

DELIVERING BETTER OUTCOMES IN EDUCATION: THE WORLD BANK'S EXPERIENCE

Arianne Wessal, Clay Wescott and Carlos Espíndola

ABSTRACT

Education is a critical public service needed to achieve shared prosperity and an end to extreme poverty. Previous research has found that monitoring and disseminating appropriate data can help improve strategic planning, resource allocation and in turn education quality and learning outcomes. There can be added benefits from using such data at the service delivery level, which can create incentives for enhancing data quality, and promote better informed decision-making. This approach requires exploitation of feedback mechanisms and learning loops, which are enabled by quality monitoring and evaluation (M&E) and communities of practice.

The following addresses the research question: Do World Bank supported education projects achieve better outcomes when there is deeper attention to designing monitoring and evaluation systems during project design?

Keywords - Education, Manage, Monitoring and Evaluation, Outcome, Quality, Rating

INTRODUCTION

The paper draws on a recently constructed database of World Bank investment projects. Ratings of education projects' M&E design and their outcomes will be compared to see whether there is a correlation between high M&E design ratings, and better outcomes. This pattern has been observed in other service areas supported by the Bank. Further analysis will look at other factors associated with better outcomes, and to determine their relative importance. The next step will be to analyze detailed project components to see if there are design features common to high performing projects with effective M&E and other outcome enabling features that are missing in weaker performing projects. Further testing will be carried out using selected case studies.

This paper draws from Arianne Wessal, Clay Wescott and Carlos Espíndola, 2014; and Arianne Wessal, Matt Treuth, and Clay Wescott, 2014.

Copyright: © 2015 Arianne Wessal, Clay Wescott and Carlos Espíndola. Copyright for this article is retained by the authors, with first publication rights granted to the International Public Management Review (IPMR). All journal content, except where otherwise noted, is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Unported License. By virtue of their appearance in this open-access journal, articles are free to use, with proper attribution, in educational and other non-commercial settings.

Corresponding Author: awessal@worldbank.org

The findings will be of interest to World Bank operational staff, education professionals, client governments, and other development partners. Evaluations of World Bank projects have shown a decline in outcome ratings in recent years. Improvements in M&E and other features may point the way to reversing this trend. We expect that when the enabling environment encourages experimentation, tight feedback mechanisms and constant communication make it possible for project managers to make real-time changes to projects throughout the project cycle. Feedback mechanisms could include such measures as evaluation committees, seminars and workshops, automated systems, reporting and follow-up procedures. It is also essential to tailor project components to local factors such as implementation capacity and political support. As problems arise, consideration could be given to concerns at the political, organizational, and project levels before deciding on a solution (Andrews, Pritchett and Woolcock, 2012). Project managers could be given greater incentives to draw on aspects of past successful projects, try new concepts, and adapt to changing conditions.¹

RESEARCH ON ACHIEVING RESULTS

The developing world has seen rapid improvements in recent years. The number of people in poverty has been cut by over half since 1990. Over the same period, 2.1 billion people gained access to improved drinking water and the child mortality rate has dropped by 41 per cent (United Nations 2013). However, rapid improvements are creating expectations for more equitable and just patterns of progress, and reducing the gap between possibilities and actual experience (Woolcock 2013). The different aspects of development are uneven, with more people in the world owning mobile phones (6 billion) than having access to toilets and latrines (4.5 billion) (UN News 2013). In addition, poverty-reduction efforts in some geographic regions have not enjoyed the same level of success as in other regions. Over the last two decades, the number of people living in extreme poverty in Sub-Saharan Africa rose from 290 million in 1990 to 414 million in 2010. The failure to achieve poverty reduction goals at the regional level raises the question of why some efforts fail, while similar efforts deliver quality results in other regions.

