

Improving Management in Education: Evidence from a Randomized Experiment in Madagascar¹

Trang Nguyen (MIT)
Gerard Lassibille (University of Bourgogne)

June 2008

Abstract

This paper evaluates several interventions in Madagascar that sought to promote top-down and local monitoring of the school to improve education quality. Randomly selected school districts and subdistricts received operational tools to facilitate the administrators' supervision tasks. Randomly selected schools in these treated districts were reinforced with teacher tools and parent-teacher meetings centered around a school report card. The other schools in treated districts did not receive any direct intervention; it was up to the administrators to distribute the report cards and tools, and implement effective monitoring in those places. We find little impact of targeting district and sub-district administrators on their behavior and on the performance of the schools under their jurisdiction. Meanwhile, the intervention implemented at the school level led to an improvement in various observable performance measures. Teachers prepared lesson plans and evaluated their students more frequently. Teacher attendance and communication with parents did not improve, consistent with predictions of the multi-tasking principal-agent problem. On balance, the effect of the school intervention on student attendance was positive: 4.3 percentage point increase from 87% mean attendance. Test scores also improved by 0.1 standard deviations after two years, particularly in math and Malagasy. These results suggest that improving management via mediated control in the hands of government bureaucrats does not do much in this context, while beneficiary monitoring leads to an improvement in teaching activities and student achievement.

¹ We thank Esther Duflo for her guidance on this project from the onset. Abhijit Banerjee and participants of the MIT development lunch have provided very helpful comments. This study would not have been possible without the excellent field work by the AGEMAD team at the Madagascar Ministry of Education. We are grateful to the World Bank and AFD for financial and technical support.

1. Introduction

Inadequate quality in education is a concern in many developing countries, despite increased funding and inputs that go into the system. In Madagascar, recent reforms since 2002 to provide free primary education as well as free inputs (school kits and basic textbooks) are still insufficient to ensure a basic education. Only 63% of Grade 5 children pass the primary-cycle exam, a minimum level of language and math knowledge presumed at this grade (Tan 2005). The political economy of public good provision argues that government bureaucrats might have better incentives to improve quantity, rather than quality, of services since the benefits of improving service quality are diffuse and harder to verify (Glewwe and Kremer 2005). The World Development Report 2004 thus points to the need of promoting local accountability, that is, enabling the beneficiaries of education to take control of service quality (World Bank 2004).

It remains an open empirical question which interventions, solely top-down approaches or with beneficiary participation, actually improve public services. A top-down approach could aid bureaucrats in their work to monitor school quality. Its impact is based on the premise that these administrators have the right incentives and would carry out effective monitoring. While Olken (2007) shows that external auditors were successful in monitoring corruption and ensuring quality in Indonesian road projects, we cannot say anything definitive about relying on the existing management system to do the same. On the other hand, promoting local accountability may work well if we know exactly what makes beneficiary participation effective. Beneficiary participation takes several forms: parents can monitor teachers either directly via a punishment and reward scheme or indirectly by complaining to the district supervisors. However, participatory programs that require joint actions by a group of parents may face coordination failures and free-riding problems.

To contribute to this debate, we present in this paper evidence from a large-scale randomized experiment in Madagascar's public primary school system. We study the impact of top-down interventions within the existing education system, alone versus in

combination with local accountability. The goal of the interventions evaluated here (called AGEMAD² program) was to facilitate the administrators' tasks in supervising their schools, on one level. On another level, the motivation was also to provide the village community with information about their school's performance and to allow them to coordinate on taking actions directly. A presumption behind AGEMAD is that lack of information and complementary inputs hold back effective control by the beneficiaries and quality service delivery by the education workers.

Our strategy involves three interventions implemented by the Ministry of Education. The first intervention provided the district administrators with operational tools for their tasks and a training on those tools. The second intervention did this, and in addition, trained and provided respective tools to the subdistrict head. Some examples of the administrators' tools are supervision forms for school visits and information on the performance and resource level at each of their schools. These operational tools are designed to help administrators better supervise the quality of education delivered at the school and allocate resources efficiently within their jurisdiction.

The third intervention had an additional treatment at some randomly selected schools to involve the parents in monitoring school activities. This school-level intervention entailed school meetings organized by the AGEMAD team to provide the parents with the school report card and allowed them to coordinate on doing something about education at the village. Even when the beneficiaries can exercise control and teachers exert more efforts, they may still lack the complementary inputs to improve teaching quality. AGEMAD therefore provided pedagogical and administrative tools for teachers. Examples of such complementary inputs are structured tools for lesson planning, records of student attendance and learning, and reports to parents and to the school director.³ The other untreated schools in treated districts and treated subdistricts benefited from the top-down interventions but did not receive the report cards and teacher tools directly. It was

² AGEMAD is the French acronym for improving management in education in Madagascar.

³ Evidence to date (Glewwe et. al (2002), Glewwe et. al (2004), and Banerjee et. al (2002)) shows that more resources alone, such as textbooks, flip charts, additional teachers, do not lead to improved learning at the school. Our conjecture is that resources directly aimed at changing the way education is delivered, such as the AGEMAD tools, can work.

the responsibility of the district and subdistrict heads to distribute them and implement a monitoring system in those schools.

We collected data on a variety of teaching practices through a school survey. We also collected student attendance data during unannounced visits and student test scores from an achievement test administered independently. We find that the interventions targeted at the district and subdistrict administrators have minimal effects on their behaviors and on the schools and students under their responsibility. This is not because the tools failed entirely: each tool is used by 90% of the subdistrict heads and more than half of the district heads on average. While treated subdistrict heads visited their schools more often, they did not actively involve the parents and the local community in monitoring the school. The teachers in these two treatment groups do not plan their lessons or evaluate their students more, and the students perform no better than those in the comparison group.

On the contrary, the third intervention, implemented at the school level, led to an improvement in several teaching behaviors and student outcomes. Teachers' lesson planning practices improved by 0.26 standard deviations on average, and student evaluation practices by 0.14 standard deviations. Meanwhile, teacher attendance and the school's communication with parents did not improve. This result is consistent with the predictions of a multi-tasking principal-agent model in which teacher efforts are directed toward activities closely tied with observable indicators of their performance, such as those in the book records. The net effect on student attendance was positive: 4.3 percentage point increase from 87% mean attendance. Test scores of students in this treatment group are 0.1 standard deviations higher than those in the comparison group, as well as those affected by the interventions at higher levels, two years after the implementation of the program.

We interpret these results as suggesting that in the Madagascar context, it is difficult to rely on the existing hierarchy to implement effective monitoring. In the literature evaluating top-down programs, the results presented here differ from Olken (2007). He

finds that top-down monitoring reduces corruption by more than the grassroots programs in his experiments. While monitoring in that case was carried out by an external audit agency, the evidence from Madagascar indicates that targeting bureaucrats within the system did not lead to more successful monitoring, especially when our top-down interventions did not introduce explicit incentives for the bureaucrats to do so. In this way, our paper resonates earlier evidence that local government bureaucrats by themselves do not succeed in implementing incentive programs in service delivery (Kremer and Chen (2002), Banerjee, Duflo, and Glennerster (2008)).

The grassroots approach with beneficiary control worked better than the top-down approach in our context. Comparing the results here with the literature, we argue that it was important to arrive at a specific action plan in designing this participatory program. The existing set of field experiments to evaluate the impact of accountability in developing countries gives mixed evidence. Some participatory programs do not work while others do, with specific steps for the beneficiaries to take actions. Kremer and Vermeersch (2005) does not find any improvement in test scores or teacher attendance when parent associations in Kenya are allowed to choose and implement a reward program for teachers. Banerjee et. al. (2008) finds no effect of participatory programs in India, which organized a village meeting to inform the village community about the poor status of learning there and about the control power of the village education committee. The school-level intervention in Madagascar is quite similar to this information project in India with respect to the parent meeting and the school report card components. One difference is that AGEMAD suggested in the school meeting's agenda what the parents can do, aside from the contextual distinction. The AGEMAD meetings all yielded in an action plan with specific goals and tasks for the school, as well as the actors in charge of doing and monitoring each task. This activity at least ensures no coordination failure at the early stage for the large group of parents to agree on what to do. In this sense, our results are similar to Bjorkman and Svensson (2006), which finds that accountability improves health services in Uganda. The community meetings there to discuss the status of public health services were structured to lead to an action plan. Duflo et. al (2007) also finds support for the use of parental supervision in monitoring teacher practices and

improving children's learning. They study an intervention in Kenya that empowered school committees in randomly selected schools with training in how to monitor teachers, and with a formal committee meeting to evaluate teachers' performance. They find that teachers in those schools attend class and teach more often, and their students have higher test scores than those in the schools without empowered parental involvement. In both of these latter studies as well as ours, the parents had specific steps toward taking actions.

The organization of this paper is as follows. Section 2 provides a sketch of the institutional context in Madagascar. Section 3 describes the package of interventions and the evaluation design. In section 4, we discuss data collection and experimental validity. Section 5 presents the empirical strategy and results of the experiment. The final section concludes.

