

# **Decentralization of Health and Education in Developing Countries: A Quality-Adjusted Review of the Empirical Literature**

Anila Channa  
Jean-Paul Faguet

*This Version: July 11, 2012*

## **Abstract**

We review empirical evidence on the ability of decentralization to enhance preference matching and technical efficiency in the provision of health and education in developing countries. Many influential surveys have found that the empirical evidence of decentralization's effects on service delivery is weak, incomplete and often contradictory. Our own unweighted reading of the literature concurs. But when we organize the evidence first by substantive theme, and then – crucially – by empirical quality and the credibility of its identification strategy, clear patterns emerge. Higher quality evidence indicates that decentralization increases technical efficiency across a variety of public services, from student test scores to infant mortality rates. Decentralization also improves preference matching in education, and can do so in health under certain conditions, although there is less evidence for both. We discuss individual studies in some detail. Weighting by quality is especially important when evidence informs policy-making. Firmer conclusions will require an increased focus on research design, and a deeper examination into the prerequisites and mechanisms of successful reforms.

**Keywords:** *Decentralization, School-Based Management, Education, Health, Service Delivery, Developing Countries, Preference Matching, Technical Efficiency*

## 1. Introduction

In the late 1990s it was estimated that 80 percent of the world's countries were experimenting with one form or another of decentralization (Manor 1999). Since then, new or deepening reforms have been announced in nations as diverse as Bolivia, Cambodia, Egypt, Ethiopia, France, Indonesia, Japan, Mexico, South Korea, Turkey, and many others. By now it is safe to say that experiments with, and enthusiasm for, decentralization are essentially ubiquitous across the globe.

Theories underpinning such enthusiasm are compelling and argue that by taking the government "closer to the people", decentralization can improve the responsiveness and accountability of the state, decrease corruption, increase the political voice and participation of ordinary citizens, and also reduce bureaucracy and lower the unit costs of government expenditure. The slogan "closer to the people" can be decomposed into three underlying analytical advantages that local governments have over central government: (a) superior information on local conditions and needs, (b) greater participation of citizens in decision making and the production of local services, and (c) greater accountability of public officials to voters. The local governments possessing such advantages preside over jurisdictions that are smaller and more homogeneous than those of national government. Local governments' decision making will thus be facilitated by not having to cater to a more diverse set of needs and wants. With superior information, participation, accountability, and policy challenges that are less onerous, it follows logically that decentralization should improve public services.

Yet the many surveys of the literature overwhelmingly agree that empirical evidence is inconclusive. In one of the earliest reviews, for instance, Rondinelli et al. (1983) note that decentralization seldom, if ever, lived up to its promise. Shah et al. (2004) concur in a review of 56 studies published since the late 1990's, chronicling that decentralization in some cases improved, and in others worsened, service delivery, corruption and growth across a large range of countries. Treisman's (2007) more recent survey is bleaker still. "To date," he says, "there are almost no solidly established, general empirical findings about the consequences of decentralization" (p. 250). The lack of consensus on decentralization's effects over 25 years and literally hundreds of studies is striking.

One of the main challenges faced by such review efforts is the sheer size and diversity of scholarship. The empirical literature on decentralization originates from a variety of disciplines, including public economics, development studies and comparative

politics, to name just a few. Evaluations of reforms are done in markedly different ways and focus on very different outcomes, ranging from service delivery to corruption to macro-economic stability and happiness. Any attempt to review these results as a whole quickly loses the forest for the trees in a confusion of particular findings that may appear contradictory, but are more often simply different. To draw firmer conclusions from this vast literature, we argue, a clearer organizing principle is required – a principle that allows students of decentralization to neatly compare the *causal* effects of a similar kind of reform on similar predefined outcomes.

In this review we apply such an organizing principle to assess decentralization’s ability to enhance service delivery in developing nations. Decentralization, defined here as “the transfer of authority for decision-making, finance, and management to quasi-autonomous units of local governments” (Litvack and Seddon 1999: pp.3), is probably the single most advocated measure for improving the provision of health and education in the developing world. This popularity is not surprising. Of the many arguments in favor of decentralization, the most important is that devolving power and resources to local governments can increase the accountability of public servants, and hence the responsiveness of public services to citizens’ needs (Faguet 2012 and 2008). While decentralizers’ motives have no doubt differed across different countries, improved delivery of public goods has been at least an implicit goal of most reforms, and usually an explicit one.

To ensure that our conclusions on decentralized public provision of health and education are not influenced by an arbitrarily selected group of studies, we use predefined criteria to identify papers for inclusion in this survey. Specifically, we focus on empirical evidence in the economics literature from the last 20 years that evaluates the causal effect of decentralization on service delivery in developing nations. We group these studies according to the main substantive themes they address as follows: 1) **Preference matching**<sup>1</sup>, defined as the extent to which public goods provided by local governments match citizens’ preferences or demands, and/or 2) **Technical efficiency**, meaning the production of more or better public goods by a decentralized government for a given set of inputs. In the latter theme, we further segregate the evidence into sub-categories based on whether it concerns the provision of (i) health, (ii) education to

---

<sup>1</sup> In the decentralization policy literature, the term allocative efficiency is often used to refer to this same concept.

lower tiers of the government, or (iii) education to schools in what is commonly referred to as a School-Based Management (SBM) reform.

Within this thematic classification, we further classify studies according to the self-reported quality of their data and credibility of their identification strategies, and place greater weight on what high-quality evidence has to say. Distinguishing between studies that are able to tease out the causal effects of decentralization more plausibly than others is the crucial step that allows us to identify patterns in the findings. Earlier empirical contributions on decentralization were commonly plagued by problems of attribution – surveys based on such evidence therefore had similar challenges in isolating the effects of reform. In recent years a deeper appreciation of the pitfalls associated with causal inference has pushed empiricists to find more credible identification strategies that use observational data to construct valid counterfactuals, and thus approximate the ‘gold standard’ of randomized experiments<sup>2</sup>. This is the higher quality literature we focus on in our review.

The following sections lay out our methodology and results, and then discuss and contrast individual studies in some detail. But it may be useful to first summarize our findings.

- ***The overall evidence base is thin, although this varies by category***

We find that the overall evidence base on decentralized health and service delivery in developing countries is thin. Only 35 studies meet the selection criteria detailed in section 2 below.

We also find that the distribution of scholarship is skewed by theme (preference matching vs. technical efficiency) and sub-category (health, education, SBM). For example, many more studies focus on how decentralization affects technical efficiency than preference matching. Likewise, education and SBM have been the subject of examination much more often than health. The papers reviewed are summarized by theme and sub-category in Table 1.

---

<sup>2</sup> See Angrist and Pischke 2010 for a good discussion on identification strategies

- ***The econometric techniques employed by studies are less sophisticated than we would prefer***

Fewer than a third of the papers reviewed can be classified as having a highly credible identification strategy. Our categorization hinges on the ability of the methodologies employed to mitigate endogeneity concerns, in accordance with the established hierarchy of econometric techniques. Hence, for example, randomized and quasi-randomized evidence are considered to have stronger identification strategies than cross-sectional work. Table 2 describes how we categorize the credibility of a study's empirical design in detail.

In this particular sense, the "quality" of the studies reviewed also varies substantially by theme and sub-category. Papers in the technical efficiency theme, and specifically studies investigating school decentralization reforms, appear to have a greater number of high quality contributions. By contrast, contributions in the preference matching category are not only fewer but also less rigorous, making the task of drawing conclusions from this group difficult.

- ***Externalities in health drive pessimism in the preference matching theme***

Our review indicates that pessimism in the small preference matching literature is due primarily to the externalities that characterize the health sector.

Decentralized local governments often match local preferences more efficiently while ignoring spillover effects on neighboring regions, as some of the classic public economics literature predicts (Oates 1972; Rubinfeld 1987), thus reducing overall social welfare. The evidence of preference matching in education delivery, on the other hand, appears to be somewhat positive. But the small size of this body of work limits firmer conclusions.

- ***Higher quality work on technical efficiency appears to be favorable***

Importantly, evidence on technical efficiency is on the whole optimistic. This optimism rises with the quality of the evidence. The highest quality empirics show that decentralization can enhance a variety of service delivery outcomes, from student test scores to infant mortality rates.

Although such results are not conclusive, they do demonstrate the potential of decentralization to enhance service delivery in developing countries. Stronger conclusions are not possible until the field sees a more general shift towards better research design, and the development of a deeper understanding of the prerequisites and mechanisms of successful reforms.

The rest of this paper proceeds as follows. Section 2 describes the criteria used for including and classifying studies in this review. Section 3 discusses the papers included by theme. We conclude by comparing our findings to broad surveys of the literature and suggesting priorities for future research. The papers reviewed are summarized in Tables 1 to 4. Our aim is to provide insights into patterns of findings on one piece of the larger decentralization puzzle. What follows hopefully helps to answer some important question surrounding decentralized service delivery of health and education in developing countries. Even so, we do not pretend that this survey can cover more than a fraction of a huge literature.

## **2. Scope and Methodology of Review**

This section describes the steps we undertook to identify, organize and classify studies from this vast literature for our review. Our strategy was to conduct a wide search and then systematically filter papers that met predefined criteria for relevance. Following this, we arranged the literature first by substantive theme, and then – crucially – by quality of the evidence. The latter step allows us to see empirical patterns in the scholarship that broad decentralization reviews have not previously uncovered.

### **2.1 Identifying and Organizing the Literature**

We began our search by focusing on published and unpublished working papers found on EconLit, an established bibliography of economics literature. Using the key words “decentralization” or “devolution” in conjunction with either “education”, “health” or “service delivery”, we first conducted our search in October 2010, and then subsequently updated it in December 2011.

The EconLit database yielded approximately 350 studies on decentralization and health, and some 350 more on decentralization and education. To ensure thorough coverage we then cast a wider net, consulting other key resources such as Google

Scholar, JSTOR, SCOPUS and Web of Knowledge. We also referred to major publications by international organizations such as the World Bank, used citation indices, and reviewed reference lists in identified papers to confirm that no critical contributions were omitted from our review. This pushed the total number of studies above 1,000.

Based on the information contained in their abstracts, we selected those studies that a) were of an empirical nature, b) dealt specifically with decentralization of service delivery of health and/or education in developing nations<sup>3</sup>, and c) date from 1992 or later. This reduced the sample dramatically to 35 contributions. We then organized this short-listed body of evidence into our two themes: preference-matching, and technical efficiency. The body of scholarship on technical efficiency is larger, and therefore we further segregated the papers examined into sub-categories based on whether they address: (1) decentralization of health, (2) decentralization of education to lower tiers of governments, or (3) decentralization of education to schools or School-Based Management.

Table 1 summarizes the evidence by theme, sub-category, author, publication type, and countries covered.

---

<sup>3</sup> For a good review of empirical evidence from OECD and other developed nations, refer to Ahmad, Brosio and Tanzi (2008) and Ahmad and Brosio (2009).

