From the holdings of the Department of Special Collections , Princeton University Library, Princeton, NJ

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyrightlaw. We do not charge any permission or usage fees for the publication of images of material inour collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff.For more information regarding copyright, credit, and citation please review the information onthe Copyright, Credit and Citation Guidelines page of our website.
https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. https://library.princeton.edu/special-collections/ask-us

---

Doctoral Dissertations, 2011-2026 Public and International Affairs

2013

Essays on The Political Economy of Service Delivery

Molina, Ezequiel

Princeton, NJ : Princeton University

---

http://arks.princeton.edu/ark:/88435/dsp01pz50gw223

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. https://www.copyright.gov/title17/title17.pdf

Essays on the Political Economy of Service Delivery

Ezequiel Molina

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Woodrow Wilson School
Adviser: Carles Boix

November 2013

© Copyright by Ezequiel Molina, 2013.

All rights reserved.

Abstract

This dissertation comprises three articles that analyze the theoretical and empirical implications of social accountability reforms. In the first article, I argue that there is a channel, often forgotten in the literature, through which the effect of community monitoring interventions could be diminished. Citizens may refuse to take advantage of the opportunity to monitor the government and service providers if they believe that the chances to succeed are low. In particular, I model this argument formally and find that under certain conditions a case of ***self-fulfilling prophecy in service delivery*** can occur. This means that citizen's low expectations on leaders, officials, or service providers' responsiveness may generate low community participation in monitoring the project. This can reduce project visibility, as fewer citizens are able to disseminate information about the program to other citizens. As a result, if participation in monitoring activities and project visibility is low, policymakers and providers have fewer incentives to spend time and effort improving project performance, as their efforts will not be recognized. This reinforces citizens' prior beliefs and generates a vicious circle.

The second article makes three contributions to the literature. First, it provides new evidence of the impact of community monitoring interventions using a unique dataset from the Citizen Visible Audit (CVA) program in Colombia. In particular, this article studies the effect of social audits on citizen's assessment of service delivery performance. The second contribution is the introduction a theoretical framework to understand the pathway of change, the necessary building blocks that are needed for social audits to be effective. Using this framework, the third contribution of this article is answering the following questions: (a) under what conditions do citizens decide to monitor government activity and (b) under what conditions do governments

facilitate citizen engagement and become more accountable.

For the third article of my dissertation I conducted a synthetic review on community monitoring interventions in low-and-middle income countries. The article assesses under what conditions community monitoring interventions help to improve service delivery.

Acknowledgements

I would like to express my sincere gratitude to Carles Boix, Thomas Fujiwara and Leonard Wantchekon for their support, thoughtful advice and wisdom.

Additionally, I wish to thank the Research Program in Political Economy and Thomas Romer, its Director, for financing a trip to collect the necessary data to design the impact evaluation of 'Auditorias Visibles'. I also want to thank Edgardo Mosqueira who had faith in my ability to design and carry out the social audit evaluation. Also I would like to thank Arturo Herrera, as well as the LCSPS Team at the World Bank who financed the evaluation of 'Auditorias Visibles' and allowed me to use the data for my thesis. In addition, I want to thank the team working on the coordination of the Citizen's Visible Audit program at Direccion Nacional de Planeación (DNP) in Colombia. In the same spirit, I would like to thank the Research Department at the Inter-American Development Bank for their hospitality during 2012-2013. In particular, to Virginia Oliveros, Carlos Scartascini, Ernesto Stein, Ana Corbacho and Dan Gingerich. Finally, I would like to thank 3ie for supporting chapter 3 of my dissertation.

From the administration, I wish to extend my heartfelt gratitude to Karen McGuinness, Ann Lengyel and the whole team at Woodrow Wilson School for going beyond their call of duty to make my life easier during my time at Princeton.

I save for last the most important acknowledgement. I would like to thank my wife, Eliana, for her unconditional support. Without her, this journey would have never been possible. Finally, I thank my parents, Celia and Raul, to whom I owe everything.

To my wife Eliana and my parents, Celia and Raul.

Contents

Abstract iii

Acknowledgements v

1 A Theory of Self-Fulfilling Prophecy in Service Delivery 1

1.1 Introduction 1

1.2 Illustrative Model 6

2 Can Bottom-Up Institutional Reform Improve Service Delivery? 17

2.1 Introduction 17

2.2 CVA Program 26

2.3 Theoretical Framework 30

2.4 Identification Strategy 39

2.5 Empirical Results 41

2.5.1 CVA Program Impact on Project Performance 43

2.5.2 Uncovering CVA Program Mechanisms 48

2.6 Conclusions 78

3 Exploring External Validity 81

3.1 Introduction 81

3.2 Description of the Intervention 84

3.2.1 Relevance of this Review 89

3.3 How Can Community Monitoring be Fostered? 94

3.3.1 Competing Theories 98
3.4 Objectives of the Review 101
3.5 Criteria for Including Studies in the Review 102
3.5.1 Types of Studies 102
3.5.2 Types of Participants 103
3.5.3 Types of Interventions 103
3.5.4 Outcomes Measures 104
3.6 Search Strategy and Methods of Analysis 106
3.7 Search Results 108
3.7.1 Description of Included Studies 108
3.7.2 Study Designs 109
3.7.3 Population Locations 111
3.7.4 Participants 111
3.8 Intervention Characteristics 111
3.9 Empirical Results: Statistical Meta-Analysis of Effect Estimates . . . 113
3.10 Causal Chain Analysis 117
3.10.1 Does Facilitated Contact Increase Citizens Participation? . . . 120
3.10.2 Does Facilitated Contact Improve Providers' Responsiveness? 122
3.11 Conclusions 124

A Appendix **126**

Bibliography **129**

Chapter 1

A Theory of Self-Fulfilling Prophecy in Service Delivery

1.1 Introduction

Governments, NGOs and international development organizations have embraced the idea that the way to improve living standards in developing countries is to encourage citizens to apply bottom-up pressure on their government. As a result, billions of dollars have been spent on programs aimed at encouraging community to monitor service delivery (World-Bank, 2003). However, these interventions have not been as successful as expected (Banerjee et al., 2010; Devarajan, Khemani and Walton, 2011; Lieberman, Posner and Tsai, 2013; Mansuri and Rao, 2004, 2012; Olken, 2007) and the bottlenecks that prevent them from being successful have not been clearly identify in the literature.

The idea that the community is the residual claimant in this social contract and, as a result, its members have incentives to monitor providers and demand for better services (Stiglitz, 2002) led practitioners to believe that allowing communities to have

monitoring power over providers could be beneficial in the short and in the 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 change political norms, establishing a transparent and accessible channel of communication for the community to provide feedback to providers and politicians on a regular basis. However, for that prediction to be accurate, several assumptions are necessary, i.e. citizens need to participate in monitoring activities, and politicians and providers need to be accountable. Though the literature is still in its infancy, scholars have argued the existence of potential bottlenecks which may prevent community monitoring interventions from having an impact on service delivery.

(a) *Information Asymmetries*

Scholars and policymakers have long argued that programs often fell short of their expectations because of information problems (Ringold et al., 2012). In the case of the community monitoring interventions (CMIs) there are two important potential problems: (a) the information may not have been properly disseminated, and/or (b) the information on how to monitor the project was either not provided or not understood by the citizens. This is relevant as if citizens do not know the service is important or how it has performed, it will be impossible for them to monitor it. If citizens do not know how to monitor the project, they may lose interest in participating in monitoring activities. And even if citizens decide to participate, their effectiveness will be limited.

(b) *Lack of Attention Span or Rational Inattention*

Even if information is provided it may fail to have the anticipated outcome on citizens. A factor that conditions its success is what information is to be disclosed (content), and how it is to be presented (vehicle). For instance, Bertrand et al. (2010) find precisely this for a field experiment in South Africa, which

randomized the advertising content of a loan offer. They show that including a photo of an attractive woman increases loan demand by about as much as a 25% reduction in the interest rate. 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. For instance, in a CMI to improve educational quality in India, parents were more preoccupied with more immediate needs, children's scholarships and a midday meal program, than with questioning village education committees and local government officials about education (Khemani, 2007). Citizens could also face fatigue. For example, a field experiment in Indonesia found that community-based targeting of a cash transfer program worsened as the meeting progressed (Alatas et al., 2012).

The fact that citizens are unaware of the necessary information to monitor providers might be due to rational inattention. This theory holds that individuals have a limited capacity to pay attention, and must therefore choose what to pay attention to. Consequently, it might be optimum behavior for people not to pay attention to certain information, given that it is irrelevant and competes with other, relatively more important, information. As a consequence, introducing new information does not always lead to new beliefs or changes in behavior. In order for information to change behavior, it must win many battles for the user's attention before it can inform decision making. According to Lupia (2012), there are certain necessary conditions before information can improve civic understanding. For this to occur, disclosed information must be: (1) previously unknown by the individuals, (2) sufficiently relevant and reliable so that individuals pay attention and retain it in their memory, (3) incorporated into the individual's long-term memory, so that it can be recalled later, when necessary, and (4) utilized at the opportune moment for decision making.

(c) *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 (Ringold et al., 2012). Even if citizens have all the relevant information there is an opportunity cost of participation in monitoring activities. Members of poor households may not have time for participating in local governance because of more pressing priorities such as securing food and meeting other basic needs (Banerjee and Mullainathan, 2008). In addition, citizens may also refuse to participate if they believe the CMI is a tax on their time that undermines their ability to spend time in other equally or more important perceived activities such as working, relaxing and spending time with their family.

(c) *Collective Action Failure*

Scholars have emphasized the collective action problems that can arise in the presence of a non-excludable local public good (Olson, 1971), such as community monitoring. In other words, if community members believe other fellow citizens will contribute to monitor the project, they may decide not to participate.

(d) *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 program.

(e) *Politicians and Providers Accountability*

Under this heading I present bottlenecks that refer to a lack of responsiveness on the part of the politicians and providers. The literature has found many reasons why politicians and providers may not be accountable to their principals.

(1) 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.

(2) 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).

(3) 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.

(4) Finally, often the political system does not work properly and institutions do not help translate the preference of people into policy (Acemoglu and Robinson, 2008; Adsera, Boix and Payne, 2003).

This paper contributes to the literature by providing a formal model of another potential mechanism, often forgotten in the literature, that can reduce the effectiveness of CMI to improve service delivery. If community members believe politicians and providers will not be responsive they will have fewer incentives to monitor the project. In order words, citizens may refuse to take advantage of the opportunity to influence politicians and providers if they believe the chances of success are low. At the same time, if participation in monitoring activities and as a result, project visibility, is in fact low, politicians will have fewer incentives to spend time and effort improving project performance. This can potentially degenerate into a case where citizens decide not to participate in monitoring the project and providers

and politicians exert no effort. This is what I have named a case of ***self-fulfilling prophecy*** in service delivery.

1.2 Illustrative Model

This exploratory model builds on the literature on principal-agent relationships (Ferejohn, 1986; Persson and Tabellini, 2002). In particular, this model builds on Adsera, Boix and Payne (2003) and Persson and Tabellini (2002).

Consider a society inhabited by a large number (formally continuum) of citizens, where I normalize the size (mass) of the population to unity. These citizens are of different types indexes by i. Each type i has the same basic and quasi-linear preferences over private consumption c and publicly provided goods g, which is given by

$$w^i = c^i + H(g) \tag{1.1}$$

where $H(.)$ is a concave and increasing function. Implicit is the unrealistic assumption that government spending and taxes cannot be targeted to specific groups but instead must be provided in the same, nonnegative, amount to everyone: $g^i = g \geq 0$. Let's interpret g as the services provided by the government in education, health, public infrastructure, etc. Government spending is financed by taxing the income of every individual at a common rate τ, bounded by $0 \leq \tau \leq 1$.

Income, and consequently, consumption are

$$c^i = (1 - \tau)y^i \tag{1.2}$$

Since voter and candidate heterogeneity are unrelated to the argument here, I assume that income distribution is degenerated, with $y^i = y$ for all i. The government budget constrain is given by

$$\tau y = \theta g + r \tag{1.3}$$

where θ is random variable that represents the cost of transforming private output into public goods and is exogenously given. To simplify, θ can only take two values, one low and one high, such that $\theta = \{\underline{\theta}, \bar{\theta}\}$. A higher value of θ, ceteris paribus, implies a higher cost of transforming private input into public good.

Voters cannot observe θ, but can observe the resulting level of taxes and public expenditure. Furthermore, they can observe θ if they participate in monitoring service delivery at a cost of ψ^2.

The budget constraint makes clear that tax revenues can be used for government spending, but also can be used for "rent" for politicians. This can be conceptualized in a variety of ways, from party finance to outright diversion of resources for private use in connection with the production of public goods. Whatever the interpretation, I assume that r is nonnegative and bounded: $0 \leq r \leq \bar{r} = y$. This means that the only constraint on rent is the available tax base.

I assume a political incumbent value both the expected value of staying in office, $p_I R$, and the full discretion over current rents, r, introduced above. Thus, the objective function of the incumbent is

$$E(\upsilon_I) = \gamma r + p_I R \tag{1.4}$$

where E denotes the expectations operator, with expectation taking over their utility. The parameter γ measures the transaction costs associated with rent appropriation. This is an important concept in our analysis. This variable bundles many potential costs associated with rent extraction, i.e. party reputation cost and institutional checks and balances that prevent such appropriation. I assume that $0 \leq \gamma \leq 1$. The higher γ, the lower are the transaction costs for rent appropriation. p_I measures the probability of reelection of the incumbent and R the future rents. I also assume the contender is identical in all respects.

The model's timing is the following. I assume that (1) citizens decide to participate or not in monitoring service delivery. If they do, they incur in a cost ψ^2; (2) Citizens that participate observe the realization of θ and citizens that do not participate only observe public expenditure; (3) Citizens decide their voting strategy; (4) Politicians choose policy (g and τ); (5) Elections are held in which the voters choose between the incumbent and an opponent. The opponent running against the incumbent is identical in all respects from the viewpoint of the voters. Thus the only reason for not reappointing the incumbent is to punish him ex post, and since the opponent is identical, it is indeed (weakly) optimal for voters to carry out this punishment.

I assume that voters coordinate on the same retrospective voting strategy, punishing the incumbent for bad behavior and rewarding him for good behavior.[1]
Let's start by assuming citizens decide to participate in monitoring service delivery. The voting strategy boils down to setting the reelection probability p_I as follows:

$$p_I = \begin{cases} 1 & \text{if } W(g(\theta), r(\theta)) \geq \bar{\omega}(\theta) \\ 0 & \text{otherwise ,} \end{cases} \tag{1.5}$$

[1]Here I am assuming away the potential collective action problems.

where the voters' reservation utility $\bar{\omega}(\theta)$ is conditioned on the realized (and observable) θ.

This voting strategy creates a trade-off for the incumbent. When setting a policy at stage (3), he has two alternatives. The first one is to please voters to earn reelection. In this case he maximizes his rents subject to the constraint of generating $p_I = 1$. Solving this problem, using the definition of $W(.)$ and the government budget constraint, the optimally chosen rent become

$$r(\theta) = y - \bar{\omega}(\theta) + H(g^*(\theta)) - \theta g^*(\theta) \tag{1.6}$$

In other words, the incumbent satisfies the voters' demands by choosing g and τ in an optimal way contingent on θ, giving citizens exactly the utility they require to vote for his reelection, namely $W(g(\theta), r(\theta)) = \bar{\omega}(\theta)$. He keeps any remaining tax revenue as rent for himself. The incumbent's second alternative is not to satisfy the voters, thus, foregoing reelection. When deviating, the best policy is to follow the Leviathan-like policy, earning current rents corresponding to $r = y$. Then, the incumbent prefer to please the voters if

$$\gamma r(\theta) + R \geq \gamma y \tag{1.7}$$

$$r(\theta) = y - \frac{R}{\gamma} = r^* \tag{1.8}$$

As a result, voters will choose $\bar{\omega}(\theta)$ and r^* so that the incumbent please them and he is reelected. In particular,

$$p_I = \begin{cases} 1 & \text{if } \omega \geq \bar{\omega} \\ 0 & \text{otherwise .} \end{cases} \tag{1.9}$$

Since voters can observe θ, there are no information rents to be gained by the incumbent. The only equilibrium rents comes from the fact that he can credible threat voters with following Leviathan-like policies, $r^* = Max\{0, y - \frac{R}{\gamma}\}$. Having zero information rents comes at a price for voters, as they need to monitor service delivery at a cost ψ^2.

If voters do not participate in monitoring service delivery, they will not observe the realization of θ. In this case, in the favorable state (low θ), the elected incumbent can collect additional information rent by delivering public good efficiently, raising taxes and collecting high rents as if θ were high, when in fact it is low. Since voters do not observe θ, the best they can do is to choose a non-state-contingent cutoff level for their utility.

$$p_I = \begin{cases} 1 & \text{if } \omega \geq \bar{\omega} \\ 0 & \text{otherwise ,} \end{cases} \tag{1.10}$$

Faced with such a cutoff rule, the incumbent chooses just to satisfy the voters in order to gain reelection, when it is cheap to do so, namely when θ is low. But when θ is high, satisfying the voters becomes too expensive relative to exploiting his discretionary short-run power. Thus each level of $\bar{\omega}$ implies a critical state θ. In other words, policy is set such that

$$\omega = \begin{cases} \bar{\omega}(\theta^*) & \text{for } \theta \leq \theta^* \\ 0 & \text{for } \theta > \theta^*, \end{cases} \tag{1.11}$$

What is the best voting rule for the voters?

The best voting strategy is to condition the reelection on policy outcomes. In particular, since $\theta = \{\underline{\theta}, \bar{\theta}\}$ voters could require the optimal level of pub-

lic goods and taxes, both constant at the level associated with the cutoff stage $g(\theta) = g^*(\theta^*) = g^*(\bar{\theta})$, $\tau(\theta) = \tau^*(\theta^*) = \tau^*(\bar{\theta})$ or $g(\theta) = g^*(\theta^*) = g^*(\underline{\theta})$, $\tau(\theta) = \tau^*(\theta^*) = \tau^*(\underline{\theta})$.

With strategy (1) $g(\theta) = g^*(\theta^*) = g^*(\bar{\theta})$, $\tau(\theta) = \tau^*(\theta^*) = \tau^*(\bar{\theta})$, information rents are zero and total necessary rents for the incumbent to provide services are equal to the case when θ is observable if $\theta = \bar{\theta}$. If $\theta = \underline{\theta}$, information rents would be positive $(\bar{\theta} - \underline{\theta})g^*(\bar{\theta})$. However, in this case, total rent would be $Max = \{(\bar{\theta} - \underline{\theta})g^*(\bar{\theta}), r^* = y - \frac{R}{\gamma}\}$.

With strategy (2) $g(\theta) = g^*(\theta^*) = g^*(\underline{\theta})$, $\tau(\theta) = \tau^*(\theta^*) = \tau^*(\underline{\theta})$, necessary rents for the incumbent to provide services are equal to the case when θ is observable if $\theta = \underline{\theta}$, but if $\theta = \bar{\theta}$ there is no service provision which lowers voters' utility to zero. In this case rents would be equal to γy.

I assume that $\frac{Pr(\theta=\underline{\theta})}{1-Pr(\theta=\underline{\theta})} < \frac{\bar{\omega}(\theta)}{(\bar{\theta}-\underline{\theta})g^*(\bar{\theta})-(y-\frac{R}{\gamma})}$. As a direct consequence voters will choose strategy (1) and incumbents would please them and get reelected[2].

Then, will citizens participate in monitoring service delivery? That will depend on the benefits and cost of participation.

$$Participation = \begin{cases} \text{Benefits} & P_{\underline{\theta}}\Big[\Delta(\theta)g^*(\bar{\theta}) - (y - \frac{R}{\gamma})\Big] \\ \text{Cost} & \psi^2 \end{cases} \tag{1.12}$$

where $Pr(\theta = \underline{\theta}) = P_{\underline{\theta}}$ and $(\bar{\theta} - \underline{\theta}) = \Delta(\theta)$.

Citizens will decide to participate when benefits are larger than their cost

[2]Results do not qualitatively change if I choose exactly the opposite assumption

$$P_{\underline{\theta}}\Big[\Delta(\theta)g^*(\bar{\theta}) - (y - \frac{R}{\gamma})\Big] \geq \psi^2 \tag{1.13}$$

$$\Delta(\theta)g^*(\bar{\theta}) \geq \frac{\psi^2}{P_{\underline{\theta}}} + y - \frac{R}{\gamma} \tag{1.14}$$

(a) The higher the cost of participation in monitoring service delivery, ψ^2, the lower probability of participation. The cost of participation in monitoring, ψ^2, represents the physical, monetary and phycological cost of actually participating in monitoring the project.

(b) Higher information rent $\Delta(\theta)g^*(\bar{\theta})$ leads to higher probability of participation.

(c) Higher probability of a favorable state of the world $P_{\underline{\theta}}$ leads to higher probability of participation in monitoring, as monitoring becomes more valuable. This is contingent on $\Delta(\theta)g^*(\bar{\theta}) > (y - \frac{R}{\gamma})$. If $\Delta(\theta)g^*(\bar{\theta}) \leq (y - \frac{R}{\gamma})$, $P_{\underline{\theta}}$ would play no role in the decision of whether or not to participate in monitoring service delivery.

(d) The higher tax base y, the lower probability of participation. A higher y means that politicians have more power, as they can potentially keep more rent from citizens. As a result, their bargaining power is increased for all $\gamma > 0$.

(e) Higher present value of staying in office (R) means a higher probability of participation. This is the result of the fact that a higher R means that politicians bargaining power is reduced.

(f) The lower institutional strength of the society, a high value of γ, the lower the probability of participation monitoring service delivery. If γ is high, the value of identifying corruption and inefficiency is lowered, as politicians outside option is biding. As a result, politicians need not be accountable in order to get reelected.

In places where institutions are weak, the expected impact of the monitoring program will be lower.

The corollary from the analysis is that community monitoring will have less impact on environments where the institutional system does not punish those politicians that are found to deviate from the rule of law. This is explained by the fact that citizens need to tolerate high rents because politician can credibly commit to expropriate a large portion of people's taxes. Under that scenario, even after identifying, through community monitoring activities, deviations from the rule of law, there is no way of punishing politicians.

Until now, I have assumed voters have every necessary tool to anticipate politician's actions. In particular, they knew γ and R. With this information it was clear for the voters the exact benefits and cost of monitoring providers. Now I introduce an additional problem, as voters receive only a signal of γ. Remember that the parameter γ measures the transaction costs associated with rent appropriation. This variable bundles many potential costs associated with rent extraction, i.e. party reputation cost and institutional checks and balances that prevent such appropriation. As a result, the game has a new timing. For the sake of simplicity, I assume that the signal could only take one of two values $\{\underline{\gamma}, \bar{\gamma}\}$, where $\underline{\gamma} < \frac{R}{\frac{\psi^2}{P_{\underline{\theta}}}+y-\Delta(\theta)g^*(\bar{\theta})}$. This implies that transaction costs for rent appropriation are high and it is difficult for politicians to credibly commit to provide zero public good and keep the taxes. On the other hand, $\bar{\gamma} > \frac{R}{\frac{\psi^2}{P_{\underline{\theta}}}+y-\Delta(\theta)g^*(\bar{\theta})}$ implies the opposite. Transaction costs for rent appropriation are low.

This means that now there is uncertainty about γ. As a consequence, citizens do not know for certain the value of $y - \frac{R}{\gamma}$, which is the outside option of the politician. If citizens believe that outside option is high, the probability of monitoring to affect

the final outcome is reduced, ceteris paribus. Then, citizens will have fewer incentives to monitor service delivery.

The timing of the modified game is the following: I assume that (1) citizens receive the signal of γ, either $\{\underline{\gamma}, \bar{\gamma}\}$. This signal could be understood as a synthesis of citizens previous history with politicians' accountability.[3] The signal is public knowledge; (2) Citizens decide to participate or not in monitoring service delivery. If they do, they incur in a cost ψ^2; (3) Citizens that participate observe the realization of θ and citizens that do not participate only observe public expenditure; (4) Citizens decide their voting strategy. (5) Politicians choose policy (g and τ); (5) Elections are held in which the voters choose between the incumbent and an opponent. The opponent running against the incumbent is identical in all respects from the viewpoint of the voters. Thus the only reason for not reappointing the incumbent is to punish him ex post, and since the opponent is identical it is indeed (weakly) optimal for voters to carry out this punishment.

Solving this model, when citizens receive the signal of $\bar{\gamma}$ they will decide not to monitor politicians. This is easy to see from the equations above. For values of $\gamma > \frac{R}{\frac{\psi^2}{P_{\underline{\theta}}}+y-\Delta(\theta)g^*(\bar{\theta})}$ it is always better for citizens not to participate in monitoring service delivery. Note that I showed above that when politicians are not monitored, it is in their interest to act as if $\theta = \underline{\theta}$. Following the same logic they will also act as if $\gamma = \bar{\gamma}$. This assures them a higher payoff and sice voters will see exactly what they were expecting given the signal, they have no incentive not to reelect the government. As a result, I found that in this case there is a **self-fulfilling prophecy**. Citizens, based on their low expectation of providers responsiveness, decide not to monitor service delivery. At the same time, providers respond to citizens by providing the

[3] In this paper I will not deal with how those belief were formed or potential strategic consideration of the politicians that would be present in a repeated or a long run game.

lowest amount of public good possible, confirming citizens expectations.

When voters received a signal of $\underline{\gamma}$ the resulting model is very similar to the model I detailed above. Citizens will analyze which is their best strategy based on their benefits and cost.

$$Participation = \begin{cases} \text{Benefits} & P_{\underline{\theta}}\Big[\Delta(\theta)g^*(\bar{\theta}) - (y - \frac{R}{\underline{\gamma}})\Big] \\ \text{Cost} & \psi^2 \end{cases} \tag{1.15}$$

where $Pr(\theta = \underline{\theta}) = P_{\underline{\theta}}$ and $(\bar{\theta} - \underline{\theta}) = \Delta(\theta)$.

Citizens will decide to participate when benefits are larger than their cost of participation

$$P_{\underline{\theta}}\Big[\Delta(\theta)g^*(\bar{\theta}) - (\frac{R}{\underline{\gamma}})\Big] \geq \psi^2 \tag{1.16}$$

$$\Delta(\theta)g^*(\bar{\theta}) \geq \frac{\psi^2}{P_{\underline{\theta}}} + y - \frac{R}{\underline{\gamma}} \tag{1.17}$$

Again, if citizens decide not to participate, politicians would follow the same strategy as I detailed above in the first model.

The conclusion from this very simple and illustrative model is that in environments where citizens' perceived government accountability to be weaker, the impact of the CMIs interventions will be weaker, driven by lack of community participation. Note again that this is independent of the actual value of γ. This means that there exists an equilibrium with self-fulfilling prophecies. If citizens believe γ is high, they will

not monitor service delivery. As a direct consequence, politicians will take advantage and will provide less public good.