2. Primary Education System in Madagascar

On the supply side, the administrative structure of education in Madagascar is a typical hierarchy from the Ministry of Education (MENRS) to districts, subdistricts, and schools. Overall, roughly 2.7 million students are enrolled in 15,000 public primary schools. These schools are grouped into subdistrict zones called ZAPs whereby the ZAP head provides frontline administrative and pedagogical support. These officers report to one of 111 school district (CISCO) heads, who in turn report to MENRS. The CISCO and ZAP heads manage resource flows to schools, oversee teaching and learning practices as well as the collection of school statistics. They also administer the annual national examination at the end of Grade 5 (CEPE exam). Some specific examples of their work include visiting schools for supervision, disbursing teacher paychecks, processing their in-service training and transfer requests, organizing school building and maintenance projects, and distributing books and school grants. The selection of CISCO heads is now nominally competitive, with performance-based reviews after each three-year contract. ZAP heads are usually retired teachers; they rarely face any firing threat.

On the beneficiary side, the main channel of community involvement in primary education in Madagascar has been mostly informal through local parents' associations (FRAM). Recognizing the informal existence of the FRAMs, the government of Madagascar created formal school boards (FAF) starting in 2002-2003. These committees are responsible exclusively for managing capitation grants and not focused on education quality. School boards consist of the village head, the school director, the subdistrict head (rarely there), teacher representatives and the president of FRAM. Together, the FRAM and the school board could exercise control power to pressure the teachers to deliver quality services.

In theory, FRAMs have both direct (social pressure, hire and fire contract teachers⁴) and indirect control (report school problems or complain to higher authorities). In practice, each FRAM is usually a small group consisted of a few nominated parents. First, to influence the school may require joint actions from all the parents rather than just this small group. Coordination failures and freeriding problems may lead to limited participatory actions from the existing FRAMs. Second, they seem to know little about how children at the school are doing, and thus not doing enough to influence the school's operations. Given the possible lack of knowledge of the villagers about the school's performance, providing them with school report cards and organizing local accountability meetings to help parents coordinate on doing something about education might in turn improve incentives for school teachers.

But even if the community puts pressure on the teachers or complains to district and subdistrict heads, these actors themselves do not always have the proper means to actually do their work better. Our observations from the field suggest that teachers do not dispose of basic and standardized tools for their routine tasks to ensure a quality education. Higher-level administrators do not have the right information and tools to monitor, allocate resources, and provide support optimally across the schools under them.

⁴ These contract teachers usually hold only a lower secondary or high school degree, hired directly by FRAM and paid on average half the salary of the regular teachers.

It is then possible to improve the quality of education by providing actors at all administrative levels with operational tools to improve their productivity.

3. The AGEMAD Interventions

3.1 Description and Structure of the Program

The AGEMAD program evaluated in this study was implemented by the Madagascar Ministry of Education (MENRS), with support from the World Bank and the French Development Agency. It is one of MENRS's strategic policies to develop efficient management within the primary education system, with the first 2 school years from September 2005 to June 2007⁵ as a pilot study for evaluation. We will refer to 2005-6 as "year 1" and 2006-7 as "year 2."

In 2005, MENRS initiated its AGEMAD program in 3774 public primary schools of 30 rural CISCOS. This purposive sample represents all geographical areas of the country, but focuses on those CISCOS with relatively high repetition rates. For the purpose of the experiment, we excluded schools that are too far away and extremely difficult to access.

As we investigate whether top-down interventions work to improve public services or beneficiary monitoring is also necessary, the AGEMAD initiative took place at three levels in the administrative system: CISCO, ZAP, and school. To give an overview, the CISCO and ZAP interventions provided operational tools for administrators to monitor and support their schools. Meanwhile, the school-level intervention has a "bottom-up" component on top of these treatments, giving parents a way to act directly. First, 15 CISCOS were randomly chosen to receive the CISCO-level treatment; no contact was made in the remaining 15 control CISCOS. The CISCO team received a training and sufficient operational tools to facilitate their routine tasks during the school year. Tools for CISCO heads include forms for supervision visits to the school or procurement sheets for school supplies and school grants. See Appendix Table 4 for the full list of CISCO

⁵ The school year in Madagascar runs from September to June.

tools. These are printed templates for the users to fill in according to the attached instructions. One exception is the CISCO report card that each treated CISCO received, which actually has information on the performance (dropout rate, CEPE exam pass rate, and repetition rate) and resource level of all the ZAPs under this CISCO. The CISCO head can find this tool useful in evaluating and monitoring subdistricts.

Second, some ZAPs were selected randomly within treated CISCOs to receive the ZAP-level intervention, which is quite similar to that at the CISCO but with ZAP tools and ZAP report cards. Examples of tools for ZAP heads include forms for supervision visits to the school, community meeting forms, or procurement sheets for school supplies and school grants. The full list is displayed in Appendix Table 3. The ZAP report card has information on the performance and resource level of all the schools under this ZAP. The CISCO and ZAP interventions are two levels of “top-down” programs to facilitate administrators’ work in monitoring schools under their jurisdiction. Tool distribution was accompanied by user trainings in September of each year. During training sessions, MENRS representatives emphasized on the core responsibilities of each actor, and explained how to use the tools in performing these responsibilities.

Third, within treated ZAPs, some schools were randomly chosen to receive the school treatment, in addition to the interventions at higher administration. The school-level intervention aimed to improve accountability by sharing information and mobilizing parents to coordinate on an action plan. In particular, MENRS distributed directly to the treated schools their school report card with performance indicators from the previous academic year, and organized accountability meetings for the parents to discuss it. The three simple indicators presented on the report card should be comprehensible to illiterate parents: dropout rate, CEPE exam pass rate, and repetition rate. All statistics were calculated using data from the national school census. Based on these statistics, MENRS classified schools into four categories of performance relative to resource endowment (resource is proxied by the student-teacher ratio). “Exceptional” represents schools with relatively few resources but perform well in terms of dropout rate, CEPE pass rate, or repetition rate. “Satisfactory” implies schools with good resources and good

performance. “Difficulty” means having few resources and performing poorly. “Disappointing” schools are those with good resources but performing poorly. A school may be satisfactory in terms of exam pass rate but disappointing in terms of repetition rate. An example of the school report card (translated from the original version in French and Malagasy) is included in Appendix Figure 1.

To help parents focus attention and action on results as shown in the report card, AGEMAD also organized two accountability meetings between the treated school and the parents as well as the local community. One took place at the beginning and the other near the end of the school year. MENRS suggested an agenda for these half-day school meetings and requested that an action plan was set up at the first meeting of the year and evaluated at the last meeting, but the actual content of discussion was not mandated and left to the discretion of the participants.⁶ Proposed items on the agenda include a review of the current situation at the school, its progress with information available from the report card, discussion on possible solutions to existing problems, and setting goals in an actual action plan. To arrive at a specific plan, the participants first identified the issues hampering academic achievements and then prioritized the activities to overcome these problems. The action plan had to be approved by the general audience at the parent meeting. While the level of details varies from school to school, it typically has a list of objectives or desired improvement in certain general activities such as teaching. For each objective, several tasks were identified to be performed over a specified time period, along with the actors responsible for them. One example of the goals is to increase the exam pass rate at the school by 5 percentage points by the end of the academic year. The common tasks specified for the teachers include lesson planning and student evaluation every few weeks. Each parent and the parent association are expected to monitor the student evaluation reports that the teachers are supposed to communicate to them.

⁶ MENRS did not send any of its staff to the school meetings. About half of the times, the ZAP head is present, but that is part of his/her job and not really the presence of a new authority figure. MENRS had also hired facilitators to ensure that school meetings indeed took place and to collect information on the meetings. One facilitator would attend each meeting, but would usually not intervene.

These treated schools also received director and teacher tools directly. Teacher tools are, for example, lesson-planning notebooks, class attendance registers, and grade registers. The director would have the registration roster, summary of student test scores, and community meeting forms. Appendix Tables 1-2 list all these tools, designed to be user-friendly and time-efficient. They were produced as physical sheets of paper or notebooks, for the users to fill in throughout the school year. The presumed benefit of the tools is to provide inputs complementary to education workers' effort in their work, in turn improving school quality.

The other schools in the treated ZAPs and 15 treated CISCOs did not receive any direct school-level intervention. Report cards and sample teacher tools were created for them, but it was left to the hands of the CISCO and ZAP heads to distribute, implement the tools, and monitor school activities. There was no option to have a treatment group with school report cards alone since it would have been practically difficult to work with government schools in Madagascar without going through the district and subdistrict heads.

The treatment design with the relevant samples of observations is demonstrated in Table 1. Treatment assignment was random, stratified by baseline repetition rate and school size (and geographical region for the CISCO randomization). The above levels of treatment intensity generate four types of schools. The control group includes all the schools in the 15 control CISCOs. The first treatment group refers to those schools influenced by the program only at their CISCO (schools in AGEMAD CISCOs). The second group of treated schools is influenced by interventions at both the CISCO and ZAP levels (schools in AGEMAD ZAPs). These two treatment groups represent a top-down approach to improving school quality. It is the responsibility of the CISCO and ZAP administrators to monitor the schools. The third type of treatment schools (called AGEMAD schools) received all three levels of interventions, CISCO/ZAP/school. The additional bottom-up component here engaged parents into acting directly rather than depending on the bureaucrats. Due to budget constraints, we randomly selected 303

schools in each of the four groups, i.e. 1212 schools total, as the follow-up sample for data collection.⁷

This design allows us to learn which approach in the AGEMAD interventions effectively improve the way schools function and student performance. The top-down approach may work if the CISCO-level and ZAP-level interventions produced sufficient impact on service provider behaviors and on key schooling outcomes. On the other hand, such interventions might be ineffective unless reinforced by direct involvement of the parents through accountability meetings at the school.

