Msila, V., 2005. The education exodus: The flight from township schools. Africa Educ. Rev. 2, 173–188. doi:10.1080/18146620508566299

Pew Research Center, 2016. In Key African Nations, Widespread Discontent With Economy, Corruption.

Read, L., Atinc, T.M., 2017. Information for accountability: Transparency and citizen engagement for improved service delivery in education systems.

Reddy, V., Visser, M., Winnar, L., Arends, F., Prinsloo, C., Isdale, K., 2016. TIMSS 2015 Highlights of Mathematics and Science Achievement of Grade 9 South African Learners. Pretoria.

Reddy, V., Zuze, T., Visser, M., Winnaar, L., Juan, A., Prinsloo, C., Arends, F., Rogers, S., 2015. Beyond Benchmarks. What twenty years of TIMSS data tell us about South African education. Human Sciences Research Council, Pretoria.

Reeves, C., 2017. The Early Grade Reading Study (EGRS): In depth Case Studies of Home Language Literacy Practices in Four Grade 2 Classrooms in Treatment 1 and 2 Schools (October 2016-January 2017). Pretoria.

Republic of South Africa, 2000. Education laws Amendment Act, 2000 (Act 31 of 2000). Gov. Gaz. 5, 21052.

Ross, L., 2016. The Prevalence of Cheating in National Standardised Assessments in South African Schools. University of Cape Town.

Schneider, M., Elacqua, G., Buckley, J., 2006. School choice in Chile: Is it class or the classroom? J. Policy Anal. Manag. 25, 577–601. doi:10.1002/pam.20192

Shepherd, D., 2015. Balancing Act: A Semi-Parametric Approach for Determining the Local Treatment Effect of School Type with an Application to South Africa. (Unpublished). Unpubl. Pap. Present. Quant. Appl. Educ. Conf. STIAS, Stellenbosch.

Spaull, N., van der Berg, S., Wills, G., Kotze, J., 2016. Laying Firm Foundations - Getting Reading Right. Final Report to the Zenex Foundation on poor student performance in Foundation Phase literacy. Stellenbosch.

Taylor, N., Draper, K., 2014. NEEDU National Report 2013: Teaching and learning in rural primary schools. Pretoria.

Taylor, S., Yu, D., 2009. The importance of socio-economic status in determining educational achievement in South Africa (No. 01–2009), Stellenbosch Economic Working Papers. Stellenbosch.

Van der Berg, S., 2015. What the annual national assessments can tell us about learning deficits over the education system and the school career year. South African J. Child. Educ. 5, 28–43.

Van der Berg, S., van Wyk, C., Burger, R., Kotzé, J., Piek, M., Rich, K., 2017. The performance of low fee independent schools in South Africa - What can available data tell? (No. 01/2017), Stellenbosch Economic Working Papers. Stellenbosch.

Von Fintel, M., 2015. Social Mobility and Cohesion in post-apartheid South Africa. Stellenbosch University.

Von Fintel, M., Van Der Berg, S., 2017. What a difference a good school makes! Persistence in academic performance and the impact of school quality (No. 07/2017), Stellenbosch Working Paper Series.

Wills, G., 2016. Limited support for the foundation phase: A misallocation of district resources. Policy brief, May 2016. Stellenbosch.

Wills, G., 2015. Informing principal policy reforms in South Africa through data-based evidence. South African J. Child. Educ. 5, 95–122.

Woolman, S., Fleisch, B., 2006. South Africa's unintended experiment in school choice: how the National Education Policy Act, the South African Schools Act and the Employment of Educators Act create the enabling conditions for quasi-markets in schools. Educ. Law 18, 31–75. doi:10.1080/09539960600787317

Zoch, A., 2016. Measuring social and economic mobility in South Africa: New approaches to well-known survey data concerns. Stellenbosch University.

Appendix 1: About the ESRC/DFID funded project entitled "Identifying exceptionalism and resilience among township and rural primary schools in South Africa"

"Identifying exceptionalism and resilience among township and rural primary schools in South Africa" more affectionately known as "Leadership for literacy" is an education research project lead by a multi-disciplinary team of investigators across Stellenbosch University, UCT, JET, UNISA and the Department of Basic Education. The project is funded by the Economic and Social Research Council (ESRC) and the Department for International Development (DFID) in the United Kingdom and runs over the period September 2016 to September 2018.

The research team is headed by Professor Servaas van der Berg at Stellenbosch University, who are joined by researchers Dr Nicholas Spaull, Professor Ursula Hoadley, Jaamia Galant (both at UCT), Dr Nick Taylor (JET), Dr Gabrielle Wills (Stellenbosch University) and David Carel (Stellenbosch University). The project is funded by the Economic and Social Research Council (ESRC) and the Department for International Development (DFID) in the United Kingdom and runs over the period September 2016 to September 2018.

