

IMPROVING TEACHERS' PROFESSIONALISM ON AN EXPERIMENTAL BASIS

LASSIBILLE, Gérard*

Abstract

While education economists have paid considerable attention to the performance of the education sectors, they have placed little emphasis on the management of the pedagogical process. In the context of developing countries, there are clear signs of weak management and too many school personnel neglect tasks deemed essential for student learning. Using data from a large-scale experiment conducted recently in Madagascar, the article analyzes the variations in impact of a set of management interventions designed to improve the engagement at work of public primary school teachers. While school-level interventions raise the teachers' level of professionalism on average, the results show that there is considerable heterogeneity in the impacts of such interventions. By contrast, the results indicate that targeting district and subdistrict administrators have little effect, overall and no matter the teachers' characteristics.

Keywords: teachers; management; primary education; Madagascar

JEL Codes: I21; I28; C93

1. Introduction

Most low-income countries recognize that investing in education, particularly at the primary level, is the bedrock of economic and social development. In past decades governments have emphasized the expansion of enrollment, but as coverage has risen the problems of inadequate student learning have come to the fore (see, for example, Lockheed, Verspoor, and Associates 1991). The large disparities in performance across schools and the fact that so many schools lie well inside the production frontier are telltale signs of weak management not only of resource allocation but also of pedagogical processes inside the “black box” of individual schools and classrooms (Lassibille and Tan 2003).

*Lassibille, Gérard. Institut de Recherche sur l'Economie de l'Education y Centre National de la Recherche Scientifique. gerard.lassibille@u-bourgogne.fr.

The experiment was conducted in collaboration with Jee-Peng Tan and Cornelia Jesse. Esther Duflo provided invaluable help and guidance, particularly in setting up the experimental design for this impact evaluation. The Madagascar Ministry of Education, led at the time of this experiment by Tahinarinoro Razafindramary and Paul Randrianirina, as well as the staff of Aide et Action, a nongovernmental organization, helped in conceptualizing the workflow tools, implementing the experiment, and collecting the data. Pierre-Emmanuel Couralet, Trang Van Nguyen, and Erika Strand provided on-the-ground supervision and technical support during the experiment as well as Mathieu Laroche and Muriel Nicot-Guillourel. Sajitha Bashir, Benu Bidani, Robert Blake, Deon Filmer, Elizabeth King, Arianna Legovini, Robert Prouty, Lina Rajonhson, Patrick Ramanantoanina, and Venkatesh Sundararaman provided input, advice, and comments to guide the design of the experiment. The experiment received financial support from the World Bank and the governments of France, Ireland, Madagascar, and Norway as well as from the donor partners of Education for All Fast Track Initiative through the Education Program Development Fund. The author acknowledges funding from the Government of Andalusia (grants P09SEJ-4859 and SEJ-0157) for this paper. The opinions expressed herein are those of the author alone and should not be attributed to the institutions with which he is associated or to any of the governments listed above.

As part of its strategy to improve the situation, the government of Madagascar implemented the AGEMAD (*Amélioration de la Gestion de l'Éducation à Madagascar*) program on an experimental basis. This program is the first randomized experiment to evaluate the impact of interventions specifically designed to improve management of the teaching process in a developing country. The experiment was designed to tighten the management of pedagogical workflow processes at each point along the service delivery chain of the public primary education sector (teachers, school heads, subdistrict, and district officers) in Madagascar by (a) making explicit to the actors their responsibilities and supporting them with tools and procedures to accomplish their tasks; (b) inserting supervision and follow-up at key points in the administrative hierarchy; and (c) facilitating school-community interactions and promoting accountability for results with the use of school report cards.

While Lassibille et al. (2010) have estimated the mean impacts of the program on service providers' behaviors (teacher absenteeism and the way school personnel execute tasks deemed essential to the mission of managing for results) and on a wide range of school outcomes (student attendance, dropout and repetition rates, pass rates on the primary school leaving exam and student achievement test scores), Glewwe and Maïga (2011) have examined whether the impacts of the program on students' test scores varied by the type of teachers. This paper goes further by exploring heterogeneity in impacts of the AGEMAD program on teachers' behavior at work as a function of observable characteristics of teachers and schools. The results after two years of experimentation show that intensive and direct interventions succeeded overall in changing the behavior of teachers towards better management. However, new evidence in this paper show that the mean impacts of the program hide considerable heterogeneity with regard to teacher and school characteristics. To illustrate, direct interventions to schools, reinforced by interventions at the subdistrict and district levels, had a significantly larger impact among teachers teaching small groups of students as well as among those working in a favorable environment and had no significant effect among teachers placed in a multigrade setting. The results also show that the interventions had a larger impact on contract teachers and native teachers than on their counterparts, but the gains were not statistically significant. By contrast, interventions limited to the district and subdistrict levels proved to be largely ineffective overall, with no detectable differences in impact across teacher and school characteristics. These findings have important implications for replicating the AGEMAD interventions.

The remainder of the paper proceeds as follows. Section 2 describes the context of the education sector in Madagascar. Section 3 briefly explains the AGEMAD program and the design of the experiment. Section 4 presents the empirical methodology. Section 5 discusses the estimation results. Section 6 concludes with a discussion of the policy implications of the findings.

2. Sector context

In Madagascar, as in most other low-income countries, the education sector has recently received priority attention in the country's poverty reduction strategy, and the

government has begun to channel more resources to the primary education sector (Government of Madagascar 2004a). At the time of the experiment, primary education absorbed 59 percent of public resources allocated to education, 20 percent more than in 2002. In tandem, the government introduced several major reforms in the primary education sector: abolishing school fees, distributing school grants, providing pedagogical supplies to students, introducing new pedagogical approaches, and restructuring basic education.

The most visible signs of progress were a large increase in coverage. At the time of the experiment, the education system enrolled some 3.8 million pupils in both public and private schools—more than twice the enrollment in 1996. The number of primary school teachers increased dramatically in the public sector, from some 28,000 teachers in 1996 to 59,000 in 2006. More than 4,000 new primary public schools were created during the last 10 years.