Our tracking data suggests that the interventions were well implemented with good take-up. All the AGEMAD schools that were intended actually received the tools and organized the meetings. Accountability meetings were well attended with 64 participants on average, half of whom were women. 30% of the attendees spoke up and asked questions at least once. 94% of these accountability meetings discussed the school report cards and students' learning results as intended. On average, 85% of the teachers in AGEMAD schools reported using several tools, with the pedagogical and evaluation tools used by almost everyone. Tool usage rate is lower for the director (65%). Of those who used the tools, close to 100% found them useful.

3.2 Why Might We Expect These Interventions to Work?

Teacher effort is an important input in students' learning, so higher teacher effort would improve the quality of education at school. In theory, we would expect that most means of monitoring could potentially give teachers better incentives to perform their duty, thus leading to higher test score achievement for students.

The top-down interventions need not always produce an impact. If the CISCO and ZAP administrators in treated districts actually monitor school activities, well-audited schools

⁷ Power calculations showed that this sample size is sufficient to detect an increase of 0.25 standard deviations in school-level average test scores with 95% confidence. Since we could later collect individual-level test scores, this sample of schools should allow us to detect a smaller treatment effect.

may perform better. The question is whether we can rely on the existing hierarchy to carry out effective monitoring. These interventions were implemented within the current institutions, almost entirely by the Ministry rather than a substitute system of management. The ZAP and CISCO heads do not face direct reward and punishment mechanisms for monitoring their schools well. It is ambiguous if the ZAP and CISCO interventions bring about strong changes in behaviors and student outcomes.

Alternatively, the school intervention provides information on school performance and allows the parents to coordinate on taking actions to monitor service quality. School report cards were rare in rural Madagascar, and centrally-generated report cards were non-existent before the program. This component of AGEMAD therefore provided probably new information on the quality of current services. With that, parents can exercise social pressure on the teachers or implement explicit incentive schemes so that the school does a better job of teaching. The school-level treatment is expected to affect teacher behaviors if the parents register the information provided on the report card, believe that they can exercise their control power, and successfully coordinate. This is likely the case in our intervention. Most parents attended the school meetings and participated actively in the discussion of the report card and of an action plan.

4. Evaluation

4.1 Data Collection

We expect the interventions to affect intermediate outcomes such as behaviors of education workers and the functioning of the school. This data comes mainly from the school questionnaire in 1212 schools of the follow-up sample. During each of the two school years between January and May, MENRS hired an NGO (Aide et Action) to administer a school questionnaire during randomly timed, unannounced visits to the schools in the follow-up sample. In addition to spot-checking teacher and student absenteeism, the surveyor interviewed the school director and teachers about their usage of common tools and their administrative and pedagogical practices. For example, some

self-reported measures include: pedagogical organization, student assessment and attendance registration, and communication with the parents and local school community. The year 1 questionnaire took place right after the interventions while teachers were just familiarizing themselves with the tools. Since we have better data in year 2, we will report the results from this latter school questionnaire.

In year 2, we also surveyed the ZAP heads on their interactions with their schools. In particular, we will see if treated ZAP heads pay more visits and attend more meetings at schools in difficulty as notified in the report card.

The next outcome of interest, student attendance, was collected by surveyors during unannounced school visits. From this data, we constructed three measures of attendance rates. First, “classroom count” is the ratio of student attendance in each classroom to its enrollment. This number is reported during individual teacher interviews with the surveyor. Second, “surveyor reported” is the ratio of how many students in the whole school the surveyor can quickly see upon arrival, divided by the enrollment of classes in session. This data is not the most precise since the surveyor did not have enough time to do a headcount, so he or she would give a rough guess instead. Third, “director’s book record” is attendance rate from the school director’s official record. Dropout and repetition rates are also available in the year-end statistical form (*Fiche de Fin d’Année*) administered nationally by MENRS.

We are interested in whether and what levels of interventions ultimately improve learning. Our primary data on learning comes from achievement test scores collected in both years of the experiment. MENRS administered the year 1 test to Grade 3 students in February 2006 (close to baseline). The endline test was administered to the same children in May 2007 (then in Grade 4). The testing instrument was based on PASEC tests⁸ and already well-developed with effective testing procedures. These two tests are intended to measure children’s competency in three materials: math, French, and

⁸ PASEC (Programme d’Analyse des Systèmes Educatifs de la CONFEMEN) is a program in 15 francophone countries that studies elements of learning for students.

Malagasy. They include basic calculations and grammar questions at the level that Grade 3 and Grade 4 students, respectively, are supposed to master. We normalize the endline test scores by subtracting the mean, then dividing by the standard deviation of the control group.

Moreover, the national school census serves as a secondary source of learning data. CEPE (end-of-primary-cycle national examination) exam results for Grade 5 students are available at the school level.

4.2 Descriptive Statistics

Table 2 presents descriptive statistics for a range of school characteristics prior to the interventions, and also confirms that the treatment and control groups do not differ systematically, as expected given the random assignment. The average school has more than 200 students, with roughly 54 students per teacher. 66% of Grade 5 students in these schools pass the minimum learning requirement, the CEPE exam. 18% of students repeat a grade, and 6% of students enrolled at the beginning of the school year drop out by the end. Columns 2-4 show that the treatment groups are statistically similar to the comparison group in terms of these characteristics.

4.3 Attrition

To minimize the potential bias caused by differential attrition, we tried to measure outcome variables for all the original participants of the program. The endline test was administered to as many of the baseline children as possible. The school director had asked children who no longer attended school to come at the day of the test; home visits were made to a random subset of the attriters.

Table 3 shows attrition rates by treatment assignment. The first row suggests that attrition is roughly 13% and is similar across treatment status. A further check is on the type of children who have attrited from our sample. If we look at their baseline test scores and

their gender, differences between the attriters and stayers are small and again similar in our treatment and control groups.

5. Estimation Strategy and Results

We first discuss the interventions' impact on intermediate outcomes such as teacher behaviors, director behaviors and general functioning of the school, as well as behaviors of the ZAP and CISCO heads. Then we report treatment effects on student attendance, dropout rate, and learning.

The main estimating equation used throughout this analysis is

$$(1) \quad Y_{is} = \alpha + \beta_1 * TG1_s + \beta_2 * TG2_s + \beta_3 * TG3_s + \delta X_{is} + \varepsilon_{is}$$

where Y_{is} represents outcome variables, on behaviors and on students' performance, of individual i in school s . TG 's denote treatment group indicators, with the control group being the omitted category. TG1 is an indicator for whether the school is in a treated CISCO. TG2 is an indicator for whether the school is in a treated ZAP. TG3 is an indicator for whether the school itself received the AGEMAD interventions. Given the randomization, our coefficients of interest β 's are interpretable as the causal impact of different levels of the AGEMAD treatment. We control for various baseline school or individual characteristics X . All the regressions using the school questionnaire data control for the month of random school visits to account for seasonal variation.

We also run an alternative regression with a dummy for any treatment at all (1, 2, or 3) in addition to the treatment 3 dummy. We can interpret β as the effect of Treatments 1 or 2 versus the control group, and β_4 as the effect of Treatment 3 versus Treatments 1 or 2.

$$(2) \quad Y_{is} = \alpha + \beta * AnyTreatment_s + \beta_4 * TG3_s + \delta X_{is} + \varepsilon_{is}$$

Both equations are estimated using OLS, with standard errors clustered at the CISCO level since treatments were implemented at this level.⁹ Observations are weighted by the probability of selection.

Since there are many intermediate outcomes on behaviors that may have been affected by AGEMAD, we adopt the approach from Katz, Kling, and Liebman (2007) when appropriate. We group outcomes into categories of related variables and report the average effect of each category: lesson planning, student evaluation practices, communication and meetings to discuss student matters, and so on. For example, the average standardized effect of Treatment 3 for a category of outcomes, each indexed by k , is defined as $\beta_3 = \frac{1}{k} \sum_{k=1}^K \frac{\beta_{3k}}{\sigma_k}$ where β_{3k} is the coefficient from equation (1) for outcome k , and σ_k is the standard deviation of this outcome. We calculate the standard errors of this average effect by estimating a system of seemingly unrelated regressions (SUR) for the outcomes in each category.

5.1 Intermediate Outcomes at School

5.1.1 Impact of CISCO and ZAP Interventions on Intermediate Outcomes

The first set of intermediate outcomes at the school is teacher behaviors, as measured in the school survey. Table 4 reports treatment effects on teacher attendance and pedagogical practices. The first set of *Regression (1)* reports OLS results of equation (1) for various outcome variables of teacher behaviors, listed on the leftmost column. The second set of *Regression (2)* runs equation (2) and gives quite similar results with more precise estimates. These results are robust to a varying set of teacher and school control variables.

