What Education Policies and Programs Affect Learning and Time in School in Developing Countries?

Methodology for Review of Academic and Aid Agency
Evaluations

Amy Damon Paul Glewwe Suzanne Wisniewski and Bixuan Sun

May 11, 2016



Main contribution of this report: review of the evidence on the impacts of different types of education policies and programs on student learning and time in school.

- Main contribution of this report: review of the evidence on the impacts of different types of education policies and programs on student learning and time in school.
- A main challenge: credible causal identification.

- Main contribution of this report: review of the evidence on the impacts of different types of education policies and programs on student learning and time in school.
- A main challenge: credible causal identification.
- Strategy: limit our synthesis of the evidence to 114 high quality studies that were conducted from 1990 to 2014.

- Main contribution of this report: review of the evidence on the impacts of different types of education policies and programs on student learning and time in school.
- A main challenge: credible causal identification.
- ▶ Strategy: limit our synthesis of the evidence to 114 high quality studies that were conducted from 1990 to 2014.
- ➤ To identify evaluations that have produced credible impacts of education programs or policies, the following four-step selection process was implemented.

Review Steps

Table 1- Steps Used to Select Papers Reviewed in This Report

Review Step 1	Procedures Used Search EconLit and ERIC databases.	Number of Papers 13,437
	Review abstracts to eliminate duplicate papers and papers that did not estimate the impacts of school or teacher characteristics for a developing country.	1,017
2	Review full papers, eliminate papers based on lack of relevance or lack of quantitative analysis.	320
3	Exclude papers that are not "high quality" (RCT, RDD, DD).	115
4	Number of high quality studies that are RCT studies	76

 Search was conducted on a wide variety of sources. Studies must be

- Search was conducted on a wide variety of sources. Studies must be
 - published in (peer-reviewed) academic journals from 1990 to 2014, inclusive

- Search was conducted on a wide variety of sources. Studies must be
 - published in (peer-reviewed) academic journals from 1990 to 2014, inclusive
 - unpublished academic working papers from 2010 to 2014

- Search was conducted on a wide variety of sources. Studies must be
 - published in (peer-reviewed) academic journals from 1990 to 2014, inclusive
 - unpublished academic working papers from 2010 to 2014
- Two search engines for the economics and education literatures: EconLit & the Education Resources Information Center (ERIC). Limited papers to:

- Search was conducted on a wide variety of sources. Studies must be
 - published in (peer-reviewed) academic journals from 1990 to 2014, inclusive
 - unpublished academic working papers from 2010 to 2014
- Two search engines for the economics and education literatures: EconLit & the Education Resources Information Center (ERIC). Limited papers to:
 - Papers that list both "education" as a key word & any one 124 educational programs/policy key-words

- Search was conducted on a wide variety of sources. Studies must be
 - published in (peer-reviewed) academic journals from 1990 to 2014, inclusive
 - unpublished academic working papers from 2010 to 2014
- Two search engines for the economics and education literatures: EconLit & the Education Resources Information Center (ERIC). Limited papers to:
 - Papers that list both "education" as a key word & any one 124 educational programs/policy key-words
 - Papers that include the name of at least one developing country or the term "developing country" or "developing countries" in the abstract

▶ This initial search yielded 13,437 publications.

- ▶ This initial search yielded 13,437 publications.
 - Information from the abstract was then used to limit the studies to those that appear to be potentially relevant.

- ▶ This initial search yielded 13,437 publications.
 - Information from the abstract was then used to limit the studies to those that appear to be potentially relevant.
 - We eliminated evaluations that did not focus on developing countries or did not provide a quantitative estimate of the impact of a program/policy on education outcomes

A search was conducted of several prominent working papers series: National Bureau of Economic Research (NBER); World Bank Policy Research; the Institute for the Study of Labor (IZA); the Center for Economic and Policy Research (CEPR); the CESIfo Research Network; the Rural Education Action Project (REAP) at Stanford University; Young Lives Working Papers; Abdul Latif Jameel Poverty Action Lab

- A search was conducted of several prominent working papers series: National Bureau of Economic Research (NBER); World Bank Policy Research; the Institute for the Study of Labor (IZA); the Center for Economic and Policy Research (CEPR); the CESIfo Research Network; the Rural Education Action Project (REAP) at Stanford University; Young Lives Working Papers; Abdul Latif Jameel Poverty Action Lab
 - Working papers before 2010 were excluded; high quality working papers written before 2010 should have been published by 2014