To address these challenges, many developing country governments are trying to understand why the policies put in place to reduce poverty and build prosperity are not leading to the results they want. One way forward could be a new form of knowledge, the “science of delivery”. This concept is borrowed from the healthcare field, where the previous emphasis on understanding the causes and consequences of health issues is shifting to give more attention to organizing, managing and financing health promotion (Catford 2009). Applied to the field of public management, a “science of delivery” should provide mechanism-based explanations of how and why the implementation capability of countries varies, as well as a guide to action. (Woolcock 2013) This approach differs from the institutional reform model that currently dominates the public management field. In the institutional reform model, “best practice” solutions are often chosen without significant consideration being given to their external validity. In this model, the focus is on inputs delivered rather than on outputs obtained. Another pitfall

is that reforms try to take on too much, and are stymied by the complexity, lack of clarity, uncertainty and unintended consequences of efforts to change socio-political structures (Schuck 2014: 372). The result is that projects frequently fail to achieve their goals, while the specific reasons for this failure remain hard to pin down.

In order to remedy these issues, the “science of delivery” tailors project components based on local factors such as implementation capacity and political support. As problems arise, consideration is given to concerns at the political, organizational, and project levels before deciding on a solution. Project managers are encouraged to draw on processes linked to successful projects, try new concepts, and adapt to changing conditions. The “science of delivery” approach requires experimentation, intensive field research documented in accessible case studies, improved data collection at the project level through the use of good monitoring systems, and the diffusion of ideas to enable these changes in implementation and management. The result of using a “science of delivery” approach is the incremental creation of localized projects that provide both impactful results to the target community as well as useful data and information to the public. These data give project managers the ability to understand how and why a project was effective rather than just whether it was or not. The “science of delivery” allows project managers in a region to better understand why their projects fail to achieve their desired impact, as well as give them the ability to draw on lessons learned from successful projects in other regions. At the same time, there have been recent theoretical advances in many scholarly fields ranging from systems engineering, medicine, economics, and public management that are being exploited to help countries organize the emerging evidence on successful delivery to help them improve development results (Kim 2012). These new sources of knowledge help aid managers in adapting their projects to local conditions, ultimately resulting in a higher level of success.

The World Bank and other development partners can point to many examples of delivery success, drawing on a treasure trove of evidence obtained using a mix of qualitative and quantitative methods linking successful delivery of interventions with local politics, culture, capacity, and other factors that affect delivery outcomes. However, some of this experience is not easily accessible, buried in lengthy reports, files, datasets, and as tacit knowledge in the heads of staff and evaluators.

RESEARCH ON ACHIEVING RESULTS IN EDUCATION

How can a science of delivery approach improve schooling? Research in developing countries shows that education can enable growth and poverty reduction by, inter alia, driving innovation, raising worker productivity, empowering women, and enabling high functioning institutions (Schultz 1993; Bloom, Hartley and Rosovsky 2006). However, this causal chain is not always straightforward: a Pakistan study, for example, found that educating women does not systematically affect labor productivity and allocation, pointing to the marginal role women play in Pakistan's economy (Fafchamps and Quisumbing 1998). Other research finds that learning outcomes of poor children may be held back by health, nutrition and social factors during their first five years of age, such as stunting, inadequate cognitive stimulation, iodine deficiency and iron deficiency

anemia (Walker et al 2007). Indeed, one cross national study found no link between rising education of the labor force and growth in output per worker, perhaps due to a combination of perverse governance environments, stagnant demand for educated labor, and poor educational quality (Pritchett 2001). However, another study found conclusive evidence that increased educational attainment of women of reproductive age has contributed to fewer deaths in children younger than 5 years (Gakidou et al 2010). Another cross national study focused on cognitive skills measured by international tests, rather than school attainment, and finds a strong link to economic growth independent of other contributing factors like property rights, open markets, and effective economic institutions (Hanushek and Woessmann 2008).

Thus, attaining societal benefits depends on people acquiring knowledge and skills, not just on attending school (Boissiere 2004; Glewwe 2002). Impact evaluations point to three approaches most effective in improving learning outcomes. First of all, information reforms such as providing school and student test scores helps to make better comparisons among schools, improves student performance, and lowers school fees (Andrabi and Khwaja 2014). Second, school-based management reforms such as tracking students based on prior achievement into separate classes, increasing school autonomy and empowering parents can be effective, although there only a small number of rigorous studies, and the metrics vary in different studies. Third, teacher incentive reforms can have an impact, making teachers more accountable for results by linking tenure and/or pay to performance. Again, the evidence base on this is small but promising (Bruns, Filmer and Patrinos, 2011; Duflo, Dupas and Kremer, 2008).