Chapter 2

Can Bottom-Up Institutional Reform Improve Service Delivery?

2.1 Introduction

The evidence is clear. Public service delivery in developing countries is highly inadequate. In India, one quarter of all the teachers in public schools and more than a third of nurses and doctors were absent when enumerators arrived at their respective schools and hospitals (Chaudhury et al., 2006). Not only teachers, nurses, and doctors but also public resources are "missing in action". A landmark study in Uganda found that only 13 percent of public funds assigned by the central government to the school system reached its intended destination (Reinikka and Svensson, 2004, 2011). In addition, leakages are a problem in Tanzania, where elected officials receive more than half of the total amount of the price vouchers for subsidized fertilizer (Pan and Christiaensen, 2012). Further examples of corruption and inefficient allocation of resources in service delivery abound in underdeveloped countries (World-Bank, 2003). Crucial resources required to increase opportunities for citizens, improve infrastructure, and reduce the learning gap are lost every day because of pockets of

inefficiency and corruption.

Scholars have argued that accountability failure is the fundamental problem giving rise to inefficient delivery of basic services, such as the abovementioned teacher and nurses absenteeism, leakage, and missing public funds. The argument is succinctly explained by Devarajan, Khemani and Walton (2011). In a private market transaction (e.g., paying for a haircut), the provider has direct accountability to the client or consumer. The client pays the vendor directly; he or she can observe not only the haircut being received but also the quality of the service. Furthermore, if the market is reasonably competitive, the client can go elsewhere to get a haircut, and both the salon and hairdressers know that. The 2004 World Development Report named this relationship "the short route of accountability." However, for markets characterized by natural monopolies such as the provision of electricity and sanitation as well as other strategic markets such as education and health, the government is involved in service provision. As a result, the relationship between the client and service provider is no longer a direct one, but an indirect one, intermediated by the state. Citizens, in their role as voters, hold politicians accountable for allocating resources and regulating these service provisions. In turn, politicians and policymakers rely on managers, who in turn rely on front-line providers such as teachers and policemen to deliver the service. This is the "long route of accountability."

Problems arise because the formal institutions embedded in the "long route of accountability" are often weak in developing countries (Bardhan, 2002). Patronage, lack of state capacity, or even weak electoral institutions may prevent citizens from holding officials accountable for the quality of service provision (Adsera, Boix and Payne, 2003; Besley and Persson, 2011; Persson and Tabellini, 2002).

To address these challenges, many policies such as information campaigns, scorecards, social audits, and grievance redress mechanisms have been proposed to increase beneficiariesâĂŹ involvement in service provision. Table 2.1 describes these policies. The idea that community members have incentives to monitor providers and demand better services (Stiglitz, 2002) has led practitioners to believe that allowing communities to have the power to monitor providers could be beneficial in both the short and long terms. In the short term, it could improve outcomes by identifying pockets of corruption and inefficiency in service provision. In the long term, it may contribute to changes in political norms by establishing transparent and accessible channels of communication for the community to provide feedback to providers and politicians on a regular basis.

Table 2.1: Policies Aimed at Increasing Civic Participation in Monitoring Public Officials and Providers

Policy	**Description**
Information Campaign	These campaigns include a) information on how to monitor providers and b) information on the performance of providers and/or the importance of the service.
Scorecard	This is a quantitative survey that assesses 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 cross-check the information the provider makes available against information collected from users of the service (Ringold et al., 2012).

(Continued on next page)

Table 2.1: (continued)

Policies	Description
Grievance Redress Mechanisms	These are mechanisms that provide citizens with opportunities to use information to influence service delivery and provide feedback on government programs and services, including complaint hotlines, informal dispute resolution mechanisms, and courts (Ringold et al., 2012).

Determine whether such programs actually affect the behavior of beneficiaries, providers and politicians and in turn improve outcomes is an open empirical question[1] (Mansuri and Rao, 2004). In addition, we still know little about the barriers and facilitators through which these interventions have an impact (or lack thereof).

This study aims to address this gap. In particular, using a unique dataset from the Citizen Visible Audit (CVA) program in Colombia, the study provides new evidence regarding the effect of social audits on citizens' satisfaction with infrastructure projects as well as subjective measures of the efficiency of the execution process. This study introduces a theoretical framework to understand the pathway of change, the necessary building blocks required for social audits to be effective. This approach is often ignored in literature.[2] Using this framework, this study provides an answer to both (a) under what conditions citizens decide to monitor government activity and (b) under what conditions governments facilitate citizen engagement and become more accountable.

[1] See chapter 3 for a synthetic review on whether community monitoring interventions help improve outcomes and curb corruption in service delivery in low-income communities.

[2] One notable exception is Lieberman, Posner and Tsai (2013), who develop a causal chain for information campaign interventions.

The answers to these questions are important for both practitioners and academics.

Despite the lack of rigorous evidence, local participatory and community monitoring interventions are widely accepted by all major players in the practitioner world, such as governments, non-governmental organizations (NGOs), and the donor community, as a tool to improve transparency and accountability. Increasing citizen participation in government is the main objective behind the Open Government Partnership (OGP), a global consortium of governments. Through OGP, more than 50 countries have already agreed to a range of goals related to transparency and citizen participation. International aid agencies increasingly require development projects to include "beneficiary participation" components. In the past decade, the World Bank alone has channeled $85 billion to local participatory development initiatives (Mansuri and Rao, 2012). NGOs with a focus on increasing government accountability through citizen participation continue to expand around the globe, and these manage increasing amounts of resources. For instance, Twaweza in East Africa has an annual budget of $17 million and engages in building citizen capacity to make governments accountable across the region.

From a theoretical perspective, there is no clear prediction about the impact of these programs. On one hand, some authors have identified reasons to expect community monitoring interventions to have a positive effect on improving service delivery and reducing corruption:

(a) Community members have better incentives than bureaucrats to monitor a program that may improve their quality of life (Stiglitz, 2002).

(b) Community scrutiny and monitoring may alter the incentives of representatives and providers (i.e. bureaucrats and firms) either through reputational repercussions or simply awareness of being observed (Ringold et al., 2012).

(c) Community gatherings bring together politicians, bureaucrats, firms, and citizens in a manner that may generate a closer connection between the representatives and community. This in turn may affect politicians' information sets as understanding their communities' preferences and assessing the performance of service providers' becomes less costly. (Molina, 2013*b*).

On the other hand, some authors have argued that successful implementation of community monitoring interventions might prove more difficult than expected:

(a) Monitoring public projects is a public good, so there may be a free-rider problem (Olson, 1971).

(b) Community monitoring may also be prone to being captured by local elites (Bardhan, 2002; Bardhan and Mookherjee, 2006; Olken, 2007).

(c) Citizens' beliefs regarding the probability of success of the community monitoring exercise may affect their participation rate, which in turn, affects the program's visibility. When public officials and providers observe weak community monitoring and low visibility, their incentives to spend time and effort on improving service delivery are diminished (See chapter 1).

Furthermore, empirical studies have not been able to provide a clear answer regarding the impact of community monitoring programs. Table 2.2 summarizes various evaluations of community monitoring interventions.

Evaluations have found what at first appears to be conflicting evidence regarding the effect of community monitoring schemes on service delivery outcomes. Björkman

and Svensson (2009) found that use of community scorecards in Uganda significantly increased the quality and quantity of primary health care provision. This evidence suggests that the primary causal mechanism behind this positive outcome is providers' behavioral change.

Banerjee et al. (2010), however, found the opposite result when testing the effect of an information campaign in India. They reported that neither giving citizens information on how to use existing institutions to monitor schools nor training them in a testing tool to monitor children's learning had any statistical impact on children's learning performance. Only the intervention not involving government action, training volunteers to help children learn to read, had a positive effect on children's reading skills (3%-8%).

Table 2.2: Evaluations of Community Monitoring Interventions

Author/s	**Country**	**Identification Strategy**	**Intervention**	**Results**
Banerjee et al. (2010)	India	Randomized Control Trial (RCT)	Empower parents in three ways: a) providing information, b) providing a testing tool for children, and c) training volunteers to help teach children.	A) and b) has no impact on community involvement or learning outcomes. c) improved children's reading skills (3%-8%).

(Continued on next page)

Table 2.2: (continued)

Author/s	Country	Identification Strategy	Intervention	Results
Björkman and Svensson (2009)	Uganda	RCT	NGOs encouraged communities - to be more involved with the state of health service provision and strengthened their capacity - to hold them accountable.	Increased monitoring, increased effort from service providers. Significant reduction in mortality and increased child weight.
Lieberman, Posner and Tsai (2013)	Kenya	RCT	NGOs encouraged individuals - to be more involved with the-state-of-children's learning and strengthened their capacity to improve performance.	No impact on private or collective actionable measures.
Olken (2007)	Indonesia	RCT	a)Increasing probability of top down audit and b)additional grassroots participation	a)Reduced corruption by 8%. b)Reduced corruption by less than 3%.

Similar to Banerjee et al. (2010), Lieberman, Posner and Tsai (2013) analyzed *Uwezo*, a program in Kenya that provides parents with information on their children's reading and numeracy ability as well as information on how to become involved in their children's learning. The results indicate that the information did not generate any improvement in either citizens' private actions intended to improve their children's performance or communities' collective action to demand improvements from their representatives and service providers.

Only Olken (2007) has assessed the impact of a social audit, although not a traditional one. The author studied the effect of grassroots participation in In-

donesia. This seminal contribution analyzed the effect of additional participation in the community monitoring process on unaccounted expenditure in infrastructure projects funds. In Olken's study, invitations to community meetings were randomly distributed throughout the village to encourage participation in the monitoring process. By sending a formal invitation to citizens in randomly selected villages, Olken (2007) studied the impact of having more attendees and higher representation in the community meetings on unaccounted expenditures. The author found that increasing grassroots participation had little impact, reducing unaccounted expenditures only in situations with limited free-rider problems and limited elite capture.

This study makes three contributions to literature. First, it provides new evidence of the impact of community monitoring interventions using a unique dataset from the CVA program in Colombia. In contrast with Olken (2007), where the community had the ability to manage the funds for a given project, in the present study, I assess the impact of a program aimed at empowering citizens through increasing the opportunity to hold providers and local governments accountable and not a group of peers. This means that citizens do not have decision-making power, but can only monitor infrastructure projects. This reduces the problems related to elite capture that may arise when a community has decision-making power over spending public resources. The second contribution is the introduction of a theoretical framework to understand the pathway of change in the CVA program. Using this framework, the third contribution of this study is to provide an answer to both (a) under what conditions citizens decide to monitor government activity and (b) under what conditions governments facilitate citizen engagement and become more accountable.

Finally, to the best of my knowledge, this is the first study to evaluate the impact of a community monitoring program administered by a government instead of an

NGO. This increases the external validity of the results as NGOs do not have neither the logistical power nor the mandate to administer a program if scaling it up becomes necessary.

The remainder of the paper is organized as follows. The next section provides background information on the CVA program. Section 3 introduces the pathway of change and section 4 describes the identification strategy. Section 5 presents the impact of the CVA program on project performance. In addition, using the framework introduced in section 3, section 5 provides evidence of (a) under what conditions citizens decide to monitor government activity and (b) under what conditions governments facilitate citizen engagement and become more accountable. The paper ends with the conclusions.

2.2 CVA Program

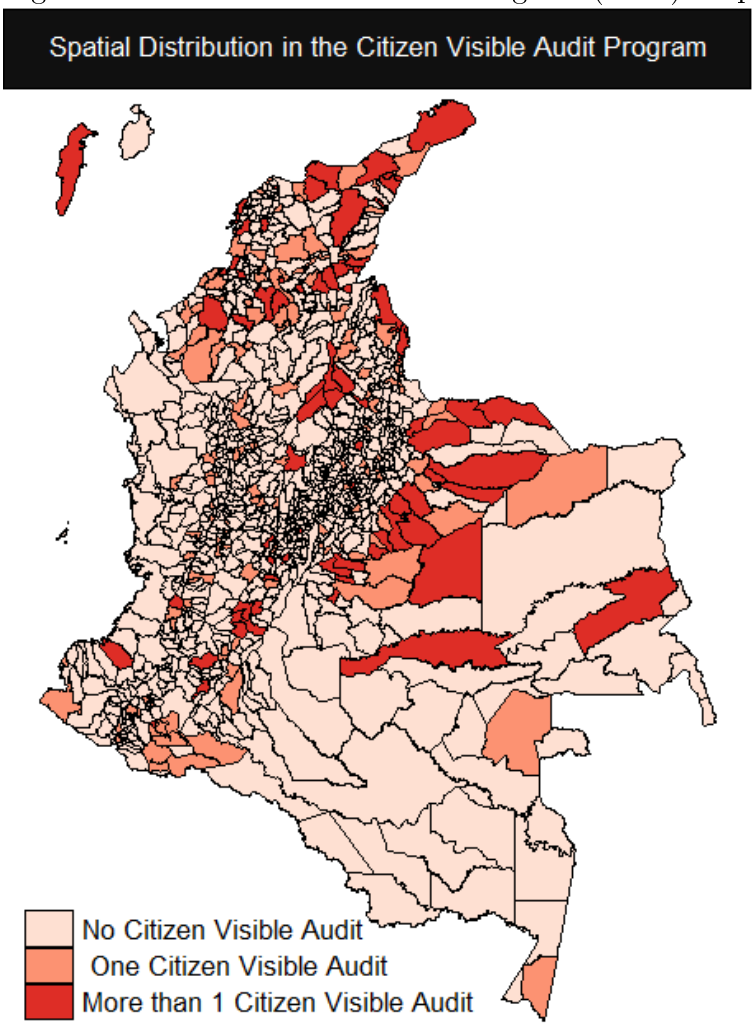
The Colombia's constitution and laws require mining royalties to be transferred to sub-national governments and allocated to public works in areas such as education, health, nutrition, and water. To prevent corruption in the use of royalties, the Colombian Anti-Corruption Presidential Commission launched the Citizen Visible Audit (CVA) program in 2008 to promote transparency and citizen participation in the management of public investments. To date, the CVA program has been implemented in over 400 communities across Colombia. Furthermore, since 2012, the CVA program has been a central piece of the Colombian government's agreement with OPG.

The selection process of specific projects to be audited by the community is based on the amount of royalties received, the number of previous irregularities in royalty

management, and the project's expected social impact. Therefore, CVA tends to focus on infrastructure projects designed to help the provision of local public services (i.e. access to water and sanitation, education, energy, and housing) for a relatively small group of people living in the nearby community. Almost 40% of the projects were related to access to water and sanitation. Another 35% were devoted to building educational infrastructure and housing. The 'average' project that underwent a CVA lasted 335 days, had 2 community forums, with 50 citizens participating in each forum. In addition, the amount spent on each project was, on average, 6240 million Colombian pesos (around $3.5 million), with more than 80% financed by royalties.

Figure 2.1 shows the spatial distribution of the CVA program.

Figure 2.1: Citizen Visible Audit Program (CVA) Map



Notes: This is my own elaboration based on data from *Direccion Nacional de Planeación* (DNP), Colombia. The figure shows the number of CVAs by municipality.

Once a specific project is selected in a given municipality, the CVA program consists of the following activities:

(a) Disseminating information about the CVA program in the neighborhoods where the project takes place using radio, newspapers, invitations, and local TV.

(b) During the first public forum, the infrastructure project is introduced to the community. Citizens are informed of their rights and entitlements, including activities they can carry out to monitor the project and responsibilities of the executing firm. A group of beneficiaries composed of interested citizens is constituted and trained to carry out community monitoring activities.

(c) Periodic public forums are held, bringing together local authorities, neighbors, and representatives from the firm carrying out the specific project. In these public forums, the projectâĂŹs current status is explained in detail to the community, which in turn might provide its suggestions and recommendations. Commitments are made by the firm, the local government, and project supervisor to solve the problems that arise during construction of the project. These commitments are monitored by the community, facilitators from the central government (DNP), and project supervisor. If a commitment is not honored, facilitators and supervisors intervene to inform the local government. If the problem persists, administrative complaints are submitted to the Supreme Audit Body in the central administration.

(d) In between public forums, the beneficiary group monitors the project and collects information on whether commitments are being honored as well as on any other new problems that may arise.

(e) Before making final payment to the executing firm, the finalized project is presented to the community. The audit results are shared with all interested and concerned stake-holders.

2.3 Theoretical Framework

Figure 2.2 presents a stylized pathway of change for the CVA program, clarifying the mechanisms through which the CVA program is expected to impact project performance.

The CVA program begins by attempting to increase local knowledge of both the focus project and monitor program. A communication campaign (building block 1) through radio, newspapers, and local TV is launched to increase participation in the first public forum. This first contact with the community allows facilitators to provide citizens with information on the project, such as what the contract stipulates that the firm has to deliver, how to monitor the executing firm, and to whom they should direct enquiries about deficiencies in the project. This is described in Figure 2.2 as capacity building (building block 2). Empowered with this new information, citizens are expected to solve the collective action problem and invest their time and effort monitoring the project (building block 3). Participation in monitoring activities could take many forms. As an organized group, citizens can take turns to visit the construction site and collect information on problems. They can also use their voice to contact the project supervisor, the representatives of the firm (building block 6), or elected officials (building block 4) to complain about the project and/or provide valuable information on specific problems facing the project. Citizens can even condition their votes on project performance or threaten governments with expensive protests. In addition, citizens can share information collected on the project with neighbors who did not take part in monitoring activities (building block 5) in order to increase visibility of the CVA program. Finally, citizens may also contact or threaten to contact opposition party leaders and the media in order to incentivize elected officials, firms and project supervisors to act.

In addition, the CVA program provides public forums where project supervisors meet with representatives acting on behalf of local government, the executing firm, and the central government. This allows citizens to make their voices heard by supervisors as well as local officials and providers and reduces the time and effort citizens would need to invest to get an appointment with these officials. Moreover, public forums reduce the cost of local officials meeting with supervisors and central government representatives. In addition, the costs for local officials to take action to solve problems arising during the construction of infrastructure projects, such as lack of planning, lack of resources to finish the project, and acts of corruption, are reduced. Finally, the public forum may signal to politicians and the executing firm of the importance of performing well on this project, as citizens are paying extra attention.

For the CVA program to generate a change in politicians' behavior at least one of the following conditions must be present:

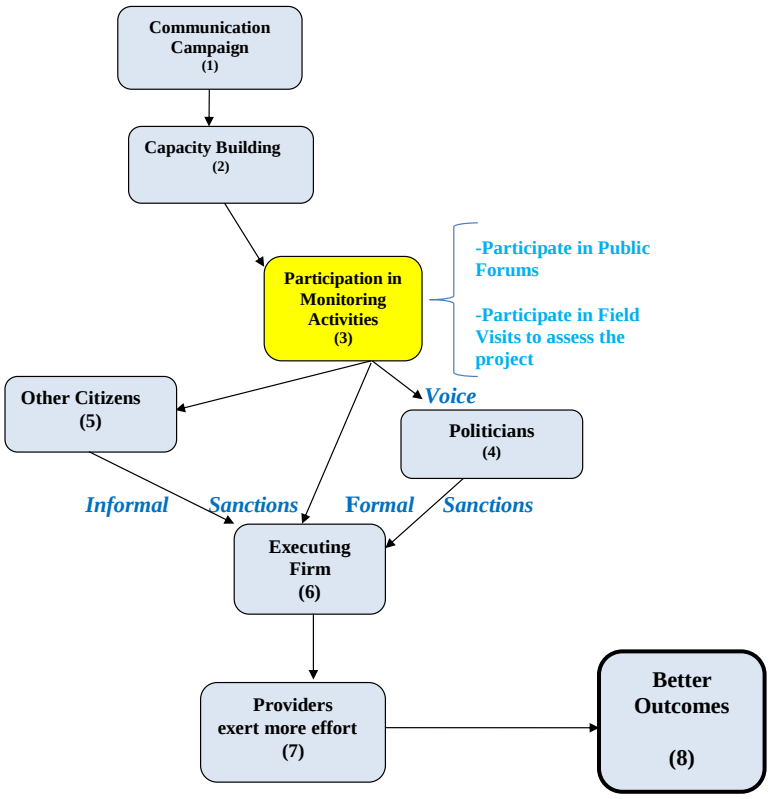
(i) It becomes easier for politicians to identify pockets of corruption and inefficiency due to increase citizen's monitoring effort.

(ii) It becomes more valuable for politicians to improve project performance, as citizens are better able to recognize whether elected officials are making an effort to improve project performance.

These conditions provide incentives for politicians and policymakers to put more pressure on providers to improve project performance and to allocate more resources, if necessary. Furthermore, politicians could enact new laws to change the way firms are selected and projects are executed.

Facing formal sanctions by politicians and/or informal sanctions by citizens, providers are more likely to exert effort (building block 7). As a result, the CVA program may improve project performance (building block 8).

Figure 2.2: Pathway of Change for the CVA Program



Notes: This figure shows the mechanism through which the CVA program is expected to impact project performance.

However, for this path of change to accurately describe the mechanisms through which the CVA program impacts project performance, several assumptions are necessary: citizens need to have adequate information on how to monitor the project, they should be able to pay the opportunity cost and coordinate their actions to monitor the project, they should believe the program has the potential to be successful,

and politicians and providers need to be accountable. When these assumptions are not met, the program suffers from bottlenecks. Those bottlenecks may prevent the CVA program from having an impact on project performance. On the basis of these bottlenecks, there are four alternative theories to explain why the program might not result in the expected outcome.

The first two theories, deficient implementation theory and civic participation failure, refer to bottlenecks that prevent citizens from actively engaging in monitoring the project. The third theory, non-accountable governments, refers to a lack of response on the part of the politicians and providers and is independent of whether civic participation in monitoring the project takes place. Finally, there is the self-fulfilling prophecy theory that consider both sides of the CVA's pathway of change: citizens' decisions to participate in project monitoring and the government's willingness to facilitate that engagement and become accountable. Below I present each theory, its empirical implications and the measurement strategy.

(a) ***Deficient Implementation Theory:*** Scholars and policymakers have long argued that programs often fall short of their expectations because of implementation failures (for a recent review see Pritchett, Woolcock and Andrews, 2013). The CVA program faces two potential deficiencies: (a.1) the information on the project may not have been properly disseminated (building block 1), and/or (a.2) information on how to monitor the project was either not provided or not understood by citizens (building block 2). This is relevant as if citizens are unaware that a project exists (a.1), citizen monitor will be impossible. If citizens do not know how to monitor the project (a.2), they may lose interest in participating in monitoring activities. Even if citizens do decide to participate,

their effectiveness will be limited.

There are two main empirical implications for the case of the CVA program:

(a.1) If information on the project is not properly disseminated, citizens will not participate in monitoring activities.

(a.2) The probability of citizen participation in monitoring activities will be a function of the quality of the information they possess.

I test (a.1) by focusing on the percentage of citizens in each community who are informed about the project's existence, and test (a.2) by measuring the percentage of citizens in each community that evaluate the quality of information received on the project as being adequate or very adequate.

(b) ***Civic Participation Failure Theory:*** One potential concern with the CVA program, and social accountability mechanisms in general, is citizens' failure to participate in monitoring the project. Under this heading, I group many bottlenecks that prevent citizens from participating in monitoring activities, which in turn reduces the potential impact of the program. In particular, if community monitoring activities are not carried out at all or are only carried out by a few citizens, their ability to uncover problems and pressure the government for accountability is significantly reduced. I identify four potential bottlenecks. (b.1) Citizens may not believe the infrastructure project is a priority for the community. (b.2) Even if the project is a community priority, opportunity costs may prevent citizens' participation. In particular, this can occur if citizens believe the CVA program is a tax on their time that undermines their ability to engage in other activities perceived as equally or more important, such as working, relaxing and spending time with their family. (b.3) If citizens were already

monitoring the project in the absence of the CVA program, formal participation in the CVA program becomes irrelevant. (b.4) Scholars have emphasized the collective action problems that can arise in the presence of a non-excludable local public good (Olson, 1971), such as monitoring a project that would benefit the community. In other words, if community members believe other fellow citizens will monitor the project, they may decide not to participate[3]. The literature has found that collective action problems of the type described above can be ameliorated when citizens expect others to reciprocate for any cooperative behavior they perform (Putnam, 1993; Svendsen and Svendsen, 2009). As a result, proxies for citizens' expectation of reciprocating cooperative behavior, often denoted as social capital, are used to study collective action problems. The empirical implications of these bottlenecks are as follows:

(b.1) If citizens perceive the objective of the project as irrelevant for the community, they will not participate in community monitoring activities.

(b.2) If the opportunity cost of time is high, participants in community monitoring activities will have a higher probability of being individuals who are inactive, unemployed, or with lower incomes.

(b.3) If the types of actions citizens are encouraged to take as a result of participation in the CVA program are not different from those they are already doing, they will not participate in community monitoring activities.

(b.4) If citizens expect that other citizens will free-ride on their efforts to monitor the project, the probability and intensity of participation will fall.

To test (b.1), I estimate the percentage of citizens that believe the project is a priority for the community. Regarding (b.2), I test whether CVA participants

[3]For an in-depth analysis of coordination problems as well as potential solutions in the context of civic participation see Bardhan and Mookherjee (2006); Mansuri and Rao (2012).

are statistically different from non participants in the following characteristics: employment status and income level. In the case of (b.3), I estimate the difference in participation rates and time spent in community monitoring activities for both groups, i.e. citizens in the control group and citizens in the treated group who participated in at least one community forum of the CVA program. To test (b.4), I use several measures. Because there is no ideal proxy for citizens' expectations of reciprocal cooperative behavior, following the literature I use willingness to trust neighbors as a proxy for social capital. In addition, I use fractionalization indexes (Alesina and Wacziarg, 2003) as predictors of a lack of social capital. In particular, I measure ethnicity, religion, and language fractionalization by community, and test whether these measures can contribute to explain variations in the outcome of interest, time spent in monitoring activities in each community. Finally, I use *ethnic outsider*, a measure suggested by Lieberman, Posner and Tsai (2013). This is a dummy variable equal to 1 if a respondent's ethnicity differs from that of the majority of respondents in that community, and 0 otherwise.

(c) ***Nonaccountable Governments Theory:*** The literature has identified many reasons why politicians and, in turn, their agents, the providers, may not be accountable to principals. Citizens may not be pivotal for politician's electoral strategy (Downs, 1957; Hotelling, 1929; Persson and Tabellini, 2002) or the political system does not work properly so institutions do not help translate citizens' preferences into policy (Acemoglu and Robinson, 2008; Adsera, Boix and Payne, 2003). Another potential problem could be that politicians and providers might change their behavior as a result of the intervention but not in the desired way.[4] The empirical implication is that if governments are not

[4]Gavazza and Lizzeri (2007) as well as Prat (2005) provide several examples that might be understood from this perspective. For example, they point to transparency policies aimed at disclosing how much time it takes for government agencies to carry out the simplest procedure, such as provid-

accountable to their citizens, I should find that the performance of politicians and providers does not increase as a result of the CVA program. I test the non-accountability theory by estimating the impact of the CVA program on citizens' evaluations of politicians and providers performance on the project.