- A search was conducted of several prominent working papers series: National Bureau of Economic Research (NBER); World Bank Policy Research; the Institute for the Study of Labor (IZA); the Center for Economic and Policy Research (CEPR); the CESIfo Research Network; the Rural Education Action Project (REAP) at Stanford University; Young Lives Working Papers; Abdul Latif Jameel Poverty Action Lab
 - Working papers before 2010 were excluded; high quality working papers written before 2010 should have been published by 2014
 - We eliminated evaluations that did not focus on developing countries or did not provide a quantitative estimate of the impact of a program/policy on education outcomes

- A search was conducted of several prominent working papers series: National Bureau of Economic Research (NBER); World Bank Policy Research; the Institute for the Study of Labor (IZA); the Center for Economic and Policy Research (CEPR); the CESIfo Research Network; the Rural Education Action Project (REAP) at Stanford University; Young Lives Working Papers; Abdul Latif Jameel Poverty Action Lab
 - Working papers before 2010 were excluded; high quality working papers written before 2010 should have been published by 2014
 - We eliminated evaluations that did not focus on developing countries or did not provide a quantitative estimate of the impact of a program/policy on education outcomes
- ► After this step, 1,017 publications remained

 All 1,017 of the evaluations were reviewed to obtain further information about each study

- All 1,017 of the evaluations were reviewed to obtain further information about each study
 - Papers were subsequently eliminated for lack of relevance, which was not evident from reading the abstracts

- All 1,017 of the evaluations were reviewed to obtain further information about each study
 - Papers were subsequently eliminated for lack of relevance, which was not evident from reading the abstracts
 - Possible reasons for lack of relevance: (1) The evaluation did not focus on a developing country, (2) The paper did not evaluate any type of education policy or program, and (3) The paper did not include quantitative analysis of the impact of an education policy or program on educational outcomes

- All 1,017 of the evaluations were reviewed to obtain further information about each study
 - Papers were subsequently eliminated for lack of relevance, which was not evident from reading the abstracts
 - Possible reasons for lack of relevance: (1) The evaluation did not focus on a developing country, (2) The paper did not evaluate any type of education policy or program, and (3) The paper did not include quantitative analysis of the impact of an education policy or program on educational outcomes
- 320 papers remained

Remaining papers were reviewed for their quality

- Remaining papers were reviewed for their quality
- Regression analysis is typically used to estimate the impact of a policy/program on an educational outcome

- Remaining papers were reviewed for their quality
- Regression analysis is typically used to estimate the impact of a policy/program on an educational outcome
 - Problem of omitted variable bias: some factors that have a causal impact on the education outcome variable of interest are unlikely to be available in the data, which can lead to bias in regression estimates of the impacts of education policies and programs, and thus to misleading results.

- Remaining papers were reviewed for their quality
- Regression analysis is typically used to estimate the impact of a policy/program on an educational outcome
 - Problem of omitted variable bias: some factors that have a causal impact on the education outcome variable of interest are unlikely to be available in the data, which can lead to bias in regression estimates of the impacts of education policies and programs, and thus to misleading results.
 - Another problem: regressions often included many school and teacher characteristics as control variables which may not be of the primary interest of the researcher; they should not be interpreted as estimates of the causal impacts of those variables

- Remaining papers were reviewed for their quality
- Regression analysis is typically used to estimate the impact of a policy/program on an educational outcome
 - Problem of omitted variable bias: some factors that have a causal impact on the education outcome variable of interest are unlikely to be available in the data, which can lead to bias in regression estimates of the impacts of education policies and programs, and thus to misleading results.
 - Another problem: regressions often included many school and teacher characteristics as control variables which may not be of the primary interest of the researcher; they should not be interpreted as estimates of the causal impacts of those variables
 - Any study based on the above methods alone was deemed not to be high quality studies

- Remaining papers were reviewed for their quality
- Regression analysis is typically used to estimate the impact of a policy/program on an educational outcome
 - Problem of omitted variable bias: some factors that have a causal impact on the education outcome variable of interest are unlikely to be available in the data, which can lead to bias in regression estimates of the impacts of education policies and programs, and thus to misleading results.
 - Another problem: regressions often included many school and teacher characteristics as control variables which may not be of the primary interest of the researcher; they should not be interpreted as estimates of the causal impacts of those variables
 - Any study based on the above methods alone was deemed not to be high quality studies
 - Studies based on matching methods were also excluded