Table 1: Summary of Evidence Reviewed												
Region		Author	By Theme		By Sub-Category			By Type of Publication				
			PM	TE	Health	Educ.	SBM	Journal	Book	Working Paper	Other Unpublished	Description
<b>Africa</b>												
1	Kenya	Dufo et al (2007)		X			X				X	Unpublished manuscript
2	Madagascar	Glewwe and Maiga (2011)		X			X	X				Journal of Development Effectiveness
3	Uganda	Akin et al (2005)	X		X			X				Journal of Development Studies
<b>Asia</b>												
4	China	Uchimura and Jutting (2009)		X	X			X				World Development
5	India	Asfaw et al (2007)		X	X			X				Journal of Developing Areas
6	Indonesia	Skoufias et al (2011)	X		X	X				X		World Bank Policy Research Paper Series
7	Pakistan	Hasnain (2008)	X		X	X				X		World Bank Policy Research Paper Series
8	Pakistan	Aslam and Yilmaz (2011)		X		X		X				Public Administration and Development
9	Philippines	Schwartz et al (2002)	X		X						X	MEASURE Evaluation, University of North Carolina at Chapel Hill
10	Philippines	Jimenez and Paqueo (1996)		X		X		X				Economics of Education Review
11	Philippines	Lockheed and Zhao (1993)		X		X		X				International Journal of Educational Development
12	Philippines	Khattri et al (2010)		X			X			X		World Bank Policy Research Paper Series
13	Russia	Frienkman and Plekhanov (2009)	X	X		X		X				World Development
<b>Latin &amp; Central America</b>												
14	Argentina	Habibi et al (2003)		X	X	X		X				Journal of Human Development
15	Argentina	Galiani et al (2008)		X		X		X				Journal of Public Economics
16	Argentina	Eskeland and Filmer (2007)		X			X			X		World Bank Policy Research Paper Series
17	Bolivia	Faguet (2012)	X		X	X			X			Book
18	Bolivia	Faguet (2004)	X		X	X		X				Journal of Public Economics
19	Bolivia	Inchauste (2009)		X	X	X			X			Book Chapter
20	Brazil	Paes de Barros and Mendonca (1998)		X			X	X				Organization matters: agency problems in health and education in Latin America
21	Chile	Di Gropello (2002)		X		X		X				World Bank Economists' Forum
22	Colombia	Faguet and Sanchez (2008)		X		X		X				World Development
23	El Salvador	Jimenez and Sawada (1999)		X			X	X				The World Bank Economic Review
24	El Salvador	Sawada and Ragatz (2005)		X			X		X			Book Chapter in World Bank Publication
25	Honduras	Di Gropello and Marshall (2005)		X			X		X			Book Chapter in World Bank Publication
26	Mexico	Gertler et al (2011)		X			X	X				Journal of Development Economics
27	Mexico	Skoufias and Shapiro (2006)		X			X			X		World Bank Policy Research Paper Series
28	Nicaragua	King and Ozler (2000)		X			X			X		World Bank Policy Research Paper Series
29	Nicaragua	Parker (2005)		X			X		X			Book Chapter in World Bank Publication
<b>Cross-country</b>												
30	Various	Arze del Granado et al (2005)	X		X	X				X		Georgia State University
31	Various	Gunnarsson et al (2009)		X		X		X				Economic Development and Cultural Change
32	Various	Hanushek et al (2011)		X			X			X		NBER Working Paper Series
33	Various	Khaleghian (2004)		X	X			X				Social Science and Medicine
34	Various	Robalino et al (2001)		X		X				X		World Bank Policy Research Paper Series
35	Various	Treisman (2002)		X	X	X					X	Unpublished manuscript
<b>Total</b>			<b>7</b>	<b>28</b>	<b>12</b>	<b>16</b>	<b>14</b>	<b>18</b>	<b>4</b>	<b>9</b>	<b>3</b>	

**Notes**

PM: Preference Matching

TE: Technical Efficiency

SBM: School-based Management



### 2.3 Quality of the Evidence

Next we evaluated the quality of the evidence. We do so in the knowledge that researchers attempting to assess the effects of decentralization on education and health services face a number of difficult challenges. These include the difficulty of disaggregating decentralization's effects from (a) the other reforms that often accompany it, and – more importantly – (b) from country politics, which necessarily plays a crucial role in both decentralization's impetus as well as its eventual effects. Together these impose sizeable data demands on researchers. Add to these problems the time it takes for service delivery outcomes to change, and the difficulty of conducting randomized decentralization experiments, and the varying quality of the evidence seems fully justified.

To classify the persuasiveness of each paper's identification strategy, we use a four point scale of *Very Strongly Credible*, *Strongly Credible*, *Somewhat Credible*, and *Less Credible*. In our categorization, the primary consideration is the nature of the data available and the identification strategy this permits. In effect, we rank papers' empirical methodologies according to their widely accepted abilities to mitigate endogeneity concerns and identify causal effects. For this we rely on the established hierarchy of identification strategies in economics as widely taught in graduate programs today. We supplement our quality distinctions by reviewing the covariates included in analysis, the measures used for decentralization<sup>4</sup>, the self-reported quality of data, and the nature of robustness checks performed in the paper. This scale, along with a snapshot of how papers in this survey have been classified, is presented in Table 2. The categorization is adapted from a similar typology by Santibañez (2006).

---

<sup>4</sup> For a good discussion on challenges of measuring decentralization, see Ebel and Yilmaz (2002)

**Table 2: Quality Distinctions**

Scale	Criteria	Preference Matching			Technical efficiency	
			Generally Positive Findings	Generally Insignificant or Negative Findings	Generally Positive Findings	Generally Insignificant or Negative Findings
<b>Less Credible Identification Strategy</b>	<ul style="list-style-type: none"> <li>Research that bases findings on self-selected populations, and makes little to no effort to produce a valid comparison group.</li> <li>Work that is likely to suffer from serious omitted variable bias and other endogeneity issues such as those related to measurement error due to self-reported poor quality of data.</li> <li>Most cross-sectional work that does not use any other sophisticated methodology to address endogeneity will fall in this category.</li> </ul>	<b>Health</b>	-	<ul style="list-style-type: none"> <li>Hasnain (2008)</li> <li>Schwartz et al (2002)</li> </ul>	-	<ul style="list-style-type: none"> <li>Treisman (2002)</li> </ul>
		<b>Education</b>	-	<ul style="list-style-type: none"> <li>Hasnain (2008)</li> <li>Freinkman and Plekhanov (2009)</li> </ul>	<ul style="list-style-type: none"> <li>Di Gropello (2002)</li> <li>Jimenez and Paqueo (1996)</li> </ul>	<ul style="list-style-type: none"> <li>Lockheed and Zhao (1993)</li> <li>Treisman (2002)</li> </ul>
		<b>SBM</b>	-	-	<ul style="list-style-type: none"> <li>Eskeland and Filmer (2007)</li> </ul>	
<b>Somewhat Credible Identification Strategy</b>	<ul style="list-style-type: none"> <li>Research that attempts to construct a valid comparison group but does so with limited success.</li> <li>Work that is likely to continue to suffer from some endogeneity biases in spite of efforts at mitigation.</li> <li>Cross-sectional work that uses matching techniques, for instance, falls in this group. Other studies using panel estimation may also fall in this group if they use random effects or between effects. Papers using difference in differences but without providing any support of its key identifying assumption of parallel trends, papers using IV that are not considered particularly strong and papers using fixed effects but with very limited covariates also fall here.</li> <li>Finally, preference matching studies that only establish a change in allocation patterns but not any enhanced alignment to citizen preferences also are in this category.</li> <li>Compared to previous category, work in this group generally attempts to validate the measure of decentralization through use of multiple measures or qualitative validation.</li> </ul>	<b>Health</b>	<ul style="list-style-type: none"> <li>Arze del Granado et al (2005)</li> </ul>	<ul style="list-style-type: none"> <li>Akin et al (2005)</li> </ul>	<ul style="list-style-type: none"> <li>Asfaw et al (2007)</li> <li>Robalino et al (2001)</li> <li>Habibi et al (2003)</li> </ul>	<ul style="list-style-type: none"> <li>Inchauste (2009)</li> <li>Khaleghian (2003)</li> </ul>
		<b>Education</b>	<ul style="list-style-type: none"> <li>Arze del Granado et al (2005)</li> </ul>	-	<ul style="list-style-type: none"> <li>Aslam and Yilmaz (2011)</li> <li>Freinkman and Plekhanov (2009)</li> </ul>	<ul style="list-style-type: none"> <li>Inchauste (2009)</li> </ul>
		<b>SBM</b>	-	-	<ul style="list-style-type: none"> <li>Paes de Barros and Mendonca (1998)</li> <li>King and Ozler (2000)</li> <li>Khattari et al (2010)</li> <li>Parker (2005)</li> <li>Di Gropello and Marshall (2005)</li> </ul>	<ul style="list-style-type: none"> <li>Gunnarsson et al (2009)</li> <li>Jimenez and Sawada (1999)</li> <li>Sawada and Ragataz (2005)</li> </ul>

<b>Strongly Credible Identification Strategy</b>	<ul style="list-style-type: none"> <li>Research that is able to construct a reasonable comparison group.</li> <li>Work that specifically attempts to address sources of endogeneity and is mostly successful in its attempt.</li> <li>Most of the studies in this category quasi-experimental designs such as difference in differences and instrumental variables. Papers providing panel estimates in a fixed effects model while controlling for more than one socio-economic covariate and more than one covariate from the health/education production function also fall here.</li> </ul>	<b>Health</b>	-	<ul style="list-style-type: none"> <li>Faguet (2004)</li> <li>Faguet (2012)</li> <li>Skoufias et al (2011)</li> </ul>	<ul style="list-style-type: none"> <li>Uchimura and Jutting (2009)</li> </ul>	
		<b>Education</b>	<ul style="list-style-type: none"> <li>Faguet (2004)</li> <li>Faguet (2012)</li> <li>Skoufias et al (2011)</li> </ul>	-	<ul style="list-style-type: none"> <li>Galiani et al (2008)</li> <li>Faguet and Sanchez (2008)</li> </ul>	
		<b>SBM</b>	-	-	<ul style="list-style-type: none"> <li>Skoufias and Shapiro (2006)</li> <li>Gertler et al (2011)</li> <li>Hanushek et al (2011)</li> </ul>	
<b>Very Strongly Credible Identification Strategy</b>	<ul style="list-style-type: none"> <li>Studies with very strong research design such as randomized control trials that have a valid control group fall here.</li> <li>Work that is likely to have limited endogeneity concerns.</li> </ul>	<b>Health</b>	-	-	-	-
		<b>Education</b>	-	-	-	-
		<b>SBM</b>	-	-	<ul style="list-style-type: none"> <li>Duflo et al (2007)</li> </ul>	<ul style="list-style-type: none"> <li>Glewwe and Maiga (2011)</li> </ul>

Our top category, *Very Strongly Credible*, consists of randomized control trials (RCTs), the ‘gold standard’ for identifying causal effects. At the other end of the spectrum, work that relies on simpler quantitative methods such as ordinary least squares (OLS), and fails to employ any more sophisticated methodology to control for endogeneity bias, is categorized as having a design that is *Less Credible* in drawing causal inferences. The papers we place in this category are mostly cross-sectional OLS analyses of observational data. Because studies here attempt to draw findings from self-selected populations without being able to construct a valid comparison group, their ability to make causal claims is limited when compared to studies that use more sophisticated econometric methods.

How do we place work into the two middle groupings? The *Strongly Credible* category consists of research that is reasonably successful in producing a valid comparison group. Much of this literature uses quasi-experimental techniques such as instrumental variables (IV) or difference in differences (DID) approaches. The key issue for being classed as *Strongly Credible* is how persuasive studies are in communicating a thorough understanding of the institutional environment and then – importantly – using this understanding to design their empirical strategy.

So for instance, *Strongly Credible* papers using IV techniques can make plausible claims for the relevance and exogeneity of their instruments. Studies using difference in differences can persuade that the treatment is what is responsible for altering a trend between treatment and control groups. For this reason, the category also contains some panel data estimations using fixed effects and a set of relevant covariates, but only where the case for limited endogeneity based on knowledge of confounding factors is particularly convincing.

The remaining studies are classed as *Somewhat Credible*. In our view, studies in this category are less persuasive in addressing endogeneity than those that are *Strongly Credible*, but more convincing than the *Less Credible* set due to their use of various kinds of comparison groups. This category thus houses diverse econometric methods, from matching to instrumental variables.

It is worth underlining what this survey does not seek to do. We recognize that the identification strategies employed by researchers are largely determined by a combination of the data available, the nature of the reforms implemented, and the nuanced questions they seek to answer. Hence we make no attempt to rank papers’ broader quality *as pieces of research*, nor comment on the analytical skills of their

authors. What we do seek to do, rather, is recognize that there is an established hierarchy of rigor in econometric identification, and apply that hierarchy to the evidence that the literature provides. This allows us to roughly categorize how convincing studies' results are, where credibility is principally determined by what data is available, and hence how we should weight evidence when making policy.

Henceforth we focus on the results of studies falling into the three stronger quality categories. Studies in the *Less Credible* category are occasionally highlighted when they present specific policy ideas that relate to studies in the more credible groups.