(d) ***Self-Fulfilling Prophecy Theory:*** In chapter 1 I argue that there is a channel, often forgotten in literature, through which the effect of a community monitor intervention could be diminished. If community members believe that after paying the opportunity cost of monitoring, coordinating their actions with fellow neighbors, and identifying a projects' problems, politicians and providers will not be responsive, citizens may not have an incentive to monitor the project in the first place. In order words, citizens may refuse to take advantage of the opportunity to influence government and providers if they believe the chances of success are low. In particular, I argue that under certain circumstances a case of ***self-fulfilling prophecy*** can occur.

The timing of the theory is as follows:

(Part I) In the beginning, citizens' low expectations regarding the accountability of leaders, officials, or service providers' may generate low levels of community participation in monitoring the project. As a result, project visibility is reduced because fewer citizens are disseminating the project's information to fellow citizens.

ing a driver's license. The problem with this information campaign that encourages individuals to use the most efficient agency is that it can generate excess demand without providing a commensurate change in the public employee compensation system. Thus, anticipating this, the staff of these agencies might reduce their efficiency once the policy is put in place. In general, therefore, in those markets where there are price rigidities (in this case, public sector salaries), transparency policies can be counterproductive.

(Part II) Thereafter, politicians observe community monitoring participation and project visibility. If participation in monitoring activities and project visibility is low, politicians have fewer incentives to spend time and effort improving project performance.

Note that this theory is the only one to engage with both sides of the pathway of change, i.e., the conditions under which citizens decide to monitor the project and under which politicians react by exerting more effort and improving performance.

The empirical implications are as follows:

(d.1) Citizens who perceive themselves as influential have a higher a probability of participating in the CVA public forums and community monitoring activities.

(d.2) Citizens who participate in the CVA public forums disseminate CVA information to other non-participating citizens.

(d.3) In projects receiving higher participation in community monitoring activities and higher visibility, politicians' performance is higher.

To test the empirical validity of these claims I follow the structure of the theory. For part I, I estimate a model at the individual level for the effect of citizens' perception regarding their influence on the time they devoted to monitoring the project and their participation in the CVA program's public forums. Thereafter, I test whether citizens who participate in the CVA public forums disseminate that information to fellow citizens who did not participate in the public forums. To test part II, I estimate a model of the percentage of citizens who are informed

about the project and the time spent monitoring the project at the community level on politicians's and project's performance.

2.4 Identification Strategy

To assess the impact of the CVA project on citizens' satisfaction performance and explore its determinants, this study conducts a retrospective evaluation and uses indicators derived from a household survey instrument. For each project with the CVA program I looked for similar projects without the program. In particular, I looked for projects within the same sector (education, health, and water and sanitation), having similarities in terms of spatial concentrations of population, initial estimated project timelines, and similar resources. Moreover, I selected projects conducted in non-contiguous communities from the same municipality in order to guarantee the same administrative procedures and level of responsibility from the local government.

Using this methodology, I found matches for 10 CVA projects out of a pool of 400 CVA projects. I expanded the search for similar projects in similar municipalities to add three additional pairs to the final sample. I created a propensity score of municipalities using the following pretreatment variables: royalties (measured in US dollars), irregularity in the management of royalties, population size, percentage of citizens with access to water and sanitation, and primary and secondary school enrollment. In the annex there is a description of each project that was surveyed and its sector, object, municipality, amount of funds and estimated completion time frame at the beginning of the project.

To understand the effect of the CVA program and its mechanisms two different random samples were collected: (a) a sample of individuals from both treated and

control projects that may or may not have participated in community monitoring activities and (b) a sample of participants in the public forums.

For (a) I used a household survey of 28 infrastructure projects, 13 that were "treated" with the CVA program and 15 that were "control projects". Each project was located on the cartographical map and the surrounding areas were sampled randomly. For each of the 13 projects, the random sample contains 30 households in the treatment group and 11 in the control group. For the two CVA projects having two controls each, each sample contains 20 households. The total sample contains 390 treated and 410 control households.

For (b), the contact information collected for each community forum by each CVA project is used. I use a random sample of 10 participants in each of the 13 treated projects.

To explore whether the matching resulted in a balanced set of respondents in the two treatment conditions, a wide range of variables revealing the social and economic characteristics of the respondents were captured, the summary statistics of which can be observed in Table 2.3. I investigate whether the populations from the treatment and control groups are statistically different in terms of observable characteristics that may affect their survey answers.

I present the results in Table[5] 2.3. While respondents from treatment and control communities are not statistically different in most characteristics, two differences are statistically significant. The population interviewed as part of the treatment group is, on average, older that the population interviewed in the control group. In addition, the population in the treatment group has a lower percentage of individuals who have received a post-secondary education. I use this information to control for those characteristics as well as other variables in the estimations of the program's impact.

2.5 Empirical Results

This section discusses the main findings. The first subsection presents the results of the impact of the CVA program on measures of satisfaction with the project and subjective measures of the efficiency of the execution process. Using the pathway of change developed in section 2.3, the second subsection examines the mechanism behind the CVA program's functioning. In particular, I examine the empirical validity of the four alternative theories described in the previous section: deficient implementation, civic participation failure, non-accountable governments and self-fulfilling prophecy. Examining these theories allows me to provide an answer to both (a) under what conditions citizens decide to monitor government activity and (b) under what conditions governments facilitate citizen engagement and become more accountable.

[5] *Asset Index* is a variable created using principal components. I use a list of 12 assets: television, refrigerator, landline telephone, cellular phone, automobile, washing machine, microwave oven, motorcycle, indoor plumbing, indoor bathroom, computer, and flat panel TV. The principal component yields an asset index that assigns a larger weight to assets that vary the most across households. *Political News (Daily)* is a binary variable equal to 1 when respondents' answer *daily* to the question: 'How often do you pay attention to the news, whether on TV, the radio, newspapers or the internet?'. *Political Knowledge (Departments)* is a binary variable equal to 1 when respondents' provide a correct answer to the question: 'How many departments does the country have?'. *Political Knowledge (Presidential Period)* is a binary variable equal to 1 when respondents' provide a correct answer to the question: How long is the presidential term of office in Colombia?.

Table 2.3: Summary Statistics - Covariate Balance Analysis

Variables	(1) Treatment (Mean)	(2) Control (Mean)	(3) Difference
1.Socioeconomics			
1.1 Income Brackets (1 to 16)	6.87	7.05	-0.19
1.2 Asset Index	-0.18	-0.01	-0.16
2.Demographics			
2.1 Gender (Male=1)	0.38	0.45	-0.08
2.2 Age	46.61	43.46	3.14***
2.3 Married/Relationship	0.65	0.70	-0.06
2.4 Number of children	3.12	2.85	0.27
3.Ethnicity			
3.1 White	0.29	0.30	-0.01
3.2 Mixed (Mestizo)	0.53	0.49	0.04
3.3 Indigenous	0.08	0.12	-0.03
4.Education			
4.1 Years of Education	7.37	8.20	-0.83
4.2 Incomplete Primary	0.27	0.23	0.04
4.3 Complete Primary	0.18	0.18	0.00
4.4 Incomplete Secondary	0.17	0.14	0.04
4.5 Complete Secondary	0.26	0.24	0.03
4.6 Superior	0.12	0.21	-0.09**
5.Knowledge			
5.1 Have Internet	0.25	0.33	-0.08
5.2 Political News (Daily)	0.94	0.95	-0.01
5.3 Political Knowledge (Departments)	0.26	0.28	-0.02
5.4 Political Knowledge (Presidential Period)	0.90	0.89	0.01
6.Activity			
6.1 Inactive	0.07	0.07	0.01
6.2 Not Remunerated Work at Home	0.30	0.29	0.01
6.3 Religion (At least Attended Once a Week)	0.77	0.77	0.00
6.Social Programs			
7.1 'Familias en Accion' (Cash Transfer)	0.24	0.26	-0.03
7.2 Sisben (Health)	0.68	0.68	0.00

*** p<0.01, ** p<0.05, * p<0.1

Notes: Columns (1) and (2) present the mean of the listed variable in the treatment and control groups, respectively. Column (3) presents the difference between the treatment and control groups, estimated with robust standard errors clustered at the community level. Using stars, I report the results from a test of the null hypothesis that the listed variable is not different between the treatment and control groups. The number of observations for all variables varies between 795 and 800. See footnote 5 for an explanation of how I construct these variables.

2.5.1 CVA Program Impact on Project Performance

To measure the CVA program's impact on project performance, I use a wide variety of measures, including overall performance, compliance with established timeframe for completion, adequacy in the administration of resources, adequate inputs, adequate planning and probability of deciding to carry out the same project again. Table 2.4 describes in detail the set of variables included in the analysis.

Table 2.4: Measures to Evaluate Project Performance

Measures	**Description**
Overall Performance	A binary variable, where one refers to citizens who evaluate the project above or well above similar infrastructure projects in the community.
Compliance with Established Timesframe for Completion	A binary variable, where one refers to citizens who report the project was finished on schedule.
Adequacy in the Administration of Resources	A binary variable, where one refers to citizens who report the project's resources were adequately managed.
Adequate Inputs	A binary variable, where one refers to citizens who report adequate quality of input to build the project.
Adequate Planning	A binary variable, where one refers to citizens who report adequate planing for the project.
Carry Out the Same Project	A binary variable, where one refers to citizens who, knowing what they know now, would have carried out this project rather than an alternative one.

Using this set of measures, I assess the CVA program's effect on project performance. Figure 2.3 below presents the risk difference for each measure in Table 2.4. The risk difference is the difference between the proportions of individuals having the outcome of interest between the treatment and control groups. Each circle represents the estimated effect of the program using each measure, and the horizontal lines

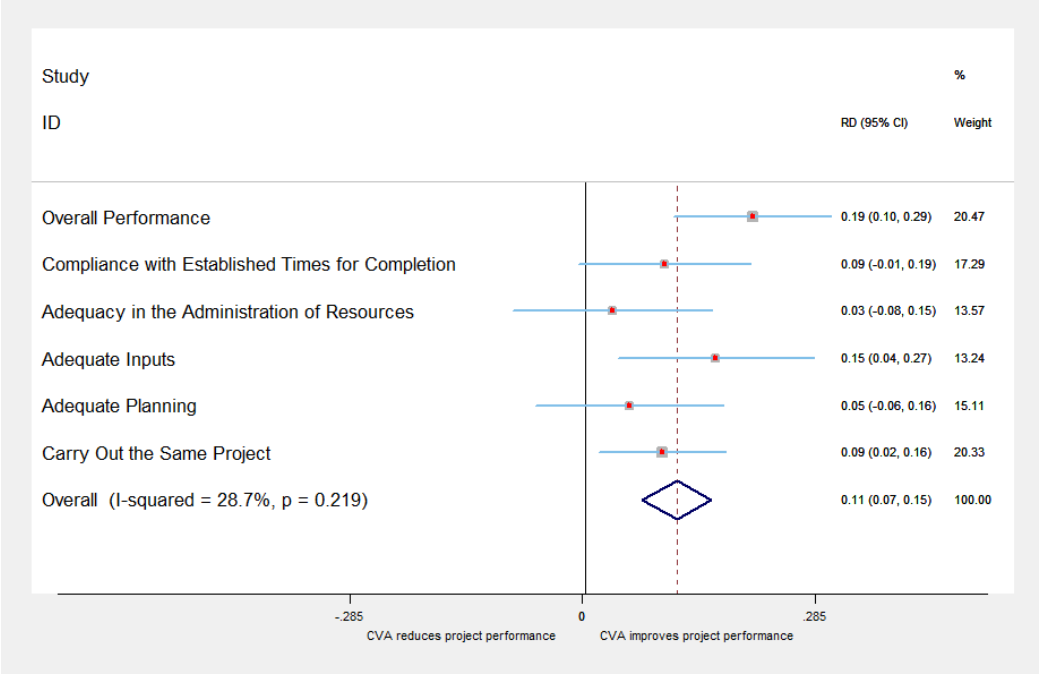
represent the 95% confidence intervals.

All the measures show a positive impact for the CVA program, and three out of six measures are statistically significantly different from zero. The magnitude of changes in citizens' satisfaction is substantial. To interpret these magnitudes, note that the CVA program resulted in an increase of 19 percentage points in terms of citizens who report being satisfied with the projects, as well as a 15 percentage point increase in citizens reporting that the project used adequate inputs and a 9 percentage point increase in citizens who would carry out the same project again.

One concern raised in literature regarding the use of perception measures deals with the possibility of introducing uninformative noise (Olken, 2009). To ameliorate these concerns, I aggregate the measures using meta-analysis. The weight used to aggregate each measure is proportional to its precision, which is defined as the inverse of its squared standard error. This is represented in the figure by the size of the square associated with the estimated effect of each measure. The aggregate effect is represented by the dashed line and the lateral diamond tips represent the 95% confidence intervals. The aggregate effect is positive and statistically significant. This analysis shows that the positive result does not rely on only one measure, but that several complementary measures point to the CVA program having a positive impact on project performance.

To test the robustness of the results shown in Figure 2.3, I estimate equation 2.1 using a probit model, where h is a household in a given community, which is denoted as c . *TREATED* is a binary variable that equals 1 for respondents in communities with CVA projects and 0 otherwise. $OUTCOME$ refers to the variables explained in Table 2.4.

Figure 2.3: Impact of the CVA Program on Project Performance



Notes: The figure shows the effect of the CVA program on various project performance measures. Each circle represents the estimated effect for each measure, and the horizontal lines represent the 95% confidence intervals. The vertical dashed line represents the aggregated effect of the program and the diamond's tips represent the confidence intervals. The size of the squares represents the weight of each measure in the aggregation procedure.

$$\frac{\partial OUTCOME_{ch}}{\partial TREATED_{ch}} = \frac{\partial \Phi(x_{ch}\beta)}{\partial TREATED_{ch}} = \phi(x_{ch}\beta)\beta_{ch} \tag{2.1}$$

Columns (1)-(3) of Table 2.5 present the results. Column (1) shows the number of observations for each specification. Column (2) shows coefficients for the marginal effect. The standard errors are adjusted for clustering at the community level c, and I include a set of control variables: gender, age group dummies, income, education dummies and ethnicity dummies. Column (3) shows the p-value for the null hypothesis of no effect.

As shown in Table 2.5, the CVA program resulted in greater community satisfaction across a wide variety of measures. Among communities having CVA projects, citizens were more likely to report both higher project performance as well as having more adequate inputs for building the project. Once I account for all controls, I found that the point estimates are higher and more significant than in Figure 2.3. The overall performance for CVA communities increases from 19 to 22.8 percentage points, adequate inputs from 15 to 20.9 percentage points, and citizens reporting that they would carry out the same project increases from 9 to 10.8 percentage points.

For columns (4) and (5) of Table 2.5 I use the same variables from Table 2.4 but a different scale. Each variable is used as asked on the survey, on a 5-point scale from 1 (best) to 5 (worst). As a result, negative coefficients will indicate an improvement in project performance. For ease of interpretation, I estimate the regression in equation 2.2 using ordinary least squares (OLS).

$$OUTCOME_{ch} = \alpha + \beta TREATED_c + X'_{ch}\gamma + \varepsilon_{ch} \tag{2.2}$$

Again, I cluster standard errors by community to take consider the fact that there are multiple respondents h in each community c. The vector X represents a set of respondent control variables: gender, age group dummies, income, education dummies and ethnicity dummies. Regression results confirm the patterns shown in Figure 2.3 and in columns (2) and (3) of Table 2.5. Overall, the point estimates are now generally negative and statistically significant. Estimations using ordered probit and logit models produce qualitatively similar results.

One potential concern with the accuracy of the results is whether the overall positive average treatment effect on the treated is masking differences across pair of

Table 2.5: Impact of the CVA Program on Project Performance

		Probit		OLS	
	(1)	(2)	(3)	(4)	(5)
Measures	No of obs.	Coef.	p-value	Coef.	p-value
Overall Performance	422	0.228	0.000	-0.441	0.002
Compliance with Established Timeframe for Completion	356	0.078	0.321	-0.217	0.250
Adequacy in the Administration of Resources	279	0.015	0.854	0.019	0.919
Adequate Inputs	272	0.209	0.017	-0.192	0.050
Adequate Planning	316	0.089	0.129	-0.082	0.153
Carry Out the Same Project	420	0.108	0.009	–	–

Notes: Each cell in columns (2) and (4) is the coefficient of the treatment dummy from a different regression. In columns (2) and (3), I use the binary version of the variables explained in Table 3. All regressions are estimated by computing the marginal effect using a probit model with robust standard errors, adjusted for clustering at the community level. To compute the marginal effect, I evaluate all other variables at their means. The results are qualitatively similar using OLS and logit models as well as when evaluating the other variables at zero or one when computing the marginal effects. In columns (4) and (5) all questions are multiple-ordered response questions on a scale from 1 (best) to 5 (worst). All regressions are estimated using OLS with robust standard errors, adjusted for clustering at the community level; the results are qualitatively similar using ordered probit and probit models. All regressions include controls for gender, age, income, ethnicity and educational background.

projects. In other words, can a CVA program improve project performance under any context? Have all CVA programs been equally successful? If the CVA program was to be scaled up, would it be reasonable to expect that every infrastructure project increases its project performance between 7 and 15 percentage points, at least as evaluated by citizens? To determine if treatment effects were heterogeneous across projects, I use the fact that for each CVA project, I have a similar project without an attached CVA program to decompose the overall performance measure by pairs of projects.[6]

[6]I found similar results for the other measures.

Although samples within each pair are too small to yield conclusive results, the evidence from Figure 2.4 is striking. While the CVA program have positive and statistically significant impacts on average, not all projects that underwent the CVA program were equally successful. In particular, the CVA program's estimated effect is discouraging in Yopal, Arauca, Maicao, and Cienaga. Though not statistically significant[7], the results suggest that the CVA program worked very differently in those places than in Neiva or Tuchin. Furthermore, even in the same municipalities, the results differed. While the water and sanitation infrastructure project in Yopal did not appear to have improved as a result of the CVA, the educational facilities building projects showed a highly positive effect. This result suggests that more information is required to understand why a program succeeds or fails. In particular, understanding CVA program's pathway of change is necessary. In the next section, I explore bottlenecks that can explain why the program was not equally successful across projects.

2.5.2 Uncovering CVA Program Mechanisms

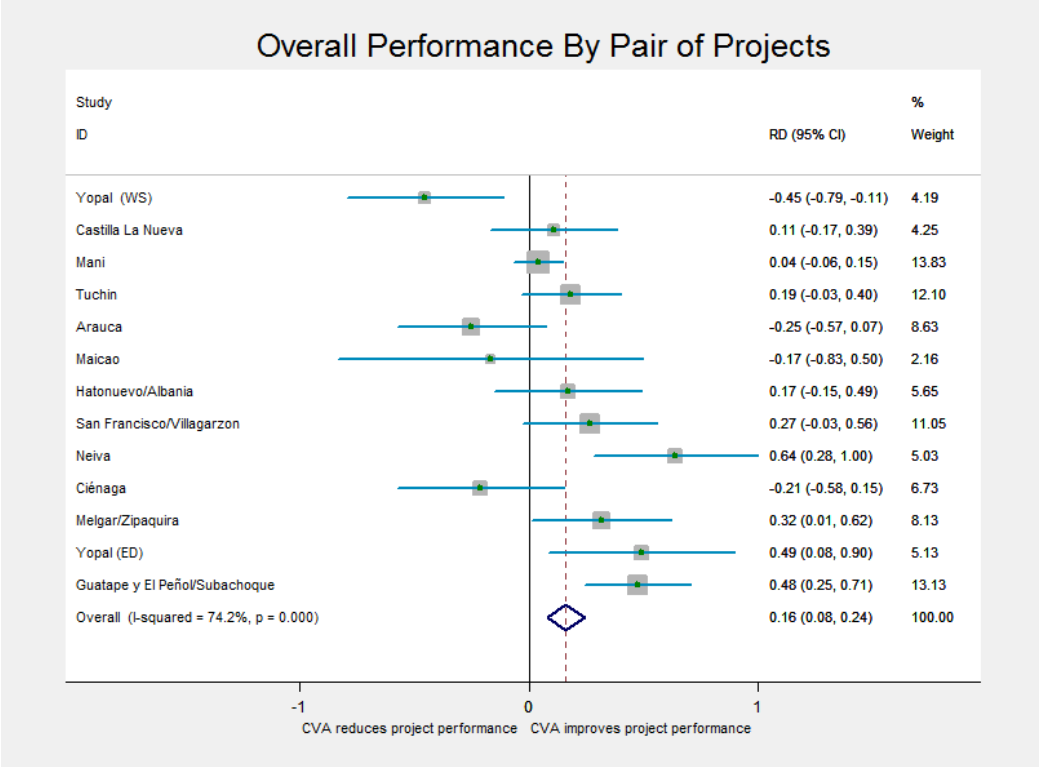
In section 2.3 I presented the CVA program's pathway of change and discussed four alternative theories. In this section I use this framework to shed light on two important questions to which literature has not been able to provide a clear answer:

(a) Under what conditions do citizens decide to monitor the project to improve accountability?

(b) Under what conditions do governments facilitate citizen engagement and become more accountable?

[7]Yopal is the only case in which the difference is statistically significant. Once we control for respondents' socio-economics characteristics, Yopal remains negative but is no longer statistically significant.

Figure 2.4: Decomposing the Average Treatment Effect of the CVA Program



Notes: The figure shows the heterogeneity of citizens satisfaction with the program by pairs of projects. Each circle represents the estimated effect and the horizontal lines represents the 95% confidence intervals. The vertical dashed line indicates the aggregated effect of the program and the diamonds' tips represent the confidence intervals.

Table 2.6 presents the four alternative theories discussed in section 2.3 and re-introduces the empirical implication of each theory.

Table 2.6: Alternative Theories and Empirical Implications

Theory	**Focus**	**Empirical Implications**
(a) Deficient Implementation Theory	C	**(a.1)** If information on the project is not properly disseminated, citizens will not participate in monitoring activities. **(a.2)** Citizens probability of participation will be a function of the quality of the information they possess.
(b) Civic Participation Failure	C	**(b.1)** If citizens perceive the objective of the project as being irrelevant for the community, they will not participate in community monitoring activities. **(b.2)** If the opportunity cost of time is high, participants in community monitoring activities will have a higher probability of being individuals who are inactive, unemployed, with a lower income. **(b.3)** If the types of actions citizens are encouraged to take as a result of participation in the CVA program do not differ from those they were already carrying out, they will not participate in community monitoring activities. **(b.4)** If citizens expect other citizens will free-ride on their efforts to monitor the project, the probability and intensity of participation will be lower.
(c) Non-Accountable Politicians and Providers	G	If governments are not accountable to their citizens, politicians' and providers' performance should not increase as a result of the CVA program.
(d) Self-Fulfilling Prophecy	B	**(d.1)** Citizens who perceived themselves as influential have a higher probability of participating in the CVA public forums and community monitoring activities. **(d.2)** Citizens who participated in the CVA public forums disseminate information of the CVA to other non-participating citizens. **(d.3)** In projects having more participation in community monitoring activities and higher visibility politicians' performance is higher.

Notes: Focus refers to which segment of the pathway of change is emphasized, citizen ability to monitor project performance (C), governments accountability (G), or both (B).

Below I assess the empirical merits of each theory.

Deficient Implementation Theory

I test (a.1) by focusing on the percentage of citizens in each community who are informed about the project's existence. The results shown in Figure 2.5 indicate that for many CVA projects (Yopal: Water and Sanitation project (WS), Castilla La Nueva, Melgar, Maicao), less than 50% of the potential beneficiaries, i.e., citizens who live close to the location of the project, know of its existence. I test whether this could be explained by recollection bias, a bias that can arise when respondents' answers are a function of the time that has passed since the project was built. The evidence shows that the percentage of citizens who report knowing about a given project is not correlated with when the project began or finished.[8]

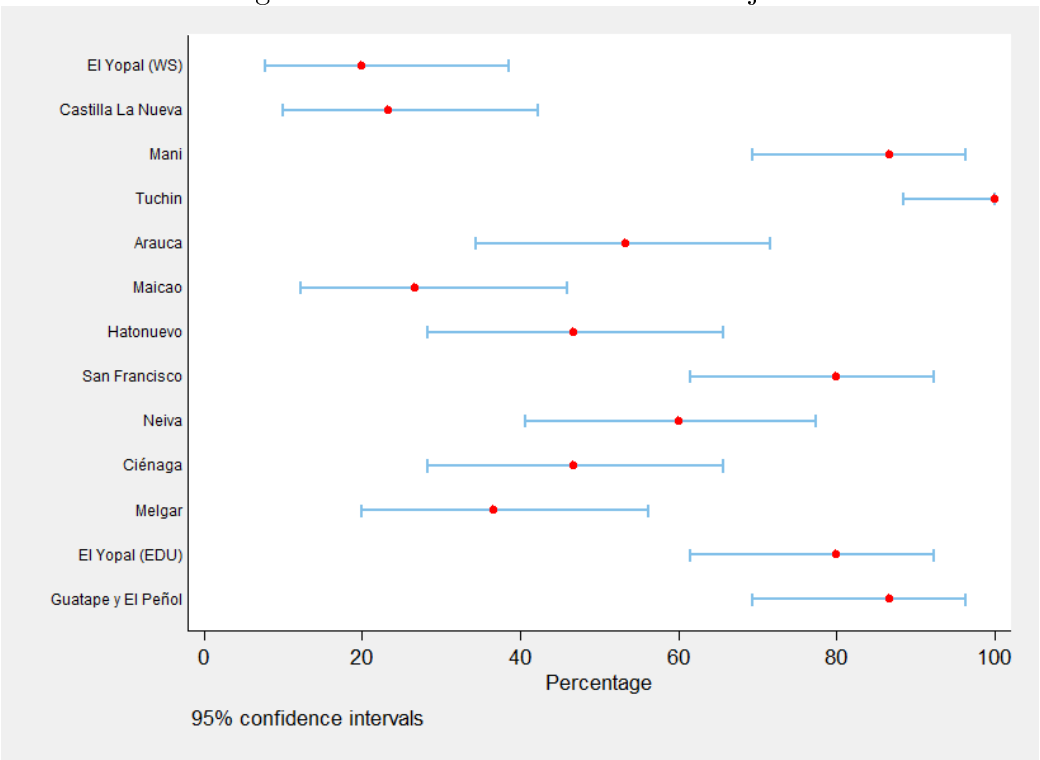
If citizens are ignorant of a project's existence, it will be impossible for them to monitor it.

