

ARTICLE



Community monitoring interventions to curb corruption and increase access and quality in service delivery: a systematic review

Ezequiel Molina^a, Laura Carella^b, Ana Pacheco^b, Guillermo Cruces^b and Leonardo Gasparini^b

^aWorld Bank, Washington DC, USA; ^bCEDLAS & Universidad Nacional de La Plata, La Plata, Argentina

ABSTRACT

There is a belief that allowing communities monitoring power over providers could be beneficial for improving service delivery and reducing corruption in service delivery. In community monitoring interventions (CMIs), the community is given the opportunity to observe and assess providers' performance and provide feedback to providers and politicians. This systematic review and meta-analysis appraises and synthesises evidence on the effects of CMIs on access and quality of service delivery and corruption outcomes in low and middle-income countries. The results indicate evidence of beneficial effects of CMIs on service delivery quality and on helping to curb corruption. The potential benefits of CMIs on access to and quality of services are likely to be higher when interventions are designed so that contact between both actors are promoted, and tools for citizens to monitor agents' performance are provided. However, more rigorous research is needed to address this hypothesis.

ARTICLE HISTORY

Received 31 August 2017
Accepted 31 August 2017

KEYWORDS

Community monitoring; community-based; corruption; service delivery; systematic review; education

1. Introduction

Corruption and inefficient allocation of resources in service delivery are widespread in low and middle-income countries (Pande and Olken 2011). There is increasing evidence that corruption holds back countries' economic development and erodes their citizens' quality of life (Mauro 1995; Svensson 2005; Singer 2013). Millions of people around the world encounter administrative corruption in their daily interactions with public services (Transparency International 2013). The World Bank Institute estimates that total bribes in a year amount to about 1 trillion USD (Rose-Ackerman 2004), making corruption account for around 3 per cent of world GDP (Svensson 2005). Bribes are used to influence the actions of public officials, either to performed their duties, distort the duties or to prevent them from performing their duties.

In many countries, corruption is widespread throughout the public sector, not only among high level public officials. Administrative corruption imposes a heavy burden on citizens' and firms' time and resources. For example, Olken and Barron (2009) estimate that 13 per cent of the cost of a truck driver's trip in Indonesia is allocated to pay bribes to police officials that they encounter on their journey. In Uganda, bribes represent 8 per cent of firms' production costs (Svensson 2003). Corruption creates discontent with public services, undermines trust in public institutions (Sacks and Larizza 2012; Singer 2013) and stifles business growth and investment.

Often, it is the poor and the vulnerable who suffer the most from public sector corruption (Olken 2006; Sukhtankar 2011). A landmark study in Uganda found that only 13 per cent of the public funds that the central government had assigned to the school system reached the intended destination (Reinikka and Svensson 2004; Reinikka and Svensson 2005; Reinikka and Svensson 2011). In India, the lack of monitoring and accountability has resulted in high levels of public sector absenteeism, with one quarter of all the teachers in public schools and more than a third of nurses and doctors being absent from their duties (Chaudhury et al. 2006).

Community monitoring interventions (CMIs) aim to address such issues and have become common in recent years. Such programmes seek to involve communities in the monitoring of public service providers to increase their accountability to users. However, their effectiveness in reducing corruption and improving access and quality of services remains unclear. A number of well-known studies of CMIs purport to demonstrate the effectiveness of CMIs.

Over the last decade, the World Bank alone has channelled USD 85 billion to local participatory development (Mansuri and Rao 2012). The United Nations has set increasing citizen participation as its main strategy to achieve good governance and human rights (UN 2008). Nongovernmental organisations (NGOs) with a focus on increasing government accountability through citizen participation continue to expand around the globe, managing increasing amounts of resources.¹

Reports from major donors acknowledge that the lack of systematic evidence limits our ability to make precise claims regarding the relationship between CMIs, corruption and service delivery outcomes (Brinkerhoff and Azfar 2006). There are a number of nonsystematic literature reviews on community interventions or interventions for increasing citizen engagement in service delivery, for example Devarajan, Khemani and Walton (2011), Gaventa and Barrett (2012), Mansuri and Rao (2012), Pande and Olken (2011) and Ringold et al. (2012). This is the first systematic review of the evidence on CMIs. We used systematic methods to search, appraise and synthesise evidence, reaching lessons for policy transparently. Related systematic reviews have examined specific interventions aimed at reducing corruption (Hanna et al. 2011) or improving social cohesion (King, Samii, and Snilstveit 2010).

The paper is structured as follows. In Section 2, we present the intervention background and the theory of change. Section 3 presents objectives and the review approach. Section 4 presents search results, study descriptives and bias assessment. Section 5 gives results for primary and secondary outcomes. Finally, Sections 6 and 7 provide a discussion and conclusions for policy, programmes and research.

2. Community monitoring interventions

The idea that community members have incentives to monitor providers and demand better services (Stiglitz 2002) led practitioners to believe that allowing communities to have monitoring power over providers could be beneficial for improving service delivery and reducing corruption in both the short and long term. In the short term, it could improve outcomes by identifying pockets of corruption and inefficiency in service delivery. In the long term, it may contribute to changes in political norms and to establishing a transparent and accessible channel of communication for the community to provide feedback to providers and politicians on a regular basis.

Over the last two decades, programmes aimed at encouraging community monitoring have been introduced in Albania, Argentina, Brazil, Cambodia, Cameroon, Colombia, Kenya, India, Indonesia, Malawi, Philippines, South Africa and Uganda, among others (Reinikka and Svensson 2004; Reinikka and Svensson 2005; Reinikka and Svensson 2011; Pan and Christiaensen 2012; Tosi 2012; Ferraz, Finan, and Moreira 2012; Capuno and Garcia 2010; Ringold et al. 2012).

CMIs, often referred to as social accountability mechanisms, are approaches where the community is given the opportunity to participate in the process of monitoring service delivery, where monitoring means being able to observe and assess providers' performance and provide feedback to providers and politicians. The Association for the Empowerment of Workers and Farmers in India

Table 1. Interventions to increase civic participation in monitoring.

Intervention	Description
Information campaign	These are efforts to inform citizens about their rights to services, quality standards and performance campaigns. In particular, it can include information on the importance of the service, on providers' performance and on how to monitor providers
Scorecards and citizen report cards	These involve quantitative surveys that assess users' satisfaction and experiences with various dimensions of service delivery. It often involves a meeting between the recipients of services and providers to discuss the findings of the survey and to develop a follow-up plan (Ringold et al. 2012)
Social audit	Social audits allow citizens receiving a specific service to examine and crosscheck the information the provider makes available against information collected from users of the service (Ringold et al. 2012)
Grievance redress mechanisms	These are mechanisms that provide citizens with opportunities to use information redress to influence service delivery and give feedback on government programmes and services, mechanisms including complaint hotlines, informal dispute resolution mechanisms and courts (Ringold et al. 2012)

was the first organisation to introduce a social accountability initiative, through social audits in the early 1990s (Maru 2010).

The four major categories of such interventions are information campaigns, scorecards and citizen report cards, social audits and grievance redress mechanisms (GRMs) (Table 1). These four categories share two common elements: a clear objective of reducing corruption and improving service delivery and using encouragement of the community to monitor service delivery as a key intervention instrument.

Information campaigns are one of the most common interventions to encourage participation and interest in service delivery monitoring. They usually involve provision of information on the benefits of the service to be delivered (health, education, police and so on) and the current state of the service in the community. The information could be provided door to door, in public gatherings aided by local leaders, through radio, newspapers or other means (for example Keefer and Khemani 2011). Information campaigns can also include information on how to monitor providers (for example Banerjee et al. 2010).

Scorecards, often referred to as citizen report cards, encourage participation in monitoring service delivery by giving citizens a voice to demand better services. An important difference between information campaigns and scorecards is that the latter can include an interaction between citizens and providers, while the former does not include a forum for such interaction. For example, in Uganda, NGOs facilitated village and service provider staff meetings in which community members discussed information on the status of health service delivery and were encouraged to develop plans identifying steps providers should take to improve service delivery (Björkman and Svensson 2009).

Social audits involve interactions not only between citizens and providers but also with politicians, as for instance in Colombia's Citizens Visible Audit programme which worked with public infrastructure projects in water and sanitation infrastructure, schools and hospitals (Molina 2013b). A social audit involves dissemination of information through radio, newspapers and local TV in the neighbourhoods where the project is due to take place; a public forum where citizens are introduced to the project and told about their rights and entitlements, including the activities they can do to monitor the project and the responsibilities of the executing firm; establishment and training of a group of beneficiaries to carry out community monitoring activities; periodic public meetings, bringing together local authorities, neighbours and representatives from the firm carrying out the specific project, to discuss the state of the project and enable community members to voice concerns and the firm and local government to make commitments to solve the problems; regular monitoring of the project by the beneficiary group and collection of information on whether commitments are being honoured; and presentation of the finalised project to the community before making the final payment to the executing firm and sharing of the audit results with all interested and concerned stakeholders.

Social audits can also give citizens the power to make decisions about the project. An example of a CMI where citizens had decision-making power is the Kecamatan Development Program (KDP) in Indonesia, a community-driven development programme that funded projects in about 15,000 villages each year (Olken 2007). Funds were paid to village ‘implementation teams’ in three instalments. To receive the second and third payments, the teams had to make accountability reports at an open village meeting.

GRMs provide people with opportunities to use information to influence service delivery. GRMs capture different mechanisms that provide citizens with opportunities to use information redress to influence service delivery and give feedback on government programmes and services. Such mechanisms include complaint hotlines, informal dispute resolution mechanisms and courts (Ringold et al. 2012).

We developed a stylized causal chain (Figure 1), which presents the processes involved in a typical CMI and the mechanisms through which the programme is expected to have an impact on corruption and service delivery outcomes, where service delivery is defined as the process through which basic services, such as education, health and security, are delivered to communities.² The causal chain suggests

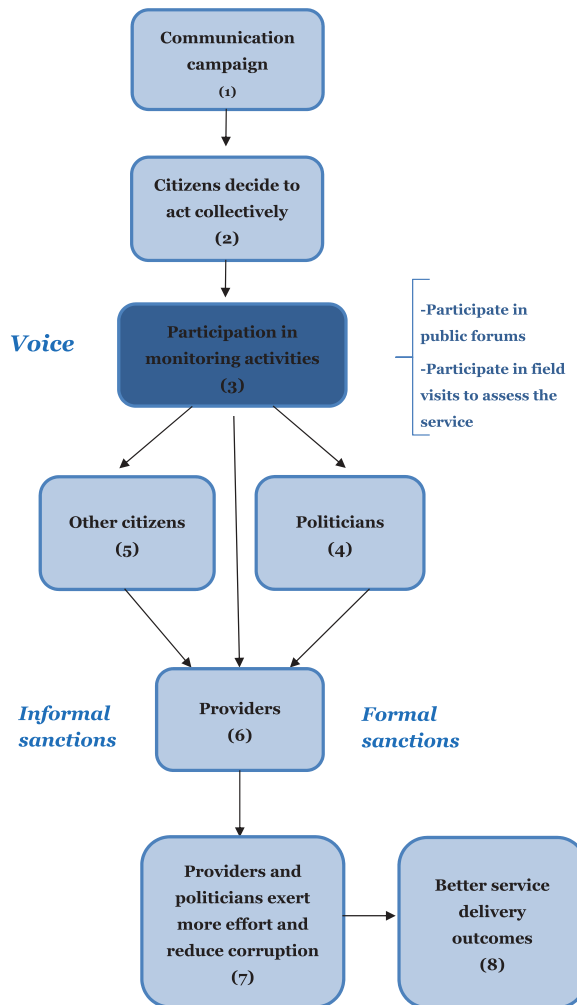


Figure 1. Causal chain for community monitoring interventions.

- CMI will increase the quantity and the degree to which citizens are involved in monitoring service providers.
- As a result of the CMI, politicians and providers will exert more effort and improve their performance in relation to service delivery.
- CMI will reduce the probability of corruption.
- CMI will improve access and quality of the service provided.

A typical CMI begins by attempting to make the project or service that it aims to monitor salient in the community. This is usually done through a communication campaign (building block 1) using as many mediums as possible, such as radio, newspapers, door-to-door campaigns and local TV. The campaign's primary objective is to increase citizen knowledge of (a) the performance of the service to be monitored and/or (b) the importance of the service or project for the community.

Equipped with this information, citizens can engage in different activities. For instance, they might change their private actions or contact fellow community members to collectively pressure providers and politicians to improve the quality of the service through monitoring activities (building block 2). To encourage citizens to monitor service providers, CMI usually include activities to build the capacity of beneficiaries to monitor providers.

Empowered with information, citizens are expected to solve the collective action problem and invest their time and effort to participate in monitoring service delivery (building block 3). Participation in monitoring activities could take many forms, depending on the specific CMI. For instance, social audits have public forums and scorecards and can include meetings between providers and citizens.

As an organised group, citizens can take turns to visit the place where the service or project takes place, such as a school, construction site or hospital, and collect information on its problems, for example absenteeism, use of low-quality inputs in the construction process, and unresponsive front-line providers. Citizens can then contact providers (building block 6) and/or elected officials (building block 4) to file complaints about the service and provide information on the specific problems the service is facing. In addition, citizens are expected to share the information collected by monitoring providers with their fellow neighbours who did not take part in monitoring activities (building block 5), to increase visibility of the CMI and put pressure on providers and politicians.

Citizens' participation in the programme may reduce the cost of monitoring front-line providers for politicians and managers. Citizens' monitoring activities also increase both visibility and citizens' ability to recognise whether elected officials are making an effort to reduce corruption and improve service delivery. As a result, there may be a greater incentive for politicians and policy-makers to achieve better results and to put more pressure on providers to improve service delivery (building block 7). The threat of formal sanctions by politicians and/or informal sanctions by citizens is assumed to motivate service providers into exerting greater effort.

Many of these processes are mediated by local norms and context. Participation in the CMI will be influenced by the strength of the community to act collectively. For example, communities with a history of grassroots participation are expected to organise more rapidly and more efficiently (Björkman and Svensson 2010). History can play an important role here. For example, in parts of Africa, the history of slave trade left an imprint in cultural norms and beliefs which arguably diminished the trust among fellow citizens and reduced the strength of the community to act collectively (Nunn and Wantchekon 2011). The independence and strength of the local media is also assumed to impact upon the visibility of the project (Reinikka and Svensson 2005; Ringold et al. 2012). Finally, this is a dynamic process, which makes understanding the specific history of service delivery, citizen engagement and political accountability in the community where the intervention took place crucial (Grandvoininnet, Aslam, and Raha 2015).

There are also some features that are specific to each intervention. Scorecards have an added accountability mechanism through which citizens meet with service providers to discuss how to improve the service. This face-to-face interaction introduces intrinsic motivation arguments for the service providers, which may contribute to improving their performance. This will be moderated by

whether it is credible for a given community to establish an informal system of rewards and sanctions. Additionally, the meeting could result in new ideas for providers and citizens on how to use and manage the service in a more efficient way.

Social audits contain an additional component in the form of public forums that allow citizens to make their voice heard by local officials and providers, simultaneously reducing the time and effort citizens would need to invest to get an appointment with officials, costs for central government representatives to be heard by local officials and costs of local officials to take actions to solve problems that arise during the implementation of the projects.

Several assumptions underlie the causal chain. Citizens need to participate in monitoring activities and politicians and providers need to be accountable. For citizens to participate, they need to have adequate information on how to monitor the project, be able to pay the opportunity cost of participation and coordinate their actions to monitor the project. Finally, citizens should believe that the programme has the potential to be successful, pay attention and be able to understand the information provided. Providers and politicians need to gain popularity, increased salary and/or social recognition, obtain re-election or avoid social disapproval or an indictment. Whether or not these assumptions are met can affect citizens' decision on whether to monitor government activity and the governments' willingness to facilitate citizen engagement and become more accountable.