### **3. The Effects of Decentralization on Education and Health**

We now move on to the heart of our review, and describe the papers included in this survey in some detail. We first discuss preference matching and then turn to technical efficiency. Throughout this section, we make use of our quality distinctions when describing key papers in order to allow readers to understand how our conclusions are drawn.

#### **3.1 Preference Matching**

Although preference matching is one of the classic arguments posited in favor of decentralization (see Oates 1972), the empirical evidence devoted exclusively to testing this proposition is surprisingly small. It also produces somewhat contradictory results for the service delivery of education on one hand, and the provision of health on the other.

The book and two papers we review with *Strongly Credible* identification strategies in the theme of preference matching yield somewhat contradictory findings. In one of the first papers to employ a before and after estimation strategy to examine preference matching, Faguet (2004; see also Faguet 2012 for a substantial update) finds favorable evidence from Bolivia. Bolivia undertook devolution in 1994, and as part of the reform moved responsibility of key public services to local governments. The shift in responsibility was accompanied by two other critical changes – the doubling of funds available to these devolved units during this period, and the establishment of oversight bodies to monitor local spending.

Faguet examines the patterns of investment in public investment projects in a total of ten categories, including education and health, from 7 years before and 3 years after this reform. By doing so first for municipal averages, and then one by one for all 311 municipalities examined, he finds a statistically significant increase in investment in education overall, as well as a statistically significant increase in 71% of individual municipalities in just three years after devolution.

This shift in investment patterns was especially evident in poorer regions. As devolution increased funding to previously neglected regions, this finding is not necessarily an indication of greater preference matching. But Faguet then offers further evidence to support his responsiveness argument – he demonstrates that regions with high illiteracy levels, or where there seemed to be a greater need, invested more heavily in education. Regions with strong education indicators, on the other hand, prioritized

other sectors. This, he contends, “implies that local government is more sensitive to local need than central government” (p. 24). The author’s optimism is supported by similar findings in the sectors of water management and urban development, but noticeably not in our second sector of interest here – health.

However, greater spending on socially oriented sectors does not necessarily imply that preference matching has improved. This is the main contrasting finding in Skoufias et al.’s (2011) recent working paper on Indonesia. The paper exploits an arguably exogenous phasing of local direct elections to conduct a difference in differences analysis of the effect of political decentralization on the pattern of public spending.

Although fiscal and administrative devolution commenced in 1999 in the country, in 2005 Indonesia implemented electoral reforms to enhance accountability in service delivery. The date local elections were held in a particular municipality depended on when the previously appointed head of government completed their tenure. Deviations in this timing were a result of illness, death, no confidence votes and the creation of new districts and thus, claim the authors, exogenous to spending allocation.

Skoufias et al. compare changes in expenditure patterns in districts that held local elections in 2005 to patterns in districts that did not hold elections until 2008, after providing proof that pre-implementation trends in spending in both groups were similar. They find that political decentralization was associated with greater overall public spending – when disaggregated, however, they demonstrate that while there was an increase in the education sector, there was no significant difference in health spending. Skoufias et al. then follow Faguet (2004), attempting to use his methodology to establish whether these shifting patterns were based on local needs. In contrast to Faguet, however, they find no evidence to suggest an improvement in preference matching at all.

The two contributions from the *Somewhat Credible* identification group are also contrary. In the only cross-country study concerned with preference matching that we review here, Arze del Granado et al. (2005) seek to establish that Faguet’s (2004) findings on the change in functional composition post decentralization in Bolivia are “not a unique experience of a specific country” (p. 4). Employing a before and after strategy similar to Faguet, but using data for 45 developed and developing countries over 28 years, Arze del Granado et al. analyze the relationship between the ratio of local expenditure to total government expenditure as the measure of decentralization, and the ratio of health and education spending to overall spending.

The authors employ a total of five econometric models to ensure robustness of their results, including OLS with fixed and random effects. They control for per capita income, budget balance and population and use time effects in all, and country effects in some, of their specifications. They find a statistically significant relationship between decentralization and expenditure ratios. Because, they assert, “implicit in the argument that decentralization can increase allocative efficiency is the implication that decentralization is likely to alter the composition of public expenditures” (Arze del Granado et al. 2005: p. 2), they concur with Faguet (2004) and conclude in favor of the potential of decentralization to enhance preference matching.

Akin et al. (2005) take a slightly different tack. They attempt to provide a deeper understanding of spending allocations within the health sector after a decentralization reform occurred in Uganda. The authors postulate a model in which users undervalue public-type health goods such as family planning, health education, immunization and infectious disease control. Because local governments will be more responsive to the preferences revealed by their residents for private-type health goods, the authors posit that districts will under-provide public-type health care and ignore spillover effects on neighboring regions if they are not under the same jurisdiction.

Akin et al.’s theory is borne out in the district-level data they examine from Uganda. They find, after controlling for per capita income in a fixed effects model, that decentralization is associated with higher budgeting of private-type health goods. Because their main empirical strategy compares districts that decentralized earlier to those that decentralized later, the authors attempt to provide validation that the groups did not differ systematically. Moreover, Akin et al. also cite evidence in favor of crowding-out effects – districts whose neighbors budget higher amounts on public-type goods budget less on such goods themselves. On the basis of this evidence, Akin et al. (2005: p.3) pessimistically call for “A reappraisal of the central government’s role in providing public goods in developing countries”.

What is interesting, however, is that their argument is not one against the preference matching effects of decentralization per se, as they assume local governments are indeed responding to local inclinations in Uganda. Rather, their pessimism arises from the result of the responsiveness versus spillover effects trade-off. The Oates Decentralization Theorem (1972) suggests that devolution is superior only so long as there are no spillover effects. In the presence of spillover effects, the theoretical prediction for



preference matching of decentralization is ambiguous or even negative (Besley and Coate 2003; Bardhan and Mookherjee 1998).

The papers with *Less Credible* empirical designs take Akin et al.'s pessimism even further. Schwartz et al. (2002), for instance, examine the trends in spending composition of health services in 1600 regions in the Philippines to show, like Akin et al., a shift in local spending composition from public-type health services to private-type curative health care. Along a similar vein, Hasnain (2008) considers budget allocation trends in Pakistan's province of Punjab and reports that decentralized local governments are prioritizing allocations for infrastructure over those for health and education. And in sharp contrast to all of the studies above, Frienkman and Plekhanov (2009) do not find a change in allocation patterns after decentralization in Russia at all. The authors use a between effects model on cross-sectional data to conclude that fiscal decentralization is not significantly associated with an investment in education inputs.

So what, if anything, can we take away from this short review of the evidence of decentralization's ability to enhance preference matching? The literature in this theme is small, and the number of high quality contributions is even smaller. But studies across the quality distinctions appear to mostly concur that decentralization changes the patterns of local spending. On the other hand, whether or not these changes are responsive to local needs is an area where there is less agreement. While the evidence appears somewhat encouraging for enhanced preference matching in education, contributions in the area of health are decidedly pessimistic due both to a lack of visible change in allocation patterns and the possibility of externalities in the area.

**Table 3: Selected Empirical Evidence of Preference Matching**

*Source: Cited articles*

No.	Author (Date)	Country of Study	Date Implemented	Programme Description	Method of Analysis	Sample	Measure/s of Decentralization	Results	Identification Strategy	No.
<b>Health</b>										
1	Faguet (2004); Faguet (2012); Faguet and Sanchez (2008)	Bolivia	1994	Increase in devolved funds to LG, responsibility for public services, establishment of oversight committees	OLS using a fixed effects model	Universe of 311 regions over 1987 - 2007	<ul style="list-style-type: none"> <li>Binary measure of before and after D implementation</li> </ul>	<ul style="list-style-type: none"> <li>Investment in health did not change significantly post D</li> </ul>	Strongly Credible	1
2	Skoufias et al (2011)	Indonesia	1999	Increase in devolved funds to LG, responsibility for public services. In 2005, direct election of local government	Difference in differences	200 out of 400 districts during 2001 to 2006	<ul style="list-style-type: none"> <li>Binary measure of election date</li> </ul>	<ul style="list-style-type: none"> <li>Overall public expenditure increased post D</li> <li>No significant change in health spending post D</li> </ul>	Strongly Credible	2
3	Akin et al (2005)	Uganda	1993 - 1994	Fiscal decentralization, with rule-based unconditional grants given to regions	OLS using a fixed effects model	30 out of 45 regions during 1995 - 1998	<ul style="list-style-type: none"> <li>Years since unconditional grant given to LG</li> <li>Proportion of LG expenditure financed by LG revenues</li> </ul>	<ul style="list-style-type: none"> <li>Share of allocation to health decreased in local budgets, esp. in preventive and primary health care areas</li> <li>Some evidence of spillover, where neighbours of high health spending districts have lower health spending</li> </ul>	Somewhat Credible	3
4	Arze del Granado et al (2005)	Cross-country	Various	Various	Various including OLS using fixed and random, also QMLE models	45 developed and developing countries - Unbalanced panel over 1973 - 2000	<ul style="list-style-type: none"> <li>Share of LG expenditure in total government expenditure</li> </ul>	<ul style="list-style-type: none"> <li>Higher D associated with higher proportion of spending on health and education</li> </ul>	Somewhat Credible	4
5	Hasnain (2008)	Pakistan	2001	Limited financial devolution but implementation of rule-based transfers, responsibility for public services, political devolution	Allocation trend analysis	33 out of 35 districts in Punjab over 2006-2007	<ul style="list-style-type: none"> <li>Budget allocations to sectors</li> <li>Relative size of LG spend on sector</li> </ul>	<ul style="list-style-type: none"> <li>LG spend focused on infrastructure and away from health</li> <li>Provinces driving health interventions, providing incentives to shift away budgets</li> </ul>	Less Credible	5
6	Schwartz et al (2002)	Philippines	1991	Increase in devolved funds to LG, responsibility for public services	Spending trend analysis	1600 LG - period includes 4 yrs before and 6 yrs after	<ul style="list-style-type: none"> <li>Binary measure of before and after D implementation</li> </ul>	<ul style="list-style-type: none"> <li>Greater % spend allocated to health post D</li> <li>Higher spend on private health goods, explained by expenses arising from devolution of hospitals</li> </ul>	Less Credible	6

<b>Education</b>										
<b>7</b>	Faguet (2004); Faguet (2012); Faguet and Sanchez (2008)	Bolivia	1994	Increase in devolved funds to LG, responsibility for public services, establishment of oversight committees	OLS using a fixed effects model	Universe of 311 regions over 1987 - 2007	<ul style="list-style-type: none"> <li>▪ Binary measure of before and after D implementation</li> </ul>	<ul style="list-style-type: none"> <li>▪ Investment in education increases significantly post D</li> <li>▪ Investment increases are associated with illiteracy levels</li> </ul>	Strongly Credible	<b>7</b>
<b>8</b>	Skoufias et al (2011)	Indonesia	1999	Increase in devolved funds to LG, responsibility for public services. In 2005, direct election of local government	Difference in differences	200 out of 400 districts during 2001 to 2006	<ul style="list-style-type: none"> <li>▪ Binary measure of election date</li> </ul>	<ul style="list-style-type: none"> <li>▪ Overall public expenditure increased post D</li> <li>▪ Increase in spending on education post D</li> </ul>	Strongly Credible	<b>8</b>
<b>9</b>	Arze del Granado et al (2005)	Cross-country	Various	Various	Various including OLS using fixed and random, also QMLE models	45 developed and developing countries - Unbalanced panel over 1973 - 2000	<ul style="list-style-type: none"> <li>▪ Share of LG expenditure in total government expenditure</li> </ul>	<ul style="list-style-type: none"> <li>▪ Higher D associated with higher proportion of spending on health and education</li> </ul>	Somewhat Credible	<b>9</b>
<b>10</b>	Freinkman and Plekhanov (2009)	Russia	Phased beginning in 1994	Increased fiscal powers with rule-based transfers, responsibility for public services	OLS using a between effects model	73 out of 83 regions, with data collection in 2003	<ul style="list-style-type: none"> <li>▪ Share of LG education expenditure financed by own revenue</li> <li>▪ Type of decentralization arrangement</li> </ul>	<ul style="list-style-type: none"> <li>▪ No significant impact on inputs</li> </ul>	Less Credible	<b>10</b>
<b>11</b>	Hasnain (2008)	Pakistan	2001	Limited financial devolution but implementation of rule-based transfers, responsibility for public services, political devolution	Allocation trend analysis	33 out of 35 districts in Punjab over 2006-2007	<ul style="list-style-type: none"> <li>▪ Budget allocations to sectors</li> <li>▪ Relative size of LG spend on sector</li> </ul>	<ul style="list-style-type: none"> <li>▪ LG spend focused on infrastructure and away from education</li> <li>▪ Provinces driving education interventions, providing incentives to shift away budgets</li> </ul>	Less Credible	<b>11</b>

**Notes**

PM: Allocative Efficiency

TE: Technical Efficiency

D: Decentralization

LG: Local Government

## 3.2 Technical Efficiency

The body of work on the ability of decentralization to enhance technical efficiency in the delivery of education and health fortunately is much larger than that found in the previous theme. Strikingly, it is also more rigorous, and fairly optimistic of the potential of decentralization to improve service delivery.