Another potential deficit in CVA program's implementation relates to those who knew of the project but might not have a clear understanding of how to engage governments and providers in order to improve the project. Without good information regarding how to monitor the project and to hold the government accountable, even the best program can fail. In order to test (a.2), I measure the percentage of citizens in each community that evaluated the quality of information received on the project as being adequate or very adequate.

As shown in Figure 2.6 below, the results are highly heterogeneous. While in some projects, information provided to the public was adequate, in others, it appears as if it

[8]The citizens were told the project's location, name and objective to help them remember it. In addition, citizens who did not live in the community for at least two years before the project was executed were excluded from the protocol and not surveyed.

Figure 2.5: Do Citizens Know the Project?

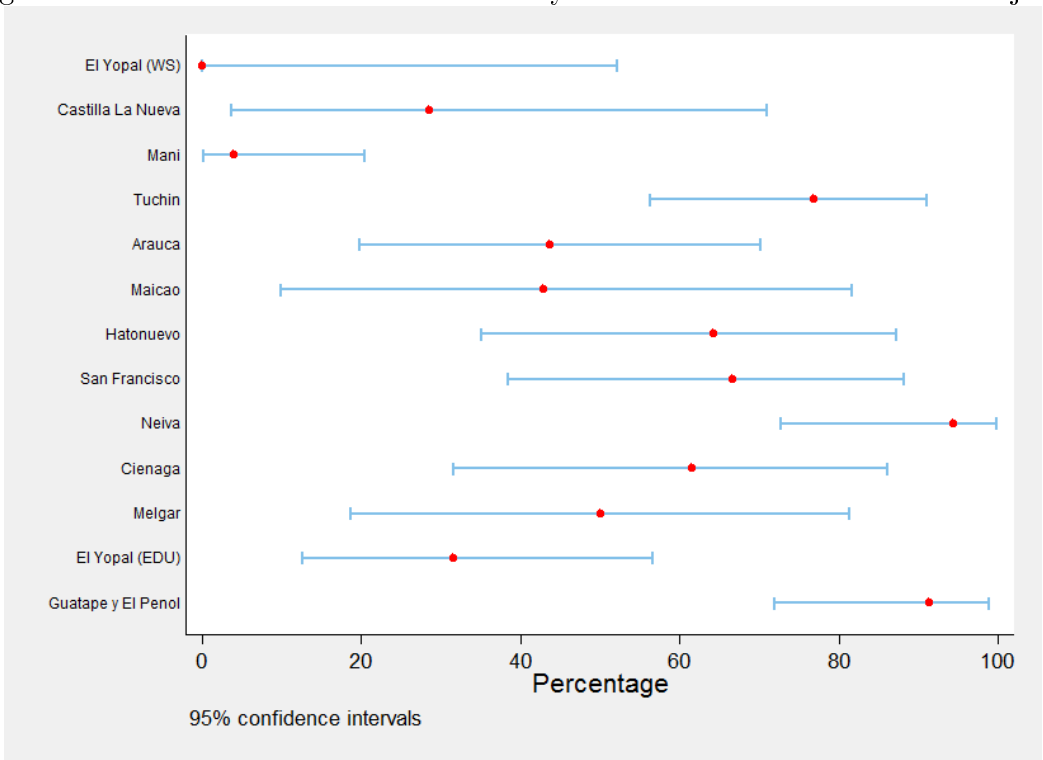


Notes: The figure shows the percentage of citizens who know of the project in each treated community. The circle represents the point estimate and the horizontal lines depicts the 95% confidence intervals.

the information did not facilitate participation in the community monitoring scheme. Again, there is evidence of problems with implementation of the CVA program. This is relevant as the evidence shows a high correlation between those reporting to have received adequate information about the project and those who spent time in monitoring activities. Perception of good quality information is associated with citizens spending an increase of almost seven [3-11] additional monitoring minutes a week.

Though the absolute percentages in some communities are low, it is relevant to determine the counterfactual to assess the degree of deficiency in CVA program implementation. The results from Figure 2.5.2 indicate that the CVA program improved the information citizens received by 12 percentage points. This magnitude is not only substantial but also statistically significant [0.04-0.21]. Moreover, the

Figure 2.6: Do Citizens Have the Necessary Information to Monitor the Project?



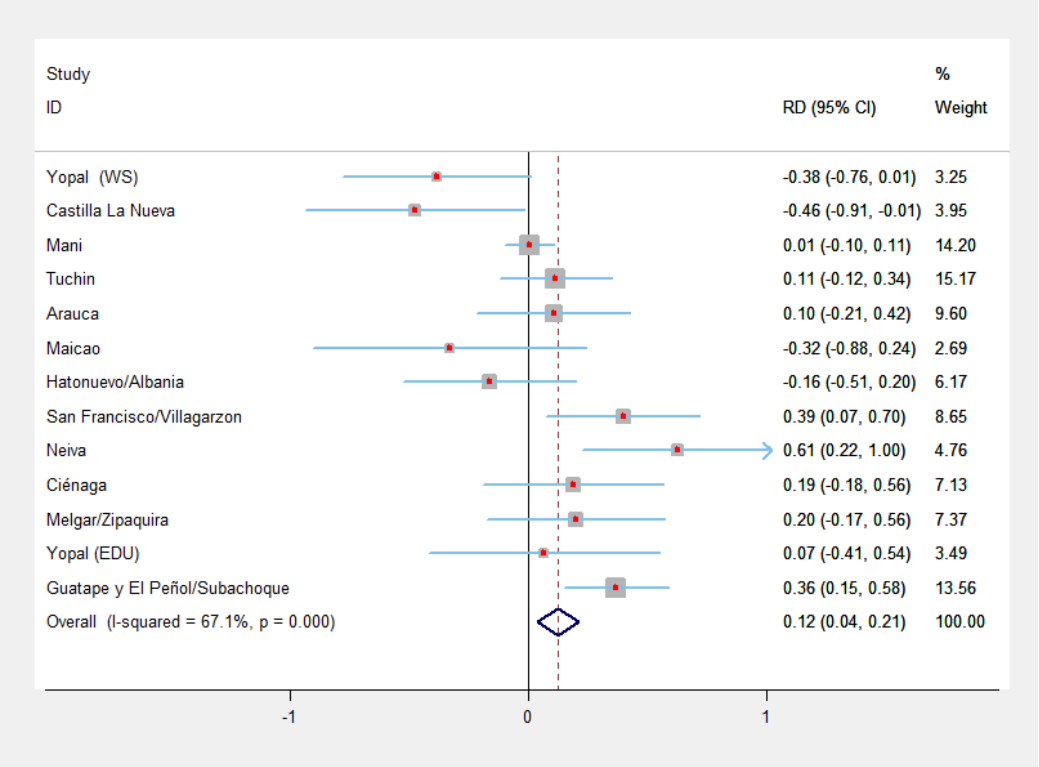
Notes: The figure shows the percentage of citizens satisfied with the quality of information received on each treated project. The circle represents the point estimate, and the horizontal lines represent the 95% confidence intervals.

results are highly heterogenous across communities.

Civic Participation Failure Theory

Because the CVA program was implemented after the infrastructure project was already selected by the government without any community participation, determining whether the project is really a priority for the community is unclear. If the project was not a priority, we would expect that the CVA program would not induce the community to monitor the project. To test (b.1), I estimate the percentage of citizens that believed the project was a priority for the community.

Figure 2.7: Does the CVA Program Improve the Information Citizens Receive?



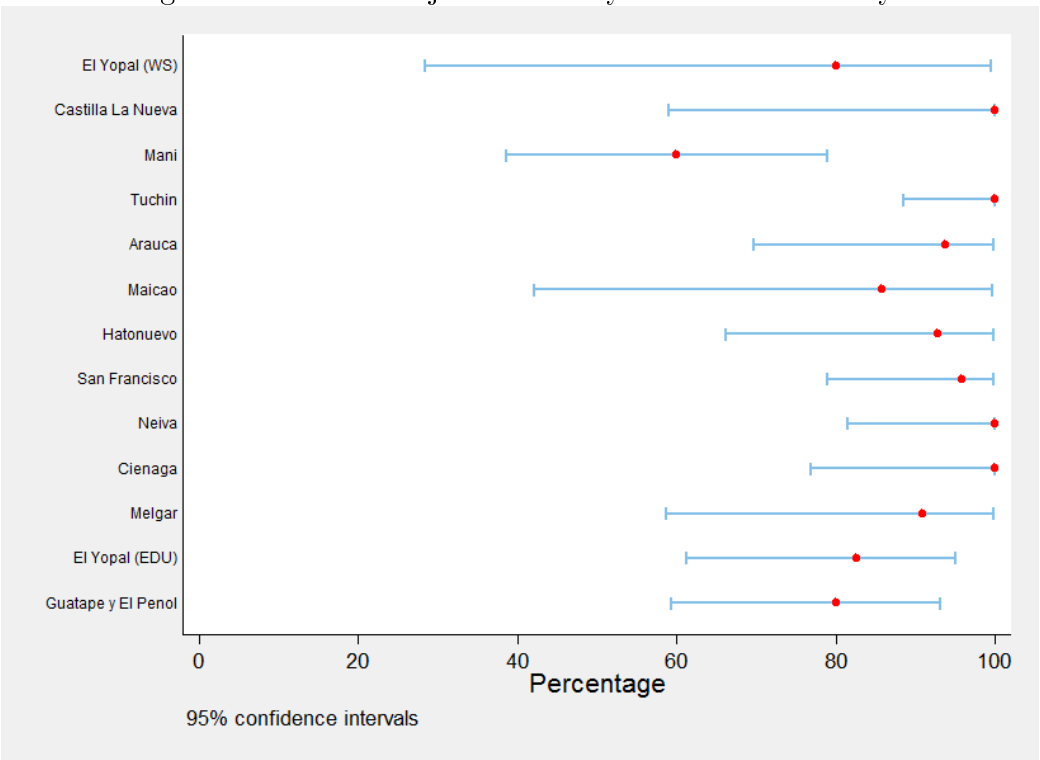
Notes: The figure shows the effect of the CVA program on the percentage of citizens satisfied with the quality of information received. Each circle represents the estimated effect and the horizontal lines represent the 95% confidence intervals. The dashed vertical line indicates the aggregated effect of the program and the diamonds' tips represent the 95% confidence intervals.

Figure 2.8 presents the evidence. On average, more than 80% of citizens believe the project was a priority for their community, which suggests that a lack of interest in the project was not a binding constraint.

To assess whether the opportunity cost of participation is a bottleneck (b.2), I explore whether CVA participants are statistically different from non participants in the following characteristics: employment status, whether they work at home or not, and income level.[9] Table 2.7 presents the results. I estimated a participation model using a probit model and robust clustered standard errors at the community

[9]Results using *asset index* provides qualitatively the same results.

Figure 2.8: Is the Project a Priority for the Community?



Notes: The figure shows the percentage of citizens who classify the project as a priority for the community. Each circle represents the point estimate and the horizontal lines represents the 95% confidence intervals.

level but could not isolate individual characteristics acting as determinants of participation in CVA public gatherings. I present the marginal effect of these variables, controlling for demographics characteristics, educational dummies, and participation in government-funded social programs. This result suggests that the opportunity cost of participation is not a crucial bottleneck as there is no statistical significant differences in employment status, or type of work, or between income levels of participants and non-participants.[10]

In the case of (b.3), if participants were encouraged to engage in the same activities they would have carried out in the absence of the program, formal participation in the CVA program may be irrelevant. I estimate the difference in participation

[10]Reports from the field indicate that the time of each community forum as well as location was agreed with the community to maximize the number of participants.

Table 2.7: Opportunity Costs and Participation in Public Forum/s of the Social Audit

Variables	Participation in Public Forum/s of the Social Audit (1) Coef.	(2) p-value
Socioeconomics		
Income Brackets (1 to 16)	0.00	0.983
Activity		
Inactive	-0.15	0.201
Not Remunerated Work at Home	-0.068	0.286
Employed	-0.08	0.315
Controls - Demographics		
Gender (Male=1)	0.08	0.175
Adult (25-40 years old)	-0.14	0.115
Middle Age (41-64 years old)	-0.135	0.138
Old Age (65+ years old)	-0.07	0.517
Married/Relationship	0.04	0.213
Controls - Education		
Complete Primary	0.00	0.998
Incomplete Secondary	0.07	0.369
Complete Secondary	0.09	0.207
Higher Education	0.17	0.057
Control - Participation in Social Programs		
'Familias en Accion' (Cash Transfer)	0.07	0.293
Sisben (Health)	0.02	0.694
Pseudo R^2	0.04	
No. of observations	516	

Notes: In columns (1) and (2) I use participation in at least one community forum, which only citizens in treated communities would be able to attend, as a dependent variable. The regression is estimated using a probit model with robust standard errors, adjusted for clustering at the community level. To compute the marginal effect, I evaluate all other variables at their means. The results are qualitatively similar using OLS, logit models, and when evaluating the other variables at zero or one when computing the marginal effects.

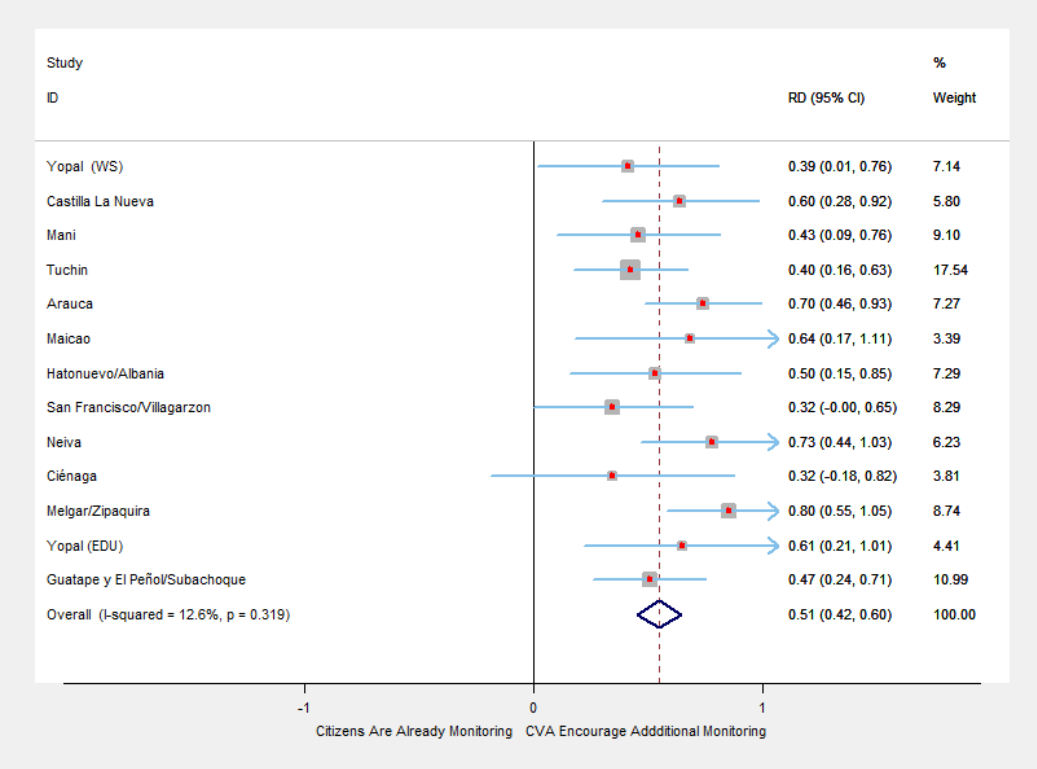
rates as well as time spent in community monitoring activities for both groups (i.e., citizens in the control group and citizens in the treatment group who participated in at least one public forum of the CVA program). The evidence from Figure 2.9 suggests that the percentage of CVA participants spending their time monitoring is higher than citizens in the control group. On average, there is a 51 percentage point increase in participation rates in monitoring activities for CVA participants. Not only do more citizens spend their time monitoring, but also spend, on average, 30 minutes conducting monitoring activities in a week, whereas citizens in the control group only spend, on average, 8 minutes a week. This does not appear to be a crucial bottleneck for success of the CVA program, because in the program's absence, most citizens are not actively engaged in community monitoring activities.

To test (b.4), collective action failures, I use several different measures. Because there is no ideal proxy for citizens' expectations of reciprocal cooperative behavior, I follow the literature in using trust in fellow neighbors as a proxy for social capital. In addition, I compute fractionalization indexes, using the traditional formula $(1 - \sum_{i=1}^{n} s_i^2)$, where s_i is the share for each group i, as predictors of lack of social capital.[11] In particular, I measure ethnicity, religion, and language fractionalization by community.

Table 2.8 present the results. Column (1) shows the percentage of citizens in each community who report trusting in their fellow neighbors and columns (2)-(4) present ethnic, religion and language fractionalization indexes for each treated community. The measure of trust in their fellow neighbors shows, on average, that 68% of citizens trust one another. This would indicate a high level of social capital, a finding that

[11] See Alesina and Wacziarg (2003) for details.

Figure 2.9: Do the Types of Actions Citizens are Encouraged to Take Differ from What They Are Already Doing?



Notes: The figure shows that the percentage of community forum participants in treated communities who spend time monitoring the project is significantly and statistically higher than this percentage among citizens in control group projects. Each circle represents the estimated effect, and the horizontal lines represent the 95% confidence intervals.

the language fractionalization index also suggests. Both ethnic and religion fractionalization show a larger range.

Next, I test whether heterogeneity in the community's willingness to monitor the infrastructure project, can be explained using the proxies for social capital (or lack of). To construct the variable *Time spent in community monitoring activities* I use responses to "Did you devote any of your time to monitor the project?" The responses are as follows: (1) No time, (2) Less than 10 minutes a week, (3) Between 10 and 30 minutes a week, (4) Between 31 and 60 minutes a week and (5) More than an hour a week. Using average time thresholds for each bracket (0, 5, 15, 45, 60), I compute

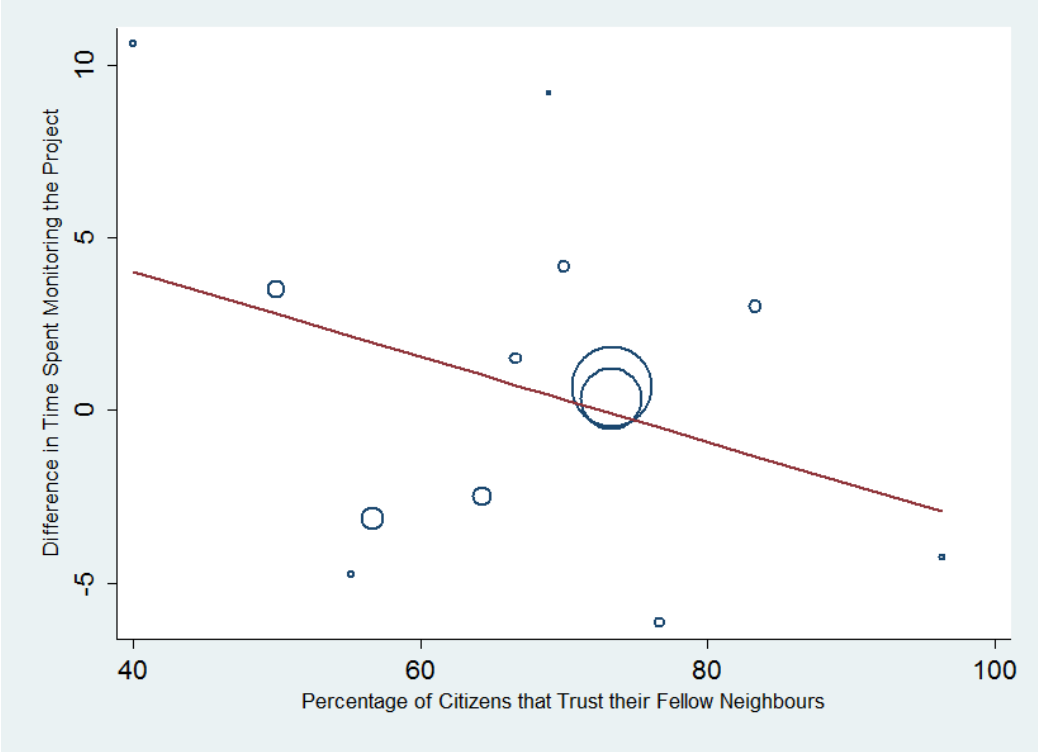
Table 2.8: Proxies for (Lack of) Social Capital

Project	(1) Trust	(2) Ethnic Fractionalization	(3) Religion Fractionalization	(4) Language Fractionalization
Yopal (WS)	0.64	0.42	0.29	0.00
Castilla La Nueva	0.50	0.50	0.38	0.00
Mani	0.83	0.26	0.39	0.00
Tuchin	0.69	0.19	0.66	0.00
Arauca	0.57	0.62	0.62	0.00
Maicao	0.67	0.67	0.50	0.00
Hatonuevo	0.70	0.71	0.47	0.00
San Francisco	0.55	0.65	0.12	0.10
Neiva	0.60	0.58	0.25	0.00
CiÃľnaga	0.77	0.53	0.43	0.00
Melgar	0.73	0.47	0.28	0.00
Yopal (EDU)	0.73	0.43	0.39	0.00
Guatape y El PeÃśol	0.96	0.57	0.06	0.00

Notes: Column (1) refers to the percentage of citizens in each community who report to have some or a lot of trust in their neighbors. Columns (2)-(4) present my own calculations for the ethnic, religion and language fractionalization indexes in each of the treated communities using the traditional formula $(1 - \sum_{i=1}^{n} s_i^2)$ where s_i is the share for each group i.

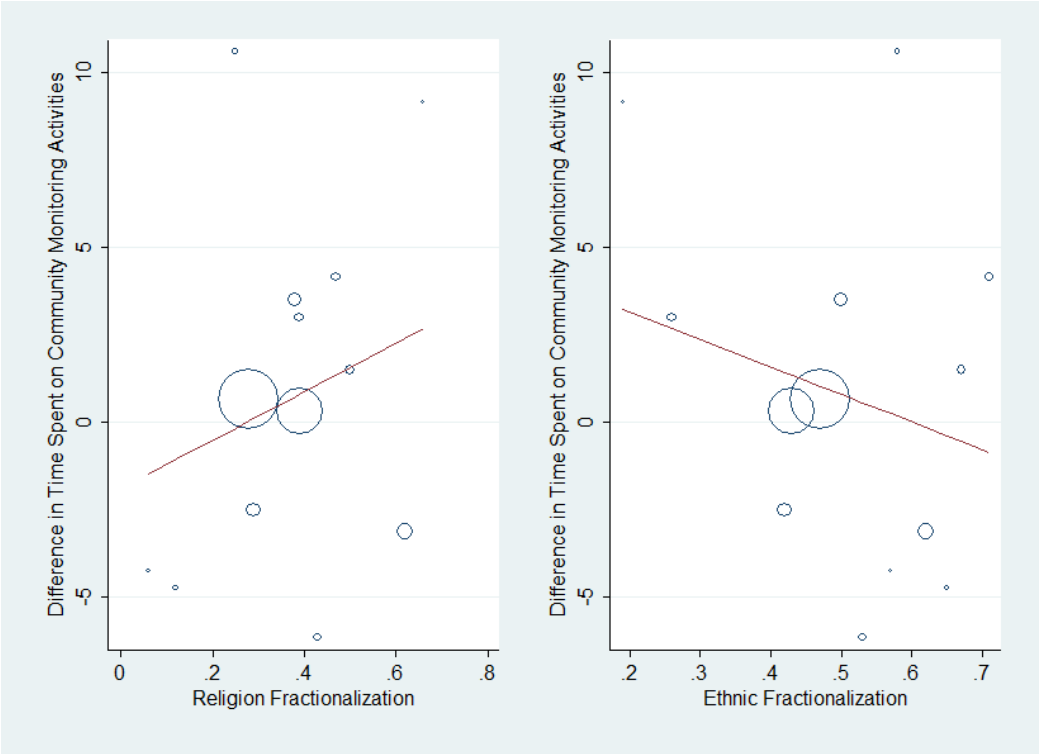
the variable for both treated and control communities. Figure 2.10 and 2.11 show the results from a meta-regression for the 13 pair of projects included in this study. I estimate a meta-regression model to account for the fact some projects only had a few individuals answer the question about the time spent in community monitoring activities. Meta-regression allows more precise estimates to have more weight in the regression, which is represented in the figure by circle size. The results indicate that neither trust nor ethnic or religion fractionalization has any statistically significant effect on the difference in time spent monitoring by each pair of communities.

Figure 2.10: Effect of Citizens' Trust on Participation in Monitoring Activities



Notes: Using meta-regression, the results shown in the figure indicate that citizens' trust in one another is not a causal factor behind the observed differentials in community monitoring activities. Each circle represents the point estimate, and the lines represents the 95% confidence intervals. To construct *Time spent in community monitoring activities* I use responses to "Did you devote any of your time to monitoring the project?" with an answer scale of (1) No time, (2) Less than 10 minutes a week, (3) Between 10 and 30 minutes a week, (4) Between 31 and 60 minutes a week and (5) More than an hour a week. Using thresholds for each bracket (0,5,15,45,60) I compute this variable for both treated and control communities.

Figure 2.11: Effect of Ethnic and Religion Fractionalization on Citizens' Participation in Monitoring Activities



Notes: The figure shows a meta-regression. The size of each circle represents the weight of each project, and the line represents the estimate of the coefficient of ethnic/religius fractionalization on difference in time spent monitoring by each pair of communities. To construct *Time spent in community monitoring activities* I use responses to "Did you devote any of your time to monitoring the project?" with an answer scale of (1) None of my time, (2) Less than 10 minutes a week, (3) Between 10 and 30 minutes a week, (4) Between 31 and 60 minutes a week, and (5) More than an hour a week. Using average time thresholds for each bracket (0, 5, 15, 45, 60) I compute the variable for both treated and control communities.

Overall, the evidence suggests that citizens' expectations for other fellow citizens reciprocal behavior cannot explain why citizens in some CVA programs decide to spend more or less time monitoring the project than their counterfactual counterparts.