Choosing the particular focus of school reform differs as an education system moves from one performance level to the next. Moving from poor to fair performance, for example, typically has a focus on gathering information, focusing on core competencies and improving organization, finances and pedagogy; moving from good to great, in contrast, focuses on shaping the career paths of the teaching profession in a clearly defined way. Even within these patterns, context is key. For example, in some regions schools share school performance data publicly, while others share such data only privately among schools (Mourshed, Chijioko and Barber 2010; Banerjee and Duflo 2011: 97-101).

DATA ANALYSIS OF WORLD BANK INVESTMENT PROJECTS

Building on this scholarly research, we now shift to analysis of the World Bank experience in supporting education, and the role of monitoring and evaluation in improving outcomes. Recent portfolio analysis indicates a decline in M&E quality across all sectors. Comparing the Bank's investment projects completed in FY07-09 with those completed in FY10-12, the number of projects rated high or substantial on M&E declined from 33 to 26 percent. Reasons included limited baseline data, unfocused indicators, and indicators measuring just outputs rather than outcomes. There were also weak monitoring organizations with high staff turnover and vacancies, unclear roles and responsibilities for data collection, weak MIS and data quality, and limited data utilization for decision making (IEG, 2014: 45). This weakening M&E performance has contribut-

ed to an overall decline in outcome across all sectors. The number of completed investment projects with moderately satisfactory outcomes or better has fallen from 79 percent in FY04-06 to 75 percent in FY08-10 to 69 percent in FY10-12 (Ibid: 35). These trends are both statistically significant at the 95 percent confidence level.

The same ratings are also declining in the Education sector, and have been since FY01. A portfolio review looked at education projects completed over the period FY01-09 to understand this declining trend.² While the decline in outcome rating went along with a decline in M&E quality, there were many other factors judged to be more important in explaining the decline in outcome. There were 15 factors found to contribute to unsatisfactory performance, about half concerning design and half implementation/supervision. Five of the seven most important factors leading to weak performance were design issues: over ambition in relation to the strength of political commitment, over ambition in relation to the time period, inadequate readiness for implementation, weaknesses in technical design, including prior analytical work, and over ambition or excessive complexity in regard to country institutional or implementing capacity. The two leading implementation weaknesses were lack of responsiveness or proactivity in supervision, and external factors such as a political or economic crisis. Weak M&E implementation, including lack of evidence of outputs or outcomes, was the least important factor leading to weak performance (IEG 2011: 36). The portfolio review implies that these factors may have been increasingly triggered since FY2001 because Bank operations tended to have more challenging objectives focusing, *inter alia*, on post-primary education, learning, efficiency, and employment outcomes, rather than primary education, access and equity outcomes (Ibid.).

ANALYSIS

Using the recently constructed Independent Evaluation Group (IEG) database of World Bank projects and their components, an initial analysis of the projects within the database was taken in order to see whether there is a correlation between high M&E ratings and high outcome ratings in World Bank education projects, as validated in Implementation Completion Report Reviews. The analysis was conducted by using all the education projects with M&E quality ratings and outcome ratings found in both the IEG components database (privately available) and in IEG's publicly available database of World Bank project ratings. Only projects with both M&E quality ratings and IEG outcome ratings that overlapped in both databases were used for the analysis. There were a total of 63 projects that fell into this category. The completion dates for the Implementation Completion Report Reviews fall between 2007 to 2012.

The results were compared using a two by two table in which projects with IEG outcome ratings which were moderately satisfactory and above were grouped and those which were moderately unsatisfactory and below were grouped. These were compared to M&E quality ratings with categories of modest and above and the category of negligible. In this case there were a total of 63 projects with both IEG outcome ratings and M&E ratings.

The analysis showed that a majority of the projects had M&E quality ratings of modest and above and outcome ratings that were moderately satisfactory or above, while less than half of projects with a modest and above M&E rating had moderately unsatisfactory and below outcome ratings.

Figure 1: Outcome and Quality Ratings

Outcome ratings	Quality ratings		Totals
	Modest and above	Negligible	
Moderately satisfactory and above	41	1	42
Moderately unsatisfactory and below	15	6	21
Totals	56	7	63

Figure 2 shows a breakdown of the 41 projects with moderately satisfactory and above ratings and modest and above M&E quality. There are only 10 projects (16%) that had an outcome rating of satisfactory and an M&E rating of substantial. Although the highest M&E rating possible is ranked high, there were no projects with a high M&E rating in the 63 analyzed.