For example, one potential concern is that citizens will fail to participate in monitoring activities (building block 3). In particular, if community monitoring activities are not carried out, or carried out by only a few citizens, their ability to uncover problems and put pressure on the government to provide accountability can be significantly reduced.

Another potential concern is that politicians and providers may not be accountable to their citizens (building blocks 4 and 6). For example, if the community is not needed for a politician to stay in power, the politician's performance may not improve as a result of the CMI. In addition, in communities where providers are not responsive to politicians, CMIs will only be effective if they are able to change provider behaviour. Communities often have low state capacity, which limits the ability of governments to monitor providers (Besley and Persson 2011). If this is the case, putting pressure on the government is likely to be ineffective and only competition or informal sanctions from the community may have an effect on providers' performance.

Citizens need to participate in monitoring activities and politicians and providers need to be accountable. For citizens to participate, they need to have adequate information on how to monitor the project, be able to pay the opportunity cost of participation and coordinate their actions to monitor the project. Finally, citizens pay attention to the information, be able to understand the information provided and believe the programme has the potential to be successful. Providers and politicians need to gain popularity, increased salary and/or social recognition, obtain re-election or avoid social disapproval or an indictment. If these assumptions are not met, the underlying programme theory of the CMI breaks down and this may prevent CMIs from having an impact on service delivery outcomes. In particular, whether or not they hold true can affect citizens' decision on whether to monitor government activity and the governments' willingness to facilitate citizen engagement and become more accountable. Below, we present the bottlenecks as well as the empirical implications.

2.1. Civic participation failure

One potential concern with CMIs is that citizens will fail to participate in monitoring activities (building block 3). We have identified six potential bottlenecks³ that could prevent citizens from participating in monitoring activities, which in turn reduces the potential impact of the programme (see Table 2). In particular, if community monitoring activities are not carried out, or carried out by only a few citizens, their ability to uncover problems and put pressure on the government to provide accountability can be significantly reduced.

Table 2. Bottlenecks preventing citizens from participating in monitoring activities.

Bottleneck	Description	Empirical implications
Information gaps	Scholars and policymakers have long argued that programmes often fell short of their expectations because of information problems (Ringold et al. 2012). In the case of the CMI, there are two important potential deficiencies: (a) the information may not have been properly disseminated (building block 1) and/or (b) information on how to monitor the project was either not provided or not understood by the citizens (building block 2)	<ul style="list-style-type: none"> • If the information is not properly disseminated, citizens will not participate in monitoring activities • Citizens' probability of participation in monitoring activities will be a function of how well they understand how to monitor providers
Lack of attention span or rational inattention	Even if information is provided, it may fail to have the anticipated outcome. A factor that conditions its success is what information is to be disclosed (content), and how it is to be presented (vehicle). In the case of CMIs, citizens' lack of attention span might prevent them from absorbing the information provided by the intervention and properly monitor providers. Citizen may also choose not to pay attention often describe as rational inattention. As a consequence, introducing new information does not always lead to new beliefs or changes in behaviour	If citizens' opportunity cost of paying attention to the information is high or they lack of attention span, their probability of participation will decrease ⁴
High opportunity cost of participation	Citizens, and particularly the poor, simply may not have the time to get informed or give feedback on service delivery because of more pressing priorities such as securing food and meeting other basic needs (Banerjee and Mullainathan 2008)	If opportunity cost of participation is high, probability of participation will be lower
Collective action failure	Scholars have emphasised the collective action problems that can arise in the presence of a nonexcludable local public good (Olson 1971), such as community monitoring. If community members believe fellow citizens will contribute to monitor the project, they may decide not to participate	If citizens expect other citizens would free ride on their efforts to monitor the project, the probability and intensity of participation will be lower
Citizens' beliefs can prevent participation	Citizens may refuse to take advantage of the opportunity to influence politicians and providers if they believe the chances of success are low. These beliefs can become a self-fulfilling prophecy where citizens refuse to participate and as a consequence, providers have fewer incentives to improve performance (Molina 2013a)	Citizens who perceived politicians and/or providers are responsive to them have higher probability of participation in community monitoring activities
Elite capture	Community monitoring may also be prone to be captured by local elites (Bardhan 2002; Bardhan and Mookherjee 2006; Olken 2007). The rationale is that when decision-making is part of the CMI, the elite would want to take advantage by capturing the monitoring process and appropriate the resources associated with the programme	If the CMI is captured by local elites, the participation will be limited to its supporters, which may affect the effectiveness of the programme. It is an empirical question whether the elite capture could improve or worsen outcomes. See Atlas et al. (2012) for an example of different types of elites

2.2. Politicians and providers' accountability

Under this heading, we present potential reasons for a lack of responsiveness on the part of the politicians and providers (See Table 3). The literature cites many reasons why politicians and providers may not be accountable to their citizens (building blocks 4 and 6). Below, we identify three potential bottlenecks.

Table 3. Bottlenecks causing a lack of responsiveness from politicians and service providers.

Bottleneck	Description	Empirical implications
Unresponsive politicians	Even in well-functioning democracies, citizens in a given community may not be pivotal for politician's electoral strategy (Downs, 1957; Hotelling, 1929; Persson and Tabellini 2002). This means that citizens' support is not needed for politicians to win elections and/or stay in power. Additionally, especially in developing countries, often the political system does not work properly and institutions do not help translate the preference of the people into policy (Acemoglu and Robinson 2008). Keefer and Khemani (2004, 2005) argue that public service providers have weak incentives to improve performance quality because their jobs are protected by political agents – politicians have stronger incentives to provide secure public sector jobs as teachers, health workers and local bureaucrats, than to pressure these job holders to improve service delivery.	If the community is not needed for the politicians to stay in power, we should find that politicians' performance does not increase as a result of the CMI, irrespective of what happens with citizen engagement in monitoring activities.
Unresponsive providers	The literature on providers' motivations to deliver services no longer assumed them to be either public-spirited altruists (knights) or passive recipients of state largesse (pawns). Instead, they are often considered to be in one way or another self-interested (knaves) (Le Grand, 2003). Communities in developing countries often have low state capacity, which limits the ability of governments to monitor self-interested providers (Besley and Persson 2011). If this is the case, putting pressure on the government will be ineffective and only competition or informal sanctions from the community may have an effect on providers' performance.	In communities where providers are not responsive to politicians, CMIs will only be effective if it changes providers' behaviour. If communities can impose a credible threat of informal social sanctions to unresponsive providers, the probability of a change in behaviour from providers will be higher, regardless of whether they are responsive to politicians. If communities can choose providers, competition among them will foster better performance ⁵ .

CMI: Community monitoring intervention.

3. Approach

We followed Campbell and Cochrane Collaboration guidelines for systematic reviews and meta-analysis (Shadish and Myers 2004; Booth 2011; Higgins and Green 2011) and were also informed by approaches developed by Zie (White 2009; Snilstveit 2012; Waddington et al. 2012).

We aimed to answer two specific review questions⁶:

- (1) What are the effects of CMIs that use encouragement of the community to monitor service delivery, on service delivery and corruption in Lower and Middle Income Countries (L&MIC)s?
- (2) What are the mechanisms through which CMIs effect a change in service delivery and corruption outcomes?

We searched for eligible studies of CMIs in any L&MIC country defined according to the World Bank definition at the time the intervention being studied was carried out. We searched for relevant studies across a broad range of online databases, websites and knowledge repositories which allowed the identification of both peer reviewed and grey literature. Keywords for searching were translated into Spanish, French and Portuguese and relevant non-English language literature was included. We also conducted reference snowballing and contacted experts and practitioners to identify additional studies. We used Endnote software to manage citations, abstracts and documents. First, stage results were screened against the inclusion criteria by two independent reviewers with additional supervision by a third.

Table 4. Forensic economic estimates of corruption.

Study	Variable definition	CMI type	Effect size	95 Per cent confidence interval		ES type
Olken (2007) – invitations	Per cent missing funds ^a	Social audit	0.08	–0.10	0.26	SMD
Olken (2007) – invitations + comment	Per cent missing funds ^a	Social audit	0.08	–0.10	0.25	SMD
Olken (2007) – all interventions			0.08	–0.08	0.24	SMD
Reinikka and Svensson (2011)	Share of funding reaching school	IC	0.22	0.05	0.40	SMD
Meta-analysis			0.15	0.01	0.29	SMD

^aCalculated as log-reported value – log actual value (major items in roads and ancillary projects).
CMI: Community monitoring intervention; SMD: standardised mean differences.

Included CMIs needed to have a clear objective of reducing corruption and/or improving service delivery, use encouragement of the community to monitor service delivery as a key intervention instrument and fall into one of the four intervention categories defined above. Eligible comparisons could be passive, such as no formal process of monitoring (for example Björkman and Svensson 2009), or active, for example CMI with no encouragement to participate (for example Olken 2007).

We included quantitative studies with either experimental or quasi-experimental design to address review question 1. In addition, both quantitative and qualitative studies were eligible for inclusion to address review question 2, provided the evidence was taken from a study eligible for inclusion for question 1 or a ‘sibling study’ (that is a study on the same intervention as those included for question 1).

Corruption and service delivery were the primary outcomes of the review. We defined corruption as dishonest or fraudulent conduct by those in power. A big issue in the literature is the difficulty in measuring corruption accurately (Pande and Olken 2011). As a consequence, we included corruption estimates from both the forensic economic literature (Zitzewitz 2012) as well as measures based on perceptions of corruption. We also defined service delivery outcomes as access to and quality of the service. For example, if the goal of the CMI was to facilitate household access to clean water, the percentage of access to clean water and water quality would be the outcome of interest. If the goal was to monitor school performance, children’s tests scores would be the desired outcome. We also collected data on intermediate outcomes measuring changes in behaviour, such as whether participants contribute to monitoring of the service or project and the behaviour and performance of providers and politicians.

Two reviewers independently coded and extracted data on study details, design and relevant results from the included studies.⁷ Studies were critically appraised for potential bias across using a tool provided by 3ie (available in Waddington et al. 2014). ‘Low risk of bias’ studies were those in which clear measurement of and control for confounding was made, where intervention and comparison groups were described adequately and risks of spillovers or contamination were small, and where reporting biases and other sources of bias were unlikely. ‘Medium risk of bias’ studies had suspected threats to validity of the attribution methodology, or possible risks of spillovers or contamination, arising from inadequate description of intervention or comparison groups or suspected reporting biases. We did not classify any included studies as being of ‘high risk of bias’.

We used a combination of statistical meta-analysis and narrative synthesis along the causal chain to pool the evidence. To prepare the data for meta-analysis, we calculated standardised mean differences (SMDs) and average standardised effects⁸ and 95 per cent confidence intervals (95% CIs) for continuous outcome variables, and risk ratios and risk differences and 95 per cent CIs for dichotomous outcome variables. All effect sizes were computed as continuous outcomes, excepting those from Pandey et al. (2007), which were computed as dichotomous outcomes.

We then synthesised results using statistical meta-analysis. Where possible, we extracted data on intermediate outcomes such as citizen participation and public officials and service providers’ responsiveness.

Where studies reported the effect of more than one intervention, we chose the interventions that could be classified as one of our four categories. In cases where more than one intervention was relevant, we pooled their effects before integrating them into meta-analysis, taking into account the correlation of the treatment and control groups between study arms to address possible dependency.⁹ This is the case of Afridi and Iversen (2013), Banerjee et al. (2010), Barr et al. (2012), Olken (2007) and Pradhan et al. (2014). Where we identified different measures for the same outcome, we followed a similar procedure. We computed a synthetic effect size, defined as the mean effect size in that study, with a variance that takes account of the correlation among the different outcomes (Borenstein et al. 2009). Finally, when we found different follow-up periods for comparable interventions, we compared them considering similar horizon time. This is the case for Björkman and Svensson (2009) and Björkman, de Walque and Svensson (2013), who report both the short and the medium-term impacts of an intervention in Uganda, and the short-term impact of another intervention in the same place. In these cases, we only run meta-analysis for short-term effects.

Further details about the methods used to undertake the review are provided in the published protocol and the full technical report (Molina et al. 2016).

4. Included study characteristics

The search process (Figure 2) returned 109,017 references, of which 36,955 were eliminated as duplicates and a further 71,283 were excluded at the title screening stage. The remaining 787 papers were included for abstract screening and 181 studies were included for full-text screening.

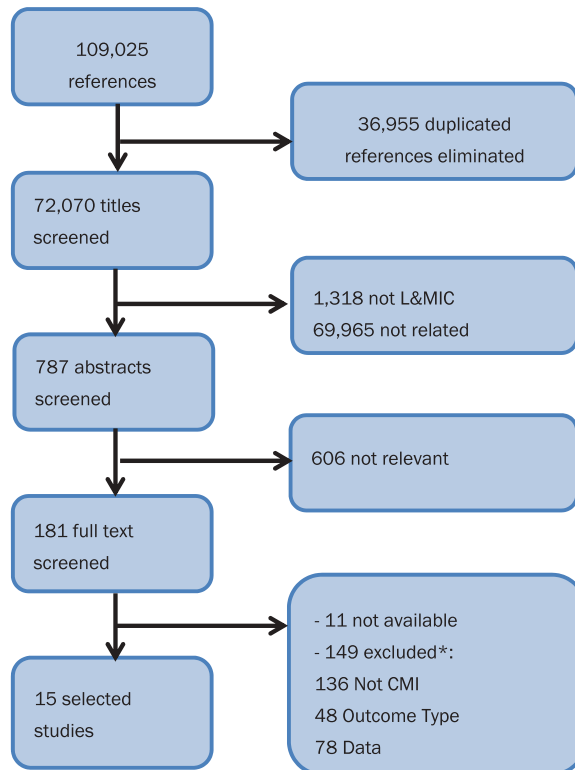


Figure 2. Search and selection process.

Fifteen studies met the inclusion criteria for addressing review question 1 and six 'sibling studies' were also included to address question 2.

Descriptive information on included studies is provided in [Appendix](#) table. Studies evaluated the effects of 23 different CMI in the areas of information campaigns (10), scorecards (3), social audits (5) and combined information campaigns and scorecards (2). Included studies were conducted in Africa (6), Asia (7) and Latin America (2). Uganda and India had the largest presence of CMI impact evaluations, with four studies conducted in each country. We also identified two studies from Indonesia and one each from Benin, Liberia, Colombia, Pakistan and Mexico. Most studies focused on interventions in the education sector (9), followed by health (3), infrastructure (2) and employment promotion (1).

Studies assessed the effects of CMIs on all primary outcomes. Eleven studies reported on improvement in prevalence condition, seven on access to service, three on perception of corruption, two on average waiting time to get the service and two on forensic economic estimates for corruption. Different outcome measures were used by sector. For example, in health, access was measured as utilisation, coverage or immunisation. Ten studies used randomised assignment and five used quasi-experimental methodologies such as instrumental variables and matching (Keefer and Khemani 2011; Molina 2013b).

Seven studies were categorised as of low risk of bias in attributing outcomes to the intervention, based on our five criteria of selection bias and confounding, spillovers, outcome reporting bias, analysis reporting bias and other sources of bias (Andrabi, Das, and Khwaja 2015; Banerjee et al. 2010; Barr et al. 2012; Björkman and Svensson 2009; Olken 2007; Pandey et al. 2007; and Pradhan et al. 2014). The remaining eight studies were classified as of medium risk (Afridi and Iversen 2013; Björkman, De Walque, and Svensson 2013; Gertler, Patrinos, and Rubio-Codina 2008; Keefer and Khemani 2011; Molina 2013b; Pandey, Goyal, and Sundararaman 2009; Pandey, Goyal, and Sundararaman 2009; Piper and Korda 2010, and Reinikka & Svensson, 2011). Thus, the overall risk of bias assessment is predominantly low, with 13 out of 15 papers having this level of risk, followed by 2 papers with medium risk of bias (Figure 3).