While there has been substantial progress in coverage, enormous challenges and problems remain as far as the performance of the education system is concerned. Several observations are particularly noteworthy in this regard. While entry rates to grade one are nearly universal, less than half of each cohort of students reaches the end of the five-year primary cycle. Despite government interventions, repetition rates are still uniformly high throughout the primary cycle, averaging about 21 percent in the public sector. With regard to student learning, the PASEC (Programme d'Analyse des Systèmes Educatifs de la Confemen) survey sponsored by Confemen (Conférence des Ministres de l'Education des Pays ayant le Français en Partage) indicates that a large number of public primary school students are not achieving the curriculum objectives, with children in grade five correctly answering only half of the test items in Malagasy and mathematics and around 30 percent in French (PASEC 2007). The fact that public schools with comparable resources achieve widely different outcomes suggests that many aspects of the pedagogical process in schools and classrooms are poorly managed and that tasks essential for student learning are often neglected (Lassibille and Tan 2003).

Results from a specific tools survey conducted in the framework of this impact evaluation are particularly telling in this regard. The data show that pupil absenteeism is poorly monitored by teachers, with attendance taken, on average, only 13 days a month. Essential pedagogical tasks are often neglected: 14 percent of teachers do not use weekly lesson plans; teachers prepare daily lesson plans for a very small portion of the school year, covering an average of 34 days out of a total of 150 days; only 15 percent consistently prepare daily and weekly lesson plans. Students' academic progress is poorly monitored too: the results of tests and quizzes are badly recorded, communication from teachers to parents on student learning via student report cards is often perfunctory, and pupil absences are rarely communicated to parents. The general impression that emerges from these observations is a lack of organization, control, and accountability—all of

which can compromise the performance of the system and the chances of success for the many ongoing reforms.¹

3. Description of the AGEMAD program and design of the experiment²

Madagascar has an extensive network of public primary schools that are funded and managed by centrally appointed civil servants. At the time of the experiment, the network of some 18,000 public primary schools—more than 90 percent of them serving rural populations—enrolled 3.1 million pupils and employed 18,000 school directors and 59,000 teachers. These schools were supervised by 1,500 subdistrict officers (*chefs ZAP*) who provide the first line of administrative and pedagogical support to the schools assigned to their care. These officers report to one of 111 district administrators (*chefs CISCO*) who, in turn, report to the Ministry of Education. The subdistrict and district administrators perform various administrative duties. These include, among other things, supervising teaching and learning practices, implementing in-service training for teachers and school directors, distributing paychecks to school personnel, organizing school building and maintenance projects, and overseeing the distribution of books and small grants to schools.

The AGEMAD program seeks to tighten the management of pedagogical workflow processes at each point along the service delivery chain. For each set of targeted actors, the AGEMAD interventions consist of (a) providing operational tools, report cards, and related instructional manuals, (b) holding meetings between school staff and the community to develop and agree on a school improvement plan, and (c) sponsoring training sessions following a set agenda that informs and motivates teachers, school directors, and subdistrict and district administrators in working towards better performance.

The set of operational tools and processes for each actor focuses attention on six broad domains of activity: pedagogy, student learning and follow-up, management of instructional time, administration, school statistics, and partnership with the local community. More than 30 tools were developed for tasks considered essential for a well-functioning system. Among them, seven were designed for use by teachers, eight by school principals, eight by subdistrict administrators, and nine by district administrators.³ Each operational tool encourages superiors to pay closer attention to the work of his or her direct subordinates (for example, requiring school directors to review information supplied by a teacher and suggest specific follow-up) and is accompanied by a manual

¹ A myriad of demand-side conditions also explains the low performance of the primary education system, among them household poverty, community characteristics, the opportunity cost of children's time, or the parental perceptions of the value of education.

² For more details, see Lassibille et al. (2010).

³ For details on the tools, their presentation, their purpose, and how they are to be used by the various actors, see Government of Madagascar (2004b).

that explains how and when each tool should be used. Training modules were specifically customized and offered to all relevant personnel.

Report cards, with customized information for each school, subdistrict, or district, were prepared using data from the annual school census. At the school level, they contain a limited set of performance indicators, information on enrollment and resource endowments, and selected comparative data showing a school's ranking on performance and resource endowments relative to that of other schools. At the subdistrict and district levels, they contain aggregate information on the same indicators as the school report cards and list the schools or subdistricts according to their performance. At the district and subdistrict levels, the report cards can serve as an instrument for self-evaluation and as a basis for designing and implementing an action plan for improvement. At the school level, they can play a useful role by providing key information to clarify and motivate decision making. For this reason, the interventions also included structured school meetings with the staff of the school, parents, and community members in order to engage the school and the community in purposeful discussion towards the development of school improvement plans based on the school report cards.

The experiment was designed to answer three important questions: (a) Do the AGEMAD interventions improve the functioning of schools and their performance in terms of key outcome indicators? (b) Would interventions at the district and subdistrict levels have an impact in the absence of school-level interventions? and (c) Would all schools in a subdistrict require direct interventions to achieve the desired impact, or could they simply learn from demonstration schools?

These questions were addressed by implementing a nested randomized experiment. The intervention design of the experiment, with the relevant number of observations, is shown in Figure 1. First, 30 relatively accessible and similar districts were selected as the universe for the experiment. Among them, 15 districts were randomly chosen to receive the interventions, while the remaining 15 districts continued to be managed as usual. The 30 districts combine a total of 573 subdistricts and 6,488 schools. For practical reasons, the most remote and inaccessible subdistricts were removed (107 subdistricts and 2,714 schools), leaving 466 subdistricts and 3,774 schools, with 259 subdistricts and 2,053 schools in the intervention districts and 207 subdistricts and 1,721 schools in the control districts. From the 3,774 schools, a follow-up sample of 1,212 schools in 179 subdistricts was randomly selected to evaluate the impact of the interventions.