### 3.2.1 Health

The lone paper with a *Strongly Credible* empirical strategy in this sub-category, for instance, provides the first piece of evidence strongly in favor of decentralization's ability to enhance technical efficiency in health delivery.

Uchimura and Jutting (2009) examine the interesting case of China, a country that has had consistently high levels of spending decentralization, but a growing recentralization of revenue decisions since 1994. Improving on previous studies that use only province-level data, Uchimura and Jutting employ data from counties in 26 provinces over a seven year period. Counties in China have responsibility for implementing health programs. However, local government officials are elected through parties, not the adult franchise, which limits political accountability of officials to citizens.

The authors determine the statistical relationship between two measures of county-level fiscal decentralization and the outcome of provincial infant mortality rates (IMR), while controlling for key health production function elements such as illiteracy rates, fertility rates and per capita income. The main specification includes fixed effects. Finding statistically significant and negative coefficients in most of the models tested, the authors conclude that counties in more fiscally decentralized provinces have lower IMR. Interactions between their two measures of decentralization – own expenditure financed and proportion of provincial expenditure - are also positive. This suggests to the authors that IMRs are lower in provinces not only where fiscal capacity is strengthened, but also where counties and provinces have a functional transfer system in place.

Two contributions from the group with *Somewhat Credible* evidence are also positive. In a study quite similar to the above, Asfaw et al. (2007) consider empirical evidence on rural infant mortality rates from India. Decentralization in India has a long history. But, it took its current form with the passing of the 1989 Panchayat Raj bill and later constitutional amendments in the early 1990s that devolved power to the traditional

village organizations or Panchayats. Panchayats now form a part of the local government, hold elections, and bear responsibility for health and education delivery. Evidence suggests, however, that different states have followed differing models of devolution, making comparative analysis of the reform difficult.

Nonetheless, Asfaw et al. attempt to estimate the role of devolution in affecting the outcome of rural infant mortality rates using data from 14 states over seven years. In order to improve the measure of fiscal decentralization, the authors use factor analysis of three different indicators to build a decentralization index, including a Panchayat's share in total expenditure, Panchayat expenditure per rural population and share of Panchayat's own revenue in Panchayat's expenditure. Controls are included for per capita income of state, percentage share of literate women, and an index of political decentralization (constructed from data on voter turnout, women's participation in polls and number of polling stations). Notably absent though are important health production function components such as fertility.

The authors demonstrate a statistically significant and negative relationship between decentralization and IMR in both their random and fixed effects models, but not in the between-effects model. Asfaw et al. (2007) conclude that having an above average decentralization index is associated with a 17.16% reduction as compared to states with below average fiscal decentralization scores. The results hold when the measure of decentralization is altered, when indices are made continuous measures and also when two year averages of IMR are used.

The final positive single country study we review here is due to Habibi et al. (2003) who consider devolution of basic health and education (see also next section) services in Argentina. In their paper, Habibi et al. use nationwide data from over a 25 year period to examine the relationship between two measures of fiscal decentralization and the infant mortality rate, while controlling for per capita income, per capita expenditure and the number of public sector employees. The authors present results from three models, including OLS, GLS and OLS with fixed effects, finding a significant and negative relationship between the parameters of interest. On the basis of these findings, the authors conclude that devolution can have positive effects on human development, especially when there is greater tax accountability in a province.

Less optimistically, Inchauste (2009) reports Bolivian evidence from the first half of the 2000s in the context of the Highly Indebted Poor Countries (HIPC) initiative, which directed resources saved from repayment of debt to local governments based on poverty

levels. Although she shows that there has been increased investment in both health and education, she does not find a significant association between the number of poor in a municipality and HIPC transfers, and argues that HIPC funds have not been targeted well.

Using a random effects model, Inchauste also examines the relationship between the change in health spending and (1) the change in share of unattended illnesses and (2) the unattended cases of respiratory diseases, finding a significant decline in the former and no significant change in the latter. The author controls for several socio-economic factors such as family size and illiteracy. But the short time period studied presents two important shortcomings: (1) the entire sample is in the post-decentralization period, and (2) it rules out the use of municipal fixed effects, which implies that her results may suffer from endogeneity despite the controls employed. Inchauste argues that there has been a lack of improvement in social indicators based on this mixed result, as well as on the results on education which we discuss in the next section. But the lack of pre-reform data imply that her results may say more about administration of the HIPC initiative than about decentralization *per se*.

What cross-country evidence do we have of associations of decentralization and health service delivery? Two notable studies over the past decade investigate the impact of decentralization on health service delivery, finding somewhat mixed results that appear to depend crucially on the level of development of a nation.

In an oft-cited paper, Robalino et al. (2001) perform a cross-country data regression, using IMR as the dependent variable, fiscal decentralization as the independent variable and a set of controls for GDP per capita and institutional capacity variables such as corruption, ethno-linguistic fractionalization and political rights. Like several other papers in this section, however, the authors miss out on controlling for fertility, which scholars argue is a key determinant of IMR (Treisman 2007). Their sample, though not expressly given in their paper, comprises between 45-70 low and high income countries and they rely on data from GFS.

Robalino et al.'s fixed effects model yields a significant and negative relationship between the key measures of interest. In their basic model, if a country with a GDP per capita of USD 2000 increases its share of expenditures managed by local governments by 10%, this would be associated with a 3.6% decrease in mortality rates. Robalino et al. also find evidence to show that the benefits associated with fiscal decentralization may have a U shaped curve with respect to GDP per capita, implying that countries with low

and high incomes are more likely to benefit from the reform than middle income countries. They conclude that decentralization benefits are “particularly important for poor countries” (Robalino et al. 2001: p. 11).

According to Khalegian (2004), on other hand, the benefit curve is L shaped for immunization. Using data on 140 low and middle income countries over 18 years, he conducts a cross-country regression of a measure of decentralization against immunization rates against measles and diphtheria. Unlike other papers in this area, Khalegian uses a political, not a fiscal measure of decentralization, sourced primarily from the Database of Political Institutions. Here, decentralization is measured as a binary variable, indicating whether or not local governments have authority for taxing, spending and regulation although measures of fiscal decentralization from GFS are also employed in the regression. The typical socio-economic controls for GDP per capita, population density, and illiteracy are included, in addition to measures of institutional quality, ethnic conflict and fractionalization, and income inequality.

The author notes positive and significant coefficients for lower income decentralized countries in his main specification which uses between effects and time dummies – decentralization is associated with 8.8% and 8.3% increase in diphtheria and measles coverage, respectively. In middle income countries, however, this reverses and decentralization is associated with a decrease in diphtheria and measles coverage of 4.9% and 5.5% respectively. Analysis indicates that the turning point is per capita GDP of 1400 (in 1995 USD), after which a negative relationship stabilizes.

Khalegian’s outlook based on these results is mixed with regards to decentralization – after exploring some channels, he proposes that the difference between the results seen in lower and middle income countries can be attributed to the possibility that poorer countries decentralize less fully than middle income ones. This implies that were complete devolution to occur, we would see a negative impact on immunization. The author uses this to encourage continued central government support of health initiatives.

In the *Less Credible* category, Treisman (2002) uses OLS regression on cross-sectional multi-country data to show that decentralization’s effects may be sensitive to the income level of a country. Using data on 166 countries, Treisman explores how having constitutional sub-national authority relates to two indicators of health care performance – the share of infants inoculated for diphtheria, tetanus and pertussis, and the share of population for which 20 essential drugs are available and affordable. Once an extensive set of socio-economic controls are added, the paper’s findings are not

significant for sub-national authority, although greater electoral accountability is correlated with better access to medicines. In his analysis, nations with per capita GNP greater than USD 5000 have worse service delivery performance than their counterparts with lower per capita GNP.

In sum, while the body of scholarship in this sub-category is larger than that in preference matching, it is still rather thin. Moreover, high credibility contributions are also rare. That said, the three country studies of *Strongly Credible* or *Somewhat Credible* empirical strategies all demonstrate the ability of health decentralization to have a positive influence on infant mortality rates. The same is not necessarily true for immunization, although we are forced to draw this conclusion on the basis of two cross-country studies with less convincing methodologies.

### **3.2.2 Education**

Amongst studies of education, two papers of high quality set the stage for prevalent optimism. In the first, Galiani et al. (2008) examine decentralization in Argentina by comparing changes in student test scores in secondary schools that have always been under provincial control to changes in schools that were under federal control until the 1991 reform. Like many other Latin American countries, Argentina undertook devolution to provinces as part of a broader structural reform, first devolving responsibility for pre-schools and primary schools, and then undertaking the same reform for secondary schools. Provinces now have authority over personnel and budgeting decisions, while schools are largely responsible for textbook selection and teaching methods.

Galiani et al. perform a difference in differences estimation while controlling for provincial per capita, unemployment and fiscal deficits. Using average school test scores from a sample of students tested in almost 99% of the secondary school universe, they compare the change in outcomes in those schools that were decentralized to changes in those schools that were always provincial. Because the impact on scores is unlikely to be immediate, the authors estimate the impact of exposure to decentralized schools for up to five years by cohort. Their results show a positive association between decentralization and Mathematics and Spanish scores - after five years, a 4.9% and 6.9% increase compared to the mean, respectively. The analysis in the paper does not include any validation of parallel trends, which is the key identifying assumption of a difference in differences estimation - however, the authors do use matching techniques to demonstrate



that the results do not change when looking at treatment and control groups with similar characteristics.

In a comparable paper with a solid methodological design, Faguet and Sánchez (2008) use changes in enrolment rates in state schools as the measure of student achievement in order to evaluate the impact of decentralization on service delivery in Colombia. They analyze the impact of a phased decentralization reform in the country, which not only left local governments responsible for provision of public services but also provided them increased fiscal powers to fulfill this responsibility.

Their empirical strategy is deemed *Strongly Credible* for a number of reasons. First, the authors, like some of the other papers we review in this theme, overcome the challenges of measuring decentralization by using three different indicators. This reduces the possibility that results are driven primarily by choice of measure. Second, they include relevant controls for resources that may influence school attendance rates independently, in addition to using many of the standard local government level controls for wealth and unemployment, limiting the possibility of omitted variable bias. Finally, in order to address the possibility of reverse causality, i.e. an increase in enrolment rates causing the increase in local government spending, the authors instrument for decentralization using t-1 and t-2 lagged local per capita tax revenues. These instruments should be correlated to the decentralization ratio of own resources to total expenditure, but are unlikely to influence increases in student attendance directly.

Using both OLS and IV, they find that measures of decentralization have a significant and positive correlation with changes in student enrolment, with the effects being larger for smaller municipalities.

The larger *Somewhat Credible* category fuels further enthusiasm for decentralized education delivery. In a paper discussed earlier under Health Technical Efficiency, Habibi et al. (2003) report the empirical relationship between fiscal decentralization and the ratio of students enrolled in secondary school per 1000 primary students. Using Argentine data from 1970-1994 in a fixed effects model, the authors find that their measure of decentralization - own resources to total resources - has a positive and significant association with their measure of education output.