The majority of studies (11 out of 15) were adequately protected against performance bias as the units of treatment were located far from the control units. While in some cases, the comparison group was selected from villages where the intervention was not carried out but was located near villages that had received the intervention,¹⁰ and in other cases, the comparison group received a different treatment or the same intervention with a different degree of intensity (for example Afridi and Iversen 2013; Olken 2007), the authors took that into consideration while designing the intervention and selecting cluster for their standard errors. We found just one case of potential outcome reporting bias, where the outcome reported was a new type of literacy test and the

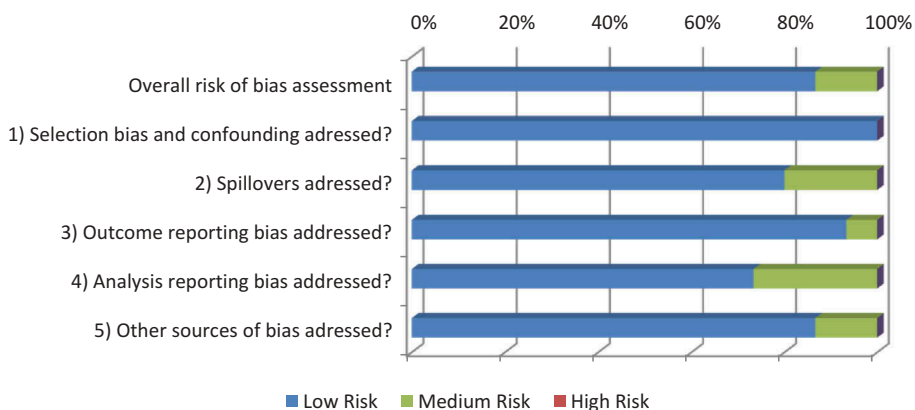


Figure 3. Summary of critical appraisal of studies of effects.

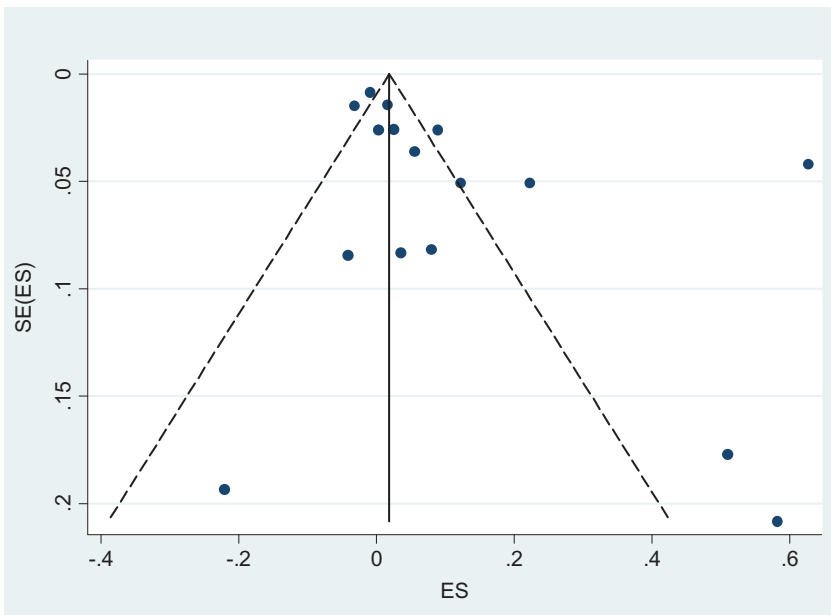


Figure 4. Funnel plot showing pseudo 95 per cent confidence limits.

authors had not clearly justified the reason for using the measure over standard ones (Keefer and Khemani 2011). With regards to analysis reporting bias, in most of the included studies, different measures for the same outcome are reported or different specification and estimation methods are applied. There was no evidence that outcomes or specifications methods were selectively reported and the authors used 'common' methods of estimation.

We also assessed publication bias in the body of evidence, using tests for small study effects. We pooled all SMD effect sizes (reversing the sign when needed, so all measured positive effects in the same direction) and present the results in Figure 4. The asymmetry in the plot, together with the result of the Egger et al. (1997) test (p -value = 0.058), suggests possible evidence of underreporting of small sample studies, which is often considered as evidence of publication bias. However, the low power of the tests prevents us from drawing strong conclusions.

5. Impacts of CMI (review question 1)

In this section, we present the results of synthesis of effects to assess whether the included interventions reduced corruption and improved service delivery outcomes (review question 1). The included studies use a range of different measures to assess primary outcomes and it would not be sensible to pool conceptually different outcomes in our meta-analysis. To avoid this problem, we grouped studies only when the outcome variables represent a similar broad outcome construct and the intervention is implemented in the same sector. In the case of service delivery, we differentiated access from quality and performed separate analysis by sector and outcome.

5.1. Corruption outcomes

We identified few studies assessing the effect of CMI on corruption outcomes, using either forensic estimates (two studies, two interventions) or perception measures (three studies, four interventions). This limits our ability to generalise these results.

Included studies on the effects of CMI on corruption outcomes were implemented in infrastructure, education and employment assistance programmes. Two studies used forensic economic estimates to collect outcomes data, and three studies (comprising four interventions) measured perception of corruption.

We looked for different measures of corruption in the papers considered, with the aim of extracting measures based on the application of economics to the detection and quantification of behaviour (Zitzewitz 2012), in this case, corruption. With this purpose, we extracted all measures that we could identify for each intervention. Table 2 lists the measures reported in each case.¹¹ We identified two studies reporting forensic measures of corruption (Olken 2007 and Reinikka and Svensson, 2011). Olken (2007) evaluates the impact of increasing citizen participation in social audits in Indonesia on corruption, with two different treatment arms testing different variations of social audits. Villagers were invited to participate in social audits ('accountability meetings') in both treatment arms, but in one group, the invitation was accompanied by an anonymous comment form, which could be submitted in a sealed box. The results of this exercise were summarised in the accountability meetings.¹² The study reports a forensic measure of corruption which is the percentage of missing funds in roads and ancillary projects. The effect of social audits only was SMD 0.08 (95% CI [-0.10, 0.26]) while the effect of social audits with anonymous comment form was also 0.08 (95% CI [-0.10, 0.25]). The combined effect for both treatment arms was SMD 0.08 (95% CI [-0.08–0.24]).¹³

Reinikka and Svensson (2011) evaluate the effect of systematic publication of monthly financial transfers to schools in Uganda. They found that a school close to a newspaper outlet suffers less from the capture of funds as compared to a school away from a newspaper outlet (Reinikka and Svensson, 2011).¹⁴ The SMD shows that schools where the intervention took place had 22 per cent less corruption than the others.

Meta-analysis suggests that overall effect of these interventions is positive, improving corruption outcomes by 0.15 SD, as it shown in Table 4 and Figure 5. Since Olken (2007) finds no statistically significant effects, this result is probably driven by Reinikka and Svensson (2011), who did find a positive and statistically significant effect.

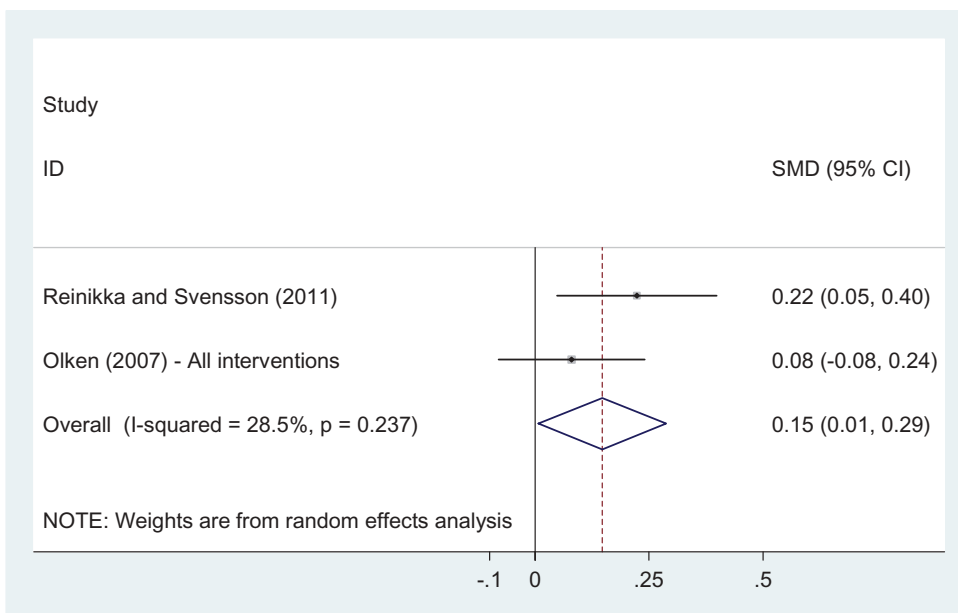


Figure 5. Corruption – forensic economic estimates.

Table 5. Perception measures of corruption.

Study	Variable definition	CMI type	Effect size	95 Per cent confidence interval			ES type
Molina (2013b)	Adequacy in the administration of resources	Social audit	0.03	-0.08	0.15		RD
Pandey et al. (2007)	Percentage of household reporting that development work has been performed in the village	IC	0.09	0.03	0.15		RD
Meta-analysis			0.08	0.02	0.13		RD
Afridi and Iversen (2013) – second audit	Total number of irregularities ^a	Social audit	-0.22	-0.39	-0.05		SMD
Afridi and Iversen (2013) – third audit	Total number of irregularities ^a	Social audit	-0.23	-0.43	-0.04		SMD
Afridi and Iversen (2013) – All interventions			-0.23	-0.38	-0.07		SMD

^aReversed sign for this variable.

CMI: Community monitoring intervention; SMD: standardised mean differences.

It is difficult to detect and measure corruption objectively and because of that, we included more subjective perception measures. Table 5 lists the outcome measures reported in the three studies (four interventions) that we have included in this category. We were able to compute RD for the first two studies and SMD for the third one, so we analyse them by separate.

We identified two studies assessing the effect of CMI on corruption perception measures for which comparable effect size were available. Molina (2013b) found evidence that a social audit improved the perception of the administration of resources in Colombia, while Pandey et al. (2007) suggest that an information campaign carried out in India increased the probability of households reporting that development work took place in the villages, which can be interpreted as a reduction in corruption. The overall effect is RD 0.08 (95% CI [0.02, 0.13]), as can be seen from the forest plot presented in Figure 6. The CIs are overlapping and the test of homogeneity does not suggest between study variability. Both studies report a reduction in the perception of corruption among beneficiary households.

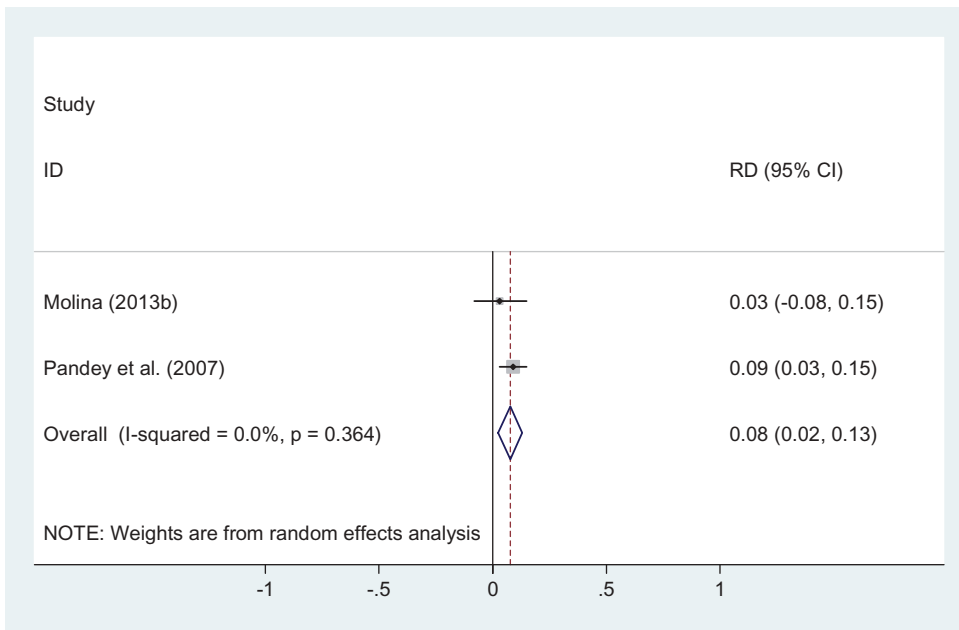


Figure 6. Corruption – perception measures.

Afridi and Iversen (2013) assess the effect of social audits in India. They estimate the effect on reported irregularities in Gram Panchayats with one audit as compared to those with two and three audits, respectively. The effect after two audits was to reduce irregularities by 0.22 SD (95% CI [-0.39, -0.05]) and after three audits by 0.23 SD (95% CI [-0.43, -0.04]), suggesting increased corruption in *Gram Panchayats* after these interventions, and a worsening effect with two or more social audits, as compared to those with one audit only. The authors explain that this is likely due to increasing complaints over time as participants become more aware of their entitlements, or have greater confidence in the integrity of the audit process, or if the quality of the social audit improves to increase the number of irregularities identified (Afridi and Iversen 2013).

5.2. Access to services

A number of different outcome measures were included as proxies for access to services. In this section, we begin with health services and present the results for utilisation, immunisation and other measures of access, followed by results for enrolment and dropout rates in the education sector.

5.2.1 Health

We identified two studies that assessed at the impact of CMI on utilisation. Björkman and Svensson (2009) evaluate the same intervention, a combination of a scorecard with an information campaign both in the short and in the medium term,¹⁵ using the same group of 50 facilities/communities that were randomly assigned to treatment and control group.

In addition Björkman, de Walque and Svensson (2013) assess a new intervention, an information campaign with new treatment and control groups in which 25 new facilities were randomly assigned to a treatment group (13 units) and control group (12 units). This intervention differs from the previous one since it does not include a scorecard with relevant information about the health service provision.

In these studies, authors report an 'average standardised treatment effect' following Kling et al. (2004)'s methodology that can be combined into one meta-analysis given their homogeneity and given that they are comparable as they all imply better access to health services. Table 6 presents the effect size for each intervention regarding utilisation of health services.¹⁶

Looking at the short run, both interventions show an increase in the access to health services, although for the second one, the result is statistically no significant. This suggests that effects are stronger when the information campaign is coupled with a scorecard, which is consistent with the authors' findings, who hint that without information, the process of stimulating participation and engagement had little impact on health workers' performance or the quality of health care

Table 6. Utilisation.

Study	Variable definition	CMI type	Effect size	95 Per cent confidence interval		ES type
Björkman and Svensson (2009) – short term	Utilisation/Coverage ^a	Scorecard + IC	2.13	0.79	3.47	ASE
Björkman, De Walque and Svensson (2013) – short term	Utilisation/Coverage ^a	IC	0.04	-0.41	0.49	ASE
Meta-analysis			0.99	-1.05	3.02	ASE
Björkman and Svensson (2009) – medium term	Utilisation/Coverage ^a	Scorecard + IC	0.34	0.12	0.55	ASE

^aWeighted mean of average number of patients visiting the facility per month for out-patient care, average number of deliveries at the facility per month, share of visits to the project facility of all health visits averaged over catchment area and share of visits to traditional healers averaged over catchment area.

CMI: Community monitoring intervention; ASE: average standardised effect.

Table 7. Immunisation.