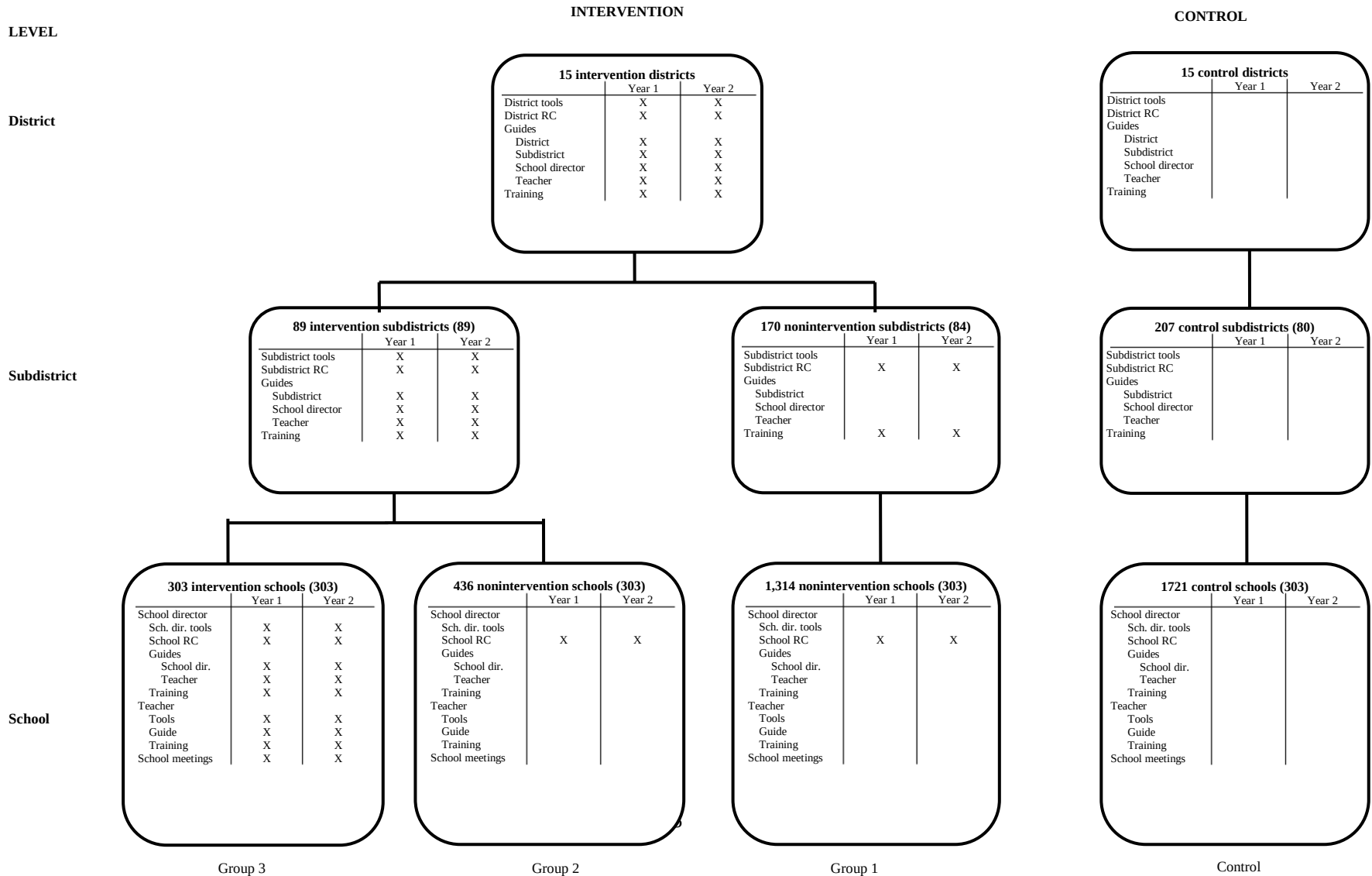
In the intervention districts, the relevant administrators received operational tools, guidebooks explaining the use of the tools, and training relevant to their tasks, along with a district report card. Within the intervention districts, subdistricts were randomly selected to receive operational tools relevant to subdistrict-level functions, corresponding guidebooks, specific training, and a subdistrict report card. The other subdistricts received only training and a report card; schools in these subdistricts received only a report card. A sample of schools in the intervention subdistricts was randomly selected to receive the operational tools designed for use by school heads and teachers, corresponding guidebooks, training, and a school report card to be discussed in meetings with the

community. For the other schools, subdistrict administrators received school report cards for onward distribution and templates of the school-level tools for replication.

To summarize, the nested design of the experiment created four types of schools: schools in control districts receiving no intervention at any level in the system (control group), schools in intervention districts and in non-intervention subdistricts receiving no school-level interventions (group 1), schools in intervention districts and subdistricts receiving no school-level interventions (group 2), and schools in intervention districts and subdistricts directly receiving school-level interventions (group 3). The diffusion of interventions to schools in groups 1 and 2 was left to the initiative of the subdistrict or district administrator, following the training-of-trainers model used extensively in Madagascar (*formation en cascade*). By contrast, schools in group 3 received the interventions directly, and these school-level processes were reinforced at the subdistrict and district levels. The comparison of these four groups of schools makes it possible to answer key questions about the interventions. Comparing group 1 and the control group reveals the effect of a training-of-trainers or “cascade” model that involves only the district-level administrators in implementing the program. Comparing group 2 and the control group reveals the impact of a more intensive form of the cascade model that involves both subdistrict- and district-level interventions. Comparing group 3 with the control group reveals the maximum impact of the full set of interventions at the school, subdistrict, and district levels. Comparing groups 2 and 3 reveals the impact of the most intensive package of interventions, which directly targeted teachers and school directors along with subdistrict and district administrators, over the second-most intensive package, which directly targeted only subdistrict and district administrators.

Because schooling processes take time to evolve, the experiment took place over two school years ending in June 2007. To assess the impacts of the program on behaviors, a school questionnaire was administered during randomly timed and impromptu visits to the schools in the follow-up sample. The survey provided detailed information on the pedagogical and administrative organization of the schools and tracked teachers’ behaviors to see how they perform their job. It also included separate sections to gather individual data on teacher characteristics and basic skills and to check teacher and student absenteeism. The survey was conducted at the beginning and end of the experiment; it gathered information on some 4,000 teachers in the 1,200 schools involved in the impact evaluation.