Freinkman and Plenakanov's (2009) examination of the impact of fiscal decentralization on student scores in Russia presents the only comparison of preference matching and technical efficiency of decentralization that we have come across in this review. The authors evaluate the statistical relationship between test scores of students

from 73 regions in Russia tested in 2004 and 2005 and fiscal decentralization of a region. The 1994 reforms passed responsibility of key public services to local governments, giving them control over 80% of social spending on health and education. The authors exploit regional variances to estimate the impact of decentralization in a between effects model, employing many of the conventional controls, but also including one for initial stock of inputs. Their data on household factors, however, is fairly limited.

Despite this, their findings are rather interesting – they find that the change in spending on education was marginal with no significant impact of decentralization variables on computers, pre-school years or student teacher ratios. However, the relationship between student outcomes, as measured by an average of Language and Mathematics test scores, and decentralization is consistently positive in all of their specifications. A 10 percentage point increase in own revenues of municipalities is significantly correlated with 30 percent of one standard deviation improvement in secondary school exam scores. Taken together, they propose that their results are consistent with a technical efficiency argument arising from accountability and local official incentives, rather than allocative efficiency of increased inputs into the education production process.

Aslam and Yilmaz (2011) are similarly positive about decentralization, and support their arguments with analysis conducted on a unique dataset collected from 183 randomly selected villages in 5 purposively chosen districts in Pakistan. Pakistan embarked on an ambitious decentralization program in 2001, which left local governments responsible for basic service delivery, although many scholars note that devolution over fiscal and personnel management was limited.

The authors construct a measure of education service delivery by collecting retrospective data from villagers on changes in capital improvements, school maintenance and education services during the period 1995 to 2007. They validate this data on historical changes through on-site documentation. They then regress a composite measure of these indicators on a dummy variable for decentralization using a fixed effects model, to find that provision of education increased dramatically after the introduction of the decentralization reform. Of some worry, however, are data limitations due to which the only control they can include is an estimate of village population.

Can we conclude that decentralization enhances technical efficiency in education based on the above? Inchauste (2009) would disagree. Her examination of the relationship between changes in education allocations, and children not attending school

and un-enrolled children in Bolivia, shows limited support for devolution. Using data over a slightly longer period than her analysis of health indicators (see above), her results for funds made available to local governments through the HIPC initiative are again mixed. She finds that increases in education transfers were associated with a decrease in children not attending school in the 1999-2002 period, but an increase in unenrolled children in 2002-2005. The impact of education spending in both periods, and on other intermediate education indicators, is not significant.

The lower quality evidence, primarily simple regression analysis on cross-sectional data, is likewise mixed. Some of the contributions, nonetheless, do present noteworthy findings. One example is Di Gropello (2002), which shows conflicting results on the impact of municipality level and school level devolution on student test scores in Chile. Using an education production function design for testing conducted in 1996, Di Gropello regresses the outcome of student test scores on a measure of fiscal decentralization, school level parameters of autonomy and participation, and some controls commonly found in the education economics literature. Both devolved wage incentives and training expenditure at the municipal level are associated with higher test scores, while greater financial autonomy as measured by municipal own funds spent on education to total funds spent on education is not. At the school level, Di Gropello finds that coefficients of involvement in financial and pedagogical decision-making are significant and positively associated with student test scores.

By contrast, Lockheed and Zhao (1993)'s review from the Philippines is decisively negative. By comparing national, private and municipal or *barangay*-run schools and controlling for socio-economic background, they find no significant difference in attitudes or achievement in science or mathematics. They argue that this is due to little actual control and resources being devolved to local schools, presenting the "empty opportunity of decentralization" in the country. Treisman (2002) is equally pessimistic in the only cross-country evidence we review in this sub-category. He finds that the presence of constitutional autonomy and electoral accountability at the local level were both associated with a higher level of youth illiteracy in data from up to 166 nations. In line with his findings in the area of health, the negative associations he reports are stronger for countries with GNP per capita greater than USD 5,000.

In summary, this sub-category is not only larger than the previous ones, but also significantly more positive about decentralization's effects, especially when adjusted for quality of evidence. Although the marginal increase in number of contributions comes

from studies falling in the middle quality distinction, they are almost unanimous in their support of decentralization's ability to enhance both the quality and quantity of education.

### ***3.2.3 School-Based Management (SBM)***

As indicated by a number of recent reviews, the literature in this category has made considerable progress. We provide a basic examination of the most important pieces of works in this literature below, but also refer readers to Galiani and Perez-Truglia (2011); Bruns et al (2011); and Barrera-Osorio et al. (2009) for more comprehensive reviews<sup>5</sup>.

What does 'gold standard' evidence tell us about the effectiveness of SBM reforms? Two recent experimental contributions, and the only ones to be categorized as having *Very Strongly Credible* research designs in our review, investigate SBM's potential in enhancing student attainment. They yield contradictory findings.

The more optimistic evidence comes from Duflo et al.'s (2007) paper on a randomized control trial in Western Kenya. The trial tested a number of interventions on a total of 210 primary schools, one of which involved an SBM component that empowered school councils to hire and monitor contract teachers. Duflo et al. compare the SBM groups to their counterparts in the control group, to show that students in the treatment cell scored 0.18 and 0.24 standard deviations higher in Mathematics and Language than their non-treated counterparts two years following the intervention.

On the other hand, Glewwe and Maiga (2011) present less optimistic experimental results. They examine a randomized trial in Madagascar, which involved management reforms at three levels – district, sub-district and school. In a sample of 30 districts, sub-districts and schools were randomly sorted into treatment and control groups. Glewwe and Maiga document some school improvements in the first six months, but by the end of two years find no discernible impact on aggregated test scores. They conclude not against the reform per se, arguing instead that results may be driven by the short time since intervention. Their conclusion is consistent with suggestions in the SBM literature from the US that reforms may take up to five years to affect student test scores (see Borman et al. 2003).

---

<sup>5</sup> For good reviews of this literature from developed countries, see Summers and Johnson (1994) and Borman et al. (2003). See also a related and relatively more rigorous literature on charter and grant-maintained schools from the US and UK respectively e.g. Abdulkadiroglu et al. (2011) and Clark (2009)

The four studies that rely on the quasi-experimental technique of difference in differences, on the other hand, are unanimously favorable. The former two we classify under the *Strongly Credible* quality distinction, while the latter two are deemed to have *Somewhat Credible* identification strategies primarily as a result of challenges they face with the key parallel trend assumption required in a through DID analysis.

The strongest paper of this type is due to Gertler et al. (2011), who consider the AGE (*Apoyo a la Gestión Escolar*) intervention from Mexico. AGE is an SBM reform that provides training and small grants to parent associations in disadvantaged schools to invest in infrastructure and materials. Gertler et al. exploit the phased implementation of the AGE program to achieve identification, comparing schools that adopted AGE earlier to those that adopted it later.

The key concern is if these two groups differed systematically from each other, and these differences were correlated with school performance. To address this, the authors first provide proof that trends in both early and late adopters were parallel. They contend that because many of the unobservable factors that confound identification tend to be fixed over time, difference in differences is an appropriate strategy that allows the elimination of not only pre-treatment differences, but also all other time invariant differences between the treated and control groups. To further limit chances of endogeneity, the authors use a vector of relevant time-varying school characteristics.

Their analysis suggests that participation in the AGE program is associated with a 0.6 and 0.4 percentage point reduction in failure and repetition rates, respectively. This translates to a 4% and 5.4% decrease in these respective indicators. The authors find no significant association between AGE and intra-year drop-out rates. Through qualitative research, the paper also suggests that the channel for improvement is the increased participation of parents in decision-making.

In the second study of this type, Skoufias and Shapiro (2006) also use a difference in differences method but combine it with a matching technique to examine a different intervention from Mexico. They consider the PEC (*Programa Escuelas de Calidad*) program, another SBM type reform which provides annual grants to disadvantaged schools to improve education quality. The program gave up to a five-year USD 15k grant to the 20,000 schools, or 10% of the schooling system, that volunteered to participate. Like other SBM reforms, school councils participated in the design, implementation and monitoring of the improvement plans that the grants financed.

The authors use data for approximately 75,000 schools to first conduct a simpler OLS, and then a much improved difference in differences with matching estimation, to address concerns of lack of counterfactual data, and self-selection. Their use of propensity score matching attempts to correct for self-selection by allowing comparison of PEC and non-PEC schools that are as similar as possible. This similarity is captured by an index or propensity score that is based on observable characteristics of schools. In addition, the authors control for a variety of school and municipal characteristics.

Employing this mix of methods, they find that PEC participation is significantly associated with a 0.24, 0.24 and 0.31 reduction in dropout, failure and repetition rate, respectively. The impact, it should be noted, is marginal and represents a 6% to 8% reduction relative to the baseline means. Due to data limitations, the authors do not provide pre-intervention trend validation.

Paes de Barros and Mendonca (1998)'s study of the three key SBM changes in Brazil of financial autonomy of schools, head teacher election and establishment of school councils has similar limitations where lack of pre-implementation trend validation is concerned. Using one of the first difference in differences methodologies seen in the field, Paes de Barros and Mendonca examine changes in a series of outputs by using data from education censuses during the 1981-1993 period. Their methodology relies on estimating, over this period of time, the change in states that received an innovation against the change in groups that did not receive an innovation. To address time-varying characteristics, the authors include a set of controls for per capita income and indicators of teacher quality.

Their findings suggest that financial autonomy is associated with a significant drop in repetition rates. Similarly positive trends are evident when the authors examine mean level lags in grades attended and share of children with lag – the former is negatively and significantly associated with school financial autonomy and the presence of school councils, while the latter has a negative and significant correlation with school council presence only. To perform their analysis, however, Paes de Barros and Mendonca use state-level averages, which may mask important variation at lower levels of government.

Also supportive is the only investigation from Asia that we review in this sub-category. In a recent World Bank working paper, Khattri et al (2010) evaluate the impact of school-based management reforms implemented in 2003 in 23 districts in the Philippines. The program, sponsored by the World Bank, involved providing training to principals and parents in designing School Improvement Plans, in addition to direct

funding for the improvements planned. Implementation of the program was in three phases, with early selection based primarily on a perception of school capability.

Using data from the period 2003 to 2005, Khattri et al. compare the test score performance of students in schools that implemented the intervention in the first phase to those that implemented the intervention in a later phase. Due to the nature of the implementation strategy, however, the authors note chances of bias. These chances are highlighted by their analysis of pre-intervention trends, which shows that attainment trends in the treatment and comparison groups were not similar. To mitigate this concern, Khattri et al. combine their DID strategy with matching. However, they concede that their analysis may be more useful for noting trends as they “cannot make any claims of causality” (p. 11).

In comparison to the control group, the treatment group showed a 1.45 percentage point improvement in overall student attainment. For the subjects of Science, English and Mathematics, the improvements were 1.45, 1.32 and 1.88 percentage points, respectively.

In contrast to the above papers, the rest of the *Somewhat Credible* identification strategy category has mixed findings. The two cross-country studies on SBM are negative. Gunnarsson et al. (2009) evaluate the effects of School-Based Management reforms in 10 Latin American countries. School autonomy and participation, two indicators that have been popularly used by many other authors working on SBM, are measured for the various countries using survey data from 1997 and then quantified using factor analysis. The authors support their OLS analysis by instrumenting for autonomy and participation using principal attributes and legal structure.

After first noting the variation in autonomy and participation across countries, they find a negative and significant association between school autonomy and test scores. They do, however, find a positive relationship of test scores with parental participation. That said, Barrera Osorio et al. (2009) indicate that Gunnarsson et al.’s instruments are not necessarily solid. Good instruments are correlated with the endogenous variable, in this instance autonomy, but have no effect on the dependent variable except through the endogenous variable. However, of concern is the fact that both principal attributes and legal structures can plausibly have an independent impact on attainment.

The second cross-country study comes from a recent contribution by Hanushek et al. (2011) who use data from four waves of PISA test scores to establish the relationship between student achievement and autonomy in curricular, personnel and budgeting

areas. Their dataset contains test scores and background data on 1 million students from 42 countries, of which 25 are classified as high income nations.