► Three types of studies that were considered to be high quality studies

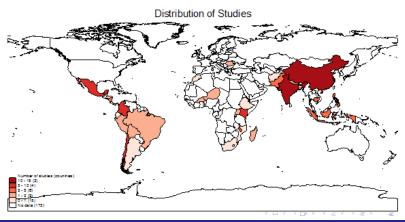
- Three types of studies that were considered to be high quality studies
 - 1. Evaluations based on a well-implemented randomized controlled trial (RCT), as these studies avoid, or at least minimize, many types of estimation problems

- Three types of studies that were considered to be high quality studies
 - 1. Evaluations based on a well-implemented randomized controlled trial (RCT), as these studies avoid, or at least minimize, many types of estimation problems
 - 2. Evaluations based on a difference in differences (DD) regression (which requires longitudinal data)

- ► Three types of studies that were considered to be high quality studies
 - 1. Evaluations based on a well-implemented randomized controlled trial (RCT), as these studies avoid, or at least minimize, many types of estimation problems
 - Evaluations based on a difference in differences (DD) regression (which requires longitudinal data)
 - 3. Evaluations based on regression discontinuity design (RDD)

- ► Three types of studies that were considered to be high quality studies
 - 1. Evaluations based on a well-implemented randomized controlled trial (RCT), as these studies avoid, or at least minimize, many types of estimation problems
 - 2. Evaluations based on a difference in differences (DD) regression (which requires longitudinal data)
 - 3. Evaluations based on regression discontinuity design (RDD)
- ▶ The set of papers that were retained 114 high quality studies

► The 114 high quality studies are from 36 different developing countries; most studies from China & India



Step 4: Identify Randomized Controlled Trial (RCT) Evaluations

 Well implemented randomized controlled trial (RCT) studies arguably have the highest credibility, when implemented correctly

Step 4: Identify Randomized Controlled Trial (RCT) Evaluations

- Well implemented randomized controlled trial (RCT) studies arguably have the highest credibility, when implemented correctly
- ► Of the 114 high quality studies, about two thirds (75) were RCTs

We also identified, through the methodology just described, the "grey" literature from several international organization and NGO sources using the same selection criteria

- We also identified, through the methodology just described, the "grey" literature from several international organization and NGO sources using the same selection criteria
 - Purpose: identify unpublished studies that may have employed the same types of high quality quantitative methods

- We also identified, through the methodology just described, the "grey" literature from several international organization and NGO sources using the same selection criteria
 - Purpose: identify unpublished studies that may have employed the same types of high quality quantitative methods
- ► The grey literature consists of papers, evaluations & reports :

- We also identified, through the methodology just described, the "grey" literature from several international organization and NGO sources using the same selection criteria
 - Purpose: identify unpublished studies that may have employed the same types of high quality quantitative methods
- ▶ The grey literature consists of papers, evaluations & reports :
 - that attempt to estimate the impact of an education program, but have not been published in any academic outlet, either as a working paper or in an academic journal

- We also identified, through the methodology just described, the "grey" literature from several international organization and NGO sources using the same selection criteria
 - Purpose: identify unpublished studies that may have employed the same types of high quality quantitative methods
- ▶ The grey literature consists of papers, evaluations & reports :
 - that attempt to estimate the impact of an education program, but have not been published in any academic outlet, either as a working paper or in an academic journal
 - written by international aid agencies (both governmental & non-governmental) or by the consulting firms that those agencies hire

To find high quality evaluations in the grey literature, we searched the websites of, and/or directly contacted, the following organizations:

Table 2. Summary of Grey Literature Search

Grey Literature Source	# of Education Related Studies or Reports	# of Quantitative Studies on Education	# of Additional High Quality Quantitative Evaluations
Mathematica Policy research	67	5	1
Save the Children Fund	66	5	1
Millennium Challenge Corp (MCC)	51	2	0
USAID	0	0	0
SIDA	75	0	0
DFID	0	0	0
Norway Agency for Devel. Cooperation	0	0	0
Oxfam	17	0	0
UNICEF	230	26	1

► We found very few studies in the grey literature that were unpublished and met the quality standards laid out above

- ▶ We found very few studies in the grey literature that were unpublished and met the quality standards laid out above
 - However, given the rising popularity of implementing RCTs for impact evaluations, this is likely to change in the coming years

- ▶ We found very few studies in the grey literature that were unpublished and met the quality standards laid out above
 - However, given the rising popularity of implementing RCTs for impact evaluations, this is likely to change in the coming years
- Increasingly NGOs and bilateral aid agencies are interested in rigorously evaluating the impact of their programs in order to efficiently and effectively direct their limited resources

- ▶ We found very few studies in the grey literature that were unpublished and met the quality standards laid out above
 - However, given the rising popularity of implementing RCTs for impact evaluations, this is likely to change in the coming years
- Increasingly NGOs and bilateral aid agencies are interested in rigorously evaluating the impact of their programs in order to efficiently and effectively direct their limited resources
- ► The more rigorous evaluations we have on diverse topics in the education sector, the clearer the answers will be for important policy concerns

Table 3. Rubric for Categorizing the Effectiveness of Studies.