Figure 1: Description of the interventions and experimental design



4. Empirical strategy

Following Lassibille et al. (2010), this article focuses on the tasks that Malagasy educators deem essential to the mission of managing for results. Seven tasks are considered essential for teachers' role (Government of Madagascar 2004b). They include taking daily roll call, preparing daily and bimonthly lesson plans, monitoring student learning, testing pupils, helping lagging students, and discussing student learning issues with the director. A "good" teacher (or more accurately, a minimally conscientious teacher) is defined here as one who performs all seven workflow tasks that Malagasy educators consider essential to the role.

The impact of the AGEMAD interventions on the behavior of school personnel, at the end of the two-year evaluation period, can be estimated by running the following regression:

$$y_i = \alpha + \sum_{j=1}^3 \beta_j g_{ij} + \varepsilon_i \quad (1)$$

where y_i represents the outcome variable for individual i , which takes the value of 1 if the teacher executes all the tasks that are considered essential to his role and 0 if not;⁴ g_{1i} is a dummy variable indicating whether teacher i is in a school that received interventions via district administrators only; g_{2i} is an indicator for whether teacher i is in a school that received interventions via subdistrict and district administrators; and g_{3i} is a dummy indicating whether teacher i is in a school where interventions at the subdistrict and district levels were combined with direct interventions to teachers and school director. Because treatments were random, the intervention variables g and the error term ε are not correlated, and, consequently, equation (1) can be estimated using ordinary least squares (OLS). In this randomized setting, the coefficients of interest β are interpretable as the average causal impacts of different levels of the AGEMAD treatment, and α measures the average value of the outcome variable in the control group. A similar formulation is used to evaluate the pairwise impacts of the three interventions relative to each other.⁵

To possibly increase the precision of the estimated impact of the interventions, an alternative model including other explanatory variables to the standard model (1) is also run:

$$y_i = \alpha + \sum_{j=1}^3 \beta_j g_{ij} + \gamma X_i + \varepsilon_i \quad (2)$$

where X_i is a vector of school and teacher characteristics x_{ij} .

⁴ An alternative approach is to follow the methodology of Kling, Katz, and Liebman (2007), who analyze program effects using a seemingly unrelated regression for each task—either separately or in clusters of tasks. This approach was considered for this article, but ultimately the judgment of Malagasy educators that the individual tasks are an integrated package of closely connected actions required for managing the teaching and learning process was accepted.

⁵ Lassibille et al (2010) use a similar model to estimate the mean impacts of the program. However, their results do not account for seasonal variation (see section 5).

Finally, in order to measure the heterogeneous impacts of the AGEMAD program, the three intervention variables are interacted with each school and teacher characteristic in the following:

$$y_i = \alpha + \sum_{j=1}^3 \beta_j g_{ij} + \eta C_i + \sum_{j=1}^3 \delta_j g_{ij} C_i + \gamma X_i + \varepsilon_i \quad (3)$$

where C_i denotes an individual or a school characteristic that is expected to change the impact of the program on the outcome variable.

For a dummy variable C , the interpretation of the coefficients is as follows: η is an estimate of the difference in the outcome between individuals in the control group who are given a value of 1 for variable C and those who belong to the benchmark group ($C = 0$); β_j is an estimate of the impact of treatment j for individuals in the benchmark (base) group, the corresponding effect for those who are assigned a value of 1 for C is given by the sum of the coefficients $\beta_j + \delta_j$; if δ_j is significantly different from 0, there is evidence of heterogeneity of treatment effects between the two subgroups of individuals. For a continuous variable C , α measures the average value of the outcome variable in the control group when $C = 0$; η represents the marginal change in the outcome variable for the control group, resulting from a marginal variation in C ; β_j is the estimated impact of treatment j when C has a value of 0; the interaction term δ_j represents the marginal change in the impact of treatment j resulting from a marginal increase in C .

Summary statistics for the variables under consideration are reported in Table 1 jointly with t-tests of the equality of means across treatment groups. The results confirm that teachers in the treatment and control groups do not differ systematically, as expected given random assignment. Furthermore, results not shown here to save space indicate that teacher and school characteristics are moderately correlated with each other, whatever the treatment group.

As shown in Table 1, the gender composition of the sample is well balanced, with about two-thirds of female teachers in each group of schools. Attachment to the local community is measured by a dummy indicating whether the teacher was born in the region where the school is located. Table 1 shows that about 75 percent of the teachers share a strong attachment to the local community. Several studies have argued that motivation and incentives to help the community are probably higher among local teachers because of their ties to the local area and greater sense of accountability (see, for example, Alcázar et al. 2006). Teachers who have maintained a firm attachment to the local community are therefore expected to be more receptive than others to interventions designed to improve the way schools in their native area function.

Table 1: Teacher and school characteristics by treatment group

	Control group ^a	Differences with respect to control group ^b			
		Direct to teachers and school directors (group 3)	Via subdistrict and district administrators (group 2)	Via district administrators (group 1)	Number of observations
Male	0.334	-0.043 (0.029)	-0.060* (0.035)	-0.020 (0.032)	4,110
Attachment to local community	0.765	-0.090 (0.094)	-0.079 (0.089)	-0.076 (0.089)	4,110
Regular teacher	0.493	0.053 (0.041)	0.020 (0.034)	0.020 (0.043)	4,110
Multigrade teacher	0.219	-0.011 (0.032)	-0.001 (0.037)	0.036 (0.032)	4,110
Number of students	39.987	-0.851 (3.017)	-1.858 (2.737)	-2.558 (2.782)	4,110
Index of school conditions	103.392	-3.273* (1.612)	-2.006 (1.841)	-3.025 (1.790)	4,110
Parents' wealth index	99.387	1.352 (1.246)	1.336 (1.175)	0.901 (1.063)	4,110

* significant at the 10 percent level.

Note:

a. Average; observations are weighted by the probability of selection of the school.

b. Mean differences across groups for all teachers in the follow-up sample. Numbers in parentheses are standard errors. Observations are weighted by the probability of selection of the school; standard errors are clustered at the CISO level.

c. Interventions at the school level are combined with those at the subdistrict and district levels.