⁹ A probit estimation instead of the linear probability version of equations (1) and (2) provides very similar results. The estimates are also similar if we restrict the sample to treatment groups 2 and 3 and cluster the standard errors at the school level.

We find that schools benefiting from the CISCO and ZAP interventions do not generally have better teacher practices. The SUR analysis for lesson planning is presented in Panel B. Columns 3, 4, and 6 show no significant improvement in the schools that received the district and subdistrict interventions. Teachers of these schools (treatment groups 1 and 2) are on average 0.1 standard deviations more likely to perform these tasks, but the estimate is statistically indistinguishable from zero. Similarly in terms of student evaluation practices, column 6 in Panel C suggests that schools in the treated districts and sub-districts do not perform these tasks better than the control schools. We will return to this point when we discuss the ZAP's behaviors.

As shown in Panel A of Table 4, the CISCO and ZAP interventions do not improve teacher attendance. We study three measures of teacher attendance. That reported by the director during the random check is attendance at the teacher level, while the other two measures are at the school level. The third measure, percentage of teachers present reported by the surveyor during his random visit, comes from an approximate headcount upon arrival and is relatively noisy. The mean attendance rate in column 1 suggests less than 10% absenteeism, in contrast to higher teacher absenteeism in Kenya--20% (Glewwe et. al. 2003) or India--25% (Chaudhury et. al. 2006). The coefficient estimates of treatment effects for treatments 1 and 2 are often negative in columns 3, 4, and 6, though very small and hardly distinguishable from zero.

The second set of intermediate outcomes at the school is the director's behaviors and school functioning. We collected some observational data at the outset of each unannounced school visit to measure general school functioning, but find no sign of significant improvement in those schools receiving the top-down interventions. According to Panels A and B of Table 5, there is no strong difference between treatment 1 or 2 and the control schools concerning whether students are playing in the hall during class time, whether the teachers are teaching, and the chance that a classroom has a teacher or a lesson in progress. The results on control and reporting practices by the director are displayed in Panel C. Our estimation of regression (2) implies zero average treatment effect for schools in the district and subdistrict interventions versus the control

group. The estimate of -0.04 standard deviations is statistically insignificant. Together, one interpretation of these no-impact findings for the top-down interventions is that the treated CISCO and ZAP heads do not by themselves implement more effective monitoring.

5.1.2 Impact of School Intervention on Intermediate Outcomes

In contrast to the above results on the CISCO and ZAP interventions, the additional school-level intervention appears to be effective in monitoring some school activities. As displayed in Table 4, teachers' practices are improved in AGEMAD schools along some dimensions. Lesson planning is performed much more frequently in AGEMAD schools than in all the others (Panel B, column 2). From a SUR analysis, we find that teachers of AGEMAD reinforced schools (treatment group 3) are on average 0.26 standard deviations more likely to perform these tasks. This estimate is statistically significant at the 5% level. Student evaluation (Panel C) has improved on average by 0.14 standard deviations in AGEMAD schools compared to control schools. This estimate has large standard errors, and is statistically significant only in regression (2). As implied in column 5, while schools in treated districts and sub-districts do not improve student evaluation practices, schools treated directly with the parents' involvement fare better by 0.07 standard deviations. These results suggest that allowing the beneficiaries to coordinate on exercising their control power to improve education leads to better teaching quality in this context.

However, teacher attendance does not improve in AGEMAD schools even though the director says that (s)he records teacher absences more frequently. The coefficient estimates on treatment 3 in Panel A of Table 4 are often negative, though very small and hardly distinguishable from zero. It may be costly for the parents to verify teacher presence. Then as long as the school director states in the official school records that teachers still show up regularly, there is no additional incentive induced by AGEMAD for teachers to come to school more regularly. Contract teachers are associated with

lower presence rate than regular teachers, but the treatments do not affect their attendance differentially.

While teachers say they talk to the director often, only about half of them communicated with some parents over the previous month. Teachers in treated schools are no more likely to communicate with the parents or with the director on student matters, as presented in Panel D of Table 4. Overall, we do see that teacher behaviors improve on dimensions that are easily recorded for future evaluation of their work (filling out lesson planning and giving exams) and do not improve on other dimensions (teacher attendance and communication with the parents and the director). This finding is consistent with a multi-tasking principal-agent model in which teacher efforts are directed toward activities closely tied with observable evaluation indicators of their performance.

Table 5 presents the impact results of the additional school-level intervention on director behaviors and school functioning. While the standard errors in regression (1) are large, we find in regression (2) an average treatment effect of 0.07 standard deviations better control and reporting practices by the director in AGEMAD schools versus those in the district and subdistrict interventions. This average effect estimate, reported in Panel C column 5, is statistically significant at the 10% level. In particular, the director claims he is 8.5 percentage points more likely to record teacher absenteeism in treated schools (column 2, significant at the 1% level). Nonetheless, as we have discussed, teacher attendance does not improve. Self-reported monitoring from the director is either inaccurate or does not appear effective.

Similar to our findings for the teachers, there is no significant improvement in terms of the director's communication with teachers and with the community on school matters. While the estimates in Panel D are not small (average effects around 0.08 standard deviations), they are imprecise. According to Panels A and B of Table 5, we do not find significant improvement in general school functioning in AGEMAD schools. In sum, the school-level intervention affects mainly the teaching activities at school and the director's monitoring practices.

5.2 Behaviors of ZAP and CISCO Heads

To better understand the contrast between the impact of top-down interventions alone versus in combination with parental involvement, this section investigates further some behaviors of the ZAP and CISCO heads. We first look at how the different interventions affect the frequency of their school supervision. Then, we study the effect of the ZAP intervention on how ZAP heads allocate their efforts across supervising different types of schools under them.

We find that all the AGEMAD interventions lead to more school visits by the ZAP head, but (s)he does not hold more meetings with the local community. Specifically, AGEMAD schools and schools in AGEMAD ZAPs receive more frequent visits from the authorities. Panel E of Table 5 presents the impact of the AGEMAD treatments on the number of days since the last visit from the ZAP and CISCO. It is not surprising that CISCO heads tend to visit schools less than ZAP heads. On average, the last visit is 0.32 standard deviations more recent in AGEMAD schools than in comparison schools, and 0.3 standard deviations more recent in schools of treated ZAPs (treatment 2). Despite this evidence of more visits from the authorities, we still find no improvement in pedagogical practices at treatment 2 schools, implying that monitoring from the ZAP head does not appear effective.

In addition, we look at a few self-reported behaviors of ZAP heads from the ZAP questionnaire. The results from this analysis suggest that treatment ZAPs (treatment groups 2 or 3) do not hold more meetings with the local community and the school board. To reconcile this result with the above finding on more ZAP visits from the school questionnaire, one possible interpretation is that treated ZAP heads exercise more school visits but ignore the accountability channel of involving the village community.

In terms of effort allocation by the ZAP head, treated ZAPs were given information on the performance of schools in their jurisdiction. As described in section 3.1, the report

cards classify schools in each ZAP into four categories: exceptional, in difficulty, satisfactory, and disappointing. We expect treated ZAPs to provide differentially more support to schools in difficulty and exceptional schools than to satisfactory schools if they want to increase overall performance in the ZAP. In contrast to the spirit promoted in AGEMAD report cards, we do not find strong evidence for better allocation of efforts by the ZAP heads toward “schools in difficulty.”

Table 6 presents the effects of the ZAP intervention on the number of meetings and visits that each school receives from the ZAP head in year 1, interacted with the school’s initial performance category. Schools are categorized by three indicators prior to the program: net exit rate,¹⁰ primary-cycle exam pass rate, and repetition rate. Each column in this table represents a regression. The estimates on the school category terms suggest that exceptional schools and schools in difficulty in the comparison group tend to receive fewer meetings and visits from the ZAP administrator than satisfactory schools.

Coefficients on the interaction terms tell us, if anything, the opposite of what we expected although the estimates are statistically insignificant. In general, treated ZAP heads have even relatively fewer interactions with exceptional and difficulty schools than with their satisfactory schools. They did not seem to act based on the information provided to them on their schools’ performance.

5.3 Student Attendance and Dropouts

With the CISCO and ZAP interventions doing little to affect routine activities at the school, it is not surprising that we find no evidence for their impact on student attendance. Student attendance during the school year was collected by surveyors during unannounced school visits. We have three measures of attendance rates: “classroom count” is at the classroom level, while the other two measures are at the school level. “Surveyor reported” attendance rate is the ratio of the number of students in the whole

¹⁰ Net exit can be understood as dropout rate. For terminology purposes, the Ministry calls this measure net exit since schools only report enrollment at the beginning and at the end of the year. We only know how many children have left each school on net, but not whether some have transferred or dropped out completely. Student transfers are rare in Madagascar, so net exit from school is almost identical to the dropout rate.

school that the surveyor can quickly see upon arrival, divided by enrollment of the classes in session. Table 7 reports the treatment effects on attendance, for each of the three measures. As shown in columns 3, 4, and 6, the district and subdistrict intervention dummies always have coefficients statistically indistinguishable from zero, and smaller than that of the school intervention.