What works	 Requires 3+ studies with positive and significant outcomes May include a few non-significant or negative estimates, but most are significantly positive.
What often works	Evidence that the intervention often works, but sometimes does not work. Before implementing, policy makers should carefully consider whether conditions in their countries are similar to those in the countries where the policy appears to have worked. Requires 3+ studies Relatively equal mix of positive and significant estimates with negative and/or nonsignificant estimates.
What is promising	Promising evidence that the intervention works, but needs additional evidence. Requires: 1-2 studies. If 2 studies, 2+ positive and significant estimates are required; may include a limited number of non-significant estimates. and even 1 negative estimate. If 1 study, only positive and at least 1 significant estimate.
What does not work	No evidence that the intervention works Requires 2+ studies Most estimates are negative or insignificant
What we don't know and needs more research	No evidence that the intervention works, in part due to too little evidence on the intervention effectiveness, needs further evaluation. Requires: 1-2 studies

 Using the rubric discussed above, we discuss the effectiveness of interventions to address common problems in the education sector.

- Using the rubric discussed above, we discuss the effectiveness of interventions to address common problems in the education sector.
- In doing so we focus on the two broadest types of goals for education interventions:

- Using the rubric discussed above, we discuss the effectiveness of interventions to address common problems in the education sector.
- In doing so we focus on the two broadest types of goals for education interventions:
 - 1. Increasing time in school

- Using the rubric discussed above, we discuss the effectiveness of interventions to address common problems in the education sector.
- In doing so we focus on the two broadest types of goals for education interventions:
 - 1. Increasing time in school
 - 2. Increasing student learning as measured by test scores.

► A household's decision on whether to send their child to school is based on comparing the (perceived) costs and benefits

- ► A household's decision on whether to send their child to school is based on comparing the (perceived) costs and benefits
 - If they think the benefits exceed the costs, then they will enroll their child for another year.

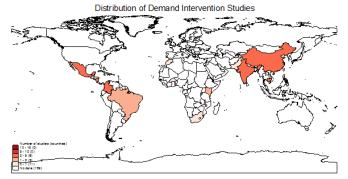
- ► A household's decision on whether to send their child to school is based on comparing the (perceived) costs and benefits
 - If they think the benefits exceed the costs, then they will enroll their child for another year.
- However, there are reasons why households may not enroll their child, even if they think the benefits exceed the costs

- ► A household's decision on whether to send their child to school is based on comparing the (perceived) costs and benefits
 - If they think the benefits exceed the costs, then they will enroll their child for another year.
- However, there are reasons why households may not enroll their child, even if they think the benefits exceed the costs
 - 1. Parents may not accurately perceive the returns (or costs) to education

- ➤ A household's decision on whether to send their child to school is based on comparing the (perceived) costs and benefits
 - If they think the benefits exceed the costs, then they will enroll their child for another year.
- However, there are reasons why households may not enroll their child, even if they think the benefits exceed the costs
 - Parents may not accurately perceive the returns (or costs) to education
 - Households may be credit constrained and thus cannot afford to send their child to school

- ➤ A household's decision on whether to send their child to school is based on comparing the (perceived) costs and benefits
 - If they think the benefits exceed the costs, then they will enroll their child for another year.
- However, there are reasons why households may not enroll their child, even if they think the benefits exceed the costs
 - Parents may not accurately perceive the returns (or costs) to education
 - Households may be credit constrained and thus cannot afford to send their child to school
 - 3. Even if parents choose the optimal level of education, they are unlikely to take into account the social benefits from their child's education. From society's viewpoint, they may under-invest in education

Studies on interventions that are intended to increase households' demand for schooling are distributed across 20 countries. Most studies are from: China, Colombia, India and Mexico



Many studies have examined the impact of inputs, broadly defined, on students' educational outcomes.

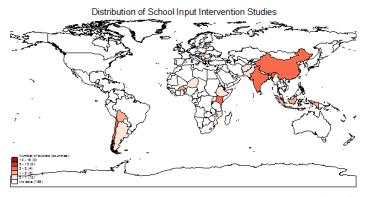
- Many studies have examined the impact of inputs, broadly defined, on students' educational outcomes.
- This reflects

- Many studies have examined the impact of inputs, broadly defined, on students' educational outcomes.
- This reflects
 - the fact that most of the costs of building and operating schools are input costs

- Many studies have examined the impact of inputs, broadly defined, on students' educational outcomes.
- This reflects
 - the fact that most of the costs of building and operating schools are input costs
 - the fact that most inputs are easier to measure than other factors that affect education outcomes. Thus data on inputs are more likely to be collected.