Source: Author's analysis of the 2007 AGEMAD school survey.

Based on their job status, there are two broad categories of teachers in Madagascar. The first group consists of civil service teachers who are recruited and paid by the government. The second group consists of FRAM teachers.⁶ In past years, FRAM teachers were hired by the communities to respond to the lack of publicly paid teachers in the country; they were paid entirely by the parents' association through fees paid by pupils' families. Since 2002–03, the government has contributed to the pay of FRAM teachers. Today, these teachers make up 50 percent of the teaching staff at the primary level, compared to only 5 percent in 1998. While contract teachers in Madagascar have more years of formal schooling than civil service teachers on average, they have little prior exposure to teacher training programs, and their professional qualifications are generally lower. They are paid considerably less than their counterparts for doing the same job. Their monthly salary is 60 percent below that of civil service teachers on

⁶ FRAM is the acronym for Fikambanan'ny Ray Amandrenin'ny Mpianatra, the school-based parent-teacher association.

average, and they are only paid for 10 months out of 12.⁷ Contract teachers may be motivated to work harder in order to increase their chance of having their contract renewed and eventually to being integrated into the regular teaching force. In this regard, the AGEMAD program could favorably affect the behavior of contract teachers by giving them the opportunities to perform better at work and thus to improve their job security and their promotion opportunities.

Multigrade teaching is common in Madagascar, with more than 50 percent of the public primary schools practicing it (World Bank 2002). Multigrade teachers represent about 20 percent of the teachers in the experimental data (Table 1). The set of AGEMAD operational tools and processes for most of the tasks for which teachers are responsible were not designed specifically to tighten the management of pedagogical workflow processes in multigrade situations. For this reason, multigrade teachers could be less responsive to the interventions than their counterparts, and improvements in teachers' behavior could be quite different between teachers teaching a multigrade class and those teaching a normal class.

By increasing monitoring, control, and follow-up actions, the AGEMAD interventions inevitably increase the workload of teachers, particularly those who are responsible for a large number of students. As a consequence, teachers teaching large classes may have less incentive to respond to the AGEMAD interventions and thus to change their behavior at work compared to teachers teaching small groups. According to Table 1, a Malagasy teacher is teaching about 40 pupils on average; however, results not shown here to save space indicate that this average pupil-teacher ratio hides substantial heterogeneity, with a standard deviation of around 17 within each group of interventions.

School conditions are measured via an asset-based index. This index is constructed using principal components analysis based on the physical conditions of the school: the structure is permanent, the number of classrooms is sufficient, and the school is equipped with electricity, water, latrines, and chairs for all pupils. The index ranges from about 126 in schools with all of these features to 75 in those with none of them. Several studies have shown that adequate working conditions are a powerful determinant of teachers' commitment (Firestone and Pennell 1993). In light of this result, one can reasonably expect that teachers may respond more positively to the treatments when physical conditions are adequate and may be more likely to change their behavior at work than teachers working in an unfavorable environment.

Parents' wealth is defined as mean parental wealth in the school and is derived from data collected on a randomized sample of students in each school. The wealth index is constructed using principal components analysis (see, for example, Filmer and Pritchett 1998). It is based on a family's ownership of the following assets: refrigerator, tap water, electricity, television, radio, phone, and fan. The index ranges from around 132 for a family who owns all the items to around 96 for a family who owns none of them.

⁷ For more details on FRAM teachers, see Government of Madagascar (2008); World Bank (2002).

According to Table 1, the treatment groups are not statistically different from the control group in terms of wealth; results not shown to save space indicate that the standard deviation of the wealth index is large and similar across groups (around 1.46 in each group), indicating an important heterogeneity of wealth across communities. As explained above, AGEMAD seeks to motivate parental participation in school management. The program facilitates school-community interactions and promotes accountability through school report cards that are the basis for meetings to discuss school improvement plans. Because wealthier parents may be better motivated to attend the meetings, better equipped to use the information provided by the report cards, and, presumably, better able to monitor school activities and to motivate teachers to change their teaching style and behavior at work, the impact of the interventions could be concentrated largely among better-off communities where literacy rates are higher.

5. Empirical results

Equations (1), (2), and (3) are estimated using OLS, with standard errors clustered at the district and subdistrict levels since treatments were implemented at these levels. Observations are weighted by the probability of selection of the school. All the regressions are controlled for the month of random school visits to account for seasonal variation.⁸

In order to make the presentation easier to follow, Table 2 shows the average impact of the treatments calculated from equations (1) and (2) and the heterogeneous impacts derived from equation (3). For each categorical variable C , the table gives the discrete first difference from the base category and shows the corresponding interaction effect. For continuous variables C , the marginal impacts and the interaction effects are calculated at the sample mean of the variable and at 0.5 standard deviation above the sample mean.

The results are presented in two parts. The first pertains to the impacts of each set of interventions on the behavior of school personnel with respect to the control group. The second pertains to the pairwise impacts of the three treatments relative to each other.

a) Impacts of each set of interventions with respect to the control group

The first three columns of Table 2 show the impacts of each set of interventions with respect to the control group. The first row reports the impacts estimated from equation (1), that is, without control variables. The results show that interventions at the school level, reinforced by interventions at the district and subdistrict levels (that is, treatment group 3), succeeded in changing the behavior of school personnel towards better management. According to Table 2, teachers in schools where the interventions targeted the entire chain of service delivery were 21 percent more likely to execute all the tasks deemed essential for good classroom management than teachers in the control group, and the difference was statistically significant at the 10 percent level. In schools where the interventions percolated down to schools indirectly through administrators at the

⁸ The full regression estimates are available from the author upon request.

subdistrict and district levels (that is, schools in groups 2 and 1), the impacts on the extent to which teachers performed their essential tasks were 11 and 12 percent, respectively, but were not statistically significant.