The school intervention, on the other hand, led to some changes in behaviors. We have seen that there are better pedagogical practices at AGEMAD schools, so we would expect the treatment effects to translate into improved student participation at school, including regular attendance and less dropout at the end of the school year, and improved learning. Student attendance rate (in percent) is overall better in AGEMAD schools and not in the other treatments. As shown in Table 7, AGEMAD schools have 4.29 percentage points higher attendance in the director's record compared to the control group average of 87%. Classroom count attendance also has a positive treatment coefficient, but not statistically significant. The regression using the surveyor's estimate of attendance also has insignificant treatment coefficients.

The general observation from this experiment is the lack of impact in the top-down interventions, but the school intervention worked. One exception to this general observation is the result on self-reported dropout rate. Madagascar's administrative data, reported by the schools themselves, reveals positive treatment effects on student performance at the end of year 1, as displayed in Table 8.¹¹ AGEMAD schools perform on average 0.17 standard deviations higher than the comparison schools, in terms of exam attendance rate, pass rate and dropout. Grade repetition in year 2 is also 5.1 percentage points lower, with both estimates statistically significant. In these regressions, the district and subdistrict interventions have an equally strong positive impact as the school-level intervention. This result diverges from the evidence on the other outcome measures we have thus far. One possible explanation is that self-reported dropout rates are relatively easy to manipulate on the statistical form submitted to the Ministry, as

¹¹ This table reports OLS results, and they are similar when we run a difference-in-differences regression by including the pre-AGEMAD values of these outcome variables. So we report OLS to be consistent with the rest of the paper.

opposed to teachers' lesson planning practices or student test scores measured by a third party. Even if there is no true reduction in dropout rates, schools in the treated districts, possibly feeling they were under extra scrutiny by the Ministry, might have reported better rates.

5.4 Learning

To investigate the program effects on the intensive margin of learning, we exploit Grade 4 student test score data from the achievement test at the end of year 2. Table 9 displays regression results of equations (1) and (2), in which the dependent variable is normalized test score. As shown in columns 2, 3, and 5, treatments solely at the CISCO or ZAP levels both have close to zero impacts, and the estimates are insignificant. This finding is not surprising given that there is no evidence for improvement in teacher behaviors in these two top-down treatment groups.

On the other hand, the magnitudes of regression (1) coefficient estimates suggest that students in the AGEMAD reinforced schools perform 0.095 standard deviations better than those in the control schools. Regression (2) attempts to alleviate the power issue by pooling treatments 1 and 2 together. The coefficient on Treatment 3 (AGEMAD schools) is significant at the 10% level. This result on learning is consistent with the general distinction in treatment effects between the top-down interventions alone and in combination with the school intervention.

This impact of the school intervention on the total score reflects impacts on math and Malagasy subject scores, with the exception of French not improving at all in AGEMAD schools. Math scores improve by 0.12 standard deviations, while Malagasy scores improve by 0.08 standard deviations. But in the Madagascar context, teachers do not master French very well. The AGEMAD interventions, as designed, do not help them teach better something they do not know so well themselves.

6. Conclusion

This study contributes to a growing quest to understand better which kinds of top-down or participatory programs might or might not work to improve service quality. We provide evidence from a large randomized experiment in Madagascar's public primary schools. This AGEMAD program had two levels of top-down interventions within the current system, where it distributed operational tools to enable better school monitoring by the district and subdistrict heads. Some randomly selected schools received the school-level treatment on top of the above. As in several other participatory programs, the school intervention here organized accountability meetings between the school and the parents to discuss information on the school report card, and encouraged the establishment of an action plan for the school. We expect the different AGEMAD interventions to influence first-hands the teachers' behaviors if monitoring by the administrators or by the parents is effective, and in turn influence students' learning.

We find that the top-down policy design with mediated control within the system by itself does not seem effective. Providing tools and school information to higher-level bureaucrats did not appear to have much impact on the schools and students under them. While the subdistrict heads claimed that they visited their schools more often, they did not hold more meetings to involve the local community into monitoring the school. There is no evidence of improvement in teaching practices and learning outcomes in these treatment groups.

On the contrary, the interventions down to the school level with beneficiary control had positive impacts. Teacher behaviors improved for some teaching tasks, such as planning lessons and giving evaluation quizzes and tests. Other tasks, less likely to show up in the book records, did not improve: teacher attendance or communication with the parents and the director. On balance, student outcomes still improved in these treated schools. In particular, test scores increased by 0.1 standard deviations after two years. This result suggests that teaching activities (lesson planning and student assessment) are important inputs in influencing student learning.

In sum, the top-down interventions in Madagascar seem ineffective in monitoring the school. While we have seen evidence that external auditors can successfully monitor service delivery, it remains a challenge to depend on bureaucrats in the system to perform that job. In this context, it is the school-level intervention with beneficiary monitoring that worked to hold the school accountable and affected teaching activities and learning.

References

- Banerjee, A., R. Banerji, E. Duflo, R. Glennerster and S. Khemani (2008). "Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India," mimeo, MIT.
- Banerjee, Abhijit, Esther Duflo, and Rachel Glennerster, (2008). "Putting a Band-Aid on a Corpse: Incentives for Nurses in the Indian Public Health Care System," *Journal of European Economic Association*, Proceedings, 6.2-3, April-May, forthcoming.
- Banerjee, A., S. Jacob and M. Kremer (2002). "Promoting School Participation in Rural Rajasthan: Results from Some Prospective Trials", mimeo, MIT.
- Björkman, M. and J. Svensson (2006). "Power to the People: Evidence from a Randomized Experiment of a Citizen Report Card Project in Uganda," mimeo, IIES, Stockholm University.
- Chaudhury, N., J. Hammer, M. Kremer, K. Muralidharan and H. Rogers (2006). "Missing in Action: Teacher and Health Worker Absence in Developing Countries." *Journal of Economic Perspectives*, 20(1), 91-116.
- Duflo, E., P. Dupas and M. Kremer (2007). "Peer Effects, Pupil-Teacher Ratios, and Teacher Incentives: Evidence from a Randomized Evaluation in Kenya." Mimeo, MIT.
- Glewwe, P. and M. Kremer (2005). "Schools, Teachers, and Education Outcomes in Developing Countries," *Handbook on the Economics of Education* (forthcoming).
- Glewwe, P., M. Kremer and S. Moulin (2002). "Textbooks and Test Scores: Evidence from a Prospective Evaluation in Kenya," BREAD Working Paper, Cambridge, MA.
- Glewwe, P., M. Kremer, S. Moulin and E. Zitzewitz. (2004). "Retrospective vs. Prospective Analysis of School Inputs: The Case of Flip Charts in Kenya." *Journal of Development Economics*, 74(1), 251-268.

Glewwe, P., N. Ilias and M. Kremer (2003). "Teacher Incentives," NBER Working Paper 9671.

Kremer, M. and D. Chen (2002). "Interim Report on a Teacher Attendance Incentive Program in Kenya", manuscript, Harvard University.

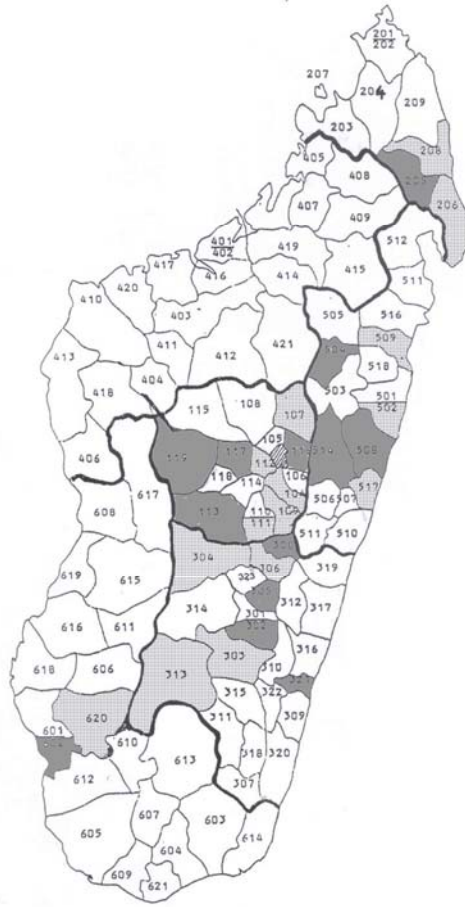
Kremer, M. and C. Vermeesh (2005). "School Committee Empowerment: Preliminary notes." Mimeo, Harvard University.

Olken, B. (2007). "Monitoring Corruption: Evidence from a Field Experiment in Indonesia," *Journal of Political Economy* 115 (2), pp. 200-249.

Tan, J.P. (2005). "Managing for Results in Primary Education in Madagascar: Evaluating the Impact of Selected Interventions," Concept Note, World Bank.

World Bank (2004), *World Development Report 2004: Making Services Work for the Poor*, New York: Oxford University Press.

Figure 1: School District Map of Madagascar



Notes: The shaded areas are the 30 school districts in our study sample. The darker ones are the treated CISCOS.