Study	Variable definition	CMI type	Effect size	95 Per cent confidence interval		ES type
Björkman, De Walque and Svensson (2013)	Immunisation ^a	IC	1.00	-0.63	2.63	RR
Pandey et al. (2007)	Vaccinations received by infants	IC	1.57	1.40	1.75	RR
Meta-analysis			1.56	1.39	1.73	RR
Björkman and Svensson (2009) – medium term	Immunisation ^a	IC	1.04	-0.52	2.61	RR

^aMean of newborns, children less than 1 year, 1-year olds, 2-year olds, 3-year olds and 4-year olds, whether the child has received at least one dose of measles, DPT, BCG and polio.

CMI: Community monitoring intervention.

(Björkman, de Walque, and Svensson 2013). The overall effect is positive but not statistically significant, and there is a large amount of between study variability ($I^2 = 88.0\%$, $p = 0.004$).

Looking at the medium term, the information campaign combined with the scorecard has a positive and statistically significant effect on service access. Regarding immunisation outcomes, Table 7 reports the short-run effects found by Björkman, de Walque and Svensson (2013) and Pandey et al. (2007). It also reports the medium-term effects for the intervention assessed in Björkman and Svensson (2009) which is not incorporated into meta-analysis given the different time horizons.

Björkman, de Walque and Svensson (2013) assess the impact of the CMI on immunisation for different age groups, while Pandey et al. (2007) compute the percentage of households where infants have received vaccinations. The overall effect is an increase in risk ratio of 1.56 (95% CI [1.39, 1.73]), implying that the effect of the CMI improved access to services by 56 per cent (Figure 7). The medium-term impact of the intervention reported by Björkman and Svensson (2009) is 1.04 SD but statistically nonsignificant.¹⁷

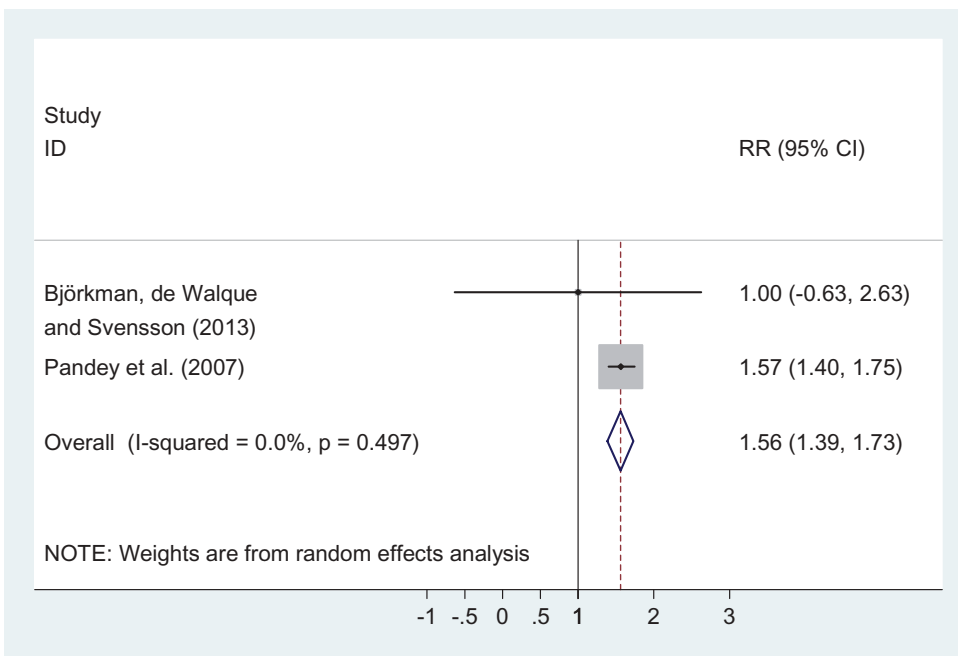


Figure 7. Immunisation.

Downloaded by [186.141.202.203] at 10:32 01 November 2017

Table 8. Other access to service outcomes.

Study	Variable definition	CMI type	Effect size	95 Per cent confidence interval			ES type
Pandey et al. (2007)	Visits by nurse midwife	IC	1.03	0.94	1.14	RR	
	Prenatal examinations		1.63	1.45	1.83	RR	
	Tetanus vaccinations		1.57	1.39	1.77	RR	
	Prenatal supplements received by pregnant women		1.45	1.29	1.64	RR	
	Vaccinations received by infants		1.57	1.40	1.75	RR	
Meta-analysis			1.43	1.29	1.58	RR	

CMI: Community monitoring intervention.

Pandey et al. (2007) also report on different measures of access to health services, specifically the percentage of households getting health services such as visits by nurse midwives, prenatal examinations, tetanus vaccinations and prenatal supplements received by pregnant women. We computed risk ratios for these outcomes, and the results are reported in Table 8. All risk ratios are above unity, with an overall effect of 1.43 (95% CI [1.29, 1.58]) implying that the intervention improved access to services in 43 per cent.

5.2.2. Education

We identified four studies evaluating effects on enrolment in six different treatment arms. Table 9 presents the effect sizes from all treatment arms. Before combining the studies into a meta-analysis, we created synthetic effect sizes for the study with multiple treatment arms (Banerjee et al. 2010) to avoid combining effects based on dependent samples. Figure 8 presents the forest plot for the meta-analysis of enrolment rates. The overall average effect of CMI on enrolment is 0.09 SD (95% CI [-0.03, 0.21]). However, it can be noted that this result is driven by the inclusion of one study for which the SMD is substantially higher than the others (Andrabi, Das, and Khwaja 2015). Also, there is significant between study heterogeneity ($I^2 = 73.6\%$, $p = 0.010$). Therefore, we performed sensitivity meta-analysis excluding this study, which yielded an overall effect of 0.05 SD (95% CI [-0.03, 0.13]).

When considering only scorecards (Gertler, Patrinos, and Rubio-Codina 2008), the overall effect is zero (SMD = 0.003; 95% CI [-0.05, 0.05]). On the other hand, information campaigns show an overall effect SMD 0.10, 95 per cent CI [0.01, 0.18], suggesting that these interventions have increased enrolment rates in 10 per cent (Figure 9).

We identified four studies measuring dropouts at schools in seven treatment arms (Table 10). For some interventions, the results suggest an increase in children out of school in the villages receiving CMI compared to those who did not receive the programme. Considering the study of Banerjee et al. (2010), the effects range from 0.02 SD improvement (95% CI [-0.01, 0.06]) for the

Table 9. Enrolment.

Study	Variable definition	CMI type	Effect size	95 Per cent confidence interval			ES type
Andrabi, Das and Khwaja (2015)	Enrolment	Scorecard	0.58	0.17	0.99	SMD	
Banerjee et al. (2010) – mobilisation	Enrolment	IC	0.059	-0.138	0.257	SMD	
Banerjee et al. (2010) – mobilisation + information	Enrolment	IC	0.05	-0.14	0.25	SMD	
Banerjee et al. (2010) – Mobilisation + information + 'Read India'	Enrolment	IC	-0.008	-0.199	0.183	SMD	
Banerjee et al. (2010) – all interventions			0.04	-0.13	0.20	SMD	
Gertler, Patrinos and Rubio-Codina (2008)	Enrolment	Scorecard	0.003	-0.048	0.054	SMD	
Reinikka and Svensson (2011)	Enrolment	IC	0.12	0.02	0.22	SMD	
Meta-analysis			0.09	-0.03	0.21	SMD	

CMI: Community monitoring intervention; SMD: standardised mean difference.

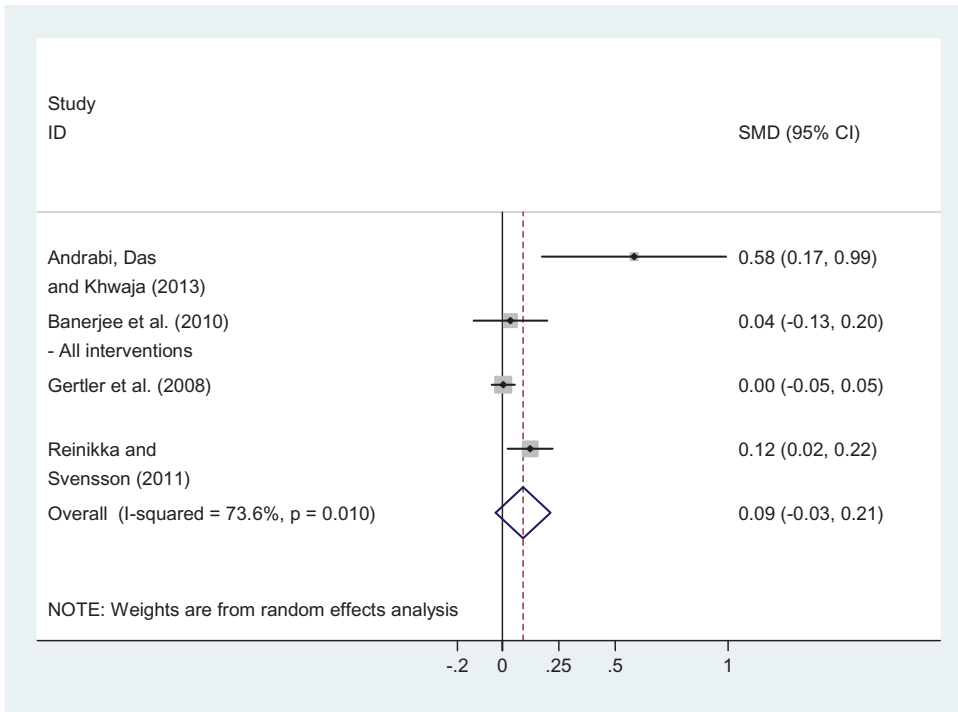


Figure 8. Enrolment.

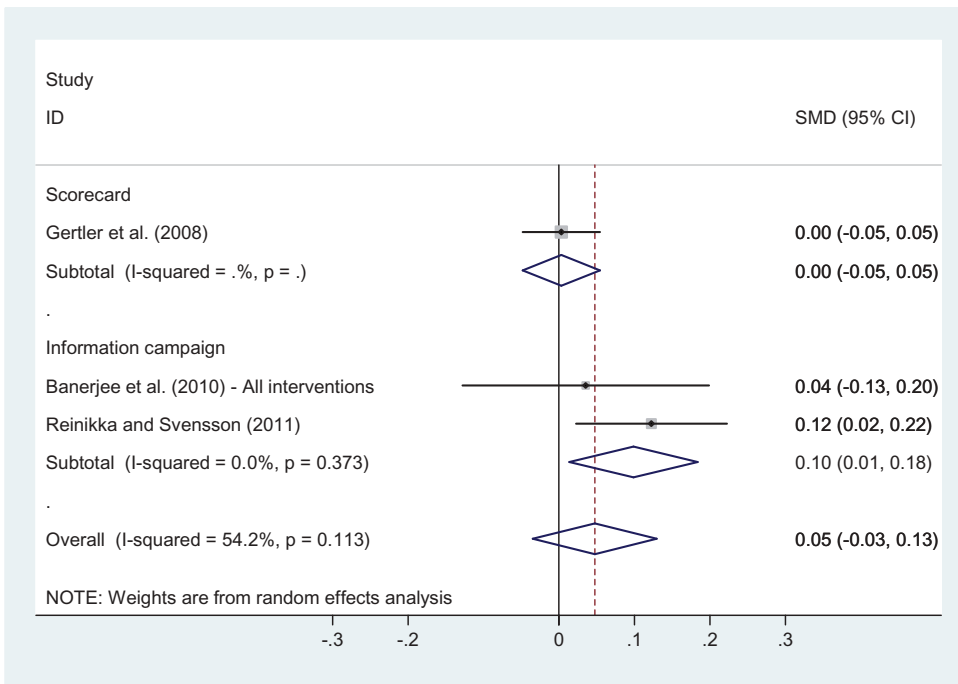


Figure 9. Enrolment – sensitivity analysis (outlier excluded).

Table 10. Dropouts.

Study	Variable definition	CMI type	Effect size	95 Per cent confidence interval		ES type
Andrabi, Das and Khwaja (2015)	Dropout rate	IC	0.220	-0.159	0.600	SMD
Banerjee et al. (2010) – mobilisation	Dropout rate	IC	0.028	-0.006	0.063	SMD
Banerjee et al. (2010) – mobilisation + information	Dropout rate	IC	0.02	-0.01	0.06	SMD
Banerjee et al. (2010) – mobilisation + information + 'Read India'	Dropout rate	IC	0.046	0.011	0.081	SMD
Banerjee et al. (2010) – all interventions			0.032	0.003	0.061	SMD
Gertler, Patrinos and Rubio-Codina (2008)	Dropout rate	IC	-0.09	-0.14	-0.04	SMD
Pradhan et al. (2014) – training	Dropout rate	IC	0.12	-0.08	0.31	SMD
Pradhan et al. (2014) – linkage	Dropout rate	IC	-0.03	-0.23	0.16	SMD
Pradhan et al. (2014) – all interventions			0.041	-0.124	0.207	SMD
Meta-analysis			0.00	-0.10	0.10	SMD

CMI: Community monitoring intervention; SMD: standardised mean difference.

treatment arm with mobilisation and information to 0.05 SD (95% [0.011, 0.081]) for the treatment arm with mobilisation and information in addition to 'Read India' – reading camps held by trained volunteers – with a combined effect of 0.03 SD (95% CI [0.003, 0.06]). The authors argue that this result is due to 'children dropping out of private or NGO schools (results omitted to save space). It may be that parents consider the reading classes to be an adequate alternative to a private school'. The CMI also resulted in an increase in dropout rates in the cases of Andrabi, Das and Khwaja (2015) and for the training intervention in Pradhan et al. (2014). On the other hand, the linkage intervention in Pradhan et al. (2014) and the study of Gertler, Patrinos and Rubio-Codina (2008) find a reduction in dropout rates after interventions.

Before performing the meta-analysis (Figure 10), we calculated a synthetic effect size for the two treatment arms included in the study of Pradhan et al. (2014) from Indonesia to avoid issues with dependent effect sizes in the meta-analysis. We did the same with the three interventions reported by Banerjee et al. (2010). Taking into account all the interventions, the overall effect is zero (SMD = 0.00; 95% CI [-0.10, 0.10]).

5.3. Quality of services

In this section, we present the analysis of effects on quality of services by sector and outcome, starting with health and followed by education.

5.3.1. Health

For health-related outcomes, we consider measures of child death and anthropometric outcomes. We identified two studies with measurements of child mortality, Björkman and Svensson (2009) and Björkman, de Walque and Svensson (2013). Table 11 shows that the short-term evaluation for the two interventions had an overall effect of RR 0.76 (95% CI [0.42, 1.11]), suggesting that child death had been reduced by 24 per cent after CMIs; however, the effect is not statistically significant. Similar conclusions apply for the medium-term effect of the information campaign combined with a scorecard where the effect is a reduction in 21 per cent in mortality.

We can also interpret an improvement in anthropometric measures as an improvement in the quality of health services provided. Table 12 reports the impact of the same two CMIs on weight-for-age scores.

In the short term, CMIs have increased weight for age scores by 20 per cent, suggesting that quality of health services has improved. The positive impact seems to be stronger in the medium term, resulting in a 29 per cent improvement. Homogeneity test suggests no variability between studies ($I^2 = 0.01\%$, $p = 0.928$).