Using a two way fixed effects model and controlling for a variety of family and school characteristics, Hanushek et al. find the relationship between the parameters of interest to be negative, albeit heterogeneous across countries based on income levels. A disaggregated analysis suggests that school autonomy is related to positive outcomes in developed and high-performing nations, but to negative ones in developing and low-performing nations. The authors perform a number of robustness checks, including one to address the concern of reverse causality – they test whether changes in autonomy in nations are predicted by previous PISA scores, and find they are not.

Hanushek et al.'s study suggests that understanding when SBM can be effective is critical. King and Ozler (2000)'s older paper on Nicaragua's reform provides an interesting answer to this question by arguing that it is *de facto* and not *de jure* school autonomy that improves student performance. Nicaragua's school autonomy intervention was implemented in 1991 and allowed Nicaraguan schools to sign contracts with the Ministry of Education to become autonomous. Autonomous schools were meant to work through school councils, which had *de jure* control to hire and fire teachers, manage school budgets and maintain infrastructure.

King and Ozler use a number of models to estimate the impact of both *de jure* autonomy measured by the signing of a contract, and *de facto* autonomy measured by factor analysis of teacher's responses. In their main specification, the authors rely on an extended production function on cross-sectional data. They attempt to mitigate challenges arising from both high student attrition in their sample and endogenous implementation of the reform, the former through estimation of a Heckman selection model and the latter through the use of a matching technique and analysis of panel data available for a smaller sample base. Their matched comparison design strategy compares similar treated and non-treated schools to show that *de facto* autonomy is associated with higher test scores in Mathematics and Spanish, but *de jure* autonomy has no significant impact. The school fixed effect strategy on the smaller sample yields consistent findings.

Parker (2005) provides more support for the case of the Nicaraguan autonomy reform, using more nationally representative data that she contends is less prone to student attrition bias than King and Ozler's (2000) analysis. To mitigate the self-selection bias that exists in her analysis, Parker also uses propensity score matching to compare autonomous and centralized schools. Her results are mixed – after controlling



for the standard components of an education production function, she finds that third graders in autonomous schools scored significantly higher than their counterparts in centralized schools in Mathematics. But for sixth graders the effect is negative for Mathematics, and in neither case are results significant for Spanish.

Also relying on cross-sectional analysis, Jimenez and Sawada (1999) study EDUCO, perhaps the most celebrated case of SBM, and find no significant difference in test scores. EDUCO (*Educación con Participación de la Comunidad*) was first implemented in 1991 and has served as a model for many of the community-run schools in the Latin American region. The program established community schools to enhance access in rural areas in El Salvador following the end of civil war. EDUCO schools are run by councils consisting of elected community members called Associations for Community Education (ACE), which have considerable authority in hiring/firing teachers, setting school curriculum and monitoring school performance.

The authors employ one of the first education production functions in the field to assess the impact of decentralization on student test scores, albeit for a fairly small sample of 600 students tested in 1996. To address endogenous program placement, Jimenez and Sawada use a Heckman two-stage procedure, exploiting the government prioritization formula as an instrument. They find no significant difference in test scores between traditional and EDUCO students. Given that EDUCO students come from disadvantaged backgrounds, the authors consider this a positive result. They also find evidence that student absenteeism is lower in EDUCO schools.

In a more recent paper, Sawada and Ragatz (2005) use propensity score matching on the same dataset, and still find no impact on student test scores. They do, however, report evidence of significantly lower teacher absenteeism in EDUCO schools. The authors propose that lower absenteeism arises as the result of improved community monitoring, and the authority of councils to hire/ fire teachers. Evidence shows that ACEs of EDUCO schools use incentives for renewable contracts to motivate this outcome among teachers. Their finding is supported by other studies that compare absenteeism rates of permanent and contract teachers, especially in India, to show that the community monitoring aspect is critical in enhancing outcomes (see for example Ramachandran et al. 2005; Banerjee and Duflo 2006).

Di Gropello and Marshall (2005) employ a methodology similar to Jimenez and Sawada (1999) to assess the impact of participating in a PROHECO (*Proyecto Hondureño de Educación Comunitaria*) community school in Honduras. PROHECO

schools were first established in 1999 in order to enhance primary school access in rural areas. Unlike more traditional schools, they are run almost entirely by school councils which are legal entities that set budgets, maintain school infrastructure and perform key personnel management functions including hiring, monitoring and paying.

The authors use a Heckman selection correction model to attempt to address endogeneity in placement of PROHECO schools into areas where parental demand is low and other community-level problems exist which would otherwise underestimate the benefits of participation. The two-step Heckit model works by first estimating a probit equation to predict participation in a PROHECO school based on exogenous factors, and then by using this probability factor in regressions against the output of drop-out rates to mitigate self-selection bias. As instruments, Di Gropello and Marshall (2005) use the presence of potable water and other community services. The authors also control for the usual education production function factors. Their results point to marginally lower dropout and repetition rates in SBM schools. Validity of their findings, however, may have been compromised due to inconsistent data collection as pointed out by the authors. Moreover, as Gertler et al (2007: p. 20) note, “because services that affect access might also affect learning, these instruments are not particularly convincing.”

The only study in this category with a *Less Credible* identification method is one from Argentina, which nonetheless demonstrates findings of interest. Eskeland and Filmer (2007) perform a simple OLS regression using an expanded education production function to investigate the impact of an education devolution reform that left many key education decisions decentralized to the municipal and school level. Exploiting cross-sectional data containing test scores of over 24,000 6<sup>th</sup> and 7<sup>th</sup> grade students across the nation, the authors present one of the first attempts to explore the interdependent nature of autonomy and participation by using an interaction model. Eskeland and Filmer find that autonomy is significantly associated with student test scores in Mathematics, but not in Language. They contend that participation has no independent effect on scores, but that its interaction with autonomy is positive and significant.

In summary, studies with credible identification strategies appear to consistently support school decentralization’s ability to improve repetition, failure and even drop-out rates. Results on student test scores, however, are mixed in both higher quality and medium quality evidence. Interestingly, authors have as a consequence tried to consider when SBM may be successful. This effort has yielded diverse results suggesting efficacy particularly in developed countries, or in schools with *de facto* autonomy, or even in the

presence of both autonomy and participation together. The scholarship on this latter aspect, however, is too small to draw firmer conclusions.

Table 4: Selected Empirical Evidence of Technical Efficiency										Source: Cited articles
No.	Author (Date)	Country of Study	Date Implemented	Programme Description	Method of Analysis	Sample	Measure/s of Decentralization	Results	Identification Strategy	No.
<b>Health</b>										
1	Uchimura and Jutting (2009)	China	1970s	Devolved expenditure but growingly centralized revenue authority, responsibility for delivery of health services, no political devolution	OLS using a fixed effects model	26 provinces over period 1995 - 2001	<ul style="list-style-type: none"> <li>• LG expenditure / LG own revenue</li> <li>• LG expenditure / provincial expenditure</li> </ul>	<ul style="list-style-type: none"> <li>• Higher LG expenditure is associated with lower IMR</li> </ul>	Strongly Credible	1
2	Asfaw et al (2007)	India	1980s	Rule-based fiscal transfers to LG, responsibility for public services, political devolution	OLS using between, fixed and random effects models	14 States over period 1990 - 1997	<ul style="list-style-type: none"> <li>• Index determined by factor analysis - share of LG expenditure in state, total LG expenditure per person, share of LG own revenue in LG expenditure</li> </ul>	<ul style="list-style-type: none"> <li>• D significantly associated with lower IMR in fixed and random effects model but not in the between effects model</li> </ul>	Somewhat Credible	2
3	Habibi et al (2003)	Argentina	1991	Decentralized financing, staff management and budgeting to LG.	OLS using a fixed effects model, GLS	23 of 23 provinces over 1970 to 1994	<ul style="list-style-type: none"> <li>• Share of resources in provincial control to total resources</li> <li>• Share of locally generated resources to locally controlled resources</li> </ul>	<ul style="list-style-type: none"> <li>• Share of locally generated to locally controlled resources is associated with lower infant mortality rates</li> </ul>	Somewhat Credible	3
4	Khaleghian (2004)	Cross-country	Various	Various	OLS using between effects model with time fixed effects	140 Low and Middle Income countries over period 1980 - 1997	<ul style="list-style-type: none"> <li>• Presence of taxing, spending or regulatory authority by LGs</li> </ul>	<ul style="list-style-type: none"> <li>• D associated with higher coverage in lower income countries, and lower coverage in higher income countries - L shape suggested</li> </ul>	Somewhat Credible	4
5	Inchauste (2009)	Bolivia	1994	Increase in transfers to LGs following HIPC initiative - 2000 onwards	OLS using a random effects model	300 municipalities over period 1999 - 2002	<ul style="list-style-type: none"> <li>• Change in spending in health</li> </ul>	<ul style="list-style-type: none"> <li>• Decrease in share of unattended illness during period</li> <li>• No significant change in cases of respiratory diseases</li> </ul>	Somewhat Credible	5
6	Robalino et al (2001)	Cross-country	Various	Various	OLS using a fixed effects model	Low and High Income countries over period 1970 - 1995	<ul style="list-style-type: none"> <li>• Expenditures managed by LG / managed by CG</li> </ul>	<ul style="list-style-type: none"> <li>• 10% increase in share of locally managed expenditures is associated with 3.6% reduction in IMR for USD2k per capita country</li> <li>• Benefits U shaped - higher for high and low income countries and low for middle income</li> </ul>	Somewhat Credible	6
7	Treisman (2002)	Cross-country	Various	Various	OLS	Up to 166 countries with cross-sectional data collected from mid-90s	<ul style="list-style-type: none"> <li>• Presence of LG exclusive authority on any one item</li> <li>• Some other measures of fiscal and political decentralization also used</li> </ul>	<ul style="list-style-type: none"> <li>• No significant association of sub-national autonomy with either indicator</li> <li>• Electoral accountability associated with greater access to medication</li> </ul>	Less Credible	7