The project was initiated in reaction to a deficit discourse where much research has focussed on the real realities of a highly underperforming schooling system in South Africa particularly among the non-fee paying public schools. Where solutions are desperately needed, less consideration has been given to exploring pockets of excellence that may exist within the non-fee paying part of the system. There has been little rigorous research on higher-performing schools in challenging contexts at least at the primary level and what may be learned from these schools, particularly with respect to how the organisation is led and managed. In this context, the objectives of this research project are fourfold.

Project objectives

The first objective of this research was to identify the number of exceptional rural and township primary schools in South Africa. This is critically important for ascertaining where access to quality school is a major determinant of success in the labour market and social mobility. Our research question in this regard was to establish how many non-fee paying rural and township schools actually exist and where they may be located? Annual National Assessment data provides a national set of data which can be analysed to answer this research question, but a decision was taken to empirically verify the performance of what may be 30 high-performing schools across a low, median and high-performing province in South Africa using tests of literacy at the grade 3 and grade 6 level.

A second objective is to gain new insights into school leadership and management practices and how these may be linked to literacy outcomes in high achieving schools relative to average or low-achieving schools in challenging contexts. This will occur through qualitative in-depth school visits (Stage 2) and the further school visits that will generate thick descriptions of management and leadership practices. It will subsequently be further informed by the quantitative investigation, after development of the SLM instrument.

A third and central objective is the development of a new School Leadership and Management instrument that captures the actual practices and behaviours of teachers and principals in challenging contexts in Africa. This will be done using the information generated by Stages 2 and 3, the available international and national evidence. This is the instrument that will be administered in the 30 school pairs to be tested.

The fourth objective of the research project is to determine how predictive this SLM instrument is of academic achievement and school functionality in these schools in challenging contexts. The relatively

small number of school (60) limits the precision with which it would be possible to estimate this, but would be enough to establish the promise of this instrument for future studies.

Project method and design

Qualitative methods are critical to achieving 2 of the 4 objectives of the project, particularly investigating if and how leadership and management processes link to improved instruction and literacy outcomes. Where learning outcomes are strongly convexly related to socio-economic compositional effects in schools, one approach to discounting for advantage in a qualitative study involves a comparative investigation in a nearby typical or underperforming school.

In this respect the mixed methods approach of the project involved at the outset a matched pairs design. Each exceptional school is paired with a nearby 'typical' or underperforming school. The matched pairs approach assumes that given a similar geographical position each school pair should have the same socioeconomic characteristics, and be influenced by similar cultural/political/local factors. In this respect in making comparisons across a high-performing and low-performing school, one can factor out some unobserved characteristics on the qualitative findings.

Final schools selected

Eventually we visited and assessed students in 61 schools in three provinces of which 32 were potential high performing schools either overall in ANA and/or in grade 6 literacy outcomes. Due to school access challenges during fieldwork we surveyed only 29 pairs rather than 30 as initially intended (10 in Gauteng, 10 in KwaZulu-Natal, 9 in Limpopo). One Limpopo school which seemed to be high performing in ANA was visited but due to bad weather and strike action we could not access its intended underperforming pair. Instead we visited 2 additional schools which we heard may be good performers.

Fieldwork and data collected

After three days of intensive training of fieldworkers (one which involved administering instruments in a pilot school in the presence of trainers), quantitative fieldwork was conducted for one day in each of the 61 schools between 6 February-March 2017 by a team of three fieldworkers.

- The first fieldworker assessed 10-15 grade 3 students in African and English language Oral Reading Fluency and word recognition. Students were sampled by i) the teacher selecting his/her two best students and the remainder were randomly selected from the class list by selecting each '7th' student down the list. This was a one-on-one test, with information captured electronically in tablets.
- The second fieldworker administered
    - the written grade 6 literacy, comprehension and vocabulary tests to an entire class. This was a pen-paper test.
    - Engaged in one-on-one reading and comprehension test with 10-15 grade 6 students from the same class (but selected in the same manner as for the grade 3 sample).
    - Ensured that the grade 6 teacher completed a teacher vocabulary test (the same as what was given to the grade 6 class).
- The third fieldworker administered a number of instruments to capture school characteristics, school climate, school functionality indicators, teacher perceptions and leadership and management practices in the school.
    - An anonymous self-administered educator survey to identify a number of factors including perceptions about management was handed out to all educators at the start

of the school day. This was completed during the course of the day, sealed in an envelope and placed in a box.

