Policy Research Working Paper 9087

Group Incentives for the Public Good

A Field Experiment on Improving the Urban Environment

Carol Newman
Tara Mitchell
Marcus Holmlund
Chloë Fernandez

WORLD BANK GROUP
Development Economics
Development Impact Evaluation Group
December 2019

Policy Research Working Paper 9087

Abstract

How to maintain communal spaces is an important concern in many developing countries, particularly in urban environments. But what strategies can communities use to overcome the public goods problems involved in maintaining their local environment? This paper investigates whether changing the incentives for a subset of the community to contribute to the public good can lead to a shift to a more efficient equilibrium for the community as a whole. The analysis uses a randomized controlled trial to test the effectiveness of a program called "Operation Clean Neighborhood," which targets established community-based organizations and encourages them, through social recognition and low-value, in-kind incentives, to work toward keeping their neighborhoods clean, with the ultimate goal of reducing flooding in these areas. The findings show that, after one year, the intervention was effective in engaging communities and improving the cleanliness of the neighborhood. There is also evidence that this leads to reduced levels of flooding. The analysis uncovers important differences in the effectiveness of the program between areas that have had increased investment in drainage infrastructure and those that have not. It also addresses the issue of spillovers, an important consideration in densely populated urban centers.

This paper is a product of the Development Impact Evaluation Group, Development Economics. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at mholmlund@worldbank.org and cfernandez2@worldbank.org.

Group Incentives for the Public Good: A Field Experiment on Improving the Urban Environment

Carol Newman[a], Tara Mitchell[a], Marcus Holmlund[b], and Chloë Fernandez[b]

[a] Department of Economics and Trinity Impact Evaluation unit (TIME), Trinity College Dublin, Ireland
[b] Development Impact Evaluation (DIME), World Bank

JEL Codes: O12, O13, O18,
Key Words: Infrastructure, Public Goods, Community-Based Organizations

Chloë Fernandez (email: cfernandez2@worldbank.org) and Marcus Holmlund (email: mholmlund@worldbank.org) are at Development Impact Evaluation (DIME), World Bank. Tara Mitchell (email: mitchet@tcd.ie) and Carol Newman (corresponding author: cnewman@tcd.ie) are at Trinity College Dublin and TIME (Trinity Impact Evaluation Unit). This research is part of the World Bank-assisted Senegal Stormwater Management and Climate Change Adaptation Project (P122845) PROGEP) project which was funded by World Bank, the Government of Senegal and the Nordic Development Fund. We are thankful to our partners at the Municipal Development Agency (Agence de Développement Municipal, ADM), Pierre Coly and Mouhamadou Ba. We are also grateful to a large number of people for their valuable comments and insights. These include Andrea Guariso, Denis Jordy, Isabelle Kane, Arianna Legovini, Gabriele Rechbauer, Fatim Seck, Eric Lancelot, Amos Abu, Farouk Banna, an anonymous DIME reviewer, and participants at conferences including the NEUDC at Tufts University, the CSAE conference Oxford and the Irish Economics Association Annual Conference, and participants at seminars in Queens University Belfast, Groningen University, University of Copenhagen, Graduate Institute Geneva, and Copenhagen Business School. We thank Molly Offer-Westort, Violaine Pierre, Felipe Dunsch, and Ababacar Gueye for invaluable assistance with project coordination and data collection. Our thanks also to our research assistants Margaryta Klymak, Michell Dong, and Qiao Wang. This work benefitted from generous funding from the UK/DfID through the i2i Trust Fund, the Bank-Netherlands Partnership Program, the Nordic Development Fund, and the World Bank. The RCT was registered on the AEA RCT Registry on July 7, 2015 (AEARCTR-0000754). Computational reproducibility was confirmed by DIME Analytics.

1. Introduction

Overcoming the public goods problem in the maintenance of communal spaces is a concern in many contexts and can be particularly challenging in urban environments. This issue is becoming increasingly important in developing countries, where communities face the threats posed by climate change and an increasing incidence of events such as flooding and drought. Increased investments in infrastructure are important to help communities face these challenges but these investments will only be sustainable if the community is engaged in behavior that supports the maintenance and function of this infrastructure. Many studies have shown that small communities can find strategies that allow them to overcome public goods problems such as the use of social pressure, reputation, and monitoring. Some key characteristics of these communities are that they are relatively small, members interact with each other regularly and members can observe each other's actions (Ostrom, 1990; Ostrom 1999, Ostrom 2008, Baland and Platteau, 1996, Besley and Ghatak, 2006). However, these strategies may be much harder to implement in densely-populated urban communities where much less is known about how best to engage communities in the provision of public goods, such as the preservation of public spaces.