Education										
8	Faguet and Sanchez (2008)	Colombia	Phased beginning in 1970s - key reform in 1991	Increase in devolved funds to LG, greater responsibility for public services, political devolution	OLS and 2SLS	90% of municipalities universe over period 1994 to 2004	<ul style="list-style-type: none"> <li>• LG own revenue sources / LG expenditure</li> <li>• Binary factor of Municipal Certification</li> <li>• Share of transfers to education expenditure</li> </ul>	• D positively associated with higher enrolment in public school	Strongly Credible	8
9	Galiani et al (2008)	Argentina	1991	Decentralized financing, staff management and budgeting to LG. Schools choose textbooks and teaching methods.	Difference in differences	Almost all secondary schools over period 1994 to 1999	<ul style="list-style-type: none"> <li>• Actual transfer from province to LG</li> </ul>	• D associated with higher Math and Spanish scores	Strongly Credible	9
10	Aslam and Yilmaz (2011)	Pakistan	2001	Limited financial devolution but implementation of rule-based transfers, responsibility for public services, political devolution	OLS using a fixed effects model	183 villages from 5 districts out of 120+. Data collected for 1995 to 2007	<ul style="list-style-type: none"> <li>• Binary measure of before and after D implementation</li> </ul>	• Magnitude of education services increases after introduction of D	Somewhat Credible	10
11	Freinkman and Plekhanov (2009)	Russia	Phased beginning in 1994	Increased fiscal powers with rule-based transfers, responsibility for public services	OLS using a between effects model	Secondary school results from 73 out of 83 regions, tested in 2004 and 2005	<ul style="list-style-type: none"> <li>• Share of LG education expenditure financed by own revenue</li> <li>• Type of decentralization arrangement</li> </ul>	• D associated with higher test scores	Somewhat Credible	11
12	Habibi et al (2003)	Argentina	1991	Decentralized financing, staff management and budgeting to LG.	OLS using a fixed effects model, GLS	23 of 23 provinces over 1970 to 1994	<ul style="list-style-type: none"> <li>• Share of resources in provincial control to total resources</li> <li>• Share of locally generated resources to locally controlled resources</li> </ul>	• Share of locally generated to locally controlled resources is associated with higher enrolment	Somewhat Credible	12
13	Inchauste (2009)	Bolivia	1994	Increase in transfers to LGs following HIPC initiative - 2000 onwards	OLS using a random effects model	300 municipalities over period 1999 - 2005	<ul style="list-style-type: none"> <li>• Change in transfers for education</li> <li>• Change in spending in education</li> </ul>	<ul style="list-style-type: none"> <li>• Increase in share of unenrolled children associated with increase in transfers</li> <li>• No significant impact of change in spending</li> </ul>	Somewhat Credible	13
14	Di Gropello (2002)	Chile	Early 1980s	Some increase in devolved funds to LG, greater responsibility for public services. In 1990s, greater pedagogical devolution to schools	OLS	50 municipalities (out of 355) - Student tests conducted in 1996	<ul style="list-style-type: none"> <li>• LG own funds / Total funds spent on education</li> <li>• School level parameters on participation, autonomy</li> </ul>	<ul style="list-style-type: none"> <li>• Municipal financial autonomy not significant</li> <li>• Municipal training spend and wage incentives positively associated with test scores</li> <li>• School involvement in financial decision-making positively associated with test scores</li> <li>• School pedagogical and curricular autonomy positively associated with test scores</li> </ul>	Less Credible	14
15	Jimenez and Paqueo (1996)	Philippines	1987	School councils raise funds, while LGs earmark taxes for supplemental school compensation	OLS	600 elementary schools over period 1982 to 1983	<ul style="list-style-type: none"> <li>• Proportion of school revenues from local sources</li> </ul>	• Locally financed schools have lower financial expenditure, indicating cost efficiency	Less Credible	15
16	Lockheed and Zhao (1993)	Philippines	1987	School councils raise funds, while LGs earmark taxes for supplemental school compensation	OLS comparison of private, national and local government schools/ HLM	8k 9th grade students in 214 schools	<ul style="list-style-type: none"> <li>• Local government school status</li> </ul>	• No significant impact	Less Credible	16
17	Treisman (2002)	Cross-country	Various	Various	OLS	Up to 166 countries with cross-sectional data collected from mid-90s	<ul style="list-style-type: none"> <li>• Presence of LG exclusive authority on any one item</li> <li>• Some other measures of fiscal and political decentralization also used</li> </ul>	• Negative and significant relationships between measures of decentralization and illiteracy	Less Credible	17

<b>School-based Management (SBM)</b>										
18	Dufo et al (2007)	Kenya	2005	Randomized trial which gave school councils money and autonomy to hire extra teachers and monitor their performance	OLS comparison of treatment and control groups. Alternate specification uses covariates	21k students from 210 schools	• Dummy variable indicating treatment group	SBM associated with • Increase in Mathematics scores of 0.24 standard deviations • Increase in Language scores of 0.18 standard deviations	Very Strongly Credible	18
19	Glewwe and Maiga (2011)	Madagascar	2005	Randomized trial in which materials, training and greater accountability is given to three levels of districts, sub-districts and schools	OLS comparison of treatment and control groups.	20k students from 30 districts over period 2006 to 2007	• Dummy variable indicating treatment group	• No significant association with test scores	Very Strongly Credible	19
20	Gertler et al (2011)	Mexico	1996	Small grants to parent councils and parental training targeted at disadvantaged areas - AGE	Difference in differences	30,000 students from 6,000 schools over 1997 - 2001	• Dummy variable indicating whether school received AGE intervention	SBM associated with • Reduction in failure rates by 4% • Reduction in repetition rates by 5.4% • No impact on drop-out rates	Strongly Credible	20
21	Hanushek et al (2011)	Cross-country	Various	Various	OLS with country fixed effects	1mn students from 42 countries - 4 waves of PISA from 2000 to 2009	• Autonomy over curriculum and pedagogy • Autonomy over personnel management • Autonomy over budgeting decision	• Overall negative association between autonomy and scores	Somewhat Credible	21
22	Skoufias and Shapiro (2006)	Mexico	2001	Annual grants of up to USD 15k given to schools/SMCs to improve education quality	OLS and Difference-in-differences with matching	75000 schools over period 2001 - 2004	• School received PEC grant in all three years • School received PEC grant in any one year	SBM associated with • 0.24 lower dropout rates • 0.24 lower failure rates • 0.31 lower repetition rates	Strongly Credible	22
23	Di Gropello and Marshall (2005)	Honduras	1999	Community schools where SMCs can hire/ fire teachers, manage school funds and maintain infrastructure	OLS with heckman correction model	200 rural schools tested in 2002 and 2003	• Probit model predicting participation in PROHECO school	SBM associated with • Higher science scores but with no change in Math or Language test scores • Marginally lower dropout rates	Somewhat Credible	23
24	Gunnarsson et al (2009)	Cross-country	Various	Various	OLS and 2SLS	17k students from 10 Latam countries - 1997 survey data	• Autonomy of school in decisions regarding hiring, budget allocation, curriculum design, etc. • Participation of parents and communities in school	• Autonomy associated with lower test scores • Participation associated with higher test scores	Somewhat Credible	24
25	Jimenez and Sawada (1999)	El Salvador	1991	Community schools where SMCs can hire/ fire teachers, manage school funds and maintain infrastructure	OLS with heckman correction model	605 3rd grade students from 162 municipalities - data from 1996	• Binary indicating whether it is an EDUCO school or not	• No association with Math or English test scores • Students in EDUCO schools have lower absenteeism	Somewhat Credible	25

26	Khattri et al (2010)	Philippines	2003	Training and direct funding for school improvement	Difference in differences with matching	5k schools from 23 districts over 2003 to 2005	• Dummy variable indicating whether school received intervention in first year	SBM associated with <ul style="list-style-type: none"> <li>• 1.45 percentage points overall improvement</li> <li>• 1.82 percentage points improvement in Science</li> <li>• 1.32 percentage points improvement in English</li> <li>• 1.88 percentage points improvement in Mathematics</li> </ul>	Somewhat Credible	26
27	King and Ozler (2000)	Nicaragua	1991	Autonomous schools with SMCs that can hire/ fire teachers, manage school funds and maintain infrastructure	Matching + validation using fixed effects and IV	3000 students from primary and secondary schools over period 1995 - 1997	• De jure autonomy - binary variable of whether the school is autonomous by law or not • De facto autonomy - actual school autonomy in various areas	• No impact of de jure Autonomy • Positive association of de facto Autonomy with Math and Spanish, no association with Language	Somewhat Credible	27
28	Paes de Barros and Mendonca (1998)	Brazil	1982	SBM with three key innovations: • Financial autonomy of schools • Ability to elect principals • Presence of school councils	Difference in differences - state-level	18 states over period 1981 - 1993	• Financial autonomy of schools • Ability to elect principals • Presence of school councils	• Lower repetition rates associated with financial autonomy • Lower mean grade level lag associated with financial autonomy and school councils • Lower proportion of students with lag associated with school council presence. • Proportion of students with lag negatively associated with principal election	Somewhat Credible	28
29	Parker (2005)	Nicaragua	1991	Autonomous schools with SMCs that can hire/ fire teachers, manage school funds and maintain infrastructure	Matching	1000 3rd and 6th grade students - tested in 2002	School autonomy in various areas	SBM associated with <ul style="list-style-type: none"> <li>• Higher third grade Math scores</li> <li>• Lower sixth grade Math scores</li> <li>• No association with Spanish scores</li> </ul>	Somewhat Credible	29
30	Sawada and Ragatz (2005)	El Salvador	1991	Community schools where SMCs can hire/ fire teachers, manage school funds and maintain infrastructure	Matching	605 3rd grade students from 162 municipalities - data from 1996	• Binary indicating whether it is an EDUCO school or not	• No association with scores • Lower teacher absenteeism in EDUCO schools	Somewhat Credible	30
31	Eskeland and Filmer (2007)	Argentina	1978	Decentralized financing, staff management and budgeting to LG. Schools choose textbooks and teaching methods.	OLS with province fixed effects	24000 6th and 7th grade students from urban schools	• Autonomy of school in various decisions • Participation of parents	SBM associated with <ul style="list-style-type: none"> <li>• Higher Math but no change in Language scores</li> <li>• Effect is stronger for poorer households</li> </ul>	Less Credible	31

**Notes**

PM: Allocative Efficiency  
TE: Technical Efficiency  
D: Decentralization  
LG: Local Government

## 4. Conclusions

How do the conclusions of our review compare with those of other broad surveys that, at various points over the past 25 years, have assessed the state of the field, attempting to reach concrete conclusions about decentralization's empirical effects? One of the broadest and most widely cited international surveys is Rondinelli, Cheema and Nellis (1983), who argue that reformers' hopes for decentralization usually outrun the reality. Serious administrative problems bedeviled implementation in most of the developing countries they survey. The few comprehensive evaluations of the costs and benefits of decentralization that had been conducted then indicated limited success in some countries and failure in others.

A decade and a half later, surveys by Manor (1999), Piriou-Sall (1998), and Smoke (2001) are only somewhat more positive, with many caveats about weakness in the evidence favoring reform. Manor notes that the evidence, though extensive, is nevertheless incomplete, while Smoke finds the evidence mixed and anecdotal, and asks whether there is empirical justification for pursuing the reform at all. Litvack et al. (1999) have a darker view: "It is not much of an exaggeration to say that one can prove, or, disprove, almost any proposition about decentralization by throwing together some set of cases or data" (p.30). Treisman's (2007) more recent review of the literature is bleakest of all. He finds results on the effects of decentralization mixed at best, and for the most part weak and incomplete. On the consequences of decentralization, he concludes "Almost nothing that is robust or general has emerged" (p.268).

Our own survey of the scholarship at first glance corroborates such conclusions. The evidence for decentralization's effects do appear to be weak, incomplete and at many times inconclusive. But when we organize the empirical literature first by theme and then – crucially – by quality of evidence provided, we are able to identify patterns of empirical results that previous surveys – including our own<sup>6</sup> – have missed. Admittedly, these patterns are not conclusive across all areas of interest. But in many instances they show that reforms can have clear, positive consequences – in some cases remarkably so, as we have attempted to outline above. This is very different from the general

---

<sup>6</sup> See for example Faguet 2004, Faguet 2008, and Faguet and Sánchez 2008.



indeterminacy that previous surveys find, and particularly important to the extent that evidence informs real world policy-making.

Still, many important questions remain unanswered. The variation in research findings in even the higher quality literature, for instance, implies that decentralization may work well in one context but not in another. Thus, one critical avenue for further research is the factors or prerequisites that enable improved outcomes after the intervention is made. A handful of studies from our review attempt to shed light on this aspect – Asfaw et al. (2007), for instance, contend that the presence of political rights enhances the consequences of decentralized health delivery, while Duflo et al. (2007) demonstrate that training and capacity building prior to implementation can deliver even better student outcomes. But a lack of focus on this crucial facet<sup>7</sup> means that to date the conditions that facilitate enhanced service delivery post reform have remained elusive.

A related area where the literature appears wanting is the channel through which improved service provision is achieved. In other words, when decentralization works, how exactly is it that it works? While there are often suggestions of better monitoring or improved incentives in the favorable evidence base, few studies provide systematic evidence to substantiate these claims. This deficiency restricts the policy implications scholarship can have for developing countries trying to implement, or even improve, their own reform.

Finally, it is difficult to step away from our review without reiterating the methodological challenges faced by empiricists in this arena. Undoubtedly, scarce data, big bang implementation and all-encompassing reform packages make it difficult to disentangle the causal effects of decentralized service delivery. The weaknesses of evidence pointed out above are not shortcomings of logic, less still of skill on the part of researchers. Rather, they are direct consequences of a combination of the data limitations that we all work under, plus the complexity of the questions we seek to answer.

The last decade has seen great improvements in identification as better data has become available, and more recently years have brought us the first randomized trials in the field. Understandably, implementing RCTs in the broader education and health decentralization reform is a much more challenging task. Yet there are other possibilities

---

<sup>7</sup> For good recent exceptions focusing on enablers, see Enikolopov and Zhuravskaya (2007) and Loayza et al. (2011)

that can add to the credibility of the research base. Early involvement of researchers together with cooperation with government agencies, for example, can allow for quasi-randomized design. And the use of panel data can help mitigate many of the econometric challenges associated with cross-sectional work. As this body of work grows, it will become possible to draw firmer conclusions on the effects of decentralization on service provision.