Nonaccountable Governments Theory

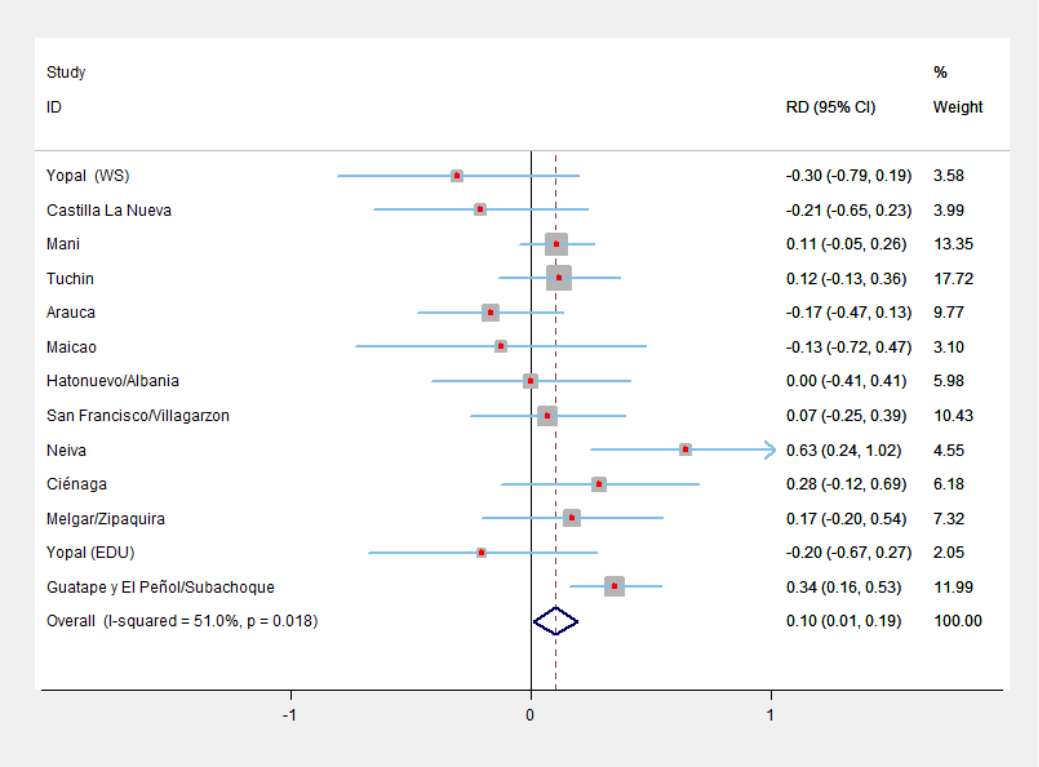
I test this theory by estimating the CVA program's impact on citizens' evaluations of politicians and providers performance on the project. In particular, I measure politicians' (providers') performance with a binary variable equal to 1 when citizens evaluate the work of local politicians (providers) in the project as good or very good, and 0 otherwise.

Figure 2.12 presents the forest plot used to test whether the CVA program has an impact on politicians' behavior. In this figure, I compute the risk difference for each pair of projects. Each circle represents the estimated effect of the program on politicians' behavior, as evaluated by citizens. The horizontal lines represent the 95% confidence interval.

Figure 2.12 shows an overall improvement in politicians' behavior as a result of the CVA program. I aggregate the measures using meta-analysis and the weight to aggregate each measure is proportional to its precision, which is defined as the inverse of its squared standard error. This is represented in the figure by the size of each squares associated with the estimated effect of each measure. The aggregate effect is represented by the dashed vertical line, and the diamonds' lateral tips represent the 95% confidence intervals. The aggregate effect shows that the CVA program improves politicians' performance by 10 percentage points, a finding that is

statistically significant $[0.01 - 0.19]$.

Figure 2.12: Impact of the CVA Program in Politicians' Performance



Notes: The figure shows the difference in percentage of citizens who evaluate the work of politicians in the project as good or very good by each pair of communities. Each circle represents the point estimate and the horizontal lines represents the 95% confidence intervals. The vertical dashed line indicates the aggregated effect of the program, and the diamonds' tips represents the 95% confidence intervals.

To estimate the robustness of the results shown in Figure 2.12, I estimate the same regression as in equation (1) but with a different dependent variable, politicians' performance. Table 2.9 presents the results. Column (1) of Table 2.9 shows the number of observations for each variable, column (2) shows the coefficient of the marginal effect and column (3) shows the p-value. The standard errors are adjusted for clustering at the community level c, and I include a set of control variables: gender, age group dummies, income, education dummies, and ethnicity dummies.

As shown in Table 2.9, the CVA program resulted in greater community satisfaction with politicians' performance. Citizens in CVA communities were more likely to report that politicians performances were good or very good. Once I had accounted for all the controls, I found that the point estimates are higher and more significant than in Figure 2.12. The evaluation of politicians performances increased from 10 to 15 percentage points for CVA communities.

To test the effect of the CVA program on providers' performance, I follow the same strategy as with politicians' performance. Table 2.9 presents the results. Although the estimated effect is positive, it is not statistically significant.

In columns (4) and (5), I use the same variables but on a different scale. Each variable is used as it was asked on the survey, on a 5-point scale from 1 (best) to 5 (worst). As a result, negative coefficients will indicate an improvement in evaluation of politicians' and providers' performances. The regression results confirm the results shown in Figure 2.12 and those from columns (2) and (3).

The evidence presented in this section indicates that the theory of non-accountable politicians is not valid in the CVA context, because I find improvements in politicians performance as a result of the project. Therefore the relevant question becomes not whether politicians are accountable, but rather under what conditions are they accountable? The next subsection sheds light on this issue.

Self-Fulfilling Prophecy Theory

To test the empirical validity of this theory I follow the structure of the theory itself. For part I, I estimate a model at the individual level of citizens' perceptions of their influence on time devoted to monitoring the project and participation in the CVA

Table 2.9: Impact of the CVA Program on Politicians' and Providers' Performances

		Probit		OLS	
	(1)	(2)	(3)	(4)	(5)
Measures	No of obs.	Coef.	p-value	Coef.	p-value
Politicians Performance	356	0.15	0.036	-0.25	0.086
Providers Performance	326	0.11	0.147	-0.13	0.239

Notes: Each cell in columns (2) and (4) is the coefficient of the treatment dummy from a different regression. In columns (2) and (3), I use a binary version of the variables. All regressions are estimated computing the marginal effect using probit model with robust standard errors, adjusted for clustering at the community level. To compute the marginal effects, I evaluate all other variables at their respective means. The results are qualitatively similar using OLS, logit models, and when evaluating the other variables at zero or one when computing the marginal effects. In columns (4) and (5), all questions are multiple-ordered response questions on a scale from 1 (best) to 5 (worst). All regressions are estimated using OLS with robust standard errors, adjusted for clustering at the community level. The results are qualitatively similar using ordered probit and probit models. All regressions include controls for gender, age, income, ethnicity and education background. I measure politicians' (providers') performance with a binary variable equal to 1 when citizens evaluate the work of local politicians (providers) in the project as good or very good, and 0 otherwise.

program's public forums.[12] Thereafter, I test whether citizens who participate in the CVA public forums disseminate the information to fellow citizens who did not participate in the public forums. To test part II, I estimate a model of the percentage of citizens who know about the project and the time spent monitoring the project at the community level on politicians's and the project's performance.

Part I

If citizens think they have no influence on government activities, they expect that even if they spend time uncovering pockets of corruption and inefficiencies, the

[12]As a robustness check I study whether citizens believe they should have an important role in monitoring public officials. If the answer to this question is negative and citizens believe monitoring public officials is not their responsibility, then the theory of self-fulfilling prophecies would not be empirically valid. This is because citizens may decide not to participate in the CVA process even if they perceived themselves as highly influential. The results, which are available upon request, indicate that more than 75% of citizens do believe they should have an important role in monitoring public officials.

probability of an improvement in project performance will be low. As a result, they may decide not to spend their time monitoring the project. In fact, on average, less than 20% of the citizens in the treated communities believe that they have some or a great deal of influence on municipal governments, which are responsible for managing the infrastructure projects.

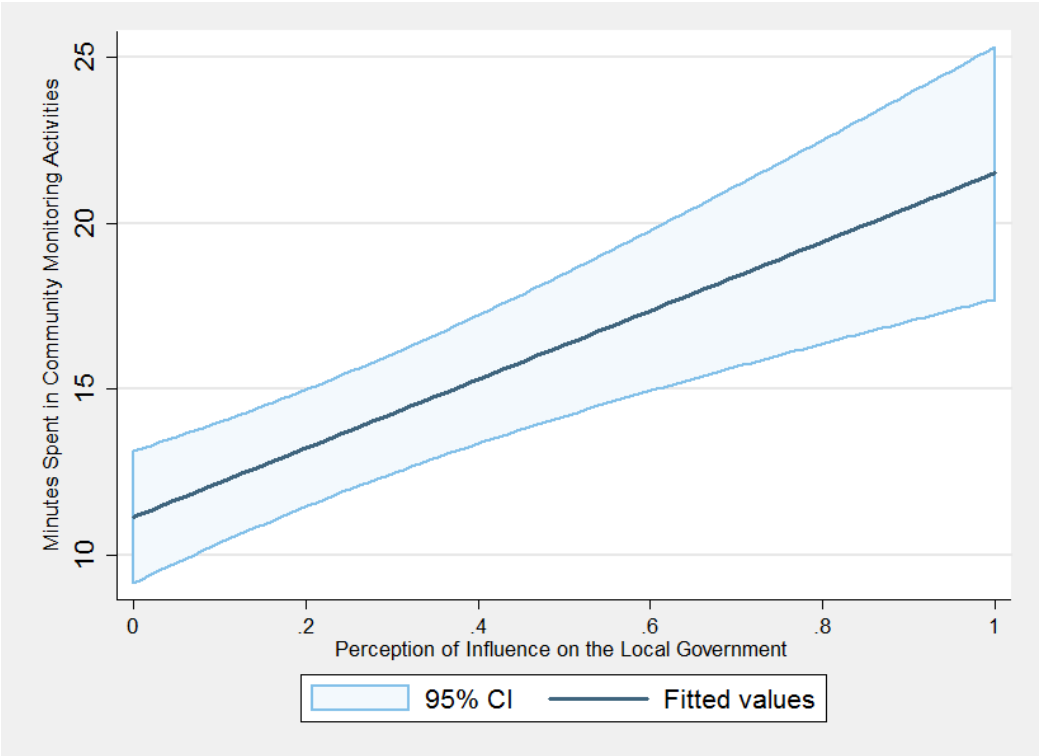
Figure 2.13 presents a graphic depiction of an estimation made using the simple linear probability model. I study whether citizens' perception of their influence on local governments is associated with community participation in monitoring activities. *Perception of influence in the local government* is a binary variable equal to 1 when citizens report to have some or a great deal of influence on the local government administration and 0 otherwise. *Time spent in community monitoring activities* measure was explained above.

The results from Figure 2.13 indicate that perception of influence is positive and highly correlated with the time citizens' spent in community monitoring activities.

Figure 2.14 displays the estimation of the relationship between perception of influence on local government and probability of participation in the social audit. The former is a binary variable equal to 1 when a citizen reports participation in at least one community forum in the CVA project, and 0 otherwise. Note that in this case, only citizens in treated communities are part of the sample I used for the estimate in this model. The results from Figure 2.14 indicate that perception of influence is positively correlated with the probability of participating in the community forum. This relationship is not only substantial but also highly significant.

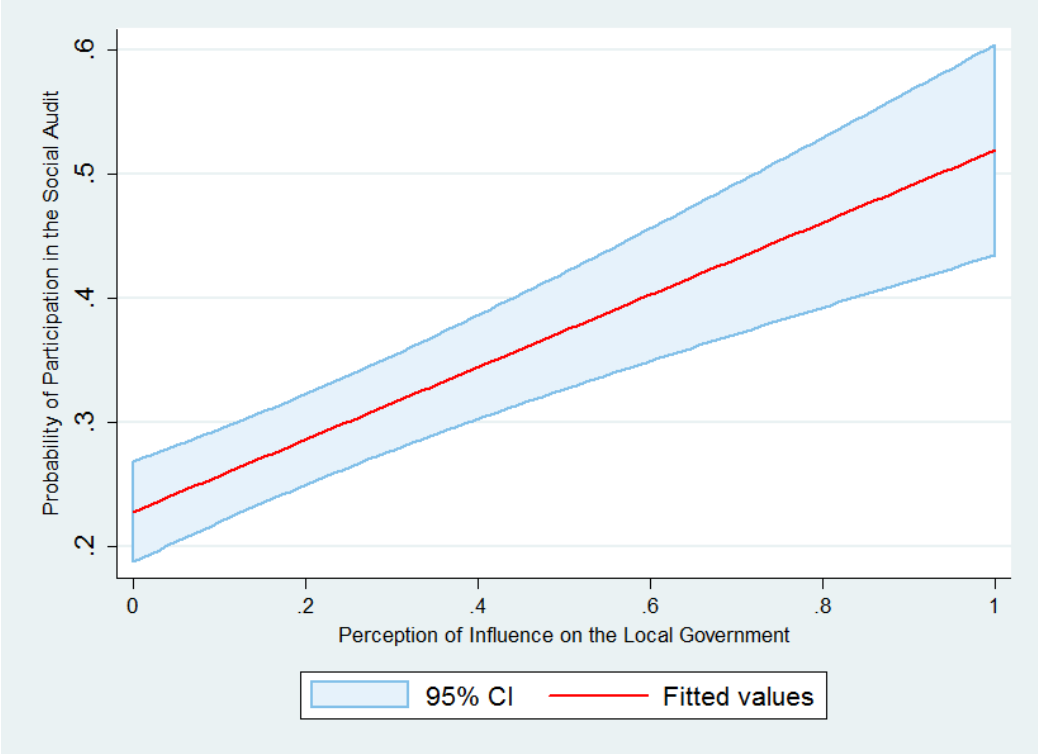
The evidence suggests that citizens' perceptions of their influence is associated with higher participation in monitoring activities in both the treated and control

Figure 2.13: Perception of Influence and Time Spent in Community Monitoring Activities



Notes: The figure shows the fitted line and the 95% confidence intervals of the effect of citizens' perceptions of their influence upon the time they spent monitoring the projects using the linear probability model. To construct *Time spent in community monitoring activities* I use responses to "Did you devote any of your time to monitor the project?" with an answer scale of (1) No time, (2) Less than 10 minutes a week, (3) Between 10 and 30 minutes a week, (4) Between 31 and 60 minutes a week and (5) More than an hour a week. Using average time thresholds for each bracket (0, 5, 15, 45, 60), I compute the variable for both treated and control communities. *Perception of influence in the local government* is a binary variable equal to 1 when citizens report to have some or a great deal of influence on the local government administration and 0 otherwise.

Figure 2.14: Perception of Influence and Participation in CVA Projects' Public Forums



Notes: The figure shows the fitted line and the 95% confidence intervals of the effect of citizens perceptions as influential on participation in CVA projects' community forums using a linear probability model. *Probability of Participation in a Social Audit* refers to a binary variable equal to 1 when a citizen reports to have participated in at least one CVA project community forum, and 0 otherwise. Note that in this case only citizens in treated communities are part of the sample I used in estimating this model. *Perception of influence in the local government* is a binary variable equal to 1 when citizens report to have some or a great deal of influence on the local government administration, and 0 otherwise.

groups, as well as with higher participation in the CVA program's public forums.

To test the robustness of the results shown in Figures 2.13 and 2.14, I estimate the regression in equations 2.3 and 2.4.

$$MONITORING_{ch} = \alpha + \beta INFLUENCE_{ch} + X'_{ch}\gamma + \varepsilon_{ch} \tag{2.3}$$

MONITORING refers to the binary variable *time spent in community monitoring activities* explained above. In addition, *INFLUENCE* refers to the binary variable *Perception of influence in the local government* explained above. For ease of interpretation, I estimate equation 2.3 using OLS. I report the results in columns (1) and (2) of Table 2.10. The regression is estimated with robust standard errors, adjusted for clustering at the community level. The vector X represents deficient implementation and civic participation failure theories discussed above and a set of control variables, including demographics and educational dummies. The results are qualitatively similar using ordered probit and logit models.

I estimate equation 2.4 using a probit model, where h is a household in a given community, denoted as c. I use participation in at least one community forum, which only citizens in treated community would be able to attend, as a dependent variable. Regression is estimated using a probit model with robust standard errors, adjusted for clustering at the community level. Columns (3) and (4) of Table 2.13 show the coefficient for the marginal effect and p-value, respectively.

$$\frac{\partial PARTICIPATION_{ch}}{\partial INFLUENCE_{ch}} = \frac{\partial \Phi(x_{ch}\beta)}{\partial INFLUENCE_{ch}} = \phi(x_{ch}\beta)\beta_{ch} \tag{2.4}$$

In the regression exercises, I control for the two theories, deficient implementation theory and civic participation failure, that attempt to explain the situations in which citizens are unable to engage with governments to monitor project performance.

On one hand, I confirmed what evidence presented above suggested: civic participation failure theory does not have any explanatory power in the CVA's context. In particular, in the case of the opportunity cost of time, income cannot explain either time spend in community monitoring activities or participation in the community forums. To test the effect of citizens' expectations of reciprocal behavior on their willingness to monitor the project, I use two measures: a measure of trust introduced above and a measure of *ethnic outsider*, suggested by Lieberman, Posner and Tsai (2013). This is a dummy variable equal to 1 if a respondent's ethnicity is different from that of the majority of respondents in that community, and 0 otherwise. Both measures fail to explain the variation in the dependent variables.

On the other hand I find that deficient program implementation and citizens' beliefs regarding their level of influence are major bottlenecks that constrain citizens' ability to engage with government to monitor project performance. In the first case, I find that having access to high quality information is crucial in the deciding whether to monitor the project. In particular, high quality information increases time spent in monitoring activities by almost seven additional minutes, and increases probability of citizen participation in the community forum by almost 16 percentage points. In the case of beliefs, perceiving oneself as somewhat or highly influential on local government increases time spent on monitoring by almost eight additional minutes as well as increases the probability of participation in the community forum by almost 27 percentage points.

These results are in line with the theory of self-fulfilling prophecies, because citizens participation in monitoring activities depend on whether they perceive themselves as influential in local government, which is a proxy for citizens' expectations of local government's accountability.

One concern with the results from Table 2.10 is the potential for a confounding effect. Participating in community monitoring activities could also affect citizens' perceptions of their influence on the local government.

If this were true and participation in community monitoring activities was crucial in changing citizens' perceptions of their influence vis à vis the municipal government, I should find that spending additional time monitoring a project increases the probability that a citizen will change his/her perception of influence. I take advantage of the fact there is a wide variation among those who spend time monitoring projects, with the coefficient of variation being $\frac{\hat{\sigma}}{\hat{\mu}} = 1.67$. I find that among this group, additional time spent in community monitoring activities has no effect on perception of influence.

Moreover, another empirical implication of the endogeneity theory is that those who participate in community monitoring activities and find the work of municipal government representatives unsatisfactory should have a lower perception of their influence. The fact that they were not able to increase politicians' accountability while participating in project monitoring result in updated priors regarding their levels of self-perceived influence. However, this is not the case, because 31% of those individuals believe they are influential, a figure not statistically different from the 34% found among those who evaluate the work of local officials as satisfactory. These numbers are substantially higher and statistically different than the 17% of citizens who perceive themselves as influential using the whole sample.

Table 2.10: When are Citizens Able to Engage with Governments to Monitor Project Performance?

	Time Spent in Community Monitoring Activities		Participation in Public Forum/s of Social Audit	
	(1)	(2)	(3)	(4)
Variables	Coef.	p-value	Coef.	p-value
Perception of Influence	7.62	0.001	0.264	0.008
Information Quality	6.55	0.009	0.159	0.000
Priority	4.51	0.160	0.191	0.080
Trust in Fellow Neighbors	-0.359	0.907	0.062	0.437
Ethnic Outsider	3.39	0.701	-0.105	0.415
Income	0.503	0.125	0.000	0.945
Controls				
Gender (Female=0)	5.50	0.066	0.051	0.544
Adult (25-40 years old)	7.25	0.028	-0.248	0.073
Middle Age (41-64 years old)	13.33	0.001	-0.235	0.105
Old Age (64+ years old)	13.11	0.076	-0.196	0.327
White	0.02	0.998	-0.226	0.151
Mixed	0.25	0.978	-0.198	0.203
Indigenous	7.45	0.476	-0.067	0.706
Complete Primary	2.30	0.523	0.105	0.233
Incomplete Secondary	5.86	0.116	0.186	0.133
Complete Secondary	3.96	0.318	0.192	0.096
Higher Education	12.50	0.024	0.241	0.064
Adj. R^2	0.15		Pseudo R^2	0.10
No. of obs	482		No. of obs	301

Notes: In columns (1) and (2) I use time spent in community monitoring activities as a dependent variable, which captures both citizens in control and treated communities. To construct this variable I use responses to "Did you devote any of your time to monitor the project?" with an answer scale of (1) No time, (2) Less than 10 minutes a week, (3) Between 10 and 30 minutes a week, (4) Between 31 and 60 minutes a week and (5) More than an hour a week. Using average time thresholds for each bracket (0, 5, 15, 45, 60) I compute this variable. *Priority* is a binary variable equal to 1 when respondents' report that the project was a priority for the community, and 0 otherwise. The regression is estimated using OLS with robust standard errors, adjusted for clustering at the community level. The results are qualitatively similar using ordered probit and probit models. In columns (3) and (4), I use participation in at least one community forum, which only citizens in treated community would be able to attend, as a dependent variable. The regression is estimated using a probit model with robust standard errors, adjusted for clustering at the community level. To compute the marginal effect, I evaluate all other variables at their means. The results are qualitatively similar using OLS, logit model, and when evaluating the other variables at 0 or 1 when computing the marginal effects.

The evidence suggests that once a person has decided to spend time monitoring the project, neither additional time spent in community monitoring activities nor their perceived performance of local officials impacts their perception of their own influence on the municipal government. As a result, it appears that perception of influence on government is not easy to change, and it is crucial to understand why citizens therefore decide to participate and spent time in community monitoring activities such as the ones encouraged by the CVA program.

To test (d.2), I assess whether an increased (decreased) participation in the CVA increases (decreases) project visibility. Using responses from participants in community forums in the CVA program I estimate the *participant multiplier effect*. For each participant, 4 $[2.4 - 5.42]$ other neighbors that did not participate in any of the public forums but were told about the CVA program and infrastructure project. This piece of evidence explains the micro-linkages between participation in the public forums and citizens knowledge about the infrastructure project being built in the community.

Part II

The evidence from Figure 2.14 and Table 2.10 above suggest that the CVA program improved politician's behavior, as evaluated by citizens. As argued above, there are two potential explanations:

(a) It becomes easier for politicians to identify pockets of corruption and inefficiency due to increase citizen's monitoring effort.

(b) Improvements in project performance becomes more valuable for politicians, as citizens are better able to recognize politicians' effort.

Table 2.11 presents the results from an examining of whether these theories have empirical validity. I estimate the regression in equation 2.5 and 2.6 using OLS,

where *OUTCOME* refers to the differential in overall performance and difference in politicians' performances[13]. In equation (5) *MONITORING* refers to time spent in monitoring activities, explained above. This variable is used to test potential explanation (a). In equation (6) *VISIBILITY* refers to the percentage of citizens in each community who know about the project. This variable is used to test theory (b).

$$OUTCOME_c = \alpha + \beta MONITORING_c + \varepsilon_c \tag{2.5}$$

$$OUTCOME_c = \alpha + \beta VISIBILITY_c + \varepsilon_c \tag{2.6}$$

Columns (6) and (7) from Table 2.11 assess the validity of potential explanation (a) using OLS and columns (8) and (9) do the same using metaregression. I use metaregression to account for the fact that variables for some communities are estimated with a small sample. Metaregression allows computation of a weight for each pair of projects based on their standard error, ameliorating the small sample problem within projects. I found that time spent monitoring projects cannot explain politicians' performance improvement in CVA projects. In particular, using OLS or meta-regression, average time spent monitoring activities by CVA project is not associated with changes in *Difference in Politicians' Performance*. Point estimates are small and have no statistical significance. Similar results are found for *Difference in overall performance* in columns (2) and (3) using OLS and in columns (4) and (5) using metaregression.

[13] *Difference in overall performance* refers to the difference between the proportion of citizens who evaluate the project performance as good or very good by pair of communities. *Difference in Politicians' Performance* refers to the difference between the proportion of citizens who evaluate the work of politicians as good or very good by pair of communities.

In addition, columns (6) and (7) from Table 2.11 assess the validity of potential explanation (b) using OLS and columns (8) and (9) using metaregression. The evidence from Table 2.14 suggests that the CVA program has stronger treatment effect on politicians' and project performance in communities where a higher percentage of citizens had knowledge of its existence. Using OLS, the point estimate indicates that increasing project visibility by 10 percentage points for CVA projects leads to an increase in *Difference in Politicians' Performance* by 3,8 percentage points. This effect is not only substantial but also statistically significant at the 93% significance level. In the case of *Difference in overall performance* in columns (2) and (3), the point estimate is stronger and significant at the 95% level. The estimated effect and its significance levels weakens when I estimate the coefficient using meta-regression in columns (4) and (5) as well as in columns (8) and (9).

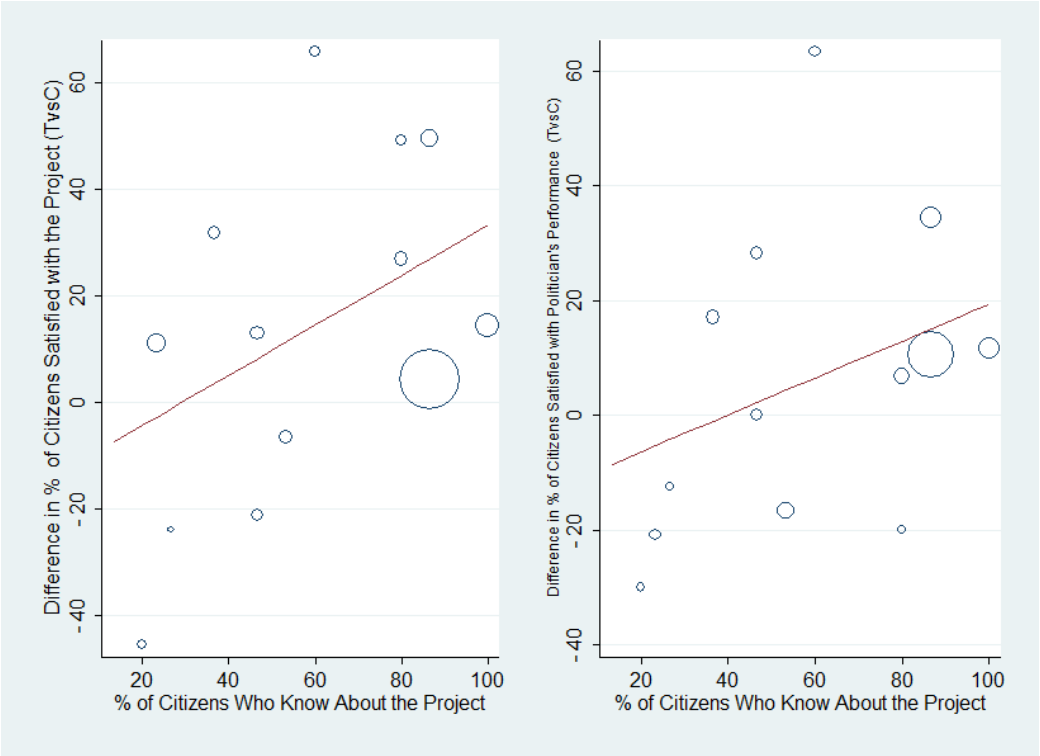
Table 2.11: Determinants of Behavioral Change in Politicians' and Project Performance

		Difference in Overall Performance				Difference in Politicians' Performance			
		OLS		Meta-regression		OLS		Meta-regression	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Measures	No of obs.	Coef.	p-value	Coef.	p-value	Coef.	p-value	Coef.	p-value
Time Spend Monitoring Projects	13	0.02	0.287	0.01	0.300	0.02	0.197	0.01	0.211
Project Visibility	13	0.64	0.048	0.47	0.136	0.38	0.069	0.32	0.194

Notes: Each cell in column (2), (4), (6) and (8) is the coefficient from a different regression. In columns (2), (3), (6), and (7) regressions are estimated using an OLS model with robust standard errors. In columns (4), (5), (8), and (9) regressions are estimated using meta-regression. The measure *overall performance* as well as *time spent monitoring projects* measure has been explained above. *Difference in overall performance* refers to the difference in overall performance by pair of communities. *Project visibility* refers to the percentage of citizens that know about the project in each community. *Difference in Politicians' Performance* refers to the percentage of citizens who evaluate the work of politicians as good or very good between the pairs of communities.