In this paper, we investigate whether changing the incentives for a subset of the community to contribute to the public good can lead to a shift to a more efficient equilibrium for the community as a whole. The setting for our study is two low-income peri-urban areas of Dakar in Senegal, Pikine and Guédiawaye, which represent about 12% of the national population. These two municipalities are highly prone to flooding, and approximately half of the residents live in flood-prone areas. This has important impacts on health, livelihoods, and assets, and disproportionately affects more vulnerable population groups. Some major investments in new drainage infrastructure have been made in these areas in recent years. However, such infrastructure has historically suffered from a lack of maintenance and misuse, including the dumping of household waste.

To overcome this public good problem, we designed an intervention called "Operation Clean Neighborhood" (*Opération Quartier Propre* or OQP) to try to shift behaviors and encourage local populations to keep their community and infrastructure clean. We use a randomized controlled trial to test the effectiveness of OQP, which targets established Community Based Organizations (CBOs) and encourages them, through social recognition and low-value in-kind incentives, to work towards keeping their neighborhoods clean.[1] We test whether community engagement in keeping the neighborhood clean is improved through OQP and whether this in turn leads to reduced flooding. After one year, we find that OQP is indeed effective in engaging communities and in improving the cleanliness of the neighborhood and also find evidence that this leads to reduced levels of flooding. We also find important differences in the effectiveness of the program between areas which have had increased investment in infrastructure and those which have not.[2] In our analysis we also address the issue of spillovers, an important consideration in densely populated urban centers.

[1] The literature on the effectiveness of financial rewards in improving the impact of community-engagement projects is mixed. While there is some evidence that they can work in certain contexts (see for example, Kremer et al. (2009) and Cappelen et al. (2013)), the evidence for using non-financial rewards is also appealing (see for example, Ashraf et al. (2013) and Ariely et al. (2013)). Moreover, in the context of our study, an intervention offering financial rewards would have been politically impossible to implement.

[2] Our intervention took place in the context of a larger infrastructure project that was ongoing in the region, meaning that some areas had received greater infrastructure investment than others at the time when our program was implemented. This was taken into consideration in the experimental design.

In our context, we have both a linear public good (cleanliness) and a threshold public good (lack of flooding). As we are most interested in a reduction in flooding, the literature on threshold public goods is particularly relevant. In the case of a threshold public good, an efficient equilibrium can exist where the public good is provided. However, an inefficient equilibrium with zero contributions to the public good also exists. In this paper, we seek to investigate whether it is possible to shift a large community that is stuck in an inefficient equilibrium towards the efficient one, by targeting a key subgroup within the community. While threshold public goods games have been tested in laboratory experiments, the literature on testing these kinds of interventions in the field is much sparser.[3]

The design of our intervention draws on a number of findings in the literature related to behavior in public goods games. Firstly, a higher threshold has been shown to discourage contributions (Cadsby and Maynes, 1999) while higher step returns to the public good have been found to increase contributions (Croson and Marks, 2000). In our intervention, by providing incentives for CBO members to achieve a clean neighborhood, the return for members to achieving the threshold is increased and also the level of threshold required to receive a reward is reduced (relative to the threshold required to reduce flooding). Secondly, other factors that have been shown to be important for successful provision are the size of the group (Croson and Marks, 2000; Feltovich and Grossman, 2015) and the degree of interconnection of group members (Ostrom, 2009). As this intervention is targeted at existing CBOs, the size of the group that must work together to contribute to the public good is smaller than the community as a whole and is also a group that has a high level of existing social capital. Finally, the existence of leaders and coalitions can help to increase the provision of the public good, although this also depends on the returns to members of the coalition, and the likelihood of free-riding by non-members (Moxnes and van der Heijden, 2003; Gächter and Renner, 2018; Bosettia et al., 2017). The CBOs in our context can play a leadership role within the community where their actions are highly observable by other community members, potentially shifting beliefs about the equilibrium and thereby encouraging contributions by non-members. We contribute to this literature by providing evidence from an RCT that these types of mechanisms can be effective at overcoming the public goods problem in a densely-populated urban setting.