Table 2: Impacts of the interventions on teachers' behaviors at work

	Differences with respect to control group			Differences between intervention groups		
	Direct to teachers and school directors ^a (group 3)	Via subdistrict and district administrators (group 2)	Via district administrators (group 1)	Group 2 versus group 1	Group 3 versus group 1	Group 3 versus group 2
Mean impacts						
Not controlled for (model 1)	0.2088* (0.112)	0.1081 (0.109)	0.1175 (0.108)	-0.0020 (0.042)	0.0895** (0.040)	0.0961*** (0.036)
Controlled for (model 2)	0.2201** (0.106)	0.1183 (0.106)	0.1319 (0.103)	-0.0069 (0.041)	0.0871** (0.037)	0.0995*** (0.036)
Heterogeneity in impacts (model 3)						
Gender						
Female	0.1998* (0.106)	0.1150 (0.107)	0.1566 (0.104)	-0.0334 (0.044)	0.0449 (0.041)	0.0857** (0.039)
Male	0.2688** (0.112)	0.1262 (0.114)	0.0777 (0.103)	0.0570 (0.051)	0.1842*** (0.046)	0.1342*** (0.045)
Male-female	0.0689 (0.046)	0.0112 (0.054)	-0.0789** (0.037)	0.0904* (0.047)	0.1394*** (0.047)	0.0485 (0.042)
Attachment to the local community						
Non-local teacher	0.1340* (0.070)	-0.0253 (0.082)	0.0196 (0.074)	-0.0393 (0.062)	0.1134* (0.062)	0.1594*** (0.051)
Local teacher	0.2474** (0.124)	0.1717 (0.122)	0.1710 (0.120)	0.0077 (0.048)	0.0749* (0.044)	0.0712 (0.045)
Local teacher–non-local teacher	0.1134 (0.104)	0.1970* (0.111)	0.1514 (0.111)	0.0470 (0.071)	-0.0386 (0.074)	-0.0882 (0.065)
Employment status						
Contract teacher	0.2553** (0.114)	0.1475 (0.114)	0.1125 (0.111)	0.0470 (0.050)	0.1416*** (0.040)	0.1048** (0.045)
Civil service	0.1878* (0.104)	0.0890 (0.103)	0.1488 (0.098)	-0.0576 (0.046)	0.0382 (0.047)	0.0947** (0.040)
Civil service – contract teacher	-0.0675 (0.052)	-0.0584 (0.045)	0.0362 (0.045)	-0.1047** (0.050)	-0.1033** (0.048)	-0.0100 (0.044)
Multigrade teacher						
Monograde teacher	0.2511** (0.103)	0.1272 (0.104)	0.1532 (0.099)	-0.0178 (0.045)	0.0973** (0.042)	0.1226*** (0.040)
Multigrade teacher	0.1074 (0.126)	0.0881 (0.123)	0.0622 (0.118)	0.0288 (0.065)	0.0529 (0.059)	0.0146 (0.061)
Multigrade – monograde teacher	-0.1436** (0.061)	-0.0391 (0.060)	-0.0911** (0.044)	0.0466 (0.069)	-0.0444 (0.064)	-0.1080 (0.069)
Number of pupils						
At sample mean	0.2209** (0.104)	0.1187 (0.104)	0.1346 (0.102)	-0.0079 (0.041)	0.0853** (0.038)	0.0997*** (0.036)
At sample mean + 0.5 sd	0.1805* (0.108)	0.0782 (0.110)	0.1174 (0.107)	-0.0258 (0.045)	0.0631 (0.043)	0.1035** (0.042)
0.5 sd change from sample mean	-0.0404*** (0.016)***	-0.0406** (0.019)	-0.0172 (0.021)	-0.0179 (0.0151)	-0.0222 (0.019)	0.0038 (0.017)

Table 2: Impacts of the interventions on teachers' behaviors at work

	Differences with respect to control group			Differences between intervention groups		
	Direct to teachers and school directors ^a (group 3)	Via subdistrict and district administrators (group 2)	Via district administrators (group 1)	Group 2 versus group 1	Group 3 versus group 1	Group 3 versus group 2
Index of school conditions						
At sample mean	0.2194** (0.104)	0.1123 (0.104)	0.1282 (0.102)	-0.0072 (0.040)	0.0872** (0.037)	0.1000*** (0.034)
At sample mean + 0.5 sd	0.2574** (0.110)	0.1008 (0.111)	0.1561 (0.110)	-0.0449 (0.044)	0.0970** (0.040)	0.1495*** (0.038)
0.5 sd change from sample mean	0.0380** (0.018)	-0.0115 (0.016)	0.0279 (0.020)	-0.038* (0.020)	0.0098 (0.019)	0.0495*** (0.020)
Parents' wealth index						
At sample mean	0.2213** (0.104)	0.1219 (0.104)	0.1339 (0.101)	-0.0063 (0.041)	0.0871** (0.037)	0.0992*** (0.036)
At sample mean + 0.5 sd	0.2344** (0.099)	0.1182 (0.101)	0.1508 (0.094)	-0.0279 (0.048)	0.0851* (0.045)	0.1162*** (0.039)
0.5 sd change from sample mean	0.0131 (0.027)	-0.0037 (0.026)	0.0168 (0.022)	-0.0216 (0.021)	-0.0020 (0.021)	0.0171 (0.015)

*** significant at the 1 percent level; ** significant at the 5 percent level; * significant at the 10 percent level.

Note: Numbers in parentheses are standard errors.

a. Interventions at the school level are combined with those at the subdistrict and district levels.

Source: Author's analysis of the 2007 AGEMAD school survey.

Adding teacher and school covariates to equation (1) does not significantly change these patterns of impacts and increases only very slightly the precision of the estimated impacts. According to row 2 of Table 2, controlling for teacher and school characteristics in treatment groups 1 and 2 resulted in an increase by about 1 percentage point in the probability of performing the essential tasks, but the estimated impacts of the interventions targeting only subdistrict and district administrators over the control group—13 and 12 percentage points, respectively—were still statistically indistinguishable from 0. In schools where the interventions targeted the entire chain of service delivery, the estimated impact on teachers' behavior increased from 20.9 to 22.0 percent, and the significance level rose from 10 to 5 percent.