Figure 2.15 complements Table 2.11 and displays graphically the metaregression estimation. Although this procedure weakens statistical significance, there is evidence of stronger treatment effects from the CVA program on politicians' and projects' performance in communities where a higher percentage of citizens knew of its existence.

Figure 2.15: Effect of Project Visibility on Government Accountability and Project Performance



Notes: The figure shows a meta-regression. The size of each circle represents the weight of each project, and the line represents the estimate of the coefficient of project visibility, as measured by the percentage of citizens who know about the project, against the difference in the percentage of citizens who evaluate the work of local politicians as good or very good by pairs of communities.

The findings from this section support the theory of self-fulfilling prophecies in chapter 1. In communities where citizens believe they have little or no influence on governmental activities, citizens' participation in the CVA program and time spent monitoring projects are both lower. This in turn reduces the project's visibility, which in turn reduces politicians' incentives to improve performance. In communities where

citizens believe they are more influential, they participate in monitoring activities more often and give higher visibility to projects, which in turn motivates politicians to change their behavior and improve overall project performances.

2.6 Conclusions

This study examines the impact of the Citizen Visible Audit (CVA), a social audit program that aims to monitor the use of royalties in Colombia, on project performance. The results suggest that the CVA program is successful in improving project performance in almost all performance measures I compute, including one aggregate measure. This suggests the program has a highly positive impact in improving citizens' satisfaction with the project as well as boosting citizen's assessment of the project's execution quality. However, treatment effects are highly heterogeneous across communities.

I developed a theoretical framework for examining social audits that allows me to identify potential bottlenecks hindering the CVA program. My theory contributes to explain both (a) under what conditions citizens decide to monitor a project and (b) under what conditions governments facilitate citizen engagement and become more accountable.

In the case of (a), I found that neither the opportunity cost of participation nor collective action bottlenecks can explain variations in citizens' willingness to participate in monitoring a project. The evidence suggests that heterogeneity in citizens monitoring efforts is the result of two major bottlenecks: (i) citizens knowledge about the program and how to monitor it and (ii) citizens' beliefs regarding the probability that participation in monitoring activities will lead to improvements in

project performance.

In the first case, I find that in some communities citizens did not know the infrastructure program existed, whereas in others, they knew the project existed but did not have access to information on how to monitor it. These information shortfalls prevent citizens from taking an active role in CVA community forums and community monitoring activities in general.

In the case of beliefs, the evidence suggests that perceiving oneself as having influence on local government is crucial for citizen's decisions regarding whether to spend time in community monitoring activities. Citizens may refuse to take advantage of the opportunity to monitor government and providers if they believe that after paying the opportunity cost of monitoring, coordinating their actions with fellow neighbors, and identifying projects' problems, politicians and providers will not be responsive to these concerns.

Regarding (b), there is supporting evidence of stronger treatment effects of the CVA program on politicians' performance on communities where the project enjoyed higher visibility among the population.

Overall, the results offer though preliminary, the results give supporting evidence for what I have called a theory of *self-fulfilling prophecy in service delivery.* In chapter 1, I argue that there is a channel, often not accounted for in literature, through which the effect of community monitoring interventions could be diminished. In particular, I argue that citizen's low expectations of accountability by leaders, officials, or service providers' may generate low community participation in monitoring a project. As a result, project visibility is reduced, because fewer citizens talk about this program to

other fellow citizens. Politicians and providers, in turn, observing weak community monitoring participation and low project visibility, may have fewer incentives to spend time and effort improving project performance.

Future work should aim to collect more and better data to allow for a better understanding of the pathway of change operating in social audits and community monitoring interventions in general.

Chapter 3

Exploring External Validity

3.1 Introduction

Corruption and inefficient allocation of resources in service delivery are widespread in low and middle income countries (Olken and Pande, 2012). There is increasing evidence that corruption holds back countries' economic development and erodes their citizens' quality of life (Mauro, 1995; Singer, 2013; Svensson, 2005). Millions of people around the world encounter administrative corruption in their daily interactions with public services. Using a 0-100 scale on perceived levels of public sector corruption, only a third of the 176 countries covered in the Transparency International Corruption 2012 Index score above 50. The World Bank Institute estimates that total bribes in a year are about 1 trillion USD (Rose-Ackerman, 1999), making corruption around 3 percent of the world GDP (Svensson, 2005). Bribes are used to prevent public officials from doing their work. For instance, under the presidency of Fujimori in Peru, there is direct evidence, in the form of signed receipts, that politicians and judges received bribes ranging from 3,000 to 50,000 USD and the media received as much as 1.5 million USD per month for turning a blind eye to

government malfeasance (Mcmillan and Zoido, 2004).

Corruption is not only a problem of a few high level public officials, but there is evidence of a more widespread problem in the public sector. Gorodnichenko and Sabirianova (2007) estimate the aggregate amount of bribes collected by low and medium level public officials in Ukraine to be between 460 and 580 million USD, about 1 percent of its GDP. Administrative corruption imposes a heavy burden on citizens' and firms' time and resources. Olken and Barron (2009) estimate that 13 percent of the cost of a truck driver trip in Indonesia is allocated to pay bribes for police officials they encounter on their routes. Svensson (2003) finds that firms in Uganda pay 8 percent of their production cost in bribes. Corruption creates discontent with public services, undermines trust in public institutions (Sacks and Larizza, 2012; Singer, 2013), and stifles business growth and investment. Khwaja and Mian (2005) find that politically connected firms receive substantially larger loans from government banks in spite of having a 50 percent higher default rates.

Resources needed to improve equality of opportunities and provide services for citizens are lost every day due to pockets of corruption and inefficiency (World-Bank, 2003), which in turn, results in inadequate provision of key services. Often, it is the poor and the vulnerable who suffer the most (Olken, 2006; Sukhtankar, 2012; World-Bank, 2003). A landmark study in Uganda finds that only 13 percent of the public funds the central government assigned to the school system reach its intended destination (Reinikka and Svensson, 2004, 2011). Similarly, leakages are also a problem for Tanzania, where elected officials are the recipients of more than half of the total amount of subsidized fertilizer's price vouchers (Pan and Christiaensen, 2012). In Indonesia, village officials hide their corruption by deflating quantities, i.e., they claim to procure enough rock, sand, and gravel to make a road that is 20

cm thick, but instead build a road that is only 10 cm or 15 cm thick. Since the roads they build are thinner than official engineering guidelines, they will not last nearly as long, and will need to be replaced sooner (Olken, 2007, 2009). In India, the lack of monitoring and accountability implies that one quarter of all the teachers in public schools and more than a third of nurses and doctors are absent from their duties (Chaudhury et al., 2006). Corruption also impacts service delivery in Brazil. Municipalities where corruption in education has been detected have test scores that are 0.35 standard deviations lower than those without corruption, as well as higher dropout and failure rates. Moreover, teachers in corrupt municipalities are 10.7 percentage points less likely to receive pedagogical training and less likely to have a computer or science lab (Ferraz, Finan and Moreira, 2012).

This chapter provides the results of a synthetic review of community monitoring interventions aimed at improving service delivery. I refer to service delivery as the process through which basic services, such as education, health, and infrastructure are delivered to communities[1]. I measure service delivery outcomes as quality of the service. For example, if the goal of the intervention is to improve education attainment, improvement in student's test scores, management practices and provider's absence rate are the outcomes of interest.

In short, I found suggestive evidence of positive impacts of community monitoring programs on services provision. Section 2 provides some background to the research question by defining community monitoring, reviewing its relationship with community participation, and overviewing the type of community monitoring interventions aimed at promoting service delivery. Section 3 lays out the theory of change for

[1]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.

community monitoring interventions. Section 4 sets out the objectives and questions addressed in this review. Sections 5 and 6 describe the methodology of the review, dealing with inclusion criteria and search strategy respectively. Section 7 describes the included studies. Section 8 outlines the interventions to improve service delivery identified in this review: information campaigns, scorecards and social audits. Section 9 reports the results of the meta-analysis of included impact evaluations. Section 10 presents a causal chain analysis. The final section serves as a conclusion and provides recommendations for moving forward with the research agenda.

3.2 Description of the Intervention

The idea that community members have incentives to monitor providers and demand for better services (Stiglitz, 2002) led practitioners to believe that allowing communities to have monitoring power over providers could be beneficial in the short and in the 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 change political norms, establishing a transparent and accessible channel of communication for the community to provide feedback to providers and politicians on a regular basis.[2]

This idea sets the stage for a move to encourage governments in developing countries to become accountable to their own citizens, in an attempt to reform institutions from the bottom up. As a consequence, over the last two decades programs

[2]I used the term politicians or policy makers interchangeably to refer to the high-level elected officials or civil servants responsible for carrying out legislative, regulatory and supervision responsibilities, including majors, governors and congressmen. On the other hand, I refer to providers as both: a) the individuals involved in the administration and delivery of services and b) the frontline providers, such as policemen, teachers, public work employees and social workers. Finally, I used the terms citizens and users of services interchangeably to refer to the people who can potentially receive and benefit from a service delivery services, such as health, education, social protection, infrastructure, and police.

aimed at encouraging community monitoring have been introduced in countries across continents and cultures, including Albania, Argentina, Brazil, Cambodia, Cameroon, Colombia, Kenya, India, Indonesia, Malawi, Philippines, South Africa, Uganda, among many others.

This idea was operationalized by the introduction of community monitoring interventions (CMIs), often referred to as social accountability mechanisms. These programs can be broadly defined as interventions where the community is given the opportunity to participate in the process of monitoring service delivery, where monitoring means being able to observe, assess providers' performance and provide feedback to providers and politicians.

The Association for the Empowerment of Workers and Farmers in India was the first to introduce a social accountability initiative, through social audits in the early 1990s (Maru, 2010).[3] Association workers read out loud government accounts and expenditure records at community meetings, and then invited villagers to testify to any discrepancies between official records and the villagers' personal experience. Since then a range of different variations of community monitoring have been implemented, with information campaigns, scorecards/citizen report cards, social audits, and grievance redress mechanisms being the four major categories of such interventions. These four sub-categories of community monitoring share two common elements: 1) a clear objective of reducing corruption and improving service delivery, and 2) using encouragement of the community to monitor service delivery as a key

[3]The word 'audit' is derived from Latin, which means 'to hear'. In ancient times, emperors used to recruit persons designated as auditors to get feedback about the activities undertaken by the kings in their kingdoms. These auditors used to go to public places to listen to citizens' opinions on various matters, like behavior of employees, incidence of tax and image of local officials (for Good-Governance, 2005).

instrument. Table 3.1 below summarizes the key components of these interventions.

Table 3.1: Interventions Aimed at Increasing Civic Participation in Monitoring Public Officials and Providers

Policies	**Description**
Information Campaign	These are efforts to inform citizens about their rights to services and quality standards. In particular, it can include information on the importance of the service, on providers' performance, and on how to monitor providers.
Scorecard	It is a quantitative survey that assesses 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 cross-check 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 to influence service delivery and give feedback on government programs and services, including complaint hotlines, informal dispute resolution mechanisms, and courts (Ringold et al., 2012).

Information campaigns are one of the most common interventions to encourage participation and interest in monitoring. It usually involves information on the benefits of the service to be delivered (health, education, police, etc.) and the current state of the service in the community. The information can be provided door to door, in public gatherings aided by local leaders, through radio, newspapers and other means. Keefer and Khemani (2011), for example, study the impact of having access to community radio programs on the benefits of educational attainment in

Benin. Information campaigns can also include information on how to monitor providers. For example, Banerjee et al. (2010) analyze a randomized evaluation of three interventions to encourage beneficiaries' participation in India's educational system. Prior to conducting the interventions, information is provided on the state of educational performance. Then, implementation team a) provide information on existing institutions, Village Education Committees (VECs), to monitor schools, b) train community members in a testing tool for children, and c) train volunteers to hold remedial reading camps for disadvantage children.

Scorecards[4], often referred as citizen report cards, are another way in which to encourage citizen to participate in monitoring service delivery. The rationale is that by giving a voice to citizens, they will be encouraged to demand better services. For example, Björkman and Svensson (2009) analyze the impact of a scorecard community monitoring intervention on primary health care in Uganda. For the intervention, a non-governmental organization (NGO) facilitated village and service provider's staff meetings in which members of the communities discussed baseline information on the status of health service delivery relative to other providers and the government standard. Community members were also encouraged to develop a plan identifying key problems and steps that providers should take to improve health service delivery. 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.

[4]Scorecards for health services were pioneered in Malawi in the early 2000s by Care International. This intervention followed the spirit of individual "citizen report cards", which were first introduced in Bangalore, India in 1993. The citizen report card revealed low levels of public satisfaction with the performance of service providers. The findings were widely publicized through the media, which created pressure among public officials to organize workshops and meeting with local civic groups and NGOs. Increased public awareness on government inefficiencies and other related concerns triggered the formation of more than 100 civic groups in different parts of India, as well as the launch of many campaigns for transparent public management (Bhatnagar, 2003).

Social Audits involved interactions not only between citizens and providers, but also with politicians. This is the case of Colombia's Citizens Visible Audit (CVA), chapter 2 of this dissertation. The social audit involves: (a) disseminating information through radio, newspapers and local TV about the CVA program in the neighborhoods where the project takes place; (b) introducing the infrastructure project to the community in a public forum. Citizens are told about their rights and entitlements, including the activities they can do to monitor the project and the responsibilities of the executing firm. A group of interested beneficiaries is constituted and trained to carry out community monitoring activities; (c) periodical public forums are held, bringing together local authorities, neighbors, and representatives from the firm that carry out the specific project. The state of the project is explained in detail to the community, who can voice concerns and recommendations. Commitments are made by the firm, the local government, and project supervisor to solve the problems that may arise during the construction of the project. These commitments are monitored by the community, the facilitators from the central government and the project supervisor. If the problem persists, administrative complaints are submitted to the Supreme Audit Body in the central administration; (d) on a regular basis the beneficiary group monitors the project and collects information on whether commitments are being honored and any other new problem that may arise; (e) the finalized project is presented to the community before making the final payment to the executing firm, and the audit results are shared with all interested and concerned stakeholders.

Social Audits can also involve citizens as decision makers. In this case citizens have the power to make actual decisions over the project. The extent to which the community has control over the decisions, however, varies. An example of a CMI where citizens had decision power is the Kecamatan Development Program (KDP) in

Indonesia (Olken, 2007). This program funded projects in about 15,000 villages each year. Each village received an average of 8800 USD, which was often used to surface existing dirt roads. To control the use of funds, checks were built into KDP. First, funds were paid to village implementation teams in three installments. To receive the second and third payments, the teams had to make accountability reports at an open village meeting. Second, each project had a 4 percent chance of being audited by an independent government agency. The study introduced two anti-corruption strategies: enhancing community participation and increasing government audits. To enhance community monitoring, invitations to the community meetings were randomly distributed throughout the village. It is important to note the community decides how to allocate the funds before monitoring the project, which differentiates it from studies on CMIs described above.

Grievance Redress Mechanisms (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 programs and services. Such mechanisms include complaint hotlines, informal dispute resolution mechanisms, and courts (Ringold et al., 2012).

3.2.1 Relevance of this Review

Community monitoring interventions have widespread acceptance as a tool to improve transparency and accountability by all the major players in the practitioners' world: governments, NGOs, and the donor community. Increasing citizen participation in the government is the main objective behind Open Government Partnership (OGP), a global consortium of governments. Through OGP, more than 50 countries

around the world have already agreed upon different goals related to transparency and citizen participation. Moreover, international aid agencies increasingly require development projects to include 'beneficiary participation' components. Over the last decade the World Bank alone has channelled 85 billion USD to local participatory development (Mansuri and Rao, 2012).

The United Nations has set increasing citizen participation as their main strategy to achieve good governance and human rights (UN, 2008), and NGOs with a focus on increasing government accountability through citizen participation continue to expand around the globe, managing increasing amounts of resources. For instance, Transparency International has an annual budget of 36 million USD, which is used to advocate for increasing citizen engagement as a necessary step for development (Transparency International, 2013). Other examples of NGOs are Twaweza and the Affiliated Network for Social Accountability (ANSA). Twaweza engages in building citizen capacity to monitor governments and foster their accountability across East Africa and has an annual budget of 17 million USD. ANSA is currently operating in East Asia and the Pacific, South Asia, Africa, Middle East and at the global level to support civil society organizations in their efforts to monitor governments in service delivery and to build demand for public accountability.

Finally, through the newly created Global Partnership for Social Accountability (GPSA) a coalition of donors, governments and civil society organizations aim 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 to 125 million USD over the next seven 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.

From a theoretical perspective, as I highlighted above, there are no clear predictions as to what should be the impact of these programs. Some authors have found reasons to expect CMIs to have a positive effect on improving service delivery and reducing corruption (Stiglitz, 2002), but others have argued that successful implementation of CMIs might prove more difficult than expected. While a number of empirical studies have been conducted in recent years, there is still no clear answer regarding the impact of community monitoring programs.

Thus, whether CMIs actually affect the behaviour of beneficiaries, providers and politicians and in turn reduce corruption and improve access to and quality of service delivery is still an open empirical question. Reports from USAID acknowledge that lack of systematic evidence limits our ability to make precise claims regarding the relationship between CMIs, corruption and service delivery outcomes (Brinkerhoff, 2006). In addition, we still know little about the mechanisms through which these interventions have an impact (or lack thereof).

There are, however, several non-systematic reviews on related issues. For example, Mansuri and Rao (2012) review the evidence on the effectiveness of local participatory programs on an array of outcomes, among them, service delivery. The study focuses mostly on large-scale interventions such as Community Driven Development (CDD), and Community Driven Reconstruction (CDR). They find that, on average, results are below the expectations for these programs, suggesting failure to build cohesive and resilient organizations to pressure the government may be a reason for this. In particular, they argue both local and national context may be a key factor to determine effectiveness, in part because not all communities

have a stock of social capital that can be readily harnessed though a participatory intervention. Finally, they argue that induced participatory interventions work best when they are supported by a responsive state and local accountability institutions are robust.

Moreover, Hanna (2011) and Olken and Pande (2012) review studies of interventions aimed at reducing corruption. However, they do not provide comprehensive reviews of the literature on effects of community monitoring and use narrative methods of synthesis rather than meta-analysis. Ringold et al. (2012) review the effects of CMIs on human development, and while it is relatively comprehensive, it is a narrative review rather than a systematic review. It identifies some key impediments to the successful implementation of CMIs: a) information asymmetries between citizens and providers, b) individuals may not use the opportunity to influence service providers, c) providers that are not amenable to change, and d) fragmented civil society and weak media.

Similarly, Devarajan, Khemani and Walton (2011) review interventions aimed at strengthening the role of civil society in service delivery and government accountability, focusing on Sub-Saharan Africa. The review, which is not systematic, finds preliminary evidence of positive effects of building organic participation, supported by existing political and civil society structures, on service delivery and government accountability. The findings are mediated by the local context, as in communities where clientelism and rent-seeking was widespread, civic participation failed to have an impact on service delivery and government accountability.

Gaventa and Barrett (2012) perform a meta-case study of 100 interventions aim to increase citizen engagement in service delivery. The search for literature was

limited to studies done at the Institute of Development Studies between 2000 and 2011. 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 percent of those 153 outcomes were positive and 45 percent 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.

Finally, high quality primary studies find what at first appears to be contradicting evidence regarding the effect of CMIs on service delivery outcomes. Björkman and Svensson (2009) find that community scorecards in Uganda significantly increased the quality and quantity of primary health care provision. Banerjee et al. (2010), however, find the opposite result when testing the effect of an information campaign in India. They report that neither giving citizens information on how to use existing institutions to monitor schools nor training them in a testing tool to monitor children's learning had any statistical impact on children's learning performance.

To date, no reviews have been conducted on the effect of CMIs on service delivery outcomes and the inconclusiveness found in theoretical and empirical work described above highlights the need for systematic evidence on the effects of CMIs. The existing reviews provide some suggestive evidence of the effects of CMI, but come to different conclusions in a hotly debated, policy important field. This synthetic review aims to shed light on this debate by providing a systematic literature search, together with a comprehensive and unbiased synthesis of the existing evidence.

3.3 How Can Community Monitoring be Fostered?

Figure 3.1 presents a stylized theory of change of a typical community monitoring program, clarifying the mechanisms through which the program is expected to have an impact on corruption and service delivery. 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 means as possible, such as radio, newspapers, door to door campaign, 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 other fellow community members to collectively pressure providers and politicians to improve the quality of the service through monitoring activities. To nudge citizens into monitoring providers, CMIs usually include activities to build the capacity of beneficiaries to monitor providers. For instance, the social audit in Colombia (see chapter 2) provides information about the contractual obligations of the provider and ways for citizens to detect problems and to identify people who are in charge of receiving inquiries about the project.

Empowered with information from building block 1 and/or 2, 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 can include meetings between providers and citizens. Participation is expected to be mediated by the strength of the community

to act collectively. Communities with a history of grassroots participation are expected to organize more rapidly and more efficiently. As an organized group, citizens can take turns to visit the place where the service or project takes place (e.g. school, construction site, hospital) and collect information on its problems (e.g. absenteeism, use of low quality inputs in the construction process, unresponsive front-line providers). Citizens can then use their voice to contact providers (building block 6) and/or elected officials (building block 4) to file complains about the service and provide valuable 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 neighbors that did not take part in monitoring activities (building block 5), to increase visibility of the community monitoring intervention and pressure on providers and politicians. Finally, the independence and strength of the local media is assumed to mediate the visibility of the project (Reinikka and Svensson, 2004; Ringold et al., 2012).

Citizens' participation in the program reduces the cost of monitoring front-line providers for politicians and managers. Citizens' monitoring activities also increase both: visibility and citizens' ability to recognize whether elected officials are making an effort to reduce corruption and improve service delivery. As a result, it is more valuable for politicians to achieve better results.

These conditions provide incentives for politicians and policymakers to put more pressure on providers to improve service delivery (building block 7). By facing formal sanctions by politicians and/or informal sanctions by citizens, providers are assumed to be more likely to exert effort (building block 7).

While the description above fits different types of interventions, some elements are idiosyncratic. Below I describe two additional components of social audits and scorecards respectively. 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 changing their actions and improving 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 produce new information for providers and citizens on how to use and manage the service in a more efficient way.

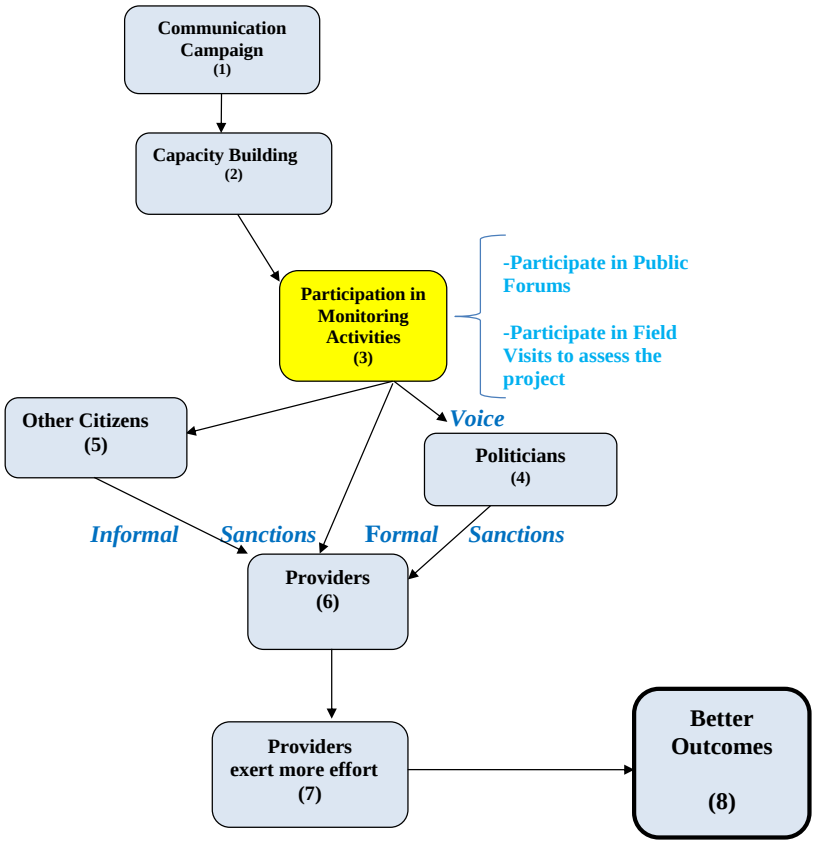
Social Audits also have an additional component in the form of public forums, where representatives from the local government, representatives from the executing firm, the central government, and the community are present. They allow citizens to make their voice heard by local officials and providers, and reduce the time and effort citizens would need to invest to get an appointment with these officials. The public forums also reduce the cost for central government representatives to be heard by local officials. Finally, the social audits reduce the cost of local officials to take actions to solve problems that arise during the delivery of service, such as lack of planning, lack of resources to finish the project, and acts of corruption. The symbolic act of the public forum may also signal to politicians and providers the importance of performing well on this project, as citizens are paying extra attention.