Table 1: Experimental Design

	Sample of observations			Description of interventions
	CISCOs	ZAPs	Schools	
Comparison	15	207	1721 (303)	No intervention
Treatment 1	15	170	1314 (303)	CISCO-level intervention
Treatment 2		89	436 (303)	CISCO + ZAP-level interventions
Treatment 3			303 (303)	CISCO + ZAP + School-level interventions

Notes: In parentheses are numbers of schools in the follow-up sample for data collection

Table 2: Baseline School Characteristics

	Comparison Group Mean	Treatment 3- Comparison	Treatment 2- Comparison	Treatment 1- Comparison	Obs.
	(1)	(2)	(3)	(4)	(5)
Panel A: Followup sample					
Enrollment	239.465	-4.785 (28.21)	-25.323 (24.89)	-39.267 (23.25)	1212
Repetition Rate	0.177	0.037 (0.02)	0.022 (0.02)	0.029 (0.02)	1212
No. of Class Sections	5.693	-0.017 (0.30)	-0.35 (0.25)	-0.366 (0.24)	1212
No. of Teachers	4.413	0.033 (0.41)	-0.383 (0.33)	-0.234 (0.34)	1212
Panel B: Sample of treated districts					
		Treatment 3 Group Mean	Treatment 2- Treatment 3	Treatment 1- Treatment 3	Obs.
	(1)	(2)	(3)	(4)	(5)
Net Exit Rate		5.872	0.931 (1.10)	0.729 (1.11)	873
CEPE Exam Pass Rate		66.260	-1.375 (3.22)	-2.572 (3.29)	707

Notes: Panel A reports mean differences across groups for all 1212 schools in the followup sample. Panel B reports mean differences in indicators for which we have baseline data for only the 15 treated CISCOs. Standard errors in parentheses, clustered at the CISCO level
* significant at 5%; ** significant at 1%

Table 3: Attrition in Endline Test Scores

	Comparison Group Mean (1)	Treatment 3- Comparison (2)	Treatment 2- Comparison (3)	Treatment 1- Comparison (4)
Percent Attrition	0.132	0.007 (0.02)	-0.005 (0.02)	0.002 (0.02)
Difference in pretest Attriters - Stayers	-0.068	-0.143 (0.08)	0.002 (0.10)	-0.012 (0.08)
Difference in proportion of males Attriters - Stayers	-0.013	0.037 (0.03)	0.011 (0.02)	0.019 (0.02)

Notes: N=24341

Standard errors in parentheses are clustered at the CISCO level.

Attrition rates in posttest data are at the individual level.

* significant at 5%; ** significant at 1%

Table 4: Teacher Attendance and Practices

	Endline	<i>Regression (1)</i>			<i>Regression (2)</i>		Obs.
	C.Group				Any		
<i>Dependent variables</i>	Mean	Treatment 3	Treatment 2	Treatment 1	Treatment 3	Treatment	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Teacher presence							
Director reported during random visit (teacher-level)	0.903 (0.008)	-0.03 (0.030)	-0.021 (0.019)	-0.007 (0.021)	-0.015 (0.022)	-0.014 (0.019)	5370
Director's book record	0.939 (0.007)	0.012 (0.020)	0.014 (0.016)	0.008 (0.018)	0.001 (0.009)	0.011 (0.017)	1061
Surveyor reported during random visit	0.974 (0.005)	-0.022 (0.018)	-0.017 (0.012)	-0.015 (0.010)	-0.007 (0.014)	-0.016 (0.009)	1069
Panel B: Teachers' lesson planning practices							
Prepared a lesson for today	0.798 (0.012)	0.093 (0.061)	0.017 (0.064)	0.05 (0.063)	0.06 (0.018)**	0.033 (0.062)	4184
Discussed lesson plan with director	0.835 (0.011)	0.065 (0.083)	0.014 (0.082)	0.037 (0.081)	0.04 (0.023)+	0.025 (0.080)	4184
Made weekly lesson plans	0.822 (0.012)	0.104 (0.040)**	0.046 (0.045)	0.04 (0.042)	0.061 (0.016)**	0.043 (0.043)	4184
Average Effect (in sd)		0.260 (0.117)*	0.077 (0.126)	0.124 (0.119)	0.160 (0.046)**	0.100 (0.121)	4184
Panel C: Teachers' student evaluation practices							
Evaluated student's learning last week	0.736 (0.014)	0.105 (0.140)	0.077 (0.137)	0.08 (0.138)	0.027 (0.017)	0.078 (0.137)	4184
Gave bimonthly exams	0.967 (0.006)	0.004 (0.011)	-0.003 (0.012)	-0.017 (0.017)	0.013 (0.011)	-0.01 (0.013)	4184
Average Effect (in sd)		0.141 (0.168)	0.086 (0.159)	0.050 (0.165)	0.073 (0.042)+	0.069 (0.161)	4184
Panel D: Teachers' communication							
Communicated with parents about student performance	0.943 (0.007)	0.017 (0.027)	-0.003 (0.029)	0.013 (0.024)	0.012 (0.02)	0.005 (0.023)	4184
Communicated with parents this month	0.580 (0.065)	-0.056 (0.059)	-0.085 (0.062)	-0.118 (0.068)+	0.045 (0.030)	-0.101 (0.063)	4184
Communicated with director about students	0.900 (0.009)	0.01 (0.017)	-0.001 (0.018)	0.019 (0.015)	0.001 (0.009)	0.009 (0.015)	4184
Communicated with director this month	0.991 (0.003)	-0.098 (0.100)	-0.119 (0.103)	-0.143 (0.105)	0.032 (-0.028)	-0.131 (0.103)	4184
Average Effect (in sd)		-0.049 (0.086)	-0.106 (0.084)	-0.095 (0.087)	0.051 (0.035)	-0.101 (0.084)	4184
Panel E: Classroom learning distribution							
Fraction of students scoring beneath the mean last exam	0.580 (0.065)	-0.185 (0.095)+	-0.164 (0.094)+	-0.131 (0.105)	-0.036 (0.025)	-0.149 (0.096)	4027

Notes: Col. (1) reports the endline mean in the comparison group.

Standard errors in parentheses are clustered at the CISCO level. Observations are weighted by the probability of selection.

+ significant at 10%; * significant at 5%; ** significant at 1%

All regressions include controls for the month during which the school was surveyed but not shown.

Regressions of teacher practices also have controls for teacher's age, age², teacher's salary, and teaching degree.

Treatment 1 is an indicator for whether the school is in a treated CISCO. Treatment 2 is an indicator for whether school is in a treated ZAP. Treatment 3 is an indicator for whether school itself received AGEMAD interventions.

Table 5: School Functioning and Director Practices

<i>Dependent variables</i>	Endline	<i>Regression (1)</i>			<i>Regression (2)</i>		Obs.
	C.Group				Any		
	Mean	Treatment 3	Treatment 2	Treatment 1	Treatment 3	Treatment	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Observations of school at visit							
Students playing in the hall	0.084 (0.017)	0.008 (0.032)	0.017 (0.029)	0.034 (0.035)	-0.018 (0.02)	0.025 (0.029)	1070
Teachers in the hall	0.096 (0.018)	-0.009 (0.029)	0.011 (0.028)	-0.003 (0.033)	-0.013 (0.02)	0.004 (0.028)	1070
Average Effect (in sd)		-0.004 (0.087)	0.050 (0.078)	0.052 (0.104)	-0.055 (0.064)	0.051 (0.083)	1070
Panel B: Observations of classrooms at visit (classroom-level)							
Classroom with a teacher	0.934 (0.007)	-0.02 (0.03)	-0.02 (0.02)	-0.013 (0.02)	-0.004 (0.02)	-0.016 (0.020)	4316
Classroom with a lesson in progress	0.843 (0.011)	-0.014 (0.06)	-0.019 (0.04)	-0.01 (0.05)	0.001 (0.03)	-0.015 (0.046)	4316
Average Effect (in sd)		-0.056 (0.097)	-0.062 (0.077)	-0.038 (0.082)	-0.006 (0.059)	-0.050 (0.076)	4316
Panel C: Director's control and reporting practices							
Teacher absences reported to Cisco	0.265 (0.026)	-0.043 (0.125)	-0.088 (0.130)	-0.076 (0.128)	0.039 (0.04)	-0.082 (0.129)	1202
Teacher absences reported to Zap	0.833 (0.022)	-0.068 (0.069)	-0.131 (0.087)	-0.027 (0.067)	0.011 (0.023)	-0.08 (0.076)	1202
Director checks student roll call at least weekly	0.231 (0.024)	0.014 (0.085)	0.027 (0.060)	-0.012 (0.062)	0.007 (0.035)	0.008 (0.060)	1202
Director records teacher absences	0.901 (0.019)	0.085 (0.034)*	0.054 (0.036)	0.07 (0.035)*	0.023 (0.012)+	0.062 (0.035)+	1202
Director records monthly student attendance	0.769 (0.026)	-0.002 (0.066)	-0.027 (0.069)	-0.08 (0.078)	0.051 (0.038)	-0.053 (0.071)	1202
Average Effect (in sd)		0.030 (0.092)	-0.057 (0.097)	-0.030 (0.090)	0.074 (0.039)+	-0.044 (0.092)	1202
Panel D: Meetings to discuss school matters							
Days elapsed since last teacher meeting	25.538 (2.710)	1.702 (3.863)	2.849 (4.131)	4.975 (4.251)	-2.212 (4.47)	3.906 (3.445)	1195
Days elapsed since last meeting with community	51.682 (5.269)	-1.79 (8.971)	-1.713 (7.852)	-6.371 (7.831)	2.291 (6.106)	-4.062 (7.678)	1195
No parent conference this past month	0.481 (0.029)	0.114 (0.090)	0.131 (0.090)	0.145 (0.091)	-0.024 (0.031)	0.138 (0.087)	1195
Average Effect (in sd)		0.079 (0.077)	0.100 (0.063)	0.099 (0.075)	-0.020 (0.060)	0.099 (0.065)	1195
Panel E: Visits from authorities							
Days elapsed since Cisco visit	294.352 (67.803)	-105.423 (68.447)	-75.272 (83.113)	-33.38 (72.301)	-51.06 (36.284)	-54.979 (71.137)	541
Days elapsed since Zap visit	88.907 (10.387)	-31.715 (8.283)**	-35.608 (10.739)**	-20.836 (10.813)+	-3.633 (5.508)	-28.208 (9.872)**	541
Average Effect (in sd)		-0.318 (0.105)**	-0.309 (0.126)**	-0.169 (0.117)	-0.080 (0.040)*	-0.240 (0.110)*	541

Notes: Col. (1) reports the endline mean in the comparison group.