Figure 2: Breakdown for the projects with moderately satisfactory and above Outcome ratings and modest and above M&E quality ratings

Outcome ratings	M&E Quality ratings		Total
	Substantial	Modest	
Satisfactory	10	6	16
Moderately Satisfactory	4	21	25
Total	14	27	41

To test the statistical significance of the correlation between the M&E quality rating and the IEG outcome rating, a chi-squared test was run. A chi-square test is a commonly used statistical test which tests the correlation between two variables of interest. The test compares if deviations in the data from theoretical proportions have occurred by chance or not. For that purpose the test uses the chi-square distribution to estimate the theoretical proportions of how the data should behave. To conduct this test we need a random sample, independent observations and over 5 observations per cell in the 2 by 2 matrix that we used. The matrix that we used for the test is shown in Figure 1. Outcomes and Quality Ratings.

The Null Hypothesis that we want to test is if these variables are independent. If we select a significant level of 0.05 we get from the test the values shown in Figure 3 be-

low. In the cells we can see the observed cell totals, the expected cell totals if the variables were independent and the Chi-square statistics for each cell.

Figure 3: Testing the Relationship between the Outcome and Quality Ratings

Outcome ratings	Quality ratings		
	Modest and above	Negligible	Totals
Moderately satisfactory and above	41 (37.33) [0.36]	1 (4.67) [2.88]	42
Moderately unsatisfactory and below	15 (18.67) [0.72]	6 (2.33) [5.76]	21
Totals	56	7	63

The Chi-square statistic for this test is 9.72 (sum of the chi-statistic for each cell). The P value for this chi-square value is 0.00182. As the p value is less than our confidence level of 0.05, we can reject the null hypothesis with a 95% confidence level. Meaning that the results are statistically significant at the >95% level and the variables are dependent.

In addition, Figure 4 presents a breakdown of the M&E quality ratings by regions. The data set showed East Asia and the Pacific, and Europe and Central Asia regions with the highest ratings.

Figure 4: ME Quality Ratings Breakdown by Region

	Africa	East Asia and the Pacific	Europe and Central Asia	Latin America and the Caribbean	Middle East and North Africa	South Asia
Modest and above	89%	100%	100%	89%	73%	92%
Negligible	5%	0%	0%	11%	18%	8%
Total	95%	100%	100%	100%	91%	100%

Of the 41 projects with moderately satisfactory and above outcome ratings and modest and above M&E quality ratings, 95% of the component records associated with these had M&E as a component or mentioned M&E explicitly as part of their component descriptions.

Common M&E features and challenges:

Of the 63 projects that had the highest M&E rating and outcome ratings – in this case a rating of Substantial and an outcome rating of Satisfactory, there were some common design features, namely that a majority of the projects with high M&E ratings contained the following elements:

- A baseline study was conducted at the outset of the program

- A clearly identified entity responsible for the monitoring and the evaluation of the project
- A thorough inventory and consideration of lessons learnt and findings from previous projects or research and literature
- Pilots and regular monitoring to allow for adaptable and flexible changes of design
- Regular surveys and evaluations included in the design of the project for mid course and correction and the acquisition of just in time information.
- Tracking of inputs outputs and outcomes
- A means of acquiring reliable data such as an EMIS or some type of Management information system.
- A thorough mid-term review, at which point in some case for some of these projects M&E systems had to be restructured

In the case of the projects that showed mixed and poor results, an analysis of ICRR M&E sections and ratings showed that the projects tended to have the following problems:

- Lack of a baseline
- Key indicators did not reflect all objectives and activities
- Too many indicators
- Poor capacity for data collection
- Implementation of project M&E fell short due to the complexity of the M&E design
- Limited use of information during project implementation
- Tracking of outputs versus outcomes
- Lack of management information system