Overall the interventions changed behavior only when targeting the entire chain of service delivery, with schools benefiting directly from the school-level interventions and indirectly through the interventions at the subdistrict and district levels. Even in this case, there is no room for complacency. After benefiting from the interventions for two years, the share of teachers who execute their duties conscientiously in this group was still only 60 percent, perhaps reflecting the fact that it takes time to change behavior. In schools where the interventions cascaded down to schools indirectly through the interventions at the subdistrict and district levels, the impact on teachers' behavior was limited. The

reasons are various, among them possibly the following: (a) lack of interest and motivation on the part of subdistrict and district administrators, which may be linked to their heavy workload and their low level of technical competence; (b) weak mechanisms for monitoring and supervising schools by subdistrict administrators, who are most directly responsible for this part of the workflow in the service delivery chain; and (c) absence of a true leadership culture among both subdistrict and district officers. Some observations are particularly telling in this regard. In districts and subdistricts where only the administrators received the AGEMAD interventions, practically none of the schools received through these administrators the workflow templates designed to tighten management of the pedagogical processes; the teachers in these schools also received no training or support from the officers;⁹ and in three-quarters of the schools, school report cards were never discussed with the communities.

Is there evidence of systematic variation in impact on teachers' behavior in AGEMAD? To answer this question, the lower panel of Table 2 displays the results for the extent to which program impacts varied with teachers and schools characteristics, obtained by estimating equation (3). The estimates show that male teachers in group 3 experienced a larger impact of the interventions than female teachers, although the difference by gender was not statistically significant. By contrast, the estimates show no detectable differences in impact among male and female teachers in schools where the interventions cascaded down to schools indirectly through subdistrict and district officers, compared with the control schools. As far as attachment to the local community is concerned, Table 2 indicates that teachers in treatment group 3 who maintained a firm attachment to the local community were 25 percent more likely to perform all the tasks deemed essential for good classroom management, compared with teachers in the control group. Full interventions were less effective among non-local teachers, with the share of minimally conscientious teachers increasing only 13 percentage points, compared with the control group. However, the difference in impact between local and non-local teachers was not statistically significant at the end of the two-year evaluation period. By contrast, when the diffusion of the program was left to the initiative of subdistrict and district administrators, the estimates do not show significant differences in the probability of performing the essential tasks for local or non-local teachers, relative to their respective peers in the control group.

With regard to teachers' employment status, results show that contract teachers in the schools receiving interventions throughout the whole service delivery chain were 26 percent more likely to execute their duties than contract teachers in the control group, and the difference was significant at the 5 percent level. Regular teachers in such schools experienced a smaller impact. According to Table 2, the share of civil servants in treatment group 3 who performed all the essential tasks exceeded the corresponding share in the control group by only 19 percentage points relative to the control group, and the gain was statistically significant at the 10 percent level. While these results tend to show that contract teachers could respond to full interventions better than civil service teachers,

⁹ Only 3 percent of the schools in group 1 and 2 received a copy of the tools from the subdistrict administrators.

the 7 percentage point difference between the two categories of teachers was nevertheless statistically insignificant. In schools where the interventions were left to the initiative of subdistrict and district administrators, the results unequivocally indicate that after two years of experimentation neither civil service teachers nor contract teachers were more conscientious in executing their duties than their peers in the control group.

Another noteworthy feature in the results is that interventions targeting school personnel, as well as the subdistrict and district administrators who supervise them, had a significant impact on teachers teaching a normal class, but were ineffective in changing the behavior of multigrade teachers, relative to the control group. According to Table 2, such interventions boosted the share of good teachers by a significant 26 percentage points for teachers teaching a normal class and resulted in an insignificant 11 percentage point gain for multigrade teachers, with the difference in impact between these two categories of teachers being statistically significant. By contrast, the results indicate that neither of these two categories of teachers responded significantly to the interventions targeting only subdistrict and district administrators. These findings have important implications, as discussed later.

The pattern of impacts also shows that interventions at the school level reinforced by those at the subdistrict and district levels had their largest effect on teachers teaching small classes. To illustrate, results in Table 2 show that a 0.5 standard deviation increase around the mean in class size was associated with a significant 4 percentage points decrease in the probability of performing the essential tasks. Similarly, the effectiveness of direct interventions varied significantly depending on teachers' working environment, suggesting that changing the behavior of school personnel towards better management is easier in schools with good physical conditions. According to Table 2, a 0.5 standard deviation increase around the mean in the index of school conditions enhanced the probability of performing the essential tasks by a significant 3.8 percentage points among teachers in treatment group 3. Parental wealth in treatment group 3 had a more mitigated effect. Results in Table 2 indicate that the wealthier the community was, the more likely were teachers in treatment group 3 to perform the essential tasks, although the interaction term between the intervention dummy and the parental wealth index was not statistically significant. In contrast, the estimates do not show detectable differences in impact across class size, school conditions, and parental wealth among teachers in treatment groups 1 and 2.

b) Pairwise impacts of the three treatments relative to each other

Comparisons across the three treatment groups can be used to evaluate the existence of spillover effects that might arise from the diffusion of practices and knowledge by administrators with responsibility for multiple schools in different locations. The pairwise impacts of the three treatments relative to each other, estimated through equations (1), (2), and (3), are shown in the three right columns of Table 2.

The results of estimating the average impacts of the interventions with and without controls unequivocally show that teachers in treatment group 2 were not significantly

more conscientious in executing their duties than those in treatment group 1. The absence of a pattern suggests that the interventions that benefited subdistrict administrators did not reinforce the impact of the interventions that benefited district administrators. The results also indicate that the impact of treatment group 2 relative to treatment group 1 did not vary among the eligible population. The analysis of subgroups finds no detectable differences in program effectiveness based on teacher and school characteristics.