There are several empirical implications from this overall theory:

- The CMIs will increase the quantity and the degree to which citizens are involved in monitoring providers.

- As a result of the CMIs, politicians and providers will exert more effort and improve their performance.
- The CMIs will reduce the probability of inefficiency and corruption.
- The CMIs will improve access and quality of the service provided.

Figure 3.1: The Pathway of Change for Community Monitoring Interventions



Notes: This figure shows the mechanism through which the CMIs program is expected to have an impact on service delivery.

However, in order for this path of change to accurately describe the mechanisms through which a CMI impacts on service delivery, several assumptions need to hold.

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, they should believe the program has the potential to be successful. In the case of providers and politicians, they need to gain popularity, receive higher salaries and/or social recognition for their responsiveness. If these assumptions are not met, the underlying program theory of the CMI breaks down and this may prevent CMIs from having an impact on service delivery. In particular, they can affect citizens' decision of whether to monitor government activity and governments' willingness to facilitate citizen engagement and become more accountable. Below we present the bottlenecks as well as its empirical implications.

3.3.1 Competing Theories

Civic Participation Failure

One potential concern with CMIs is citizens' failure to participate in monitoring activities (building block 3). We group under this heading many bottlenecks that prevent citizens to participate in monitoring activities, which in turn reduces the potential impact of the program. 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. We identify six potential bottlenecks:

Table 3.2: Competing Theories - Civic Participation Failure

Bottleneck	*Description*	*Empirical Implications*
Information Gaps	Scholars and policymakers have long argued that programs often fell short of their expectations because of information problems (Ringold et al., 2012). In the case of the CMIs 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).	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 described as rational inattention. As a consequence, introducing new information does not always lead to new beliefs or changes in behavior.	If citizens' opportunity cost of paying attention to the information is high or they lack of attention span, their probability of participation will decrease[5]
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 the opportunity cost of participation is high, the probability of participation will be lower.
Collective Action Failure	Scholars have emphasized the collective action problems that can arise in the presence of a non-excludable 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' Believe Success is Possible	Citizens may refuse to take advantage of the opportunity to influence politicians and providers if they believe the chances of success are low (See chapter 1).	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; 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 program.	If the CMI is captured by local elites its effectiveness will be substantially reduced.

Politicians and Providers Accountability

Under this heading we present potential barriers that refer to a lack of responsiveness on the part of the politicians and providers. The literature cites many reasons why politicians and providers may not be accountable to their principals (building block 4 and 6). Below we identify three bottlenecks.

Table 3.3: Competing Theories - Politicians and Providers Accountability

Bottleneck	*Description*	*Empirical Implications*
Unresponsive Politicians	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.	If the community is not pivotal, 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' behavior. If communities can impose a credible threat of informal social sanctions to unresponsive providers, the probability of a change in behavior 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.[6]
Weak Institutions Failure	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; Adsera, Boix and Payne, 2003) In particular, CMIs will work best when a robust institutional framework with clear separation of powers supports them.	In communities where the media industry is not captured by the government, CMIs will be more effective in improving politicians' performance. In communities where more citizens know about the project or service being monitored, politicians' performance will be higher.

3.4 Objectives of the Review

This paper aims to provide a synthetic review on available rigourous impact evaluations that assess the effects of CMI interventions on service delivery in low and middle income countries (LMICs). This review contributes to answer these questions by synthesizing evidence from quantitative studies and it is consistent with Campbell and Cochrane Collaboration approaches to systematic reviewing. The review is also informed by theory-based impact evaluation (White, 2009) using the theory of change (Figure 3.1) as the framework for the review, to guide the types of studies included, data collection and analysis. To ensure the review is adequately oriented towards both reporting effects and explaining the reasons for them, we will synthesize effects along the causal chain, including qualitative evidence when appropriate (Snilstveit, 2012; Snilstveit, Oliver and Vojtkova, 2012). For quantitative synthesis, where studies are judged sufficiently similar to do so, we will use meta-analysis to pool study effects.

The review aims to answer the following questions:

(1) What CMIs in low and middle income countries have been rigourously studied in terms of their effect on quality of service delivery?

(2) What did the studies find in relation to the effect of CMIs on the quality of service delivery in LMICs?

(3) Do the studies allow us to make claims about moderating mechanisms?

3.5 Criteria for Including Studies in the Review

3.5.1 Types of Studies

To review question (2) studies eligible for inclusion in effectiveness synthesis include impact evaluations based on experimental design (where randomized assignment to the intervention is made at the individual or cluster level), quasi-experimental designs (including controlled before and after (CBA) studies with contemporaneous data collection and with two or more control and intervention sites, regression discontinuity designs, interrupted time series studies (ITSs)) and ex post observational studies with non-treated comparison groups and adequate controls for confounding.

For quasi-experimental studies and observational designs with comparison groups, eligible studies must use adequate methods of analysis to match participants with non-participants, or statistical methods to account for confounding and sample selection bias. Appropriate methods of analysis to match participants and non-participants include propensity score matching (PSM) and covariate matching. Appropriate methods of analysis to control for confounding and selection bias include multivariate regression analysis using difference-in-differences (DID) estimation, instrumental variables (IV) or Heckman sample-selection correction models.

Studies that do not control for confounding factors using these methods, such as those based on reflexive comparison groups (pre-test post-test with no non-intervention comparison group), were excluded.

Studies eligible for inclusion in the synthesis of evidence answering question (3) include not only quantitative studies with the criteria mention above but also any qualitative studies of the interventions evaluated in the effectiveness studies.

3.5.2 Types of Participants

The review includes CMIs in either low or middle income countries at the time the intervention was carried out. To assess whether a country is low, middle or high income I followed the World Bank classification method. For example, for interventions carried out in 2011, to qualify as a low income group gross national income (GNI) per capita should be 1,025 USD or less; middle income, 1,026 USD - 12,475 USD; and high income, 12,476 USD or more. I include all CMIs in low and middle income countries. This review excludes CMIs in high income countries.

3.5.3 Types of Interventions

I included community monitoring interventions where the community is given the opportunity to participate in the process of monitoring service delivery. For this review, monitoring means being able to observe, assess providers' performance and provide feedback to providers and politicians. To be included interventions need to: 1) have a clear objective of reducing corruption and improving service delivery; 2) use encouragement of the community to monitor service delivery as a key intervention instrument; 3) fall into one of the following four intervention categories: information campaigns, scorecards/citizen report cards, social audits and grievance redress mechanism. These interventions have a common theory of change that exactly addresses the question of interest: whether programs that encourage community monitoring reduce corruption and improve access to and quality of service delivery.

Other interventions may include community monitoring as part of a different intervention[7]. For instance, Community Driven Development Interventions (CDDs),

[7]Access to information laws is a legal framework for public provision of information (Ringold et al., 2012). There are many laws that can potentially improve citizens' abilities to monitor service delivery, i.e. voting rights, laws that allow schools or hospitals to have users groups, the creation

Community Driven Reconstruction Interventions (CDRs), participatory budgeting, and school based management were only included if they have a clear community monitoring component. In that case, depending on the monitoring component, I classified them as information campaigns, scorecards, social audits or grievance redress mechanism. The study from Olken (2007) in Indonesia is a case in point. The monitoring program, in this case the social audit, is embedded in a larger intervention, a CDD. I included this type of interventions in our review. However, there are other CDDs and CDRs where there is no monitoring component. For instance, Casey, Glennerster and Miguel (2012), who study the impact of a CDR program in Sierra Leona, was excluded from this review. The reason is that the theory of change for these types of interventions is completely different than for CMIs. Even further, the objectives of these interventions are also different. A similar argument could be made about participative budgeting. As a result, I also excluded those interventions from this review when there is no community monitoring component.

3.5.4 Outcomes Measures

I included studies that use a range of service delivery outcomes. In particular, I looked at two ways of measuring quality of the service. Below I provide specific examples:

- *Quality of Services.* I used measures of:
  - *Improvement in prevalence condition.* For example, Björkman and Svensson (2009) capture the effect of the CMI on infant weight. Additional

of the ombudsman figure, among many others. The theory of change for these interventions is different from the one I develop for CMIs and studies assessing such interventions on their own will be excluded. Such interventions are not defined as community monitoring unless they include an additional component aimed at encouraging community monitoring. Therefore, studies assessing the effects of information campaigns which aim to induce citizens to monitor the implementation of those laws fall under our definition of community monitoring. Since these studies were classified as information campaigns they were included in the review.

measures in the health care sector could be mortality rates as well as disease prevalence in general. In Banerjee et al. (2010), there is information on student's reading and analytical ability. Additionally, information on test scores would be in this category.

  - *Perception Measures of Overall Quality.* An imperfect way to deal with the fact that it is very difficult to measure prevalence is to rely on citizen's perception measures of the overall quality of the service received.

Studies that include at least one of these outcomes are included in the systematic review. Below I specify which intermediate outcome I reviewed in question (3). These outcomes include whether participants contribute to monitor the service or project and the providers' performance.

- *Citizen's Participation in Monitoring Activities.* This could be measured by the percentage of citizens that contribute to the monitoring process.

- *Efficiency in the Provision of the Service.* I measured efficiency in two different ways:

  - *Absence Rate.* Percentage of primary service providers (i.e. teachers, nurses, etc) that are absent.

  - *Management Practices.* This refers to the quality of management practices in service delivery. The specific measures are dependant upon the specific sector and study. For example, in health, average waiting time to get the service, quality of the equipment used, and information received by patients are measured. In educational settings, student-teacher ratio and textbook-student ratio, among other, are measured. In infrastructure programs, the proportion of resources that could be accounted for when engineers inspect the construction is a way to assess management practices. For example,

in Olken (2007) this is measured by comparing the researcher's estimate of what the project actually costs to what the village reported it spent on the project on an item by item basis.

3.6 Search Strategy and Methods of Analysis

For this 'synthetic review' I searched 3 electronic databases: JSTOR, ScienceDirect and Social Sciences Citation Index.[8] In trying to avoid the bias against unpublished and non-English literature I searched in Google, Google Scholar, REPEC-IDEAS, NBER, Global Development Network, and 3ie database. I used broad keyword/topic combinations: community monitoring, social accountability, corruption, intervention, in a format applicable to each database.

I also searched using several additional methods.

- Screened the references of included studies and existing reviews for eligible studies;
- Ran citation searches of all included studies;
- Searched in conference programs and websites of key institutions; such as the World Bank, UNDP Governance Projects, Asian Development Bank, African Development Bank, Inter-American Development Bank, Research centres and

[8]This paper is part of a bigger project I am leading. I was awarded a grant from 3ie to conduct a systematic review on this topic. For the systematic review, the search will be comprehensive and involved a team of researchers. In particular we will perform searches in the Social Science Citation Index, JSTOR, International Bibliography of Social Science, EconLit, Citas Latinoamericanas en Ciencias Sociales y Humanidades (CLASE), Plataforma Open Access de Revistas Científicas Electrónicas Españolas y Latinoamericanas (e-Revist@as), Red de Revistas Científicas de América Latina y el Caribe, España y Porgugal (REDALyC), African Journals Online (AJOL), Scopus, the British Library for Development Studies (BLDS), PAIS (Public Affairs Information Service), WPSA (Worldwide Political Science Abstracts), IPSA (International Political Science Abstracts), CIAO (Columbia International Affairs Online), ABI/Inform (Ebsco), ELDIS, CAIRN and Google Scholar. Then, we would be able to test whether the extended search translate into more studies and, whether those studies change the results established in this chapter.

networks, as JPAL, MIT, IEN, Institute of Development Studies; regional, national and local non-governmental organizations; and,

- Outreached to subject-matter experts, and practitioners.
- Searched for sibling articles, doing a citation tracking of all included studies to identify any sister papers/qualitative studies, and conducting targeted searches at implementing agencies websites, Google and databases using the intervention name.

I made inclusion decisions by examining titles and abstracts. Those papers that passed the first screening, were studied in detail to assess whether they should be included based on the criteria laid above. I used Hombrados et al. (2013) quality assessment tool to assess whether the studies met the minimum quality standard. From this subset of studies I collected several groups of information on the study and about the intervention. From the study, authors, date, publication status, type of publication and language. About the intervention, I collected data on country and area, dates of the intervention and type of intervention.

I conducted statistical meta-analysis on a collection of studies with common outcomes construct (potentially measured in different ways), and similar comparison condition. For this, I extracted comparable effect size estimates from included studies. Since most experimental and quasi-experimental studies used in impact evaluation in economics and political science do not report the information required to calculate standardized mean differences, I converted all estimates to difference in rates. This way outcome measures were comparable across studies. Thus, a difference in rate greater than zero indicates an increase in the outcome under the intervention as compared to the comparison. A difference in rate less than zero indicates a reduction under the intervention as compared to the comparison. A

difference in rate equal to (or insignificantly different from) zero indicates no change in outcome over the comparison. I included one effect estimate per study in a single meta-analysis. I followed the statistical transformations for calculating difference in rates from matching-based and regression studies provided in IDCG (2012).

Finally, I conducted a synthesis of causal chain information contained in the studies. This causal chain analysis sought to decipher how interventions unfolded, if the process was consistent with theory, and, to the extent possible, what causal mechanisms were at work when interventions succeeded.

3.7 Search Results

3.7.1 Description of Included Studies

My initial search produced 1250 potentially relevant studies. Most were discarded because they did not match the subject or aims of the review. Out of the 1250 potentially relevant studies, 20 were review in full against the inclusion and exclusion criteria. Among those some did not meet the inclusion criteria, and 2 papers did not passed the quality control (Isham, Narayan and Pritchett, 1995; Narayan, 1995). Only 10 studies met the inclusion criteria, reporting on 10 interventions.

The interventions evaluated in the included studies fell into three categories discussed above: information campaigns, scorecards and social audits. While the interventions were not identical, I elaborated in the this below, they were similar enough for the purpose of this paper. Most included papers evaluate interventions that target entire communities, though the characteristics of the community vary significantly. Sample size, duration of the intervention and specific outcomes vary from study to study.

3.7.2 Study Designs

All included studies met the minimum quality requirements based on Hombrados et al. (2013) quality assessment tool. They thus all included comparison groups that did not participate in the intervention.

Out the 10 studies, 8 used randomized treatment assignment as their identification strategy (Andrabi, Das and Khwaja, 2013; Banerjee et al., 2010; Barr et al., 2012; Björkman and Svensson, 2009; Lieberman, Posner and Tsai, 2013; Olken, 2007; Pandey, Sangeeta and Sundararaman, 2009; Pradhan et al., 2013). All of the randomized field experiments reported no statistical difference between treatment and control communities in relevant pre-treatment characteristics.

The other 2 studies were not randomized, and thus employed various methods to control for potential confounders. In chapter 2 I used matched communities and matched individuals to control for selection of programs to be audited in Colombia. Treatment effects were then estimated by controlling for individual-level attributes with multiple regression models and standard errors were modified to account for clustering at project level. Keefer and Khemani (2011) relied on within-commune variation in village access to radio stations. If this is exogenous to observed and unobserved village characteristics, then identification is possible.

The conditions of the comparison groups varied by study. In all but Olken (2007) study, the comparison groups received no intervention. In the former, groups received the same intervention but less intensity. Additionally, in several studies more than one treatment was evaluated. Table 3.4 presents this information.

Table 3.4: Interventions Details

Study	*Intervention*
Andrabi, Das and Khwaja (2013)	Report Cards. Card 1 reported the score of the child on test scores and her quintile rank, average village and school score and their quintile rank. Card 2 reported the scores for all schools in the village and their quintile rank. Cards were delivered in person in discussion groups. Every group started with a 30-minute open discussion on what influences test score results followed by the card distribution. Comparison group receives no intervention.
Banerjee et al. (2010)	Three Interventions: a) Providing information on existing institutions, b) training community members in a testing tool for children, and c) training volunteers to hold remedial reading camps. Comparison group receives no intervention.
Barr et al. (2012)	Scorecard. Two Interventions: a) standard treatment. School management committees (SMC) members were trained and supported in the use of a monitoring instrument b) Participatory treatment. Differed only in that SMC members were engaged in a dialogue to design their own monitoring instrument, by defining objectives, roles, and indicators of progress. Comparison group receives no intervention.
Björkman and Svensson (2009)	Scorecard. Through two rounds of village meetings, localized NGOs encouraged communities to be more involved with the state of health service provision and strengthened their capacity to hold their local health providers to account for performance. Comparison group receives no intervention.
Pandey, Sangeeta and Sundararaman (2009)	Information Campaign. The campaign consisted of eight or nine public meetings to disseminate information to the community about its state-mandated roles and responsibilities in school management. Comparison group receives no intervention.
Pradhan et al. (2013)	Two Interventions. a) School committee training to monitor school performance. (Standard) b) It added a facilitated joint planning meetings between the school committee and the village council (linkage). Comparison group receives no intervention.
Keefer and Khemani (2011)	Information Campaign through community radio on the benefits of educational attainment. Comparison group receives no intervention.
Lieberman, Posner and Tsai (2013)	Information Campaign. The intervention provided parents with information about their children's performance and how to become more involved in improving their children's learning. Comparison group receives no intervention.
Molina (2013*a*)	Social Audit. The intervention provided citizens with information on how to conduct a social audit and facilitate periodical meetings between citizens, providers and politicians. Comparison group receives no intervention.
Olken (2007)	Intensity of a Social Audit. The intervention encouraged the local population to attend public meetings, and participate in auditing the infrastructure projects. In this case the control group also has a social audit.

3.7.3 Population Locations

All studies were based on developing countries across the world. Studies took place in Benin (Keefer and Khemani, 2011), Colombia (chapter 2), Kenya (Lieberman, Posner and Tsai, 2013), India (Banerjee et al., 2010; Pandey, Sangeeta and Sundararaman, 2009), Indonesia (Olken, 2007; Pradhan et al., 2013), Pakistan (Andrabi, Das and Khwaja, 2013), Uganda (Barr et al., 2012; Björkman and Svensson, 2009). Each study examined an intervention in just one country. These locations present wide variation across and within countries. For example, some of these communities live with less than 1 dollar a day and very low levels of quality in service delivery, while other average 3 USD and better services. Also studies also included a mix of rural and urban populations.

3.7.4 Participants

All studies measured outcomes on adult members of targeted communities. In all cases participants comprised entire communities.

3.8 Intervention Characteristics

The theory of change outlined in Figure 3.1 above is an 'ideal type', explaining the mechanisms through which CMIs are expected to have an effect on service delivery. Despite differences in intervention design, sector and location, detailed on Tables 3.4 and 3.5, the theoretical foundations of the interventions are similar. They all make explicit reference to the idea of beneficiary participation as a tool to encourage better monitoring of the service provision. Most of them share a common element: disseminating information on the benefits of the service (Keefer and Khemani, 2011) or information on the state of service provision and citizen's right and responsibilities, encouraging citizens to monitor service providers (Andrabi, Das and Khwaja, 2013;

Banerjee et al., 2010; Barr et al., 2012; Björkman and Svensson, 2009; Lieberman, Posner and Tsai, 2013; Molina, 2013*a*; Olken, 2007; Pandey, Sangeeta and Sundararaman, 2009; Pradhan et al., 2013).

Table 3.5: Intervention Basic Characteristics

Study	*Location*	*Sector*	*Research Design*	*Management of Implementation*	*Facilitate Interaction?*
Andrabi, Das and Khwaja (2013)	Pakistan	Education	RCT	NGO and RT	No
Banerjee et al. (2010)	India	Education	RCT	NGO	No
Barr et al. (2012)	Uganda	Education	RCT	NGO	Yes[a]
Björkman and Svensson (2009)	Uganda	Health	RCT	NGO	Yes
Pandey, Sangeeta and Sundararaman (2009)	India	Education	RCT	NGO	Yes
Pradhan et al. (2013)	Indonesia	Education	RCT	NGO	Yes[b]
Keefer and Khemani (2011)	Benin	Education	Quasi-experimental	Community Radio	No
Lieberman, Posner and Tsai (2013)	Kenya	Education	RCT	NGO	No
Molina (2013*a*)	Colombia	Infrastructure	Matching	Government	Yes
Olken (2007)	Indonesia	Infrastructure	RCT	NGO	Yes

Notes: *a* : yes for the participatory intervention ; *b* : yes in the linkage intervention ; RCT : randomized control trial ; RT: research team ; *Facilitate Interaction?* refers to whether there was a meeting between providers and beneficiaries as part of the intervention.

Not all projects encourage citizens with the same tools or the same intensity. In the theoretical model from section 3.3, I highlighted that interactions with front line providers could be important in parts of the world where political accountability is low. From the included studies some interventions do not facilitate the interaction between providers and the community (Andrabi, Das and Khwaja, 2013; Banerjee et al., 2010; Keefer and Khemani, 2011; Lieberman, Posner and Tsai, 2013). Among those that do, the degree and nature of interaction also varies, as detailed in Table 3.4.

Finally, most of the studies focus on the education sector, in particular, on assessing the effect of report cards on learning outcomes, while Björkman and Svensson

(2009) focus on health and Olken (2007) and Molina (2013*a*), on chapter 2, focus on infrastructure.

3.9 Empirical Results: Statistical Meta-Analysis of Effect Estimates

The CMIs interventions and their outcome measurement strategies were similar enough to permit at least a preliminary meta-analysis. Based on the primary outcomes described above, determined after having consulted the studies, I synthesized its effect estimates, displayed in Table 3.6. The effect sizes and standard errors are given in terms of difference in rate coefficients, as explained in the methodology section above. Not all studies reported effects for all measures; indeed, the results matrix is quite spotty. The right-most four columns show the results of a synthesis and homogeneity test. The synthesis provides the best guess of what is the 'actual' effect of a CMI intervention on the outcome in question given the available evidence.

The validity of a synthesized effect estimate requires that a set of homogeneity assumptions holds: a) the interventions should be the same in terms of causal mechanisms that they initiate, b) the outcome measures should be the same in terms of the phenomenon that they capture and the manner in which the phenomenon is being measured. For these CMIs interventions, neither of these homogeneity assumptions is particularly compelling in terms of face validity. The interventions were carried out by different agencies using different rules in different countries. As I highlighted above, while the theoretical model used to design this interventions is similar, the actual intervention differs. Not only because they are implemented in different countries and sectors, but also their intensity differs.

Table 3.6: Synthesis of effect estimates from CMIs interventions

Study		***Service Delivery Outcomes*** *Quality of the services* *Improvement in prevalence condition*	*Perception of overall quality*
Andrabi, Das and Khwaja (2013)	b	0.11	
	se	0.045	
Banerjee et al. (2010)	b	0.007	
	se	0.09	
Barr et al. (2012)	b	0.19	
	se	0.1	
Björkman and Svensson (2009)	b	0.12	0.25
	se	0.06	0.11
Pradhan et al. (2013)	b	0.11	
	se	0.06	
Molina (2013*a*)	b		0.22
	se		0.04
Synthesis[a]	B	0.11	0.22
	se(B)	0.03	0.04
	B/se(B)	3.87	5.93
Homogeneity[b]	p-value	0.74	0.80

Notes: For Barr et al. (2012) I use the participatory intervention and for Pradhan et al. (2013) the linkage intervention. Estimates for all studies are difference in rates. a : refers to inverse variance weighted estimates: $B = (\sum b_k * se_k^2)/\sum se_k^2$, and $se(B) = (\sum se_k^2)^{-0.5}$ for studies indexed by k. b : This gives the p value for a chi-square based homogeneity test. The homogeneity test statistic is $\sum se_k^2(b_k B)^2$, which is distributed chi-square with degrees of freedom equal to the number of studies minus one. The p value gives the probability of observing a test statistic as large as the one computed under the null hypothesis that the different studies are measuring the same underlying effect.

With respect to outcome measures, the included studies' instruments were roughly similar. For the synthesis of *Improvement in prevalence condition* I only included studies that measure test scores and objective measure of improvement in health conditions. Andrabi, Das and Khwaja (2013) administered tests of English, Mathematics and Urdu (the vernacular) as part of the their study. Criterion-referenced tests were implemented. These tests help identify whether students meet a criterion but can be less informative for those below or above the critical level. To design the tests, an extensive pilot was carried out to identify lower and upper limits of learning in the population and analyze the validity and reliability of the instrument used. The data from this phase was used to refine the final test. As a result, all three tests (English, mathematics, and Urdu) started from simple problems and gradually increased in difficulty. All non-absent children in Grade 3 (12,110 children) in the sample schools were tested in the three subjects. The tests were administered by the research team. In order to facilitate comparisons of the test over time and to better relate the test to underlying student knowledge, Andrabi, Das and Khwaja (2013) use item response theory to score the test appropriately adjusting for the difficulty of each question.

In the case of Banerjee et al. (2010) study in India (Uttar Pradesh), the data on student learning outcomes were collected using a tool developed by Pratham, the NGO that implemented the intervention. For reading evaluations, testers asked children to read a simple paragraph in Hindi. If they could read the paragraph, the tester moved up to a longer and more complex story. If the child could not read the paragraph, the tester quickly switched to a panel with a few single words on it. If the child struggled to read the words on the panel, the tester showed a panel with several letters and asked the child to identify them. A similar approach was used for math, with the levels in the math case starting from the ability to recognize numbers,

moving up to the ability of performing subtractions, and, finally, divisions. Children were 7 to 14 years old at the time.

Barr et al. (2012) used standardized public administer exam (NAPE) for children's in primary school. They measured learning outcomes by pooling NAPE exams for literacy and numeracy.

Pradhan et al. (2013) administered test in mathematics and Indonesian, designed by the Ministry of Education, to all students in grade four at baseline and grade six at endline.

Finally, Björkman and Svensson (2009) is the only study not focused on the education sector. For this study, the authors measured births, pregnancies, and deaths of children under five years, and the weight of all infants (i.e., under 18 months of age) and children (between 18 and 36 months of age) in the surveyed households. With this information at baseline and endline they estimated the impact of the program on births, pregnancies, under 5 mortality rate, child death and weight-for-age z-scores. I synthesized these measures and computed the overall effect of the program[9]. I also ran an homogeneity test for these measures and I could not reject the null that they are measuring the same underlying effect.

In the case of perception measures both studies used survey instruments to ask about the overall satisfaction with the program to targeted community .

The results seem to point to positive and statistically significant effect of the CMIs intervention in improving prevalence conditions, such as learning in school

[9]I use a common metric so 'improvement' were all transformed to positive numbers.

or health in nurseries, though the evidence does not rely in many studies. Similar results were found for the subjective perception measures on the overall quality of the CMIs intervention. Again, the matrix was very spotty in this last outcome.

The p-value of the statistical homogeneity test provides a quantification of the likelihood of homogeneity, but this test only makes sense assuming face validity of the homogeneity assumptions. In both cases, I was unable to reject that the null hypothesis that the studies are measuring the same underlying effect.