EXAMPLES FROM MEXICO, JAMAICA, BRAZIL

There are specific design features that are common to the high performing projects with effective M&E. Some possible features are evident from looking at three examples. Two are Oportunidades (formerly Progresá), a health and education conditional cash transfer program in Mexico, and the Program of Advancement through Health and Education (PATH) in Jamaica. These programs have built strong monitoring systems at the beginning of the programs with short-, medium-, and long-term outcomes identified. Programs undertake regular assessments at each step of implementation and use this in conjunction with monitoring information to make adjustments as the programs are implemented. In the case of PATH, process evaluations and spot checks are undertaken for activities being implemented. (Rawlings 2009). This enables the identification of a number of problems, including: stakeholders saw the application process as burdensome and were not clear on program rules, the system for verifying the eligibility of new beneficiaries was weak, and there was a strong unmet demand for jobs and training. Quality

data enabled management to exploit double-loop learning (Argyris and Schön, 1974) and led to a decision to revamp the management information system, revise the operations manual, use social workers as focal points to access social services, and create a “STEPS to Work” program focused on skill development and employment. Both programs demonstrate that implementation of a strong M&E system that enables the use of feedback mechanisms and learning loops to make real-time project changes can yield better development results. Evaluations of PATH showed that it was better at reaching the poor than other Jamaican safety net programs, while evaluations of Oportunidades showed the program had a significant positive impact in improving health and education. Both programs have been lauded for reaching their target populations and yielding better results than other programs.

Oportunidades is a great example of improved science of delivery through the use of both a strong M&E system and of information learned from past projects that warrants a closer look. The program began in 1997, providing monetary educational grants to poor, rural families, for each child under 22 years of age who is enrolled in school between the third grade of primary and third grade of high school. In addition to education, Oportunidades also has health and nutrition components. Government health institutions provide families with preventative health care and families also receive a fixed monthly transfer to improve food consumption, as well as nutritional supplements for young children and their mothers. Where Oportunidades truly shines is in quality at entry. At implementation, project managers planned to have an independent evaluation done by the International Food Policy Research Institute. This independent evaluation was planned with the goal of providing management with data to make real-time project changes and improve delivery. They also drew on lessons learned from past projects, recognizing that giving money to female heads of families results in better financial outcomes. These steps taken during the implementation stages translated into quality results that were reflected in the independent evaluation. The evaluation reported that improvements had been made in increasing school enrollment, nutritional quality, and access to medical care. At the time of the evaluation, Oportunidades was said to have increased secondary school enrollment rates by over 20 percent for girls and 10 percent for boys (Parker 2003). This was the first randomized controlled trial of a large scale used in developing country social policy. The results caught the eye of the Mexican federal government, although the evaluation methodology has been criticized for its sampling design, and inadequate treatment of selective attrition and sample contamination (Faulkner 2012). As of 2003, 46.5 percent of Mexico's federal annual anti-poverty budget was devoted to Oportunidades. This increase in funding allowed Oportunidades to expand to urban areas and to provide high school students with education grants. In summary, steps taken at implementation to improve the science of delivery were crucial in the success and subsequent expansion of Oportunidades. Learning from past projects and having quality external evaluation ultimately led Oportunidades to become one of the most successful conditional cash transfer programs to date. The close involvement of scholar-practitioners helped to design new conceptual approaches, ensure technical soundness and rigorous monitoring, protect the program during changes of administration, and spread the approach around the world (Lustig 2011).

In 2008, Brazil began its Second Minas Gerais Development Partnership Project, which was a sector wide approach project of over US\$1.4 billion, aimed at improving the efficiency of public resource use, supporting innovations in public management, and supporting the State Government of Minas Gerais in strengthening its M&E system (World Bank 2008). Funds were disbursed to ten eligible expenditure programs in five sectors. An extensive results monitoring framework was built into the program at implementation. Individual projects were subject to monthly monitoring and quarterly management meetings were made accessible to the press. The Government made yearly implementation data available on the web to increase program transparency and improve data dissemination. In order to increase its focus on outcomes delivered, the Bank supported the project by developing a household survey, quality assurance surveys, and a series of impact evaluations in the education, health, and transport sectors. These monitoring systems provided managers necessary feedback mechanisms and learning loops, allowing them to work towards the achievement of medium-term goals on their way to the achievement of long-term objectives. The latest Implementation Status and Results Report rated progress towards achievement of project development objective and implementation progress as satisfactory (World Bank 2013). So far, the program has succeeded in reducing the amount of time needed to start a business at Minas Facil in Belo Horizonte from 26 to 7 days. The Poverty Reduction Program has already exceeded its initial objective by benefiting over 26,000 rural families.