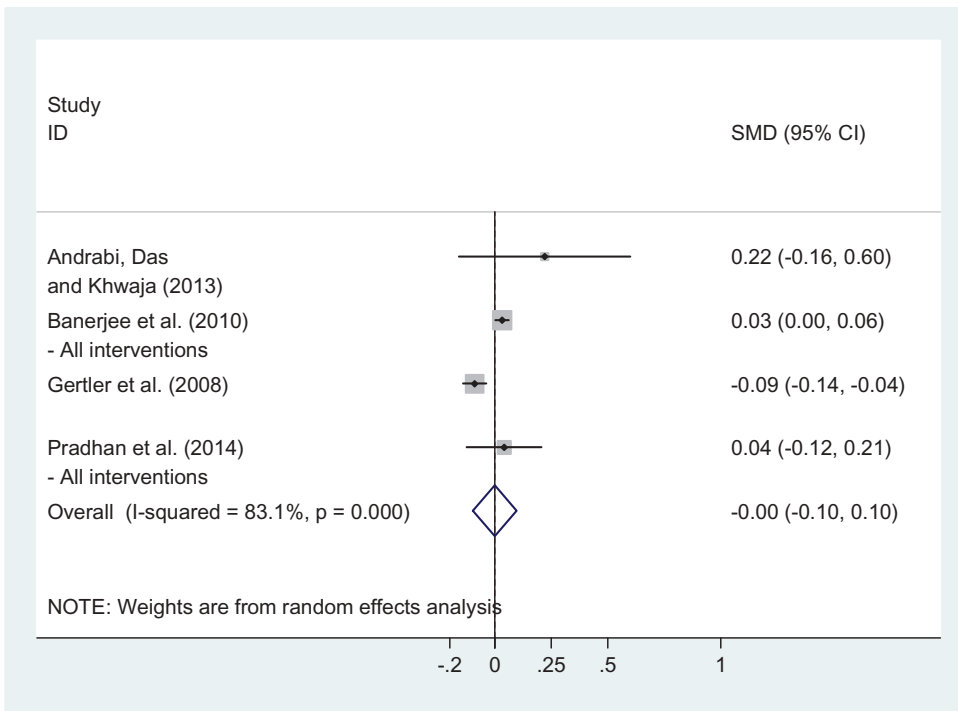


Figure 10. Dropouts.

Table 11. Child death.

Study	Variable definition	CMI type	Effect size	95 Per cent confidence interval			ES type
Björkman and Svensson (2009) – short term	Death (under five mortality rate)	Scorecard + IC	0.65	0.42	1.02		RR
Björkman, de Walque and Svensson (2013) – short term	Death (infant mortality rate)	IC	1.05	0.61	1.81		RR
Meta-analysis			0.76	0.42	1.11		RR
Björkman and Svensson (2009) – medium term	Death (under five mortality rate)	Scorecard + IC	0.79	0.57	1.08		RR

CMI: Community monitoring intervention.

Table 12. Weight for age.

Study	Variable definition	CMI type	Effect size	95 Per cent confidence interval			ES type
Björkman and Svensson (2009) – short term	Weight for age (children 0–18 months)	Scorecard + IC	1.20	1.00	1.43		RR
Björkman, de Walque and Svensson (2013) – short term	Weight for age (children 0–11 months)	IC	1.22	0.92	1.60		RR
Meta-analysis			1.20	1.02	1.38		RR
Björkman and Svensson (2009) – medium term	Weight for age (children 0–18 months)	Scorecard + IC	1.29	1.01	1.64		RR

CMI: Community monitoring intervention.

In addition to these measures of health services’ quality, these studies also report on another measure, namely average waiting time in medical facilities. The effects range from RR 0.91 (95% CI [0.81, 1.01]) to RR 1.10 (95% CI [0.81, 1.15]) and are displayed in Table 13.¹⁸ Meta-analysis for the

Table 13. Average waiting time to get the service.

Study	Variable definition	CMI type	Effect size	95 Per cent confidence interval		ES type
Björkman and Svensson (2009) – short term	Waiting time in medical services ^a	Scorecard + IC	0.91	0.81	1.01	RR
Björkman, de Walque and Svensson (2013) – short term	Waiting time in medical services ^a	IC	1.10	0.81	1.15	RR
Meta-analysis			0.99	0.80	1.17	RR
Björkman and Svensson (2009) – medium term	Waiting time in medical services ^a	Scorecard + IC	1.06	0.95	1.19	RR

^aDifference between the time the user left the facility and the time the user arrived at the facility, subtracting the examination time.

CMI: Community monitoring intervention.

short-term interventions suggests a negligible effect (1 per cent reduction in waiting time). However, there is a large between-study heterogeneity that might be driving these results ($I^2 = 70.8\%$, $p = 0.064$).

5.3.2. Education

We included six studies assessing the effect of CMI on the quality of education as measured by test scores.¹⁹ As can be seen from Table 14, three of these studies include multiple treatment arms. We calculated synthetic effect sizes combining the different treatment arms before including these studies in the meta-analysis. The overall average effect of CMI on student outcomes across these six studies is an increase of 0.16 SD (95% CI [0.04, 0.29]).²⁰

The assessment of homogeneity suggests a large amount of variability between studies. This is further supported by the forest plot in Figure 11. The effect sizes range from SMD -0.01 (95% CI $[-0.03, 0.01]$) in Uganda (Reinikka and Svensson, 2011) to SMD 0.63 (95% CI $[0.54, 0.71]$) in Liberia (Piper and Korda 2010). The CIs of these two studies do not overlap.

When we exclude the outliers (Andrabi, Das, and Khwaja 2015; Piper and Korda 2010), the overall effect is SMD 0.01 (95% CI $[-0.01, 0.03]$).

We conducted sensitivity analysis by CMI type, excluding the possible outliers (Figure 12). The overall effect for information campaigns is zero (SMD = 0.004 ; 95% CI $[-0.017, 0.024]$).

Table 14. Test scores.

Study	Variable definition	CMI type	Effect size	95 Per cent confidence interval		ES type
Andrabi, Das and Khwaja (2015)	Test score	Scorecard	0.510	0.163	0.857	SMD
Banerjee et al. (2010) – mobilisation	Test score	IC	0.01	-0.02	0.04	SMD
Banerjee et al. (2010) – mobilisation + information	Test score	IC	0.010	-0.018	0.037	SMD
Banerjee et al. (2010) – mobilisation + information + 'Read India'	Test score	IC	0.03	0.00	0.05	SMD
Banerjee et al. (2010) – all interventions			0.02	-0.01	0.04	SMD
Barr et al. (2012) – standard scorecard	Test score	Scorecard	0.03	-0.05	0.11	SMD
Barr et al. (2012) – participatory scorecard	Test score	Scorecard (participatory)	0.078	-0.002	0.158	SMD
Barr et al. (2012) – all interventions			0.056	-0.015	0.127	SMD
Piper and Korda (2010)	Test score	IC	0.63	0.54	0.71	SMD
Pradhan et al. (2014) – training	Test score	IC	-0.02	-0.09	0.04	SMD
Pradhan et al. (2014) – linkage	Test score	IC	0.07	0.02	0.13	SMD
Pradhan et al. (2014) – all interventions			0.03	-0.03	0.08	SMD
Reinikka and Svensson (2011)	Test score	IC	-0.01	-0.03	0.01	SMD
Meta-analysis			0.16	0.04	0.29	SMD

CMI: Community monitoring intervention; SMD: standardised mean difference.

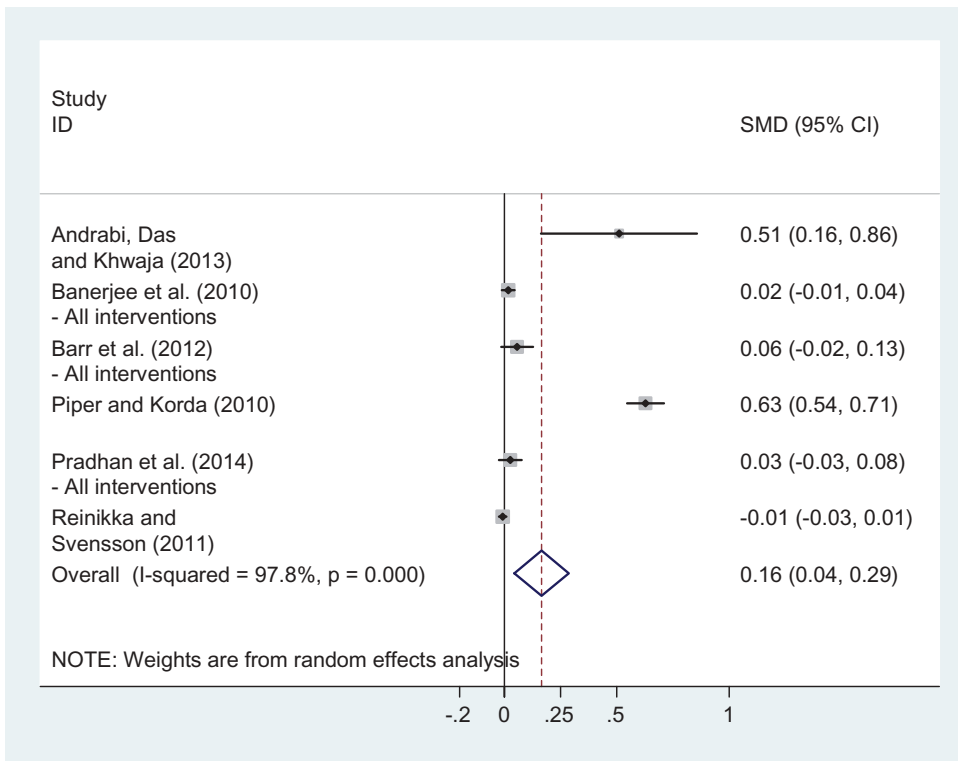


Figure 11. Test scores.

6. Analysis of causal mechanisms (review question 2)

In this section, we synthesise evidence on the causal pathways through which CMI may have an effect on corruption and service delivery outcomes, drawing on data available in studies of effects and the six 'sibling studies' we were able to identify (Table A1). Sometimes, causal pathways are explicit and other times, finding pathways meant looking for implicit assumptions and arguments. We revised the causal chain to show the assumptions we were able to identify from this causal pathway analysis (Figure A1) and present findings here according to determinants of citizen participation and politician and provider accountability.

6.1. Citizen participation

A potential concern with CMI is that citizens may fail to participate in monitoring activities (building block 3). We identified potential bottlenecks that could prevent citizens from participating in monitoring activities, which in turn reduces the potential impact of the programme. In particular, if community monitoring activities are not carried out, or carried out by only a few citizens, the likelihood they uncover problems and put pressure on the government to provide accountability can be significantly reduced. We grouped evidence on citizens' participation in monitoring activities into five areas.

First, *inadequate information on how to monitor the project* does not seem to be an important factor in explaining failures in participation in Colombia and India. The evidence also suggests that even when information is poor, providing that information is not sufficient to increase participation. In India, Banerjee et al. (2007) found that parents, teachers and VEC (local equivalent to

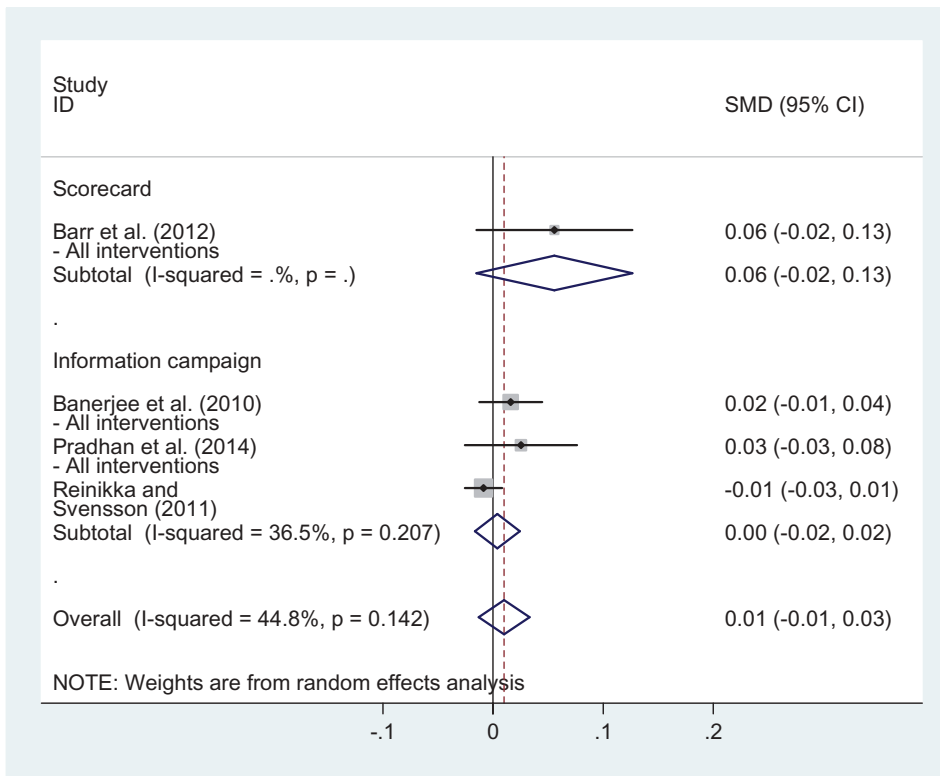


Figure 12. Test scores – sensitivity analysis (outliers excluded).

parent–teacher association) members were not fully aware of how low students' performance was or the role of local committees, and how local participation might improve school performance. Many parents were not aware of the existence of the VEC (sometimes even when they were members of it), and VEC members were unaware of the key roles they were assigned in the educational system. The authors therefore examined different information interventions to mobilise parents to monitor schools, finding that providing information was not effective in encouraging participation. However, encouraging communities to provide volunteers who were subsequently trained by Pratham staff on techniques to teach children how to read and then running after-class reading classes for them did improve reading levels by 3–8 per cent in treatment villages.

Molina (2013b) also found that in some communities, citizens were not aware of the existence of the project they were supposed to monitor, and in some others, although they knew about the project, they lacked information on how to monitor it. This prevented citizens in those communities from taking an active role in social audit community forums and community monitoring activities in general. However, this was not the only or even the most important barrier to participation in this study.

Second, evidence on whether *high opportunity costs of participation* could explain limited engagement varied. Andrabi, Das and Khwaja (2015) find that better educated parents participate more actively in monitoring activities in Pakistan, which might indicate that the opportunity cost of participation is lower for them. In contrast, Molina (2013b), studying the 'Auditorias Visibles' social audit project in Colombia, was not able to find differences between participants and nonparticipants in employment status, income level or whether they work at home or not.

Third, *pessimistic beliefs about politician and provider responsiveness* were found in studies in Colombia and India. Molina (2013b) found that perceiving oneself as influential to local government is crucial for deciding whether to spend time in community monitoring activities. Banerjee et al. (2010) suggest that citizens' perceptions about providers' responsiveness were important in determining engagement with social accountability interventions in India, since only the intervention that did not involve government action had significant effects on participation.²¹ Also in India, Afridi and Iversen (2013) believed that cases of maladministration and corruption in MGNREGA were underreported because of citizens' pessimistic beliefs about the integrity of the audit process. Singh and Vutukuru (2010), analysing the initial stages of the same intervention, also describe a situation with little political enthusiasm during the pilot phase of the social audits, but with subsequent high-level political support generating huge increases in the turnouts at the local social audit meetings.

Evidence suggested that interactions between citizens and service providers could change citizens' perceptions of low accountability. Barr et al. (2012) used the voluntary contributions mechanism game, a public goods game, played immediately after training of school management committees, to show that willingness to contribute to public goods is significantly higher for those in participatory treatment arms. However, the authors were not able to discriminate between two potential mechanisms: that the participatory treatment affects preferences or beliefs about the willingness of providers to contribute to public goods.

Woodhouse (2005) analysed the early stages of the KDP in Indonesia, finding that when villagers possessed information about their rights and, crucially, when the potential perpetrators of corruption knew that villagers had this information, it raised the perceived cost of corrupt behaviour and reduced the cost of fighting it. These pessimistic beliefs could also have been the cause of elite capture (Olken 2007).

Finally, the effect of having *beliefs that other citizens would not participate* could not be verified. Björkman and Svensson (2010) suggest that citizen participation may be threatened by differences within the community, in particular because inequality and ethnic fractionalisation may be barriers to collective action. However, the authors were unable to determine whether failures of collective action were due to lack of trust among community members, lack of trust of the community in the service providers and representatives or both.