3.10 Causal Chain Analysis

Asking why programs succeed or failed involves identifying causal pathways. Sometimes pathways are explicit and other times finding pathways means looking for implicit assumptions and arguments. There are a range of possible pathways from the CMIs process to improvement in service delivery, and determining these pathways assists us in answering how and why interventions work or not.

Through repeated reading of the included studies I extracted additional information that could help in explaining the results, paying particular attention to any data on the causal chain and how interventions unfolded in practice.

As I highlighted in the section 3.3 above, there are several potential channels through which this programs could have an impact on service delivery as well as several potential bottlenecks. Below I focus on one mediating mechanisms that has been discussed in the literature as a possible candidate to explain why some programs may fail to have an impact: interactions between beneficiaries and providers. Supporters

of this theory argue that it is plausible that interaction between beneficiaries and providers could increase citizens participation. As discussed on the explanation of the theory of change above participation could be mediated by the information citizens have on the program as well as by the belief that participation will lead to actual changes and improvement in service delivery, as opposed to a tokenism. Having providers listening to beneficiaries' complains may act a signal that citizens input will be used to help improve service delivery. At the same time, increased community participation could act as an incentive for providers to exert more effort, due to fear of formal or informal punishment. Furthermore, through providers interaction with the community, they could increase their intrinsic motivation to improve the quality of the services they provide.

Let's first assess whether the programs that foster more participatory treatments are more successful in improving service delivery. To study whether there is supporting evidence of the potential effect of interacting with providers in the included studies, I split the studies among those where the intervention involves facilitation of contacts between providers and/or politicians and the community and those that do not.

The results from Table 3.7 support the 'interaction' hypothesis. Interventions where there was substantial facilitated contact with providers show positive and statistically significant treatment effect, while studies where contacts were not facilitated show no effect of the CMIs on quality of service delivery, in this case, improvement in learning outcomes for children. The only individual study out of 4 that shows a positive and significant improvement is Andrabi, Das and Khwaja (2013).

Table 3.7: Is Facilitated Contact with Providers Important to Account for Different Findings?

Study		***Service Delivery Outcomes*** *Improvement in prevalence condition* *No Facilitated Contact with providers*	*Facilitated Contact with providers*
Andrabi, Das and Khwaja (2013)	b	0.11	
	se	0.045	
Banerjee et al. (2010)	b	0.007	
	se	0.09	
Barr et al. (2012)	b	0.08	0.19
	se	0.1	0.1
Björkman and Svensson (2009)	b		0.12
	se		0.06
Pradhan et al. (2013)	b	-0.03	0.11
	se	0.06	0.06
Synthesis[a]	B	0.05	0.12
	se(B)	0.06	0.06
	B/se(B)	0.91	1.87
Homogeneity[b]	p-value	0.23	0.75

Notes: Estimates for all studies are difference in rates. a : refers to inverse variance weighted estimates: $B = (\sum b_k * se_k^2)/\sum se_k^2$, and $se(B) = (\sum se_k^2)^{-0.5}$ for studies indexed by k. b : This gives the p value for a chi-square based homogeneity test. The homogeneity test statistic is $\sum se_k^2(b_k B)^2$, which is distributed chi-square with degrees of freedom equal to the number of studies minus one. The p value gives the probability of observing a test statistic as large as the one computed under the null hypothesis that the different studies are measuring the same underlying effect.

But the most important evidence comes from studies that implement both types of interventions (Barr et al., 2012; Pradhan et al., 2013). In both studies the standard intervention (training in the case of Pradhan et al. (2013) and an information campaign in the form of a scorecard in Barr et al. (2012)) did not produce improvements in outcomes statistically different from zero. However, the participatory variations produced positive and statistically significant results. In the case of Barr et al. (2012) school management committee members and providers were engaged in a dialogue to design their own scorecard monitoring instrument. In Pradhan et al. (2013) the intervention facilitated joint planning meetings between the school committee and

the village council, facilitating the contact with the political representatives.

3.10.1 Does Facilitated Contact Increase Citizens Participation?

As I highlighted in section 3.3 a crucial building block for CMIs to work is citizens participation and providers responsiveness. Failure to foster civic participation have long been argued as a reason why CMIs do not meet expectations in terms of their impact (Mansuri and Rao, 2004, 2012). If citizens decide not to participate in CMI, the effectiveness of the program will be limited. Lieberman, Posner and Tsai (2013) found that there are several bottlenecks that prevents targeted population from actively monitoring service providers. In their evaluation of an information campaign in Kenya they find that citizens failed to participate in the intervention. They attributed this to many bottlenecks, among them: (a) citizens do not think that it is their responsibility to monitor providers to improve service delivery, (b) citizens think they do not have the skills to make a difference, and (c) collective action problems.

The evidence from included studies indicates that participation in the program and community activity is a predictor of program success. Studies that found little effect of the program in community participation in monitoring service delivery also found no statistically significant treatment effects. For example, Banerjee et al. (2010) found that the information campaign had little impact among parent's members of VEC, the system through which parent can contribute to monitor schools performance as well as the general population of parents. The authors failed to find an impact on VEC members knowledge about the right and responsibilities as VEC members. Reasonably, the study also failed to find any impact on VEC activism, i.e.

complained to school management, raised money for the school, etc. The program only appears to have an impact on their knowledge about the state of educational performance of the village among VEC members, not among the non-VEC parents.

Barr et al. (2012) shed evidence on whether facilitated contact improve citizen participation in CMIs. The authors argue that impacts of the participatory treatment exceed those of the standard treatment primarily because of increased willingness to contribute to public goods, rather than differences in the information content of the scorecards. To address this issue, the authors used a public good game, the voluntary contributions mechanism game, which is played immediately following the training of school management committees in their respective treatments. They find that the willingness to contribute to public goods is statistically higher for participants of the participatory treatment. However, the identification strategy prevents the authors from discriminating between two potential theories. The possibility that the participatory treatment might impact outcomes in the school (test scores) and in the lab (voluntary contribution games) either by affecting preferences or by affecting beliefs about the willingness of others to contribute to public goods.

Pradhan et al. (2013) provided similar evidence of increased participation of the community in the forms of contributions (money and in-kind) to the school for the 'linkages' intervention compared to the standard (only training for the school committee) intervention.

3.10.2 Does Facilitated Contact Improve Providers' Responsiveness?

In this section I study whether facilitated interactions improves providers performance, as measured by attendance rate and efficiency in management. The first piece of evidence comes from studies where their only treatment arm did not include facilitated contact. Most studies that measure provider responsiveness found no statistically significant effect (Banerjee et al., 2010; Keefer and Khemani, 2011), while most studies where contact was facilitated found the opposite result (Barr et al., 2012; Björkman and Svensson, 2009; Molina, 2013*a*; Pradhan et al., 2013).

Björkman and Svensson (2009) found the scorecard intervention improved the type of equipment used to treat patients, reduced the average waiting time and the absence rate of staff at the nurseries, and also improved the management of the nurseries, i.e. cleaner floors and rooms, staff politeness, among others. Furthermore, using data collected through visual checks by the enumerators during the post-intervention facility survey, the authors find evidence that the program increased the opportunity the health facility gave the community to monitor them through various methods. In particular, the CMI increase the probability that the health facility had:

(i) A suggestion box for complaints and recommendations

(ii) Numbered waiting cards for its patients

(iii) A poster informing about free health services

(iv) A poster on patients' rights and obligations

The authors argue that it was precisely these improvements in the management of the health facilities that derived in better health outcomes for the targeted population. According to the authors, the reason of the program success is explained by

a change in behaviour from the health facility staff. Increased intrinsic motivation due to the interaction between the community and providers appears to be the key behind the improvement in health facility staff's efforts.

Olken (2007) study described above focused on a social audit in Indonesia. He found that the program has a slightly positive but not statistically significant effect on the management of project funds to build infrastructure. The author argues that this may reflect the fact that is hard for the community to monitor this type of expenses, as they need to have construction and procurement expertise. Khwaja (2009) found similar results for Pakistan.

Perhaps more revealing is the evidence from studies with both treatments: one with facilitated contact with providers and one without it. Those studies have a better identification strategy to answer the question of whether facilitated contact improved provider responsiveness. Barr et al. (2012) is one of such studies. They found that only the participatory treatment had a positive and statistically significant effect in reducing teachers absence rate in schools, compared to the standard treatment that did have no effect. Pradhan et al. (2013), the other intervention were both treatment were randomized, also found suggestive evidence of impact on teachers effort, though the statistical significance of the results is not present for all teacher effort outcomes, but only some of them.

The evidence in this section, though important, should be interpreted as preliminary, as the number of rigorous study is small. In order to investigate whether this is the actually the case more research is needed in the area.

3.11 Conclusions

This article examines the impact of the Community Monitoring Interventions (CMIs) on the quality of service delivery in low and middle income countries.

I identified 10 rigorous studies, reviewing 10 interventions, although they differed substantially in their intervention types and outcome measures used. As usual, reviews reflect the state of knowledge at their time of publication.

I found that CMIs have a positive significant effect on prevalence outcomes (tests scores and health status improvements) as well as perception measures of overall satisfaction with the program.

However, the devil is in the details. I found that intervention designs affect substantially the final outcomes. In particular, CMIs where there was substantial facilitated contact with providers show positive and statistically significant treatment effect, while studies where contacts were not facilitates show no effect of the CMIs on quality of service delivery. This reflects the fact that CMIs where contacts with providers is organized also have better outcomes in terms of provider responsiveness (lower absence rate and more efficiency) as well as more participation from the community (more time spent monitoring and more in-kind and monetary donations).

Unfortunately the ability to build external validity by comparing results from multiple studies is limited in this case due to the small number of rigourous studies. Increasing the number of studies, as they become available, may go some way towards further generalizing findings.

Additional attention to the details of interventions, to causal chain analysis, and to the exploration of competing causal pathways should be built into study designs. This could heighten our understanding of CMIs interventions.

Appendix A

Appendix

Table A.1: Projects - Part I

Status	*Pair*	*Object*	*Municipality*	*Sector*	*Amount* (Millions)	*Timeline*
T	1.1	Construction and Optimization of pluvial sewage (21, 22, 24, 25 st between 23 and canal)	Yopal	WS	13700	18 Months
C	1.2	Construction and Optimization of pluvial sewage (24a , 29 st between campiÃśa canal and caÃśo seco)	Yopal	WS	8450	6 Months
T	2.1	Study, Design and Construction of San Lorenzo Educational Institution	Castilla La Nueva	EDU	11352	12 Months
C	2.2	Construction of the Second Stage of Castilla La Nueva Educational Institution	Castilla La Nueva	EDU	7797	8 Months
T	3.1	Construction of aqueduct system in Las Islas Optimization of sewage networks in other places	Mani	WS	6319	18 Months
C	3.2	Construction of aqueduct in Las brisas, Guafal Pintado, and other places	Mani	WS	3000	12 Months
T	4.1	Construction of aqueduct system San Juan de la Cruz and Guaymaral and optimization for Barbacoas	Tuchin	WS	1125	7 Months
C	4.2	Rehabilitation aqueduct systems in Flecha, Optimization of Vidales and Fecha's aqueduct Guafal Pintado, and other places	Tuchin	WS	602	4 Months
C	4.3	Improvement of aqueducts of Nueva Estrella, Bella Vista, and Esmeralda sur. Construction of a drinking water plant in Nueva Estrella	Tuchin	WS	560	3 Months
T	5.1	Design and construction of canals, collectors, ditches, etc in Olaya Herrera Avenue and Libertadores neighborhood .	Arauca	WS	2850	5 Months
C	5.2	Design and construction of canals, collectors, ditches, etc in Quinta-El Dique Avenue neighborhood.	Arauca	WS	3968	6 Months
C	5.3	Design and construction of canals, collectors, ditches, etc in 16 st from CaÃśo CÃşrdoba to Humedal Pozo Las Babas and la 15 st between 15 and 16 st.	Arauca	WS	4000	8 Months
T	6.1	Construction of pluvial sewage in zone NÂř1, VÃŋncula Palacio, Almirante Padilla and zone NÂř.2, Colombia Libre, Jorge Arrieta.	Maicao	WS	4337	12 Months
C	6.2	Construction of pluvial sewage in zone NÂř1 San Antonio, Alfonso LÃşpez, Santa Fe, La UniÃşn and Rojas Pinilla, Zone 2, SimÃşn MejÃŋa	Maicao	WS	2551	3 Months
T	7.1	Design and Construction of Eje Educativo. VÃŋncula Palacio, Almirante Padilla and zone NÂř.2, Colombia Libre, Jorge Arrieta.	Hatonuevo	EDU	14998	12 Months
C	7.2	Construction of Third Stage of the Educational Institution San Rafael	Albania	EDU	8328	11 Months

Notes: *T* refers to treatment group and *C* refers to control group. *WS* refers to Water and Sanitation sector; *EDU* refers to Education sector; *GAS* refers to Natural Gas sector; *Timeline* refers to the estimated duration of the project before it is actually executed.

Table A.2: Projects - Part II

Status	*Pair*	*Object*	*Municipality*	*Sector*	*Amount* (Millions)	*Timeline*
T	8.1	Implementation of sewage master plan in San Francisco's urban region	San Francisco	WS	1807	16 Months
C	8.2	Improving Sewage system of Villagarzon's urban region	VillagarzÃşn	WS	1120	6 Months
T	9.1	Construction of a aqueduct system in San Francisco neighbourhood	Neiva	EDU	1414	10 Months
C	9.2	Construction in sewage system in Guacirco, San Luis and El Colegio	Neiva	EDU	2734	2 Months
T	10.1	Construction educational institution Manuel J Del Castillo, Isaac J Pereira and Liceo Moderno	Cienaga	EDU	6123	8 Months
C	10.2	Construction educational institution San Juan del Cordoba	Cienaga	EDU	8498	8 Months
T	11.1	Final Stage of Construction Educational Institution Sumapaz	Melgar	EDU	4988	7 Months
C	11.2	First Stage Education Institution Megacolegio	Zipaquira	EDU	5527	8 Months
T	12.1	Construction of Educational Center VÃŋncula Palacio, Almirante Padilla	Yopal	EDU	5009	10 Months
C	12.2	Construction and Improvement of Educational Institution	Yopal	EDU	4486	6 Months
T	13.1	Construction of Natural Gas Infrastructure to increase connection among low income individuals	Guatape y El Penol	GAS	6892	3 Months
C	13.2	Construction of Natural Gas Infrastructure to increase connection among low income individuals	Subachoque	GAS	1067	6 Months

Notes: *T* refers to treatment group and *C* refers to control group. *WS* refers to Water and Sanitation sector; *EDU* refers to Education sector; *GAS* refers to Natural Gas sector; *Timeline* refers to the estimated duration of the project before it is actually executed.

Bibliography

Acemoglu, D. and J. A. Robinson. 2008. "Persistence of Power, Elites, and Institutions." *American Economic Review* 98(1):267–293.

Adsera, Alicia, Carles Boix and Mark Payne. 2003. "Are You Being Served? Political Accountability and Quality of Government." *Journal of Law, Economics and Organization* 19(2):445–490.

Alatas, Vivi, Abhijit Banerjee, Rema Hanna, Benjamin A. Olken and Julia Tobias. 2012. "Targeting the Poor: Evidence from a Field Experiment in Indonesia." *American Economic Review* 102(4):1206–40.

Alesina, Alberto and Romain Wacziarg. 2003. "Fractionalization." *Journal of Economic Growth* 8(2):155–94.

Andrabi, T., J. Das and A. I. Khwaja. 2013. Report Cards: The Impact of Providing School and Child Test Scores on Educational Markets. Working paper JPAL.

Banerjee, Abhijit V., Rukmini Banerji, Esther Duflo, Rachel Glennerster and Stuti 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, Abhijit V. and Sendhil Mullainathan. 2008. "Limited Attention and Income Distribution." *American Economic Review* 98(2):489–93.

Bardhan, Pranab. 2002. "Decentralization of Governance and Development." *Journal of Economic Perspectives* 16(4):185–205.

Bardhan, Pranab and Dilip Mookherjee. 2006. "Decentralisation and Accountability in Infrastructure Delivery in Developing Countries." *Economic Journal* 116(508):101–127.

Barr, Abigail, Frederick Mugisha, Pieter Serneels and Andrew Zeitlin. 2012. Information and collective action in community-based monitoring of schools:Field and lab experimental evidence from Uganda. Working paper mimeo.

Bertrand, Marianne, Dean Karlan, Sendhil Mullainathan, Eldar Shafir and Jonathan Zinman. 2010. "What's Advertising Content Worth? Evidence from a Consumer Credit Marketing Field Experiment." *The Quarterly Journal of Economics* 125(1):263–305.

Besley, T. and T. Persson. 2011. *Pillars of Prosperity: The Political Economics of Development Clusters*. The Yrj Jahnsson Lectures Princeton University Press.

Bhatnagar, D., Dewan-A. Moreno Torres M. & Kanungo P. 2003. (2003). CitizensŠ report cards on public services: Bangalore, India. Empowerment Case Studies. Technical report World Bank.

Björkman, Martina and Jakob 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.

Brinkerhoff, D.W., & Azfar O. 2006. Decentralization and community empowerment: Does community empowerment deepen democracy and improve service delivery?. Technical report Office of Democracy and Governance, USAID. Washington, DC.

Casey, Katherine, Rachel Glennerster and Edward Miguel. 2012. "Reshaping Institutions: Evidence on Aid Impacts Using a Preanalysis Plan." *The Quarterly Journal of Economics* 127(4):1755–1812.

Chaudhury, Nazmul, Jeffrey Hammer, Michael Kremer, Karthik Muralidharan and F. Halsey Rogers. 2006. "Missing in Action: Teacher and Health Worker Absence in Developing Countries." *Journal of Economic Perspectives* 20(1):91–116.

Devarajan, Shantayanan, Stuti Khemani and Michael Walton. 2011. Civil society, public action and accountability in Africa. Policy Research Working Paper Series 5733 The World Bank.

Downs, A. 1957. *An economic theory of democracy*. Harper.

Ferejohn, John. 1986. "Incumbent performance and electoral control." *Public choice* 50(1):5–25.

Ferraz, Claudio, Frederico Finan and Diana B. Moreira. 2012. "Corrupting learning." *Journal of Public Economics* 96(9-10):712–726.

for Good-Governance, Centre. 2005. "Social audit: A toolkit. A guide for performance improvement and outcome measurement.".

Gavazza, Alessandro and Alessandro Lizzeri. 2007. "The Perils of Transparency in Bureaucracies." *American Economic Review* 97(2):300–305.

Gaventa, John and Gregory Barrett. 2012. "Mapping the Outcomes of Citizen Engagement." *World Development* 40(12):2399–2410.

Gorodnichenko, Yuriy and Peter Klara Sabirianova. 2007. "Public sector pay and corruption: Measuring bribery from micro data." *Journal of Public Economics* 91(5-6):963–991.

Hanna, R., Bishop-S. Nadel S. Scheffler G. & Durlacher K. 2011. The effectiveness of anti-corruption policy: What has worked, what hasnŠt, and what we donŠt know. Technical report DIFD Systematic Review.

Hombrados, J., B. Snilstveit, M. Vojtkova, H. Waddington and H. White. 2013. Farmer Field Schools for improving farming practices and farmer outcomes in low- and middle-income countries: a systematic review. Policy research working paper series 3ie.

Hotelling, H. 1929. "Stability in Competition." *Economic Journal* 39:41–57.

Isham, Jonathan, Deepa Narayan and Lant Pritchett. 1995. "Does participation improve performance? Establishing causality with subjective data." *The World Bank Economic Review* 9(2):175–200.

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 Series 5559 The World Bank.

Khemani, Stuti. 2007. Can Information Campaigns Overcome Political Obstacles to Serving the Poor? In The Politics of Service Delivery in Democracies: Better Access for the Poor, ed. Shantayanan Devarajan and Ingrid Widlund. Technical report.

Khwaja, A.I. 2009. "Can good projects succeed in bad communities?" *Journal of Public Economics* 93(7-8):899–916.

Khwaja, Asim Ijaz and Atif Mian. 2005. "Do Lenders Favor Politically Connected Firms? Rent Provision in an Emerging Financial Market." *The Quarterly Journal of Economics* 120(4):1371–1411.

Le Grand, Julian. 2003. *Motivation, Agency, and Public Policy: Of Knights and Knaves, Pawns and Queens*. Number 9780199266999 *in* "OUP Catalogue" Oxford University Press.

Lieberman, Evan, Daniel N. Posner and Lily Tsai. 2013. Does Information Lead to More Active Citizenship? Evidence from an Education Intervention in Rural Kenya. Technical report MIT Political Science Department.

Lupia, Arthur. 2012. Can Demand for Accountability Be Increased? Necessary Conditions for Progress. In Sina Odugbemi and Taeku Lee (eds.) Accountability through Public Opinion: From Inertia to Public Action. Policy research paper series The World Bank.

Mansuri, Ghazala and Vijayendra Rao. 2004. Community-based (and driven) development : A critical review. Policy Research Working Paper Series 3209 The World Bank.

Mansuri, Ghazala and Vijayendra Rao. 2012. *Localizing Development : Does Participation Work?* Number 11859 *in* "World Bank Publications" The World Bank.

Maru, V. 2010. "Social accountability and legal empowerment." *Health and Human Rights: An International Journal, North America.* 05:12–25.

Mauro, Paolo. 1995. "Corruption and Growth." *The Quarterly Journal of Economics* 110(3):681–712.

Mcmillan, John and Pablo Zoido. 2004. "How to Subvert Democracy: Montesinos in Peru." *Journal of Economic Perspectives* 18(4):69–92.

Molina, Ezequiel. 2013*a*. "Can Bottom-Up Institutional Reform Improve Service Delivery? Evidence from the Citizen Visible Audit program in Colombia.".

Molina, Ezequiel. 2013*b*. "A Theory of Self-fufilling Prophecy in Service Delivery.".

Narayan, Deepa. 1995. The Contribution of People's Participation: Evidence from 121 Rural Water Supply Projects. Technical report World Bank, Environmentally Sustainable Development Department.

Olken, B. 2007. "Monitoring corruption: Evidence from a field experiment in Indonesia." *Journal of Political Economy* 115:200–249.

Olken, Benjamin A. 2006. "Corruption and the costs of redistribution: Micro evidence from Indonesia." *Journal of Public Economics* 90(4-5):853–870.

Olken, Benjamin A. 2009. "Corruption perceptions vs. corruption reality." *Journal of Public Economics* 93(7-8):950–964.

Olken, Benjamin A. and Patrick Barron. 2009. "The Simple Economics of Extortion: Evidence from Trucking in Aceh." *Journal of Political Economy* 117(3):417–452.

Olken, Benjamin A. and Rohini Pande. 2012. "Corruption in Developing Countries." *Annual Review of Economics* 4(1):479–509.

Olson, Mancur. 1971. *The Logic of Collective Action : Public Goods and the Theory of Groups.* Harvard economic studies, v. 124 revised ed. Harvard University Press.

Pan, Lei and Luc Christiaensen. 2012. "Who is Vouching for the Input Voucher? Decentralized Targeting and Elite Capture in Tanzania." *World Development* 40(8):1619–1633.

Pandey, Priyanka, G. Sangeeta and V. Sundararaman. 2009. "Community participation in public schools: impact of information campaigns in three Indian states." *Education Economics* 17(3):355–375.

Persson, Torsten and Guido Tabellini. 2002. *Political Economics: Explaining Economic Policy.* Vol. 1 of *MIT Press Books* The MIT Press.

Pradhan, Menno, Daniel Suryadarma, Amanda Beatty, Maisy Wong, Armida Alisjahbana, Arya Gaduh and Rima Prama Artha. 2013. Improving Educational Quality Through Enhancing Community Participation: Results from a Randomized Field Experiment in Indonesia. Policy research working paper series The World Bank.

Prat, Andrea. 2005. "The Wrong Kind of Transparency." *American Economic Review* 95(3):862–877.

Pritchett, Lant, Michael Woolcock and Matt Andrews. 2013. "Looking Like a State: Techniques of Persistent Failure in State Capability for Implementation." *Journal of Development Studies* 49(1):1–18.

Putnam, Robert D. 1993. *Making Democracy Work: Civic Traditions in Modern Italy*. Vol. 506 Princeton University Press.

Reinikka, Ritva and Jakob Svensson. 2004. "Local Capture: Evidence From a Central Government Transfer Program in Uganda." *The Quarterly Journal of Economics* 119(2):678–704.

Reinikka, Ritva and Jakob Svensson. 2011. "The power of information in public services: Evidence from education in Uganda." *Journal of Public Economics* 95(7-8):956–966.

Ringold, Dena, Alaka Holla, Margaret Koziol and Santhosh Srinivasan. 2012. *Citizens and Service Delivery : Assessing the Use of Social Accountability Approaches in the Human Development Sectors*. Number 2377 *in* "World Bank Publications" The World Bank.

Rose-Ackerman, Susan. 1999. *Corruption and Government*. Number 9780521632935 *in* "Cambridge Books" Cambridge University Press.

Sacks, Audrey and Marco Larizza. 2012. Why quality matters : rebuilding trustworthy local government in post-conflict Sierra Leone. Policy Research Working Paper Series 6021 The World Bank.

Singer, M.M. 2013. "Bribery Diminishes Life Satisfaction in the Americas.".

Snilstveit, Birte. 2012. "Systematic reviews: from Śbare bonesŠ reviews to policy relevance." *The Journal of Development Effectiveness* 4(3):388–408.

Snilstveit, Birte, Sandy Oliver and Martina Vojtkova. 2012. "Narrative approaches to systematic review and synthesis of evidence for international development policy and practice." *The Journal of Development Effectiveness* 4(3):409–429.

Stiglitz, Joseph E. 2002. "Participation and Development: Perspectives from the Comprehensive Development Paradigm." *Review of Development Economics* 6(2):163–82.

Sukhtankar, Sandip. 2012. "Sweetening the Deal? Political Connections and Sugar Mills in India." *American Economic Journal: Applied Economics* 4(3):43–63.

Svendsen, G.T. and G.L.H. Svendsen. 2009. *Handbook of Social Capital: The Troika of Sociology, Political Science and Economics*. Elgar Original Reference Series Edward Elgar Publishing, Incorporated.

Svensson, Jakob. 2003. "Who Must Pay Bribes And How Much? Evidence From A Cross Section Of Firms." *The Quarterly Journal of Economics* 118(1):207–230.

Svensson, Jakob. 2005. "Eight Questions about Corruption." *Journal of Economic Perspectives* 19(3):19–42.

White, Howard. 2009. Theory-Based Impact Evaluation. 3ie Working papers 2009-3 International Initiative for Impact Evaluation.

World-Bank. 2003. *Making Service Work for Poor People*. World Bank.