Similar to the programs described above there were a number of projects amongst those which received the highest ratings, had similar design features to those described above. In particular the projects included pilots, which allowed for flexible reaction and redesign where needed, in addition to thorough background research, case studies and lessons learnt from other projects. Frequent reporting and field visits to identify issues, and regular M&E in order to swiftly update baseline and target values once data was available. One such example of a project that illustrates these design features well is the Guatemala – Universalization of Basic Education Project which was begun in 2001, prior to M&E being a requirement as a specific design component in World Bank projects.

This project was a continuation and deepening of the Basic Education Reform Project. The new project was focused on the expansion and enrollment capacity of primary schools through a National Community Managed Program for Educational Development (PRONADE), begun under the Basic Education Reform Project. In addition, it supported a program to enhance the goals of cultural diversity and pluralism, contained in Guatemala's constitution, the Peace Accords of 1996, and the April 2000 National Congress on Cultural Policies.

The project had four main components:

- Universalization of Primary Education
- Improving the Quality of Education
- Cultural Diversity and Pluralism
- Decentralization and Modernization.

As part of its first component, the project-included support for the expansion of enrollment of an additional 40,000 students over 4 years from 1,300 underserved communities. As part of the design regular evaluations of the program were included in order to determine the effectiveness of the support services afforded to the COEDUCAs (Parent-run school committees) and to teachers in order to help point out specific areas for immediate improvement.

In addition, the project also included pilots as part of this and other sub-components to test possible solutions. For example, the project conducted a pilot testing of two administrative models leading to the graduation of COEDUCAs in matters related to financial administration based on established criteria for graduation and periodic external monitoring and evaluation to review progress achievement difficulties and efficiency at the local and central levels. In another case, the program piloted the introduction of itinerant teachers to provide education services to 10 isolated rural communities and approximately 180 children where the cost of establishing a PRONADE (community managed schools) was too high because of the small number of children in the communities. Unfortunately, neither the Implementation Completion Report nor the ICRR noted what happened with this pilot and whether the itinerant program was institutionalized.

In 2004, a pilot for distance secondary education (Telesecundaria or video-supported teaching), was introduced in eight communities to address the need of students in rural and remote areas to access secondary education. Again the pilot experience was closely monitored and would up laying the foundation for the World Bank's Education Quality and Secondary Education Project (Loan No. 7052-GU).

In January 2008, a new administration took office and decided to institutionalize PRONADE. PRONADE schools were to become part of the regular school system, but because 2008 was a transition year, the process to absorb PRONADE teachers by the regular school system took place during 2008 and 2009. Initially PRONADE teachers were hired under temporary contracts, and then MINEDUC opened a competitive process to hire PRONADE teachers as regular teachers.

This process was about to be completed by the time of this ICR. The COEDUCAs were converted into Juntas Escolares, school boards in traditional rural schools.

CONCLUSIONS

This paper has addressed the question: Do Bank supported education projects achieve better outcomes when there is deeper attention to designing monitoring and evaluation systems during project design?

Arianne Wessal is a Consultant with the Global Partnership for Education Secretariat, World Bank.
Email: awessal@globalpartnership.org

Dr. Clay G. Wescott is Senior Consultant with the Independent Evaluation Group, World Bank, and President, International Public Management Network. Email: clay.wescott@gmail.com

Carlos Espíndola is a Managing Associate at the Infrastructure Planning and Economics group at Nathan Associates. Email: cespindola@nathaninc.com

Using a recently constructed database of World Bank projects and their components, the better performing projects do have higher M&E ratings, and the difference is statistically significant. However, even the better performing projects mainly have only a modest rating for M&E, suggesting room for improvement. Initial case study analysis found specific design features that are common to the high performing projects with effective M&E. There are also be added benefits when projects are designed encouraging experimentation and taking into account distinctive features of the local context. This analysis needs to be extended to the rest of the sample to see if these features are common, and if so, the extent to which they help to explain the improved performance.