6.2. Politician and provider accountability

Providers and politicians need to benefit (by gaining popularity, increasing or maintaining salaries and/or social recognition) in return for being responsive to citizens. If these assumptions are not met, the underlying programme theory of the social accountability information breaks down, in particular by affecting citizens' decisions on whether to monitor government activity. According to our causal chain, even if participation is high, the CMI may be ineffective if politicians can ignore the demand for accountability (for example the politicians does not need their support, clientelism, among others), or even when politicians are committed to improve service delivery, providers are nonresponsive.

In this subsection, we are interested in understanding why demand was lacking. Is it because (a) politicians would be unresponsive to demand, (b) providers would be unresponsive to increased pressure from politicians or (c) citizens believe politicians/providers would be unresponsive. It is important to note the difference between (a) and (b) with (c). While in (a) and (b), politicians and providers respectively are not responsive, in alternative (c), they are responsive but citizens believe they are not. As a result, they do not participate in monitoring and politicians and providers act as if there is no demand.

Molina (2013b) suggests that when the community increases its demands by increasing citizen participation in the social audit, the politicians respond by performing better, as evaluated by the citizens, in this case in Colombia. Keefer and Khemani (2011) found that better learning outcomes were not due to better providers' performance, but due to changes in households' behaviour, as

citizens decided to circumvent the existing institutions to monitor service providers and instead used private sector providers of education.

On the other hand, Björkman and Svensson (2009) found that scorecards were effective in improving waiting times, staff attendance and attitudes towards participants, because of specific innovations implemented to improve management as a result of the CMI. These included a suggestion box for complaints and recommendations, numbered waiting cards for patients, a poster informing about free health services and a poster on patients' rights and obligations. Hence, according to the authors, the programme was successful due to the behaviour change among health facility staff.

Perhaps, more revealing is the evidence from studies with multiple treatments, with and without facilitated (participatory) contact with providers. For example, Barr et al. (2012) found that only the facilitated contact treatment was effective in reducing teacher absence, compared to the standard treatment. Pradhan et al. (2014) also found evidence that facilitated contact between users and providers may enhance motivation for citizens to care about service outcomes and for providers to perform better, for some outcomes.

7. Discussion

Table 15 summarises the findings of the effects of CMIs. The results suggest positive effects on reducing corruption, whether measured using forensic economic estimates or through participant perception. In the case of health service delivery access, two interventions examined utilisation, one finding a moderately large effect size in the medium term. For immunisation, we identified a positive effect in the short term across two interventions, but not in the medium term.

Evidence suggests that CMIs by themselves do not improve access to services in education. Across six interventions looking at education enrolment, results indicate CMIs increase enrolment rates by a small but marginally statistically insignificant amount. Pooling across seven interventions also does not suggest that CMIs reduce school dropouts. On the other hand, evidence from 10 studies suggests that CMIs significantly improve test scores. However, the 10 CMIs to improve test scores are not exactly the same as the ones for increasing enrolment which prevents us from inferring policy conclusion regarding which outcomes are better incentives through CMIs.

Table 15. Summary of effects of CMIs.

Primary outcome	Variable definition	Number of interventions	Effect size	95 Per cent confidence interval	
Forensic economic estimates of corruption		3	0.15 (SMD) ^a	0.01	0.29
	Perception measures of corruption	2	0.08 (RD) ^a	0.02	0.13
Access to service	Utilisation (short term)	2	-0.23 (SMD) ^a	-0.38	-0.07
	Utilisation (medium term)	2	0.99 (ASE)	-1.05	3.02
	Immunisation (short term)	1	0.34 (ASE) ^a	0.12	0.55
	Immunisation (medium term)	2	1.56 (RR) ^a	1.39	1.73
	Enrolment	1	1.04 (RR)	0.52	2.61
	Dropout rate	6	0.09 (SMD)	-0.03	0.21
Improvement in prevalence condition	Child death (short term)	7	-0.00 (SMD)	-0.10	0.10
	Child death (medium term)	2	0.76 (RR)	0.42	1.11
	Weight for age (short term)	1	0.79 (RR)	0.57	1.08
	Weight for age (medium term)	2	1.20 (RR) ^a	1.02	1.38
	Test score	1	1.29 (RR) ^a	1.01	1.64
Quality of service	Average waiting time to get the service (short term)	10	0.16 (SMD) ^a	0.04	0.29
	Average waiting time to get the service (short term)	2	0.99 (RR)	0.80	1.17
	Average waiting time to get the service (medium term)	1	1.06 (RR)	0.95	1.19

^aStatistically significant at 95 per cent confidence level.

SMD: Standardised mean difference; ASE: average standardised effect.

The evidence suggests that health conditions improved for child nutrition, as measured by weight for age in two studies, but not child mortality (CMIIs are estimated to reduce mortality rates but the studies are underpowered to detect significant changes). The evidence also does not support reductions in waiting time as a result of CMIIs.

Few papers assess the effect of CMIIs on any of these outcomes, so our ability to generalise the summative findings of the review is limited.

The scant evidence found on mechanisms suggests that interventions that have modified intermediate outcomes, such as citizen participation in monitoring activities and providers and politicians' performance, have been those that include a set of tools for citizens to monitor providers or politicians and facilitate contacts between both parts. These interventions appear to be the ones that have the bigger impacts on providers' responsiveness (lower absence rates, more teachers' effort, better school inputs) and more participation of communities in monitoring activities (more time spent in monitoring, more in-kind and monetary donations).

There are many reasons why interventions may fail to increase citizen participation in monitoring activities, such as limited information or low expectations of leaders, officials or service provider accountability or about the chances of success. In addition, there may be collective action failure, where some citizens free ride on others' efforts to monitor the project. Certain groups, especially the poor, are less likely to participate in monitoring activities because they have more pressing priorities.

From the evidence included in this review, we find that in the cases where citizens decided not to monitor service provision, providers did not change their behaviour. However, for interventions to increase politician and provider performance, hence, improving service delivery, providers and politicians need to be accountable to citizens. Where enforcement mechanisms are lacking (for example if citizen support is not needed for politicians to stay in power), it is likely that CMIIs will not improve outcomes along the causal chain. If, on the other hand, citizens can impose sanctions on unresponsive providers, CMIIs are more likely to improve providers' performance.

Several studies emphasise the need to adapt interventions to local contexts including Björkman and Svensson (2010), Mansuri and Rao (2012) and Devarajan, Khemani and Walton (2011).

8. Conclusions

This review identified and analysed available evidence regarding the effects of CMIIs on both access to and quality of service delivery and on corruption outcome measures in LMICs. The results suggest that CMIIs can have positive effects on corruption measures and some service delivery measures. However, due to the small number of available studies, results should be interpreted carefully. Interventions considered to address review question (1) may be not representative since they took place mainly in Africa and Asia, in rural communities within specific contexts, so the same interventions may have different effects elsewhere.

Further studies examining the effects of CMIIs in different contexts are needed, especially factorial studies that are able to distinguish between different mechanisms of effect.

Notes

1. There is also a new Global Partnership for Social Accountability (GPSA), which is a coalition of donors, governments and civil society organisations aiming to improve development results by supporting capacity building for enhanced citizen feedback and participation to monitor service delivery. GPSA aims to reach overall funding of 75–125 million USD over the next 7 years. To date, 15 countries have joined the GPSA: Bangladesh, Belarus, Colombia, Dominican Republic, Honduras, Indonesia, Kyrgyzstan, Malawi, Moldova, Mongolia, Mozambique, Philippines, Senegal, Tajikistan and Tunisia.
2. For the purpose of this review, service delivery involves not only services but also construction of necessary infrastructure to carry out those services. As a result, we will talk indistinctly between service delivery and project performance.

3. The term bottlenecks has been used in the literature (Lieberman, Posner, and Tsai 2013) to refer to constraints that limit the effectiveness of community monitoring programmes.
4. In order to give salience to information practitioners use an array of instruments to attract the citizens' attention. We are not aware of any CMI where these incentives were embedded in the theory of change and properly assessed. This appears to be a knowledge gap for CMIs.
5. In some parts of the world, the state fails completely to provide services and to monitor illegal private service provision. Even under these environments, when citizens can choose providers overall providers' performance may increase.
6. We also aimed to answer a third review question – 'to what extent do geographic region, income level or length of exposure to interventions explains the variation in findings?' – using moderator analysis. Due to the limited number of studies, we were unable to perform planned moderator analyses by study design and length of exposure to intervention. Moderator analysis by region did not reveal big differences.
7. We critically appraised the consistency among ratings by our coders through interrater assessment (McGraw and Wong 1996). The absolute agreement intraclass correlation is 0.70 with a 95 per cent CI [0.21, 0.95]. In case of disagreement, a third researcher determined the rating.
8. ASE (Kling et al. 2004) is interpreted in the same way as SMD. It is calculated by combining several measures for the same outcome into a unique average standardised treatment effect (ASE), by estimating a seemingly unrelated regression system for K -related outcomes: $Y = [I_k \otimes (T X)] \theta + u$ where I_k is a K by K -identity matrix. The average standardised treatment effect is estimated

as

$$\tilde{\beta} = \frac{1}{K} \sum_{k=1}^K \frac{\hat{\beta}_k}{\hat{\sigma}_k}$$

where $\tilde{\beta}_k$ is the point estimate on the treatment indicator in the k th outcome regression and $\hat{\sigma}_k$ is the standard deviation of the control group for outcome k (Björkman, de Walque, and Svensson 2013).

9. In some cases, these correlations were available in the studies' databases, or where easily obtainable from tables reported in the papers. When not available, we assumed a value of 0.5 and checked whether the results changed substantially for extreme correlation values.
10. All randomised field experiments report no statistical difference between treatment and control groups.
11. In both cases, we changed the sign of the effect size so it can be interpreted properly (that is a positive effect size means that corruption has been reduced).
12. The study also evaluates the effect of external audits, which did reduced corruption, but we did not include it in the meta-analysis because it does not fall into any of our four intervention categories.
13. This finding is consistent with those reported by the author, who argues that 'increasing grassroots participation in monitoring had little average impact' (Olken 2007).
14. Finally, we did not include Banerjee et al. (2010) as it is not a measure of corruption they use, but rather they look at whether the treatments to increase community monitoring generated additional nonteaching resources for the schools. They found that none of the interventions have any effect.
15. Actually, the medium-term impact of the first intervention is assessed in Björkman, de Walque and Svensson (2013). However, to avoid confusion, we designate the latter as the main reference for the second intervention and Björkman and Svensson (2009) for the first intervention.
16. We were not able to compute neither SMD nor RR for these outcomes due to lack of information.
17. The short-term impact of this intervention is also not statistically significant, but it is not reported in the table since we were not able to compute RR.
18. It is important to note why we think this is a quality measure and not an access measure. Access is related to getting the service. However, you can get the service and the fact that you had to wait makes it less valuable and of lesser quality.
19. When different test scores were reported (for example language and math test scores), we previously pooled them following the procedure explained before.
20. It should be noted that we are excluding two studies for which we were not able to compute standardised effects (Table 12) but which found significant effects of CMIs on our outcomes of interest. Keefer and Khemani (2011) examine the proportion of children tested in the village public school who could read sentence and paragraphs (ASER literacy test), finding that the information campaign resulting from communities' access to radios enhanced literacy tests. Pandey, Goyal and Sundararaman (2009) examined the percentage of children who could pass different learning tests, including reading and writing competences and mathematics abilities.
21. Additional qualitative and quantitative evidence support this. For example, Gaventa and Barrett (2012) perform a metacase study of 100 interventions aim to increase citizen engagement in service delivery. For the 828 outcomes from the 100 reviewed case studies, only 153 came from interventions where the final goal was to strengthen the responsiveness and accountability of the state to provide services. Results indicate that 55 per cent of those 153 outcomes were positive and 45 per cent were negative. Negative results were associated with failure of citizens to participate, due in part to fear of backlash against those who speak out and a sense of tokenism in the participation mechanism.

Notes on contributors

Ezequiel Molina is an economist in the Education Global Practice of the World Bank. He works on issues of institutional reform and service delivery with research spanning the areas of governance, education, gender and poverty and inequality. Recent publications include a systematic review on the effectiveness of social accountability interventions and a paper assessing the quality of service delivery in Sub-Saharan Africa. He also has co-authored the book *Economic Polarization, Institutions and Conflict in Latin America*. Ezequiel holds a PhD in Political Economy from Princeton University, and a B.A. and M.A. in Economics from La Plata National University (Argentina).

Laura Carella is Junior researcher of the Center for Distributive, Labor and Social Studies (CEDLAS) at Universidad Nacional de La Plata (UNLP) in Argentina. She holds a Master in Economics from the UNLP. She is a Professor of the Economics Department at UNLP and teaches undergraduate courses on Econometrics and Microeconomics. She has also taught graduate courses on Microeconomics. Her major fields of research are Economics of Education, Health Economics and Social Policy.

Ana Pacheco is a Senior Researcher and Executive Coordinator for the Center for Distributive, Labor and Social Studies (CEDLAS) at Universidad Nacional de La Plata (UNLP) in Argentina. Pacheco has a Ph.D in Economics from Universidad Nacional de La Plata. She teaches undergraduate Macroeconomics, and her research is focused on development economics, with a special emphasis on applied econometrics and computed general equilibrium models.

Guillermo Cruces (PhD in Economics, LSE) is Under-Secretary of Development at the Treasury Ministry in Argentina. He is on leave as the deputy director of the Center for Distributive, Labor and Social Studies (CEDLAS) at the Universidad Nacional de La Plata, Argentina (UNLP). He is also a researcher at Argentina's National Scientific and Technical Research Council (CONICET), and a research fellow at IZA. His research is focused on labor economics and distributional analysis in Latin America and the Caribbean, and on the economics of perceptions and reference groups in general. He teaches at the graduate and undergraduate level at the Economics Department of the UNLP, and he is invited professor of labor economics at the Universidad de San Andrés (UdeSA), Argentina. He has published in journals such as the *Journal of Political Economy*, *Journal of Public Economics*, *American Economic Journal – Macroeconomics*, *Labour Economics*, *Journal of Population Economics*, *Brookings Papers on Economic Activity*, *Journal of Development Studies* and *Economia*. He has edited books and contributed to collective volumes and reports, and recently published the book *Growth, Employment and Poverty in Latin America* (Oxford University Press, 2017, with G. Fields, D. Jaume and M. Violaz). He has worked previously for the UK's Department for Work and Pensions and for the Development Studies Division of the UN's Economic Commission for Latin America and the Caribbean. He has also been a researcher at STICERD, London School of Economics and Political Science, where he obtained an MSc and a PhD in Economics, and a visiting scholar at Harvard's DRCLAS and at University of California at Berkeley.

Leonardo Gasparini is the director of the Center for Distributive, Labor and Social Studies (CEDLAS) at Universidad Nacional de La Plata (UNLP) in Argentina. Gasparini has a Ph.D in Economics from Princeton University. He teaches undergraduate and graduate courses on Income Distribution and Labor Economics at UNLP. He has published articles on income distribution, social policy and labor issues in several international journals such as the *Journal of Public Economics*, *Review of Income and Wealth*, *Social Choice and Welfare*, *Economic Development and Cultural Change*, *Journal of Economic Inequality*, *Journal of Income Distribution*, *Journal of International Development*, and others. He has written several books on development issues. Gasparini has taught courses and given seminars on distributive issues in almost all Latin American countries. Gasparini has been awarded with the Guggenheim fellowship 2009.

Acknowledgements

Many academics and practitioners read the review protocol and gave us relevant feedback. In particular, we want to thank Bénédicte de la Brière, Seollee Park and Vincenzo Di Maro for thoughtful comments. Additionally, we would like to thank reviewers from 3ie and the International Development Coordinating Group (IDCG) of the Campbell Collaboration. The usual disclaimer applies.