- Many studies have examined the impact of inputs, broadly defined, on students' educational outcomes.
- This reflects
 - the fact that most of the costs of building and operating schools are input costs
 - the fact that most inputs are easier to measure than other factors that affect education outcomes. Thus data on inputs are more likely to be collected.
- ▶ In some cases it may be that a combination of different inputs are required for effective learning. Here, the marginal effect of any one input on educational outcomes will vary according to the levels of other inputs

Studies on the impacts of school input interventions are from about 20 countries. They are concentrated in Chile, China, India and Kenya.



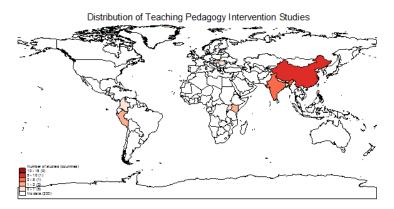
A critical determinant of the extent to which increases in schooling inputs translate into improved learning outcomes is the way in which these inputs are used in practice, and the way in which teaching and learning is organized

- A critical determinant of the extent to which increases in schooling inputs translate into improved learning outcomes is the way in which these inputs are used in practice, and the way in which teaching and learning is organized
- In many developing countries, pedagogical practices have changed very little in the past several decades

- A critical determinant of the extent to which increases in schooling inputs translate into improved learning outcomes is the way in which these inputs are used in practice, and the way in which teaching and learning is organized
- In many developing countries, pedagogical practices have changed very little in the past several decades
- There is relatively little scope for differentiating instruction to account for the actual level and variation of student ability within a classroom

- A critical determinant of the extent to which increases in schooling inputs translate into improved learning outcomes is the way in which these inputs are used in practice, and the way in which teaching and learning is organized
- In many developing countries, pedagogical practices have changed very little in the past several decades
- There is relatively little scope for differentiating instruction to account for the actual level and variation of student ability within a classroom
- ► Teaching effectively may be particularly challenging in many developing country contexts because of the high variation in the initial preparation of children when they enter school

 Studies on the impacts of teaching pedagogy are limited to 9 countries and most are from China or India

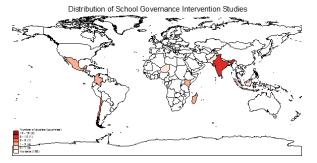


 School governance refers to a broad set of characteristics of education systems that determine how efficiently they are managed

- School governance refers to a broad set of characteristics of education systems that determine how efficiently they are managed
 - These include: goal setting, personnel policy (hiring, training, retention, and promotions), accountability and monitoring, and performance management

- School governance refers to a broad set of characteristics of education systems that determine how efficiently they are managed
 - These include: goal setting, personnel policy (hiring, training, retention, and promotions), accountability and monitoring, and performance management
 - Governance also includes: decentralization, the extent of choice and competition in school "markets", and the regulatory structure for private schools

▶ Of 24 studies in 11 countries, about 1/3 were conducted in India while none were conducted in China.



Understanding the landscape of education interventions and programming brings a for education investment spending, which is to increase the knowledge base of "what works".

- Understanding the landscape of education interventions and programming brings a for education investment spending, which is to increase the knowledge base of "what works".
 - There is no shortage of creative interventions that we have summarized here, and even more that are in the pipeline.

- Understanding the landscape of education interventions and programming brings a for education investment spending, which is to increase the knowledge base of "what works".
 - There is no shortage of creative interventions that we have summarized here, and even more that are in the pipeline.
- ► SIDA & EBA have an opportunity to be leaders in widening the scope of understanding around these creative interventions by investing in more "proof of concept" programs that are both creative as well as implemented in a way that in amenable to evaluation.

- Understanding the landscape of education interventions and programming brings a for education investment spending, which is to increase the knowledge base of "what works".
 - There is no shortage of creative interventions that we have summarized here, and even more that are in the pipeline.
- ► SIDA & EBA have an opportunity to be leaders in widening the scope of understanding around these creative interventions by investing in more "proof of concept" programs that are both creative as well as implemented in a way that in amenable to evaluation.
- While we have made headway in understanding the effectiveness of popular interventions, we still have much to learn to improve the education outcomes for many children around the world.