NOTE

- ¹ This paper has not undergone the review accorded to official World Bank publications. The findings, interpretations, and conclusions expressed herein are those of the author(s) and do not necessarily reflect the views of the International Bank for Reconstruction and Development / The World Bank and its affiliated organizations, or those of the Executive Directors of The World Bank or the governments they represent. The World Bank does not guarantee the accuracy of the data included in this work.
- ² Due to the smaller number of projects, the ratings decline for Education and most other individual sectors from FY07-09 to FY10-12 is not statistically significant (IEG, 2014). However, the decline from FY01 to FY12 is statistically significant at the 95% confidence level. The latter analysis looks at both investment and development policy operations (IEG, 2011).

REFERENCES

- Andrabi, T., J. Das, and A. I. Khwaja. 2014. "Report Cards: The Impact of Providing School and Child Test Scores on Educational Markets." http://www.hks.harvard.edu/fs/akhwaja/papers/RC_June2014.pdf
- Andrews, M., Pritchett, L. and Woolcock. 2012. Escaping Capability Traps through Problem-Driven Iterative Adaptation (PDIA). Faculty Research Working Paper Series. Cambridge, Mass.: Harvard Kennedy School.
- Argyris, C. and Schön, D. (1974) Theory in practice: Increasing professional effectiveness, San Francisco: Jossey-Bass, pp 2-3.
- Banerjee, A., and E. Duflo. 2011. Poor Economics: A Radical Rethinking of the Way to Fight Global Poverty. New York: PublicAffairs.
- Bloom, D., M. Hartley, and H. Rosovsky. 2006. "Beyond Private Gain: The Public Benefits of Higher Education." In Samuel Neaman Institute, Transition to Mass Higher Education Systems: International Comparisons and Perspectives. Haifa, Israel: Samuel Neaman Press.

- Boissiere, Maurice. 2004. "Rationale for Public Investment in Primary Education in Developing Countries." IEG Working Paper, World Bank, Washington, DC.
- Bruns, B., D. Filmer, and H. Patrinos. 2011. *Making Schools Work: New Evidence on Accountability Reforms*. Washington, DC: World Bank.
- Catford, John. 2013. "Advancing the 'science of delivery' of health promotion: not just the 'science of discovery'". *Health Promotion International*. 24 (1): 1-5, Oxford University Press, Oxford, England. <http://heapro.oxfordjournals.org/content/24/1/1.extract>
- Duflo, E., P. Dupas, and M. Kremer. 2008. "Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya." NBER Working Paper 14475. National Bureau of Economic Research, Cambridge, MA.
- Fafchamps, M., and A. R. Quisumbing. 1999. "Human Capital, Productivity, and Labor Allocation in Rural Pakistan." *Journal of Human Resources* 34 (2): 369–406.
- Faulkner, William. 2012. *Negotiated Validity: Three Perspectives on IFPRI-Progresa's sampling*. The Evaluator. London. http://wnfaulkner.files.wordpress.com/2012/08/evaluator_summer2012-final-11-14.pdf
- Gakidou, E., K. Cowling, R. Lozano, and C. Murray. 2010. "Increased Educational Attainment and its Effect on Child Mortality in 175 Countries between 1970 and 2009: A Systematic Analysis." *The Lancet* 376 (9745): 959–74.
- Glewwe, Paul. 2002. "Schools and Skills in Developing Countries: Education Policies and Socioeconomic Outcomes." *Journal of Economic Literature* 40 (2): 436–82.
- Hanushek, Eric. A., and Ludger Woessmann. 2008. "The Role of Cognitive Skills in Economic Development." *Journal of Economic Literature* 46(3): 607–68.
- Independent Evaluation Group. 2014. *Results and Performance of the World Bank Group 2013 An Independent Evaluation*. Washington, D.C.: The World Bank. <http://ieg.worldbankgroup.org/evaluations/rap2013>
- _____. 2011. *World Bank Support to Education Since 2001: A Portfolio Note*. Washington, D.C.: World Bank.
- Kim, Jim Yong. 2012. "Delivering on Development: Harnessing Knowledge to Build Prosperity and End Poverty", Keynote Speech to World Knowledge Forum, Seoul, Korea, Republic of Korea. http://www.worldbank.org/en/news/speech/2012/10/08/delivering-development-harnessing-knowledge-build-prosperity-end-poverty?goback=%252Egde_4387230_member_174342256
- Lustig, Nora Claudia. 2011. *Scholars Who Became Practitioners: The Influence of Research on the Design, Evaluation, and Political Survival of Mexico's Antipoverty Program Progres/Oportunidades*. Center for Global Development Working Paper No. 263. <http://ssrn.com/abstract=1910074> or <http://dx.doi.org/10.2139/ssrn.1910074>
- Mourshed, M., C. Chijioke, and M. Barber. 2010. *How the World's Most Improved School Systems Keep Getting Better*. McKinsey & Company.