Disclosure statement

There are no known conflicts of interest that the team is currently aware of. The team has not been part of any organisation that has implemented projects in this area nor has any interests in promoting particular findings due to personal relationships with individuals or organisations who will benefit from these. Ezequiel Molina has conducted research on community monitoring in Colombia (Community Visible Audits) as part of his dissertation work. He

studied the effects of the programme on corruption as well as political influence of the community over policy making.

Funding

3ie Systematic Review Grant provided the funding for this systematic review.

ReferencesReferences

- Acemoglu, D., and J. A. Robinson. 2008. "Persistence of Power, Elites, and Institutions." *American Economic Review* 98 (1):267–293. [10.1257/aer.98.1.267](https://doi.org/10.1257/aer.98.1.267)
- Afridi, F., and V. Iversen. 2013. *Social Audits and MGNREGA Delivery: Lessons from Andhra Pradesh*, Brookings-NCAER India Policy Forum, edited by B. Bosworth, A. Panagariya, and S. Shah. <http://www.ncaer.org/free-download.php?plD=250>
- Alatas, V., A. Banerjee, R. Hanna, B. A. Olken, and J. Tobias. 2012. "Targeting The Poor: Evidence From a Field Experiment in Indonesia." *The American Economic Review* 102 (4): 1206–40. doi:[10.1257/aer.102.4.1206](https://doi.org/10.1257/aer.102.4.1206).
- Andrabi, T., J. Das, and A. I. Khwaja. 2015. *Report cards: The impact of providing school and child test scores on educational markets*. (Policy Research Working Paper No. 7226). Washington, DC: The World Bank.
- Banerjee, A., R. Banerji, E. Dufló, R. Glennerster, D. Kenniston, S. Khemani, and M. Shotland 2007. "Can Information Campaigns Raise Awareness and Local Participation in Primary Education?" *Economic and Political Weekly*, 42(15): 1365–1372.
- Banerjee, A. V., and E. Dufló. 2011. *Poor Economics, a Radical Rethinking of the Way to Fight Global Poverty*, Public Affairs. US: Public Affairs.
- Banerjee, A. V., R. Banerji, E. Dufló, R. Glennerster, and S. Khemani 2010. "Pitfalls of Participatory Programs: Evidence from a Randomized Evaluation in Education in India." *American Economic Journal: Economic Policy*, 2(1), 1–30.
- Banerjee, A. V., and S. Mullainathan. 2008. "Limited Attention and Income Distribution." *American Economic Review*, 98 (2): 489–493. [10.1257/aer.98.2.489](https://doi.org/10.1257/aer.98.2.489)
- Bardhan, P. 2002. "Decentralization of Governance and Development." *Journal of Economic Perspectives*, 185–205. [10.1257/089533002320951037](https://doi.org/10.1257/089533002320951037)
- Bardhan, P., and D. Mookherjee. 2006. "Decentralisation and Accountability in Infrastructure Delivery in Developing Countries." *The Economic Journal*, 116(508), 101–127. [10.1111/eoj.2006.116.issue-508](https://doi.org/10.1111/eoj.2006.116.issue-508)
- Barr, A., F. Mugisha, P. Serneels, and A. Zeitlin. 2012. "Information and Collective Action in Community-Based Monitoring of Schools: Field and Lab Experimental Evidence from Uganda." Working paper mimeo.
- Besley, T., and T. Persson. 2011. *Pillars of Prosperity: The Political Economics of Development Clusters*. Princeton, NJ: Princeton University Press.
- Björkman, M., D. de Walque, and J. Svensson. 2013. "Information Is Power: Experimental Evidence of the Long Run Impact of Community Based Monitoring." World Bank Policy Research Working Series, WPS7015, Washington DC. <http://documents.worldbank.org/curated/en/739811468318587268/Information-is-power-experimental-evidence-on-the-long-run-impact-of-community-based-monitoring>
- Björkman, M., and J. Svensson. 2009. "Power to the People: Evidence from a Randomized Field Experiment on Community-Based Monitoring in Uganda." *The Quarterly Journal of Economics*, 124(2), 735–769. [10.1162/qjec.2009.124.2.735](https://doi.org/10.1162/qjec.2009.124.2.735)
- Björkman, M., and J. Svensson. 2010. "When Is Community-Based Monitoring Effective? Evidence from a Randomized Experiment in Primary Health in Uganda." *Journal of the European Economic Association*, 8, 2–3, 571–581. [10.1111/\(ISSN\)1542-4774](https://doi.org/10.1111/(ISSN)1542-4774)
- Booth, A. 2011. Searching for Studies. In *Supplementary Guidance for Inclusion of Qualitative Research in Cochrane Systematic Reviews of Interventions*. Version 1 (updated August 2011), edited by J. Noyes, A. Booth, K. Hannes, A. Harden, J. Harris, S. Lewin, & C. Lockwood. Cochrane Collaboration Qualitative Methods Group. <http://cqrmg.cochrane.org/supplemental-handbook-guidance>.
- Borenstein, M., L. V. Hedges, J. P. Higgins, and H. R. Rothstein (2009). *Introduction to Meta-Analysis*. Hoboken, NJ: Wiley.
- Brinkerhoff, D. W., and O. Azfar. 2006. *Decentralization and Community Empowerment: Does Community Empowerment Deepen Democracy and Improve Service Delivery?* Washington, DC: Office of Democracy and Governance, USAID.
- Capuno, J. J., and M. M. Garcia. 2010. "Can Information about Local Government Performance Induce Civic Participation? Evidence from the Philippines." *Journal of Development Studies*, 46(4), 624–643. doi:[10.1080/00220380903023521](https://doi.org/10.1080/00220380903023521).
- Chaudhury, N., J. Hammer, M. Kremer, K. Muralidharan, and F. Halsey Rogers. 2006. "Missing in Action: Teacher and Health Worker Absence in Developing Countries." *Journal of Economic Perspectives*, Winter 20, 91–116. doi:[10.1257/089533006776526058](https://doi.org/10.1257/089533006776526058).

- Devarajan, S., S. Khemani, and M. Walton. 2011. "Civil Society, Public Action and Accountability in Africa." HKS Faculty Research Working Paper Series RWP11-036, John F. Kennedy School of Government, Harvard University. <https://research.hks.harvard.edu/publications/getFile.aspx?id=723>
- Downs, A. 1957. *An Economic Theory Of Democracy*. Harper.
- Egger, M., G. Davey Smith, M. Schneider, and C. Minder. 1997. "Bias in Meta-Analysis Detected by a Simple, Graphical Test." *British Medical Journal*, 315, 629–634. doi:10.1136/bmj.315.7109.629.
- Ferraz, C., F. Finan, and D. B. Moreira. 2012. "Corrupting learning: Evidence from missing federal education funds in Brazil (Vol. w18150): National Bureau of Economic Research, Inc, NBER Working Papers: 18150, US."
- Gaventa, J., and G. Barrett. 2012. "Mapping the Outcomes of Citizen Engagement." *World Development* 40(12): 2399–2410. doi:10.1016/j.worlddev.2012.05.014.
- Gertler, P., H. A. Patrinos, and M. Rubio-Codina. 2008. "Empowering Parents to Improve Education: Evidence from Rural Mexico." Policy Research Working Paper 3935-IE. Revised May 2008. Washington, DC: The World Bank.
- Grandvoinet, H., G. Aslam, and S. Raha. 2015. *Opening the Black Box: The Contextual Drivers of Social Accountability*. Washington, DC: World Bank Publications.
- Hanna, R., S. Bishop, S. Nadel, G. Scheffler, and K. Durlacher. 2011. *The Effectiveness of Anti-Corruption Policy: What Has Worked, What Hasn't, and What We Don't Know*. DIFD Systematic Review. London. <https://eppi.ioe.ac.uk/cms/Portals/0/PDF%20reviews%20and%20summaries/Anti%20corruption%202011Hanna.pdf?ver=2011-10-11-142359-153>
- Higgins, J., and S. Green, eds. 2011. *Cochrane Handbook for Systematic Reviews of Interventions. (Version 5.0.2, Updated September 2009)*. Cochrane Collaboration. www.cochrane-handbook.org.
- Hotelling, H. 1929. "Stability in Competition." *Economic Journal* 39:41-57.
- Keefer, P., and S. Khemani. 2004. "Why Do the Poor Receive Poor Services?" *Economic and Political Weekly*, 39, 9, 935–943.
- Keefer, P., and S. Khemani (2005). "Democracy, Public Expenditures, and the Poor." *World Bank Research Observer*, 20, 1, 1–27. doi:10.1093/wbro/lki002.
- Keefer, P., and S. Khemani. 2011. "Mass Media and Public Services: The Effects of Radio Access on Public Education in Benin." Policy Research Working Paper Number 5559, Development Research Group. Washington, DC: The World Bank.
- King, E., C. Samii, and B. Snilstveit. 2010. "Interventions to Promote Social Cohesion in sub-Saharan Africa." *Journal of Development Effectiveness*, 2(3), 336–370. doi:10.1080/17449057.2010.504552.
- Kling, J. R., J. B. Liebman, L. F. Katz, and L. Sanbonmatsu. 2004. "Moving to Opportunity and Tranquility: Neighborhood Effects on Adult Economic Self-Sufficiency and Health from a Randomized Housing Voucher Experiment." Princeton University Working Paper No. 5. Princeton, NJ.
- Le Grand, J. (2003). *Motivation, Agency, And Public Policy: Of Knights And Knaves, Pawns And Queens*. Oxford University Press, Oxford, UK.
- Lieberman, E. S., D. N. Posner, and L. L. Tsai. 2013. "Does Information Lead to More Active Citizenship? Evidence from an Education Intervention in Rural Kenya." *World Development* 60: 69-83. ISSN 0305-750X, <https://doi.org/10.1016/j.worlddev.2014.03.014>
- Mansuri, G., and V. Rao. 2012. *Localizing Development: Does Participation Work?* Washington DC: World Bank.
- Maru, V. 2010. "Social Accountability and Legal Empowerment." *Health and Human Rights: An International Journal, North America*, 1225 05.
- Mauro, P. 1995. "Corruption and Growth." *The Quarterly Journal of Economics*, 110(3): 681–712. doi:10.2307/2946696.
- McGraw, K. O., and S. P. Wong 1996. "Forming Inferences about Some Intraclass Correlation Coefficients." *Psychological methods*, 1(1), 30. doi:10.1037/1082-989X.1.1.30.
- Molina, E. 2013a. "Community Monitoring and Self-Fulfilling Prophecies in Service Delivery." Unpublished Manuscript. Princeton University Dissertation. Princeton, NJ. <http://dataspace.princeton.edu/jspui/handle/88435/dsp01pz50gw223>
- Molina, E. 2013b. "Bottom up Institutional Reform: Evaluating the Impact of the Citizen Visible Audit Program in Colombia." Unpublished document. Princeton University Dissertation. Princeton, NJ. <http://dataspace.princeton.edu/jspui/handle/88435/dsp01pz50gw223>
- Molina, E., L. Carella, A. Pacheco, G. Cruces, and L. Gasparini 2016. "Community Monitoring Interventions to Curb Corruption and Increase Access and Quality of Service Delivery in Low- and Middle-Income Countries." *Campbell Systematic Reviews* 8 doi:10.4073/csr.2016.8.
- Nunn, N., and L. Wantchekon. 2011. "The Slave Trade and the Origins of Mistrust in Africa." *American Economic Review* 101 (7) (December): 3221–3252. doi:10.1257/aer.101.7.3221.
- Olken, B. A. 2005. Corruption Perceptions Vs. Corruption Reality. (No. W12428). US: National Bureau of Economic Research.
- Olken, B. A. 2005. Monitoring Corruption: Evidence from a Field Experiment in Indonesia (No. W11753). National Bureau of Economic Research. <http://www.nber.org/papers/w11753.pdf>
- Olken, B. A. 2006. "Corruption and the Costs of Redistribution: Micro Evidence from Indonesia." *Journal of Public Economics*, 90(4–5):853–870. doi:10.1016/j.jpubeco.2005.05.004.

- Olken, B. A. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy*, 115(2). [10.1086/517935](https://doi.org/10.1086/517935)
- Olken, B. A., and P. Barron. 2009. The Simple Economics of Extortion: Evidence from Trucking in Aceh. *Journal of Political Economy*, 117(3): 417–452. doi:[10.1086/599707](https://doi.org/10.1086/599707).
- Olson, M. 1971. *The Logic of Collective Action: Public Goods and the Theory of Groups*. Harvard Economic Studies, vol. 124 revised ed. Cambridge, MA: Harvard University Press.
- Pan, L., and L. Christiaensen. 2012. "Who Is Vouching for the Input Voucher? Decentralized Targeting and Elite Capture in Tanzania." *World Development*, Elsevier, 40(8), 1619–1633. doi:[10.1016/j.worlddev.2012.04.012](https://doi.org/10.1016/j.worlddev.2012.04.012).
- Pande, R., and B. A. Olken. 2011. *Governance Review Paper*. JPAL governance initiative. Cambridge, MA: Abdul Latif Jameel Poverty Action Lab.
- Pandey, P., A. R. Sehgal, M. Riboud, D. Levine, and M. Goyal 2007. "Informing Resource-Poor Populations and the Delivery of Entitled Health and Social Services in Rural India." *JAMA: The Journal of the American Medical Association*, 298(16), 1867–1875. [10.1001/jama.298.16.1867](https://doi.org/10.1001/jama.298.16.1867)
- Pandey, P., S. Goyal, and V. Sundararaman. 2009. "Community Participation in Public Schools: Impact of Information Campaigns in Three Indian States." *Education Economics*, 17(3), 355–375. [10.1080/09645290903157484](https://doi.org/10.1080/09645290903157484)
- Persson, T., and G. Tabellini. 2002. "Political Economics: Explaining Economic Policy." Vol. 1 of MIT Press Books. Cambridge, MA: The MIT Press.
- Piper, B., and M. Korda. 2010. "Early Grade Reading Assessment (EGRA) Plus: Liberia. Program Evaluation Report: Prepared for USAID/Liberia." Research Triangle Park, NC: RTI International.
- Pradhan, M., D. Suryadarma, A. Beatty, M. Wong, A. Gaduh, A. Alisjahbana, and R. P. Artha. 2014. "Improving Educational Quality through Enhancing Community Participation: Results from a Randomized Field Experiment in Indonesia." *American Economic Journal: Applied Economics*, 6(2), 105–126.
- Reinikka, R., and J. Svensson. 2004. "Local Capture: Evidence from a Central Government Transfer Program in Uganda." *The Quarterly Journal of Economics*, 119(2), 679–705. doi:[10.1162/0033553041382120](https://doi.org/10.1162/0033553041382120).
- Reinikka, R., and J. Svensson. 2005. "Fighting Corruption to Improve Schooling: Evidence from a Newspaper Campaign in Uganda." *Journal of the European Economic Association*, MIT Press, 3(2–3), 259–267, 04/05. doi:[10.1162/jeea.2005.3.2-3.259](https://doi.org/10.1162/jeea.2005.3.2-3.259).
- Reinikka, R., and J. Svensson. 2011. "The Power of Information in Public Services: Evidence from Education in Uganda." *Journal of Public Economics*, Elsevier, 95(7–8), 956–966. [10.1016/j.jpubeco.2011.02.006](https://doi.org/10.1016/j.jpubeco.2011.02.006)
- Ringold, D., A. Holla, M. Koziol, and S. Srinivasan. 2012. *Citizens and Service Delivery. Assessing the Use of Social Accountability Approaches in the Human Development Sectors*. Washington, DC: World Bank.
- Rose-Ackerman, S. (2004). Governance and Corruption, in *Global Crises, Global Solutions*. B. Lomborg, edited by Cambridge: Cambridge University.
- Sacks, A., and M. Larizza. 2012. "Why Quality Matters: Rebuilding Trustworthy Local Government in Post-Conflict Sierra Leone." World Bank Policy Research Working Paper No. Washington DC.
- Shadish, W., and D. Myers 2004. "Research Design Policy Brief. Campbell Collaboration: Oslo." http://www.campbellcollaboration.org/artman2/uploads/1/C2_Research_Design_Policy_Brief-2.pdf.
- Singer, M. M. 2013. "Bribery Diminishes Life Satisfaction in the Americas." Americas Barometer Insights: 2013, Number 89. Latin American Public Opinion Project, Vanderbilt University. Nashville, Tennessee. http://polisci.uconn.edu/wp-content/uploads/sites/1645/2016/02/matt_singer_cv.pdf
- Singh, R., and V. Vutukuru. 2010. *Enhancing Accountability in Public Service Delivery through Social Audits: A Case Study of Andhra Pradesh, India*. New Delhi: Accountability Initiative, Centre for Policy Research.
- Snilstveit, B. 2012. "Systematic Reviews: From 'Bare Bones' Reviews to Policy Relevance." *Journal of Development Effectiveness*, 4(3), 388–408. doi:[10.1080/19439342.2012.709875](https://doi.org/10.1080/19439342.2012.709875).
- Stiglitz, J. E. 2002. "Participation and Development: Perspectives from the Comprehensive Development Paradigm." *Review of Development Economics*, 6(2), 163–182. doi:[10.1111/rode.2002.6.issue-2](https://doi.org/10.1111/rode.2002.6.issue-2).
- Sukhtankar, S. 2011. "Sweetening the Deal?" *Political Connections and Sugar Mills in India American Economic Journal: Applied Economics*, 4, 3, 43–63, July 2012.
- Svensson, J. 2003. "Who Must Pay Bribes and How Much? Evidence from a Cross Section of Firms." *The Quarterly Journal of Economics*, 118(1): 207–230. doi:[10.1162/0033553036053180](https://doi.org/10.1162/0033553036053180).
- Svensson, J. 2005. "Eight Questions about Corruption." *Journal of Economic Perspectives*, 19 (5): 19–42. doi:[10.1257/089533005774357860](https://doi.org/10.1257/089533005774357860).
- Tosi, F. G. 2012. *Direct Popular Participation and Crises in Public Services in Argentina*. Rochester: Gordian Knot.
- Transparency International. 2013. *Policy Brief. Looking beyond 2015: A Role for Governance*. Berlin: Transparency International.
- UN. 2008. *People Matter: Civic Engagement in Public Governance*. New York, NY: United Nations Department of Economic and Social Affairs.
- Waddington, H., B. Snilstveit, J. Hombrados, M. Vojtkova, D. Phillips, P. Davies, and H. White. 2014. Farmer Field Schools for Improving Farming Practices and Farmer Outcomes: A Systematic Review *Campbell Systematic Reviews* 6 doi:[10.4073/csr.2014](https://doi.org/10.4073/csr.2014).