Standard errors in parentheses are clustered at the CISCO level.

+ significant at 10%; * significant at 5%; ** significant at 1%

All regressions include controls for the month during which the school was surveyed but not shown.

Treatment 1 is an indicator for whether the school is in a treated CISCO. Treatment 2 is an indicator for whether school is in a treated ZAP. Treatment 3 is an indicator for whether school itself received AGEMAD interventions.

Table 6: How Frequently Does the Zap Head Interact with Schools of Different Categories?

<i>Schools categorized by</i>	<i>Dependent variables</i>					
	Number of Meetings in year 1			Number of Visits in year 1		
	<i>(Comparison Group Mean = 1.582)</i>			<i>(Comparison Group Mean = 3.098)</i>		
	Net Exit	Pass Rate	Repetition Rate	Net Exit	Pass Rate	Repetition Rate
(1)	(2)	(3)	(4)	(5)	(6)	
Exceptional	0.183 (0.155)	-0.22 (0.158)	-0.072 (0.096)	-0.553 (0.479)	-1.228 (0.463)*	-0.932 (0.520)
Difficulty	-0.161 (0.162)	-0.164 (0.112)	0.011 (0.127)	-1.153 (0.426)*	-1.844 (0.695)*	-0.883 (0.673)
Disappointing	0.188 (0.316)	-0.255 (0.161)	0 (0.085)	-0.196 (0.555)	-1.397 (0.514)*	-0.394 (0.619)
Treatment Zap	0.259 (0.344)	0.238 (0.337)	0.365 (0.203)	-0.435 (0.941)	-0.821 (-1.13)	-0.347 (1.286)
Exceptional*Treatment Zap	-0.251 (0.358)	0.038 (0.296)	-0.32 (0.173)	-0.365 (0.711)	0.02 (-0.725)	-0.563 (0.989)
Difficulty*Treatment Zap	-0.172 (0.360)	-0.269 (0.360)	-0.369 (0.221)	-0.211 (0.626)	0.319 (-1.047)	-0.299 (1.091)
Disappointing*Treatment Zap	-0.247 (0.502)	-0.213 (0.435)	-0.373 (0.196)	-0.398 (0.662)	0.187 (-0.899)	-0.434 (1.069)
Constant	0.75 (0.252)*	1.066 (0.326)**	0.842 (0.263)**	2.072 (0.697)*	3.033 (0.898)**	2.194 (0.834)*
Observations	1018	785	1020	1018	785	1020

Notes: Each column is a regression of meetings or visits by ZAP head. The regressors include different categories of schools ("exceptional, difficulty, disappointing" with omitted category: satisfactory).

Schools can be categorized by net exit, exam pass rate or repetition rate, hence the corresponding columns.

Net exit can be understood as dropout rate. It is the difference between school enrollment at the beginning and at the end of the year, divided by initial enrollment.

Standard errors reported in parenthesis, clustered at the CISCO level

* significant at 5%; ** significant at 1%

All regressions control for ZAP head's years of experience, experience squared, and having an advanced teaching degree.

Treatment Zaps are those in treatment groups 2 and 3. Year 1 refers to the first year of the interventions, 2005-2006.

Table 7: Treatment Effects on Student Attendance

<i>Dependent variables</i>	Endline	<i>Regression (1)</i>			<i>Regression (2)</i>		Obs.
	C.Group				Any		
	Mean	Treatment 3	Treatment 2	Treatment 1	Treatment 3	Treatment	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Student attendance (in %)							
Surveyor reported	80.113 (1.293)	-1.156 (4.583)	-1.94 (3.645)	-1.508 (3.484)	0.567 (1.922)	-1.725 (3.524)	945
Classroom count (classroom-level)	85.333 (0.492)	2.394 (2.331)	1.897 (2.067)	1.662 (2.410)	0.613 (0.777)	1.783 (2.195)	5097
Director's book record	86.589 (0.784)	4.29 (1.96)*	1.67 (2.310)	3.12 (2.050)	1.895 (0.643)**	2.397 (2.155)	901

Notes: Col. (1) reports the endline mean in the comparison group.

Standard errors in parentheses are clustered at the CISCO level. Observations are weighted by the probability of selection.
+ significant at 10%; * significant at 5%; ** significant at 1%

All regressions include controls for the month during which the school was surveyed, and index for school infrastructure
Treatment 1 is an indicator for whether the school is in a treated CISCO. Treatment 2 is an indicator for whether school is in a treated ZAP. Treatment 3 is an indicator for whether school itself received AGEMAD interventions.

Table 8: Treatment Effects on Repetition, Dropouts and CEPE Results

<i>Dependent variables</i>	Endline	<i>Regression (1)</i>			<i>Regression (2)</i>		Obs.
	C.Group				Any		
	Mean	Treatment 3	Treatment 2	Treatment 1	Treatment 3	Treatment	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: End-year 1 school performance							
CEPE exam pass rate	0.691 (0.016)	0.004 (0.053)	0.039 (0.049)	0.061 (0.048)	-0.046 (0.020)*	0.050 (0.047)	1113
Net exit from Grade 5	0.087 (0.008)	-0.076 (0.019)**	-0.063 (0.020)**	-0.048 (0.019)*	-0.02 (0.019)	-0.056 (0.018)**	1113
Net exit from all grades	0.061 (0.007)	-0.017 (0.015)	-0.016 (0.012)	-0.024 (0.013)+	0.003 (0.006)	-0.02 (0.012)+	1113
CEPE exam attendance rate	0.899 (0.010)	0.05 (0.030)+	0.038 (0.030)	0.033 (0.028)	0.014 (0.02)	0.036 (0.027)	1113
Average Effect (in sd) (sign of net exit reversed)		0.174 (0.089)+	0.183 (0.088)*	0.206 (0.090)*	-0.021 (0.038)	0.195 (0.085)*	1113
Panel B: Grade repetition rates over time							
Year 1 grade repetition	0.220 (0.006)	-0.002 (0.022)	0.012 (0.022)	0 (0.020)	-0.008 (0.008)	0.006 (0.020)	1118
Year 2 grade repetition	0.226 (0.007)	-0.051 (0.019)*	-0.026 (0.019)	-0.045 (0.020)*	-0.016 (0.007)*	-0.035 (0.019)	1085

Notes: Col. (1) reports the endline mean in the comparison group.

Net exit can be understood as dropout rate. It is the difference between school enrollment at the beginning and at the end of the year, divided by initial enrollment.

Standard errors in parentheses are clustered at the CISCO level. Observations are weighted by the probability of selection.
+ significant at 10%; * significant at 5%; ** significant at 1%

CEPE is the primary cycle exam at the end of Grade 5

Treatment 1 is an indicator for whether the school is in a treated CISCO. Treatment 2 is an indicator for whether school is in a treated ZAP. Treatment 3 is an indicator for whether school itself received AGEMAD interventions.

Year 1 refers to the first year of the interventions, 2005-2006. Year 2 is 2006-2007.

Table 9: Treatment Effects on Endline Student Test Scores

	<i>Regression (1)</i>			<i>Regression (2)</i>	
	Treatment 3	Treatment 2	Treatment 1	Treatment 3	Any Treatment
	(1)	(2)	(3)	(4)	(5)
Dependent variables: Normalized endline test scores					
Total Score	0.095 (0.11)	0.012 (0.09)	-0.013 (0.10)	0.095 (0.050)+	0 (0.10)
French	0.035 (0.11)	-0.015 (0.10)	-0.017 (0.10)	0.051 (0.05)	-0.016 (0.10)
Malagasy	0.115 (0.10)	0.046 (0.09)	0.018 (0.10)	0.083 (0.040)*	0.032 (0.09)
Math	0.095 (0.10)	-0.006 (0.10)	-0.041 (0.10)	0.119 (0.047)*	-0.024 (0.10)

Notes: N= 21126. Test scores are normalized by subtracting the mean and divided by standard deviation of control group's baseline scores. Standard errors in parentheses are clustered at the Cisco level

Observations are weighted by the probability of selection.