- Parker, Susan. 2003. "The Oportunidades Program in Mexico." Shanghai Poverty Conference http://politiquessociales.net/IMG/pdf/Mexico_oportunidades.pdf
- Pritchett, L. 2001. "Where Has All the Education Gone?" World Bank Economic Review 15 (3): 367–91.
- Rawlings, Laura. 2009. "Monitoring and Evaluation: the Foundations for Results". Beijing, China: Impact Evaluation Workshop http://siteresources.worldbank.org/INTISPMA/Resources/383704-1184250322738/3986044-1250881992889/M&E_TheFoundationsOfResults_Rawlings_EN.pdf
- Schuck, Peter 2014. Why Government Fails So Often and how it can do better. Princeton: Princeton University Press.
- Schultz, T. Paul. 1993. "Returns to Women's Education." In Women's Education in Developing Countries: Barriers, Benefits and Policies, eds. Elizabeth M. King and M. Anne Hill. Baltimore, MD: Johns Hopkins University.
- UN News Centre, 2013. "Deputy UN Chief calls for urgent action to tackle global sanitation crisis". 21 March. <http://www.un.org/apps/news/story.asp?NewsID=44452#.UgD8GZLUmJp>
- United Nations. 2013. "The Millennium Development Goals Report" <http://unstats.un.org/unsd/mdg/Resources/Static/Products/Progress2013/English2013.pdf>
- Walker S., E. Wachs, J. Fardner, B. Lozoff, G. Wasserman, E. Pollitt, J.A. Carter, and International Child Development Steering Group. 2007. "Child Development: Risk Factors for Adverse Outcomes in Developing Countries." The Lancet 369 (9556): 145–57.
- Wessal, Arianne, Clay Wescott and Carlos Espíndola, 2014. Science of delivery: Education outcomes and the World Bank. In Andrew Robertson. Commonwealth Governance Handbook 2014/15 - Democracy, development and public administration. Cambridge, U.K.: Nexus Strategic Partnerships: 48-52.
- Wessal, Arianne, Matt Treuth, and Clay Wescott. 2014. "Science Of Delivery And Implications For Monitoring And Evaluation". In United Nations Development Program. Solutions Related To Challenges Of Independence, Credibility And Use Of Evaluation Proceedings From The Third International Conference On National Evaluation Capacities. 30 September – 2 October 2013, São Paulo, Brazil. New York: United Nations Development Program: 262-7.
- Woolcock, Michael. 2013. "The Science Of Delivery-- Taking Context Seriously: Expanding the Range of Ideas, Methods and Strategies to Enhance Development Effectiveness." Workshop presentation, World Bank Group. <http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/0,,contentMDK:23388358~pagePK:64165401~piPK:64165026~theSitePK:469382,00.html>
- World Bank. 2008. Second Minas Gerais Development Partnership Project: Project Appraisal Document (VOL 1)

_____. 2013. Minas Gerais Partnership II SWAP: Implementation Status & Results Report. Sequence 10.

About IPMR

IPMR The International Public Management Review (IPMR) is the electronic journal of the International Public Management Network (IPMN). All work published in IPMR is double blind reviewed according to standard academic journal procedures.

The purpose of the International Public Management Review is to publish manuscripts reporting original, creative research in the field of public management. Theoretical, empirical and applied work including case studies of individual nations and governments, and comparative studies are given equal weight for publication consideration.

IPMN The mission of the International Public Management Network is to provide a forum for sharing ideas, concepts and results of research and practice in the field of public management, and to stimulate critical thinking about alternative approaches to problem solving and decision making in the public sector.

IPMN includes over 1300 members representing about one hundred different countries, both practitioners and scholars, working in all aspects of public management. IPMN is a voluntary non-profit network and membership is free.

ISSN 1662-1387