## Bibliography

- Abdulkadiroglu, A., J. Angrist, S. Dynarski, T.J. Kane, and P. Pathak. 2009. "Accountability and Flexibility in Public Schools: Evidence from Boston's Charters and Pilots". Cambridge, MA: National Bureau of Economic Research.
- Ahmad, E., Brosio, G., and V. Tanzi. 2007. "Local service provision in selected OECD countries: do decentralized operations work better?" Paper presented at the International Conference on Land Policies and Fiscal Decentralization, Cambridge, Massachusetts. June 4–5.
- Ahmad, Ehtisham and Giorgio Brosio. 2009. *Does decentralization enhance service delivery and poverty reduction?* Cheltenham, UK ; Northampton, MA: Edward Elgar.
- Akin, J., P. Hutchinson and K. Strumpf. 2005. "Decentralisation and government provision of public goods: The public health sector in Uganda." *Journal of Development Studies* 41(8):1417-1443.
- Arze del Granado, FJA, J Martinez-Vazquez and R McNab. 2005. "Fiscal decentralization and the functional composition of public expenditures." International Studies Program Working Paper Series, at AYSPS, GSU.
- Asfaw, A., K. Frohberg, K. S. James and J. Jutting. 2007. "Fiscal decentralization and infant mortality: empirical evidence from rural India." *Journal of Developing Areas* 41(1):17.
- Aslam, G. and S. Yilmaz. 2011. "Impact of Decentralization Reforms in Pakistan on Service Delivery— An Empirical Study." *Public Administration and Development*. 31: 159–171.
- Angrist, J. and J. Pischke. 2010. "The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics". NBER Working Paper 15794. Cambridge, MA: National Bureau of Economic Research.
- Bardhan, P. and D. Mookherjee. 1998. "Expenditure Decentralization and the Delivery of Public Services in Developing Countries." Working Paper No. C98-104. Berkeley: Center for International and Development Economics Research.
- Barrera-Osorio, F., H. A. Patrinos and T. Fasih. 2009. *Decentralized Decision-Making in Schools*. Washington D.C.: World Bank.
- Besley, T. and S. Coate. 2003. "Centralized versus decentralized provision of local public goods: a political economy approach." *Journal of public economics* 87(12):2611-2637.
- Borman, G. D., G. M. Hewes, L. T. Overman and S. Brown. 2003. "Comprehensive school reform and achievement: A meta-analysis." *Review of Educational Research* 73(2):125.
- Bruns, B., D. Filmer and H. Patrinos. 2011. *Making schools work: New evidence on accountability reforms*. Washington, DC: The World Bank
- Cheema, G. S. and D. A. Rondinelli. (eds.). 1983. *Decentralization and Development: Policy Implementation in Developing Countries*. Beverly Hills: Sage Publications.
- Clark, D. 2009. "The performance and competitive effects of school autonomy". *Journal of Political Economy* 117 (4):745-783.
- Di Gropello, E. 2002. An assessment of the impact of decentralization on the quality of education in Chile. Washington D.C.: World Bank
- Di Gropello, E. and J. Marshall. 2005. "Teacher effort and schooling outcomes in rural Honduras." In *Incentives to Improve Teaching*, ed. E. Vegas. Washington D.C.: World Bank.

- Duflo, E., P. Dupas, and M. Kremer. 2007. "Peer effects, pupil-teacher ratios, and teacher incentives: Evidence from a randomized evaluation in Kenya". Unpublished manuscript.
- Ebel, Robert D. and Serdar Yilmaz. 2002. *On the measurement and impact of fiscal decentralization*. Washington, D.C.: World Bank Institute, Economic Policy and Poverty Reduction Division.
- Enikolopov, R and E Zhuravskaya. 2007. "Decentralization and political institutions." *Journal of public economics* 91(11-12):2261-2290.
- Eskeland, G and D. Filmer. 2007. "Autonomy, Participation and Learning: Findings from Argentine Schools, and Implications for Decentralization" *Education Economics* Vol. 15, No. 1, March pp. 103-127.
- Faguet, JP. 2012. *Decentralization and Popular Democracy: Governance from Below in Bolivia*. Ann Arbor: University of Michigan Press.
- Faguet, JP. 2008. "Decentralization's Effects on Public Investment: Evidence and Policy Lessons from Bolivia and Colombia." *Journal of Development Studies*, 44: 1100–1121.
- Faguet, J.P. 2004. "Does Decentralization Increase Responsiveness to Local Needs? Evidence from Bolivia." *Journal of Public Economics*, 88(4): 867-894.
- Faguet, J.P. and F. Sánchez. 2008. "Decentralization's Effects on Educational Outcomes in Bolivia and Colombia." *World Development*, 36(7): 1294–1316.
- Freinkman, L. and A. Plekhanov. 2009. "Fiscal decentralization in rentier regions: Evidence from Russia." *World Development* 37(2):503-512.
- Galiani, S., P. Gertler and E. Schargrotsky. 2008. "School decentralization: Helping the good get better, but leaving the poor behind." *Journal of public economics* 92(10-11):2106-2120.
- Galiani, S. and Perez Truglia. 2011. "School Management in Developing Countries". Unpublished manuscript.
- Gertler, P., M. Rubio-Codina, and H. Patrinos. 2011. "Empowering Parents to Improve Education: Evidence from Rural Mexico." *Journal of Development Studies*. In Press.
- Gertler, P., H. Patrinos and M. Rubio-Codina. 2007. "Methodological Issues in the Evaluation of School-Based Management Reforms." Unpublished manuscript, World Bank, Washington, DC.
- Glewwe, P., and E. Maïga. 2011. "The Impacts of School Management Reforms in Madagascar: Do the Impacts Vary by Teacher Type?" Available at <http://www.povertyactionlab.org/publication/impacts-school-management-reforms-madagascar-do-impacts-vary-teacher-type>
- Gunnarsson, V, PF Orazem, MA Sánchez and A Verdisco. 2009. "Does Local School Control Raise Student Outcomes? Evidence on the Roles of School Autonomy and Parental Participation." *Economic Development and Cultural Change* 58(1):25-52.
- Habibi, N., C. Huang, D. Miranda, V. Murillo, G. Ranis, M. Sarkar and F. Stewart. 2003. "Decentralization and Human Development in Argentina". *Journal of Human Development*. Vol. 4, Iss. 1.
- Hanushek, E., S. Link and L. Woessmann. 2011. "Does School Autonomy Make Sense Everywhere? Panel Estimates from PISA." NBER Working Paper 7084. Cambridge, MA: National Bureau of Economic Research.
- Hasnain, Z. 2008. "Devolution, accountability, and service delivery: some insights from Pakistan." WP 4610.

- Inchauste, G. 2009. "Decentralization in Bolivia: has it made a difference?" In *Does Decentralization Enhance Service Delivery and Poverty Reduction?*, ed. E. and Brosio Ahmad, G. Cheltenham: Edward Edgar Publishing.
- Jimenez, E and V Paqueo. 1996. "Do local contributions affect the efficiency of public primary schools?" *Economics of Education Review* 15(4):377-386.
- Jimenez, E. and Y. Sawada. 1999. "Do community-managed schools work? An evaluation of El Salvador's EDUCO program." *The World Bank economic review* 13(3):415.
- Khaleghian, P. 2004. "Decentralization and public services: the case of immunization." *Social Science & Medicine* 59(1):163-183.
- Khatti, N., C. Ling, and S. Jha. 2010. "The effects of school-based management in the Philippines: An initial assessment using administrative data." Policy Research Working Paper 5248. Washington, DC: World Bank.
- King, E and B Ozler. 2000. "What's decentralization got to do with learning? Endogenous school quality and student performance in Nicaragua." Development Research Group, World Bank, Washington, DC.
- Litvack, J., Junaid Ahmad and Richard Miller Bird. 1998. *Rethinking decentralization in developing countries*. Washington, D.C.: World Bank.
- Litvack, J., J. Seddon, J. Ahmad and Institute World Bank. 1999. *Decentralization briefing notes*. Washington D.C.: The World Bank.
- Loayza, N., J. Rigolini, V. Jamele and O. Calvo-Gonzalez. 2011. "More than you can handle: decentralization and spending ability of Peruvian municipalities." Policy Research Working Paper No 5763. Washington, D.C.: World Bank.
- Lockheed, M. E. and Q. Zhao. 1993. "The empty opportunity: Local control and secondary school achievement in the Philippines." *International Journal of Educational Development* 13(1):45-62.
- Manor, J. 1999. *The Political Economy of Democratic Decentralization*. Washington, DC: The World Bank
- Oates, W. 1972. *Fiscal Federalism*. New York: Harcourt Brace.
- Paes de Barros, R. and R. Mendonca. 1998. "The impact of three institutional innovations in Brazilian education." *Organization matters: agency problems in health and education in Latin America*:75.
- Parker, C. 2005. "Teacher Incentives and Student Achievement in Nicaraguan". In *Incentives to Improve Teaching*, edited by E. Vegas. Washington D.C.: World Bank.
- Piriou-Sall, S. 1998. "Decentralization and Rural Development: A Review of Evidence." Washington, DC: Manuscript.
- Robalino, David A., Oscar F. Picazo and A. Voetberg. 2001. Does fiscal decentralization improve health outcomes? : evidence from a cross-country analysis. Washington, D.C.: World Bank.
- Rondinelli, D. A., G. S. Cheema and J. Nellis. 1983. "Decentralization in Developing Countries: A Review of Recent Experience." World Bank Staff Working Paper No.581. Washington, DC: World Bank.
- Rubinfeld, D. 1987. "The Economics of the Local Public Sector." In Auerbach, A. and M. Feldstein (eds.), *Handbook of Public Economics*. Oxford: North-Holland.
- Santibañez, L. 2006. "School-Based Management Effects on Educational Outcomes: A Literature Review and Assessment of the Evidence Base." The World Bank, Washington, DC (processed).
- Sawada, Y. and A. Ragatz. 2005. "Decentralization of education, teacher behavior, and outcomes." In *Incentives to Improve Teaching*, ed. E. Vegas. Washington D.C.: World Bank.

- Schwartz, B. J., D. K. Guilkey, R. Racelis and Measure Evaluation University of North Carolina at Chapel Hill. 2002. Decentralization, allocative efficiency and health service outcomes in the Philippines: MEASURE Evaluation, Carolina Population Center, University of North Carolina at Chapel Hill.
- Shah, A., T. Thompson, and H.F. Zou. 2004. "The impact of decentralization on service delivery, corruption, fiscal management and growth in developing and emerging market economies: A synthesis of empirical evidence." CESifo DICE Report, 1/2004: 10-14.
- Skoufias, E. and J. Shapiro. 2006. "Evaluating the impact of Mexico's quality schools program: the pitfalls of using nonexperimental data."
- Skoufias, E., A. Narayan, K. Kaiser, and B. Dasgupta. 2011. "Electoral Accountability, Fiscal Decentralization and Service Delivery in Indonesia". Policy and Research Series Paper 5614. Washington DC: World Bank.
- Smoke, P. 2001. "Fiscal Decentralization in Developing Countries: A Review of Current Concepts and Practice." Democracy, Governance and Human Rights Programme Paper No. 2. Geneva: UNRISD.
- Summers, A.A., and A.W. Johnson. 1994. "A review of the evidence on the effects of school-based management plans".
- Treisman, D. 2002. "Decentralization and the Quality of Government." Unpublished paper, Department of Political Science, UCLA.
- Treisman, D. 2007. *The Architecture of Government: Rethinking Political Decentralization*. New York: Cambridge University Press.
- Uchimura, H and JP Jütting. 2009. "Fiscal Decentralization, Chinese Style: Good for Health Outcomes?" *World Development* 37(12):1926-1934.
- Woessmann, L. 2003. "Schooling resources, educational institutions and student performance: the international evidence." *Oxford Bulletin of Economics and Statistics* 65(2):117-170.