According to Table 2, teachers in treatment group 3 outperformed those in group 1. The specification in the first row of Table 2, which has no covariates, shows that direct and intensive interventions boosted the probability of performing the essential tasks by a significant 9 percentage points with respect to interventions involving only the district-level administrators (group 1). Controlling for teacher and school characteristics did not change the magnitude of this gain. The comparison between group 3 and group 2 is also interesting: the average impact of the full set of interventions over the second-most intensive package of interventions that directly targeted only subdistrict and district administrators (group 2), estimated with and without controls, amounted to a significant 10 percentage points. Otherwise stated, comparing groups 2 and 3 shows no evidence that the package of practices in schools that benefited directly from the interventions was diffused to other schools in the same subdistricts and districts under the supervision of the same administrators. This result implies that targeting some schools to serve as demonstration projects for replication by subdistrict and district administrators in the remaining schools under their care is unlikely to modify teachers' behavior at work. Disaggregating the average impact of treatment group 3 relative to treatment group 2 by teacher characteristics shows no significant difference in program effectiveness across gender, attachment to local community, employment status, or multigrade and monograde teachers. The estimates also show no detectable differences in impact across school characteristics, except for teachers working in schools with poor infrastructure. According to Table 2, a 0.5 standard deviation increase around the mean in the index of school conditions resulted in a significant 5 percentage point increase in the impact of treatment group 3 relative to treatment group 2.

6. Conclusions

The AGEMAD intervention program represents an important effort by the government of Madagascar to experiment with new ideas for addressing problems in the management of the primary education sector. As mentioned earlier, many aspects of the pedagogical process are poorly managed, and too many school personnel neglect tasks deemed essential for student learning. To illustrate, pupil absenteeism is poorly monitored by teachers, 20 percent of them do not prepare daily lesson plans; students' academic progress is poorly monitored, and communication from teachers to parents on student learning is often perfunctory. A large majority (around 60 percent) of the sample teachers do not consistently perform the package of seven tasks considered essential by Malagasy educators.

The randomized experiment was designed to streamline and tighten the management of pedagogical workflow processes at each point along the service delivery chain. The results show that direct and intensive interventions involving school-level personnel and subdistrict and district administrators changed the behavior of teachers towards more conscientious execution of the tasks for which they were responsible. Less costly interventions, targeting only subdistrict and district administrators—in the hope that they would take the initiative to disseminate and supervise implementation of the program—proved largely ineffective.

In addition to examining the mean impacts of the AGEMAD program, the paper analyzed impacts at the subgroup level by estimating the interactions between treatments and characteristics of interest. The results show that interventions left to the initiative of subdistrict and district administrators did not result in any significant variation in impacts across teacher and school characteristics. In this regard, the estimates show no detectable differences in the effectiveness of such interventions, based on teachers' employment status, attachment to local community, type of pedagogy, class size, working conditions, and wealth of pupils' parents.

In contrast, the results show that the average impact of intensive and direct interventions on teachers' behavior hides important heterogeneity. Interventions at the school level that were reinforced by those at the subdistrict and district levels had (a) a significantly greater impact on teachers teaching small classes and teachers teaching in schools in good physical condition and (b) no significant impact on teachers teaching multigrade classes. The results also show that, after two years of experimentation, contractual teachers experienced a greater impact of the intensive interventions compared to civil service teachers, although the difference between the two categories of teachers was not statistically significant. In the same vein, teachers with a strong attachment to the local community tended to be more conscientious in executing their duties and wealthier parents were better able to change the work behavior of teachers, although in both cases the gains were not significant.

These findings have important implications for policy development. Overall, scaling up the most effective intervention—that is, interventions at the school level, reinforced by interventions at the district and subdistrict levels—could succeed in changing the behavior of school personnel towards a more conscientious execution of the tasks for which they are responsible. However, the results from this paper clearly indicate that such interventions should initially target monograde schools only or schools with very few multigrade teachers. In the case of financial constraints, priority could be given to schools enrolling a reasonable number of students, with adequate equipment, and, to a minor extent, with a relatively large number of contract teachers.

Bibliography

Alcázar, L., Rogers, F.H., Chaudhury, N., Hammer, J., Kremer, M., and Muralidharan, K., (2006). Why are teachers absent? Probing service delivery in Peruvian primary schools. Mimeo. Washington, DC: World Bank.

Filmer, D., and Pritchett, L., (1998). *Estimating wealth effects without expenditure data—or tears: An application to educational enrollment in states of India*. Washington, DC: World Bank.

Firestone, W.A., and Pennell, J.R., (1993). Teacher commitment, working conditions, and differential incentive policies. *Review of Educational Research* 63(4), pages 489–525.

Glewwe, P., and Maïga, E., (2011). The impacts of school management reforms in Madagascar: Do the impacts vary by teacher type? *Journal of Development Economics* 3(4), pages 435–460.

Government of Madagascar, (2004a). *Rapport sur le suivi des objectifs du millénaire pour le développement à Madagascar*. Antananarivo: République de Madagascar.

Government of Madagascar, (2004b). *Amélioration de la gestion de l'éducation à Madagascar: Approche stratégique et plan d'actions*. Antananarivo: République de Madagascar.

Government of Madagascar, (2008). *Éléments de diagnostic du système éducatif Malagasy: Le besoin d'une politique éducative nouvelle pour l'atteinte des objectifs du millénaire et la réduction de la pauvreté*. Antananarivo: République de Madagascar.

Kling, J.R., J.B. Liebman, and Katz, L.F., (2007). “Experimental Analysis of Neighborhood Effects.” *Econometrica* 75 (1), pages 83–119.

Lassibille, G., and Tan, J.P., (2003). Student learning in public and private primary schools in Madagascar. *Economic Development and Cultural Change* 51(3), pages 699–717.

Lassibille, G., Tan, J.P., Jesse, C. and Van Nguyen, T., (2010). Managing for results in primary education in Madagascar: Evaluating the impact of selected workflow interventions. *World Bank Economic Review* 24(2), pages 303–329.

Lockheed, M., Verspoor, A., and Associates, (1991). *Improving primary education in developing countries*. Washington, DC: Oxford University for the World Bank.

PASEC. (2007). *Quelques pistes de réflexion pour une éducation primaire de qualité pour tous: Rapport Madagascar*. Dakar: Conférence des Ministres de l'Éducation des Pays ayant le Français en Partage.

World Bank, (2002). *Education and training in Madagascar: Toward a policy agenda for economic growth and poverty reduction*. Washington, DC: World Bank.

Journal published by the EAAEDS: <http://www.usc.es/economet/ea.htm>