+ significant at 10%; * significant at 5%; ** significant at 1%

Treatment 1 is an indicator for whether the school is in a treated CISCO. Treatment 2 is an indicator for whether school is in a treated ZAP. Treatment 3 is an indicator for whether school itself received AGEMAD interventions.

Appendix Figure 1: Sample School Report Card

MENRS

DIRESEB:

CISCO:

ZAP:

SCHOOL REPORT CARD: YEAR 2005

(Based on the information the school submitted in FFA 2004-2005)

SCHOOL:

SECTOR:

SCHOOL CODE:

ACADEMIC RESULTS

	Your school	Your CISCO
CP1		
CP2		
CE		
CM1		
CM2		
Average		

	Your school	Your CISCO
CP1		
CP2		
CE		
CM1		
CM2		
Average		




	Your school	Your CISCO
CM2 enrollment		
Present at CEPE exam		
Passed the CEPE		
CEPE pass rate		
% of CM2 passed CEPE		

STUDENTS AND TEACHERS





	Your school
Nb. Students	
Nb. Teachers	
Holding classes	
Civil service teachers	
FRAM	
Supplements	
Not holding class	
Total	

	Your school	Your CISCO
Nb. Students/teacher		
% of FRAM (contract) teachers		
% teachers not holding class		

WHERE DOES YOUR SCHOOL STAND IN TERMS OF PERFORMANCE AND RESOURCES?

	Your school relative to others in your CISCO
Rate of net exit	
Repetition rate	
CEPE results	

	Your school	Your CISCO
Rate of net exit		
Repetition rate		
CEPE results		
Nb. Students/teacher		

	Results	Nb. Students/teacher	School Category
	Better than CISCO average	Less than CISCO average	Exceptional
	Better than CISCO average	Better than CISCO average	Satisfactory
	Less than CISCO average	Less than CISCO average	In Difficulty
	Less than CISCO average	Better than CISCO average	Disappointing

Appendix Table 1: AGEMAD Tools for Use by Teachers

Objective	Activities	AGEMAD Tool			
		Code	Description	Theme	Periodicity
Provide overview of the year's curriculum and promote systematic lesson planning	Plan time use and calendar of lessons	E1A/PED	Bimonthly lesson plan for the entire school year	Pedagogy	At start of school year
	Prepare lessons Track pupils' reception of the lessons	E1B/PED	Weekly lesson plan for bimonthly periods	Pedagogy	At start of each bimonthly period
		E2/PED	Individual lesson plan	Pedagogy	Daily
Put focus on student learning	Prepare and mark tests Record test scores Identify and help lagging pupils	E3/EVA	Record of bimonthly test scores <i>(Cahier de notes de class)</i>	Evaluation	After each bimonthly test
Mobilize parental support for academic excellence	Inform parents about pupils' progress in school	E4/EVA	Individual pupil report card <i>(Bulletin individuel de notes)</i>	Evaluation	End of each bimonthly period
Reduce student absenteeism	Monitor pupils' attendance Detect possible attendance problems and their causes Take remedial action	E5/APP	Class attendance register	Time for learning	Daily
Reduce teacher absenteeism	Account for teacher absences Detect possible attendance problems and their causes Take remedial action	E6/APP	Teacher's personal leave/travel record card	Time for learning	Each occasion of absence

Appendix Table 2: AGEMAD Tools for Use by School Directors

Objective	Activities	AGEMAD tool			
		Code	Description	Theme	Periodicity
Ensure proper registration of pupils	Keep up-to-date register of pupils at the school	D1/ADM	<i>Registre matricule</i> (National Printer document)	Administration	Each time a new pupil registers
Reduce teacher absenteeism	Monitor presence of teachers	D2/APP	Attendance register	Time for learning	Daily
		D3/APP	Summary table of teacher absences	Time for learning	At end of each month
Reduce pupil absenteeism	Record summary of pupil attendance records Review attendance record and assess possible problems and their causes Take remedial action	D4/APP	Summary table of monthly record of pupil absenteeism	Time for learning	At end of each month
Focus on progress of student learning at the school	Call periodic staff meetings (<i>Conseil des Maîtres</i>) Track student learning, note strengths and weaknesses, and plan to implement remedial action as needed	D5/EVA	Summary of student test scores	Evaluation	At end of each bimonthly period
Improve school's internal efficiency	Use school's data on student flow to identify and address possible problems Ensure application of grade to grade transition criteria (i.e., <i>système de cours</i>) ^{a/}	D6/EVA	Student flow table	Evaluation	End of each school year
Strengthen school's-partnership with the local community	Organize meetings with parents and teachers, FAF and FRAM Sensitize parents of pupils at the school Work with parents to prepare a <i>contrat programme de réussite scolaire</i> (CPRS)	D7/PART	Community meeting form	Community relations	At each meeting with the community
Improve conditions for learning at the school	Assign teaching duties among staff, and allocate classrooms and teaching materials; and select teachers for in-service training.	D8/PED	Organization of pedagogical arrangements	Pedagogy	At start of school year
Improve performance of the school	Study and display in public area the school report card Update school's performance indicators Discuss the school report card with community and use it to inform development of a school improvement plan (<i>contract programme</i>) for implementation	D9/TDB	School report card (<i>Tableau de bord de l'école</i>)	Focus on results	At end of school year

a/ Automatic promotion between grades 1 and 2 and between grades 4 and 5.

Appendix Table 3: AGEMAD Tools for Use by ZAP Heads

Objective	Activities	AGEMAD Tool			
		Code	Description	Theme	Periodicity
Improve management of schools in ZAP	Keep up-to-date inventory of schools in the ZAP	Z1/ADM	List of all schools in the ZAP	Administration	At start of school year
Improve the education information system	Ensure timely and complete collection of school data on the census forms	Z2/STA	Checklist of data collection	Statistics	At start and end of school year
Improve pupils' access to pedagogical materials	Determine number of pupils in each functioning school to receive a school kit	Z3/STA	Enrollments by school and distribution of the school kits	Statistics	At start of school year
Strengthen teacher competencies	Take stock of in-service teacher training (<i>journées pédagogiques</i>)	Z4/FOR	<i>Aide mémoire</i> of in-service teacher training event	Training	After each training event
	Supervise and provide pedagogical support to teachers in schools Provide feedback to school director Track assimilation of training provided Identify potential areas of additional training to help teachers become more effective	Z5/PED	Pedagogical supervision and support form	Pedagogy	After each school visit
Improve management at school finances	Track expenditure against grants provided through the <i>Caisse Ecole</i> and <i>Caisse Cantine</i>	Z6/ADM	<i>Caisse Ecole</i> form	Administration	On each school visit
	Ensure public posting of the expenditures	Z7/ADM	<i>Caisse Cantine</i> form	Administration	On each school visit
Strengthen school-community partnership	Maintain relations with the community Visit/meet community partners for education	Z8/PART	Community meeting form	Community relations	On each school visit
Improve performance of schools in the ZAP	Take note of each school's progress from year to year Compare schools in the ZAP, identify lagging schools, and plan extra support and attention for such schools	Z9/TDB	ZAP report card (<i>Tableau de bord de la ZAP</i>)	Focus on results	At start of school year

Appendix Table 4: AGEMAD Tools for Use by CISCO Heads

Objective	Activities	AGEMAD Tool			
		Code	Description	Theme	Periodicity
Manage teacher transfers	Take stock of existing deployment of teachers Summarize and process transfer requests	C1/ADM	Teacher transfer master sheet	Administration	End of school year
Improve schooling conditions	Inform the Ministry about requirements for school kits and other school supplies Distribute school kits, textbooks and other supplies	C2/ADM	Pedagogical supplies form	Administration	Before start of school year
Improve management of school finances	Distribute the <i>Caisse Ecole</i> and <i>Caisse Cantine</i> grants Ensure follow-up on utilization of the grants	C3A/ADM	<i>Caisse Ecole</i> follow-up sheet	Administration	End of second bimonthly period
		C3B/ADM	<i>Caisse Cantine</i> follow-up sheet	Administration	End of second bimonthly period
		C4A/ADM	<i>Caisse Ecole</i> verification form	Administration	On occasion of each visit to ZAP
		C4B/ADM	<i>Caisse Cantine</i> verification form	Administration	On occasion of each visit to ZAP
Strengthen teacher competencies	Identify teacher training needs and organize training events	C5/FOR	Planning sheet for in-service teacher training (<i>journées pédagogiques</i>)	Training	Before each training event
	Implement post-training follow-up	C6/FOR	Follow-up sheet for in-service teacher training	Training	Before and after each training event
	Plan school visits to provide in-service support, giving priority to the schools in difficulty Provide feedback to school personnel based on systematic on-site recording of classroom practices	C7/PED	Class observation grid (<i>Grille d'observation de classe</i>)	Pedagogy	On each school visit
Improve performance of schools in CISCO	Evaluate outcomes across ZAPs and schools and identify lagging units for extra support and attention Analyze possible sources of difficulty in lagging units and plan remedial action for implementation	C8/TDB	CISCO report card (<i>Tableau de bord de la CISCO</i>)	Focus on results	At start of school year