- A school functionality instrument was filled out by the fieldworker, capturing amongst other things teachers' presence and activity in the classroom. Specific times were set for doing school walk arounds and observations during break time.
- A 60 minute interview was held with the school principal, and if there were any un-known answers the deputy principal was also consulted for missing answers.
- A 45 minute interview with the grade 3 home language teacher of the tested class was administered.
- A 45 minute interview with the grade 6 English language teacher of the grade 6 tested class was administered.
- Grade 3 and 6 classroom and work book observations which were captured on tablets.

Appendix 2: Project design limits the available school options for sample consideration

Establishing a final sample of schools to visit for the ESRC/DFID project was further complicated by language and grade configuration requirements for language testing, national schools' data inaccuracies as well as the matched pair design of the project. But it was the language dimensions of the project that placed the greatest squeeze on our potential outlier sample.

Language squeeze on the sample

At the onset of the project our initial plan to use 2015, 2016 and 2017 universal ANA results which tests in maths and language came to an abrupt end with union resistance to the continuation of ANA in 2015. In response to this setback and to growing concerns about abysmally low levels of basic literacy skills being acquired in both African languages and English, a decision was taken to focus on just testing *literacy* outcomes at two critical grade transition points: grade 3 and grade 6. South African basic education at the primary level is split into three phases; the foundation phase (grades 1-3), the intermediate phase (grades 4-6) and grades 7 (senior phase). Students typically learn in their home language in grades 1-3, then a language switch to English takes place in grade 4. The curriculum assumes that children have acquired basic reading skills in both their home language and English by grade 3. Testing both African language and English proficiency at this point is an important indicator or the readiness of the child to proceed to further grades and keep pace with the demands of the national curriculum.

Test development became unavoidable, particularly in testing African language literacy at grade 3, because of a lack of existing tests. For the sake of cost and time, a decision was taken to develop reading tests at the grade 3 level in only two to three of the 11 official South African languages.[14] Given the predominance of isiZulu in KwaZulu-Natal and its proliferation in many other parts of the country, including Gauteng, this was an obvious test language of choice. Anticipating a particular shortage of acceptable schools in Limpopo with its well accepted low levels of learning at the primary school level relative to other provinces, we expanded the language of testing to two areas: Sepedi and Xitsonga.

Despite testing in the most frequently occurring languages in these provinces, the best performing provincial schools in ANA are not dominated by one language group. In Limpopo, Tshivenda schools could not be considered as potential outlier schools for our sample. In the case of Gauteng we discovered that schools with an unexpected LOLT given provincial population dynamics are reflected among the best ANA performers. As an example, despite the very few number of isiXhosa speaking

14

people in Gauteng, one or two of the apparent best ANA performing non-fee paying schools have isiXhosa as their dominant LOLT.

In Gauteng with its highly diverse population dynamics in terms of home language, the consequence of migration to this economic centre, the process of identifying schools that fitted our language profile was particularly problematic. Given diverse languages and/or parent preferences for children to be taught in English, some 'outlier' township and rural non-fee paying schools have opted for English as their foundation phase LOLT. Initially we wanted to limit our sample to only schools teaching in African home language in the foundation phase. The options became so limited that we had no choice but to lift this criteria and include a few English LOLT schools in our sample, provided the dominant home language of the class matched our language testing area.

Inaccuracies in national data on schools

Determining whether schools fitted our profile was further exacerbated by data challenges, where national data on the language of learning and teaching in a school, its grade configurations (collected in the Annual Survey of Schools) and fee-paying status is often outdated or contains inaccurate information. This required hundreds of phone calls to obtain working phone numbers of schools or their principals (EMIS data is terribly inaccurate particularly in Limpopo and KZN) before verifying this information. There are some 'outlier' primary schools that only offer lower primary grades, only higher primary grades or some other grade configuration preventing testing both grade 3 and grade 6 in the same school.

Matched design limitations

The mixed methods design of the project hinged upon a matched schools approach. This meant that a selected high performer on ANA with a specific language and grade configuration would need to be matched with similar school nearby (with the same language and grade configuration). We tried as much as possible to find an under-performing or typically performing school with similar language profile of students, language of learning and teaching, fee structure and roughly comparable enrolment size; however, matching on all factors was at times impossible (no such match existed) and in a few cases involved choosing school pairs that were further away than initially intended. The success of the matching become a strong determinant when filtering down the list of potential higher performing schools to visit.

The success of a matched approach assumes that there is a random geographic spread of performance. But performance often occurs in clusters – in other words schools surrounding a potential outlier school are all doing relatively well so that there may be little visible performance differential across the schools. In Gauteng province and parts of Limpopo province this is a major issue. When the performance metric on which selection is based is potentially a noisy measure of actual learner proficiency, finding significant performance differences in learning across pairs is not guaranteed. Verifying that a school is a typical or bad performer is likely to be equally as challenging as identifying good performer.