- Waddington, H., H. White, B. Snilstveit, J. G. Hombrados, M. Vojtkova, P. Davies, and P. Tugwell. 2012. "How to Do a Good Systematic Review of Effects in International Development: A Tool Kit." *Journal of development effectiveness*, 4 (3), 359–387. doi:10.1080/19439342.2012.711765.
- White, H. 2009. "Theory-Based Impact Evaluation: Principles and Practice." Working Paper. International Initiative for Impact Evaluation: New Delhi. http://www.3ieimpact.org/admin/pdfs_papers/51.pdf.
- Woodhouse, A. 2005. "Village Corruption in Indonesia. Fighting Corruption in Indonesia's Kecamatan Development Program." World Bank Working Papers Series: Indonesian Social Development Paper No. 6, Washington, DC..
- Zitzewitz, E. 2012. "Forensic Economics." *Journal of Economic Literature*, 50(3): 731–769. doi:10.1257/jel.50.3.731.



Appendix

Table A1. Description of included studies.

Study (sibling study)	Location and treatment unit	Intervention	Target and comparison group	Study design/ attribution method	Outcomes data collected	Follow-up time period ^a	Internal validity assessment
Afridi and Iversen (2013) (Singh and Vutukuru 2010)	India: Andhra Pradesh. Gram Panchayats	Social audit (Second audit) Social audit (Third audit)	This is a panel data set which comprises official data from three rounds of social audits, with an initial sample of 300 GPs in 8 districts. It compares the results of the second (264 audits) and third audit (166) with those of the first one (284). 548 number of audits, from which 284 are first audit, and 264 are a second audit (clusters at GPs level: 300). The intervention is the Mahatma Gandhi National Rural Employment Guarantee Schemes (MGNREGA)	DID	Perception of corruption: number of irregularities	5 years	Low risk of bias
Andrabi, Das, and Khwaja (2015)	Pakistan. Schools and children	Scorecard	Villages were sampled from 3 districts: 1 each in the north, centre and south. Treatment group: Scorecards. Control group without scorecard 112 villages were chosen randomly from among those with at least one private school according to a 2000 census of private schools. First, Grade 3 children in all primary schools were tested and then, in a randomly selected 50 per cent of the villages, were disseminated report cards with the results of school and child test scores for all schools (804) and tested children (12,110)	DID	Access to service: education enrolment Changes in prevalence: test scores	1 and 2 years	Low risk of bias

(Continued)



Table A1. (Continued).

Study (sibling study)	Location and treatment unit	Intervention	Target and comparison group	Study design/ attribution method	Outcomes data collected	Follow-up time period ^a	Internal validity assessment
Banerjee et al. (2010) (Banerjee et al. 2007)	India: Uttar Pradesh. Villages	IC Treatment 1 IC Treatment 2 IC Treatment 3	The evaluation took place in 280 villages in Jaunpur district. 85 Villages as control group and 195 as target group. The final sample for the baseline survey consisted of 2800 households, 316 schools, 17,533 children (ages 7–14) tested in reading and math and 1029 VEC member interviews from the 280 villages. In the endline survey, 17,419 children were tested, a sample that includes all but 716 of the children in the baseline and, thus, very little attrition from the baseline survey (the attrition is evenly spread across the various treatment and control groups)	RCT	Access to service: education enrolment Changes in prevalence: test scores	1 year	Low risk of bias
Barr et al. (2012)	Uganda. Children	Scorecard intervention 1: standard scorecard Scorecard intervention 2: participatory scorecard	100 Rural primary schools: 30 schools were assigned to each of the standard and participatory treatment arms, with the remaining 40 serving as a control group. 3512 Students, we assume it follows the same division as the schools	RCT, DID	Changes in prevalence: test scores	2 years and 4 months	Low risk of bias
Björkman and Svensson (2009) (Björkman and Svensson 2010)	Uganda. Health facilities and healthcare users	Scorecard + information campaign	25 Facilities/communities randomly assigned as control group and 25 facilities/communities randomly assigned as target group	RCT	Access to services: utilisation/coverage, immunisation Changes in prevalence: mortality rate, weight for age Average waiting time	1 year	Low risk of bias

(Continued)



Table A1. (Continued).

Study (sibling study)	Location and treatment unit	Intervention	Target and comparison group	Study design/ attribution method	Outcomes data collected	Follow-up time period ^a	Internal validity assessment
Björkman, De Waalque, and Svensson (2013)	Uganda. Communities, health facilities and households. The first intervention was the same as described in Björkman and Svensson (2009)	Scorecard + information campaign IC intervention	25 Facilities/communities randomly assigned as control group and 25 facilities/communities randomly assigned as target group 12 Facilities/communities randomly assigned as control group and 13 facilities/communities randomly assigned as target group	Cross-section (regression), DID, Seemingly unrelated regression (Kling et al. 2004)	Access to services: utilisation/ coverage, immunisation Changes in prevalence: mortality rate, weight for age Average waiting time	2 years 1 year	Low risk of bias
Gertler, Patrinos, and Rubio-Codina (2008)	Mexico. Nonindigenous primary schools in rural areas	Scorecard	Treatment schools are those schools that first received the AGE programme at the beginning of any school year between 1998–99 and 2001–02 and had AGE continuously ever since (N = 2544). Those that had not received AGE before school year 2002–03 constitute the comparison group (N = 3483)	DID	Access to services: education enrolment and repetition	12 years	Low risk of bias
Keefer and Khemani (2011)	Benin. Households and children in second grade	IC intervention	In the target group are the households and children in the villages which access to the radio. In the control are those in villages without access to the radio. 210 Villages (4200 households) from 21 communes	Cross-section, quasi-experimental	Changes in prevalence: test scores	Not applicable	Low risk of bias
Molina (2013b)	Colombia. Projects and households	Social audit	The random sample contains 390 households for the 13 projects in the treatment group and 410 for the 11 projects in the control group	Cross-section – matching	Perception of corruption: perception of adequacy in administration of resources	Not applicable	Low risk of bias

(Continued)

Table A1. (Continued).

Study (sibling study)	Location and treatment unit	Intervention	Target and comparison group	Study design/ attribution method	Outcomes data collected	Follow-up time period ^a	Internal validity assessment
Olken (2007) (Olken 2005, Olken 2005) (Woodhouse 2005)	Indonesia. Villages	Social audit: invitations	Social audit with invitations versus social audit. 199 Villages (audit 94, control 105)	RCT, cross-section (regression)	Forensic estimates of corruption: per cent missing funds major items in roads and ancillary projects	1 year	Low risk of bias
Pandey et al. (2007)	India: Uttar Pradesh. Households	Social audit: invitations + comments IC	Social audit with invitations plus comments versus social audit 202 villages (audit 96, control 106) 105 Villages (1045 households at the baseline), from which 55 (548 households) intervention and 50 (497 households) control	DID and Cross-section	Perception of corruption: development work in villages Access to service: visits by nurse midwife; prenatal examinations, tetanus vaccinations and prenatal supplements received by pregnant women; vaccinations received by infants	1 year	Low risk of bias
Pandey, Goyal and Sundaraman (2009)	India: Uttar Pradesh, Madhya Pradesh and Karnataka. Teachers and students in villages	IC Second IC in one region	610 Villages from which 340 intervention and 270 control	RCT	Changes in prevalence: test scores	2–4 months	Medium risk

(Continued)



Table A1. (Continued).

Study (sibling study)	Location and treatment unit	Intervention	Target and comparison group	Study design/ attribution method	Outcomes data collected	Follow-up time period ^a	Internal validity assessment
Piper and Korda (2010)	Liberia. Schools	IC	Groups were randomly selected and clustered within districts, such that several nearby schools were organised together. 117 schools, from which 59 are control. The intervention was targeted at Grades 2 and 3	RCT	Changes in prevalence: test scores	1 year	Medium risk
Pradhan et al. (2014)	Indonesia. Schools	Training (T): IC Linkage (L): IC The authors also include a third treatment that we do not consider as it is not an eligible CMI for this review	2 Provinces, 9 districts, 44 subdistricts and 520 schools. Training: treatment group: 230 schools, 1060 students; control group: 190 schools and 2120 students. Linkage: treatment group: 240 schools and 893 students; control group: 180 schools and 2120 students. The authors also include a third treatment that we do not consider as it is not of the type of CMI considered in this review, the intervention introduced changes in the election of the committee. They also explore combinations of treatments given that some individuals in the control groups for each treatment had received the other treatments	DID	Changes in prevalence: test scores	1 year and 10 months	Low risk of bias
Reinikka and Svensson (2011)	Uganda. Schools	IC	Using distance to newspapers outlets, the authors construct the treatment and control group. 218 Schools for which survey data are available for the years 1991–95 and 2001, and a sample of 388 (218 + 170) schools for which survey data are available in 2001	DID, IV	Forensic economic estimates of corruption: share of grants received Access to service: education enrolment Changes in prevalence: test scores	Not applicable	Low risk of bias

IC: information campaign; SA: social audit; SC: scorecard; DID: differences-in-differences; IV: instrumental variable; OLS: ordinary least squares estimation; CMI: community monitoring intervention.
^aAverage years from intervention to endline survey.

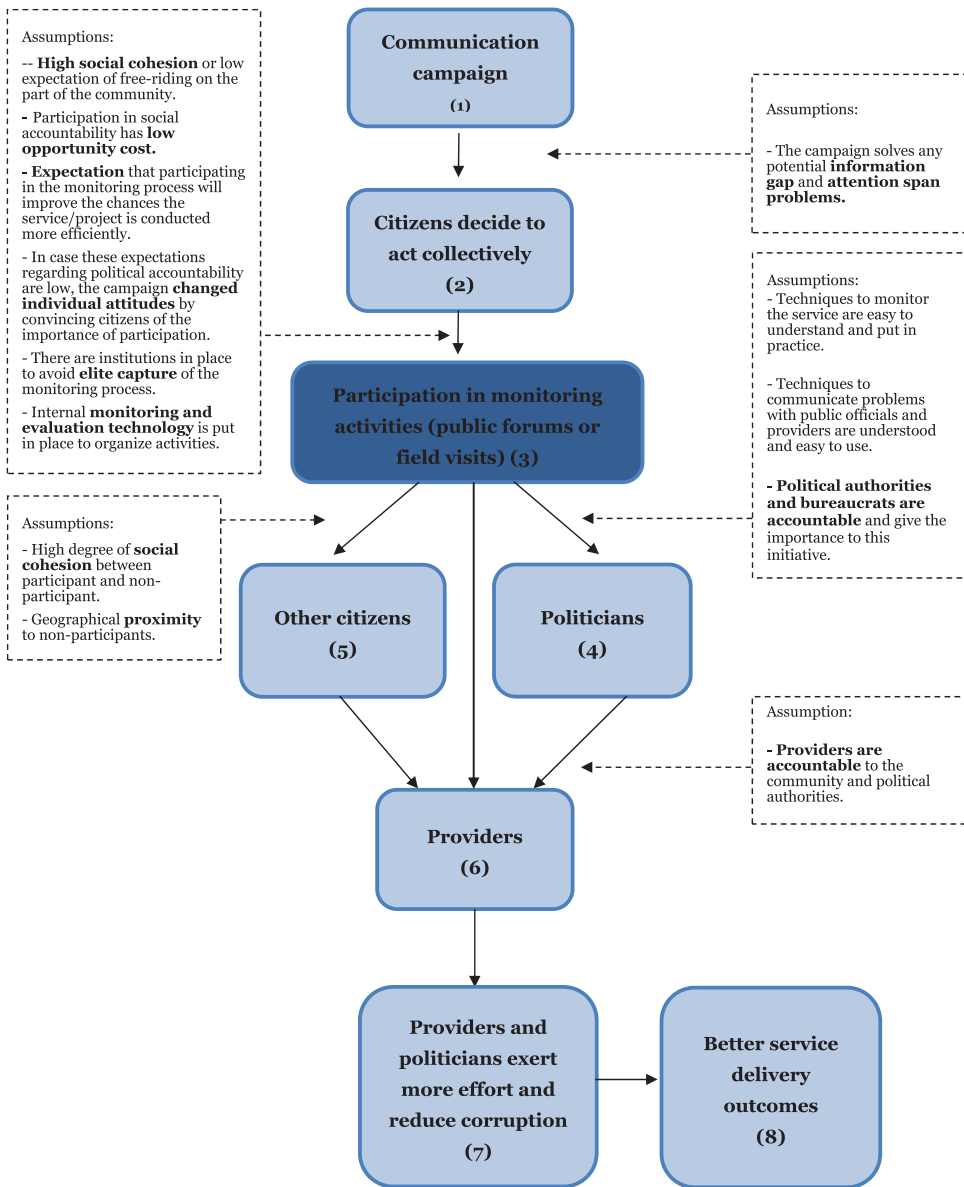


Figure A1. Revised causal chain with assumptions identified from evidence.