Our paper also contributes to the body of literature exploring the role of CBOs in the implementation of development programs, in particular in relation to public goods provision. The 2004 World Development Report (World Bank, 2004) highlighted the need for community-level engagement to solve collective action problems. In general, the evidence on the effectiveness of community-based approaches in the provision and maintenance of public goods is mixed. There are some examples of such approaches having a positive effect on access to water and sanitation (Beath et al., 2013; Guiteras et al., 2015), but evidence on projects that aim to address the maintenance of public infrastructure is scarce. As highlighted by Mansuri and Rao (2004), the success of such projects will depend on the specific context, the time horizon, and the design of the community-based scheme. Moreover, recent evidence from randomized control trials of community-based interventions suggests that the mechanisms for effective collective action will also depend on the context and type of public good in question. Support for grassroots community-level engagement as opposed to involving local authorities is provided by Sheely (2013) for the case of anti-littering in Kenya. In contrast, Jack and

[3] One exception is Carlsson et al. (2015) which investigates the impact of different strategies for eliciting contributions to a real threshold public good (a bridge) in Vietnam. They find significant effects of both social influence and defaults on the size of contributions but only when the reference or default contributions are low.

Recalde (2015) find that the involvement of elected authorities is more effective than randomly selected community members in the provision of education material for schools in Bolivia. This suggests that the type of public good may matter.

Mansuri and Rao (2013) distinguish between "induced" and "organic" forms of participation, where the former refers to external mobilization and facilitation of collective action and the latter to civic participation and collective action that arises endogenously. Our study does not attempt to create new structures, but rather examines a community-centered approach which leverages existing community groups for the management and preservation of public spaces. Ansink and Bouma (2013) present a theoretical model that shows that providing incentive payments to a coalition within the community could encourage contributions to the public good. Fafchamps (2006) also highlights the role that social capital can play in development and, in particular, discusses the role that leaders can play in mobilizing a community to contribute to a public good. The CBOs in our setting could potentially fulfill this leadership role within the community. The intervention tested in this paper is relatively light-touch and does not involve financial rewards for any community members, but it does involve providing non-monetary incentives. Our results demonstrate that incentives, including social recognition and low-value in-kind incentives, can be effective at encouraging contributions to the public good when they are targeted at a visible subgroup within the community that has high levels of existing social capital among its members.

The rest of the paper is structured as follows. Section 2 provides more background on the context of the study and describes the intervention in more detail. In section 3 we present a simple theoretical framework to guide our analysis. The experimental setup is presented in Section 4 and Section 5 describes the data and the identification strategy. Section 6 presents the results and section 7 concludes.

2. Context and OQP Intervention

Our study tests the effectiveness of an intervention we designed called Operation Clean Neighborhood (*Opération Quartier Propre*, or OQP).[4] The objective of OQP was to encourage CBOs to work within their local community to improve and maintain the cleanliness of public spaces and drainage infrastructure, with the overall aim of reducing flooding in these communities. Through the promise of public recognition, cleaning materials and small prizes such as t-shirts, CBOs were incentivized to try to achieve a certain level of cleanliness within their neighborhood. Specifically, the project was implemented in two cities, Pikine and Guediawaye, which represent about 12% of Senegal's population. Lack of urban planning and rapid dense population growth in the 1970s to 1990s has led to disastrous environmental consequences, such as desertification, trash accumulation and blockage of natural drainage systems. As a result, the region has suffered from the Sahelian Paradox, which takes the form of intense flooding during each rainy season.

This intervention took place in the context of a large infrastructure investment project, the World Bank-assisted Senegal Stormwater Management and Climate Change Adaptation Project (*Projet de Gestion des Eaux Pluviales et d'Adaptation au Changement Climatique*, or PROGEP). PROGEP involved the construction and rehabilitation of drainage infrastructure

[4] OQP was designed and implemented in partnership with the Senegalese Municipal Development Agency (Agence de Développement Municipale).

and community engagement activities to inform residents of project areas of flood risks and prevention strategies and empower them to take action on flood risk prevention and mitigation.[5] The major investment under PROGEP was the construction and rehabilitation of new drainage infrastructure in flood-prone areas. However, for such an infrastructure investment to be effective and sustainable, communities must be involved in keeping their local environment clean and in particular managing the disposal of household waste, especially in the absence of a well-functioning solid waste management system. Widespread dumping of trash in public spaces results in clogging the stormwater drainage system, reducing its effectiveness and ultimately leading to flooding in the neighborhood. Our intervention, OQP, was implemented in PROGEP areas to test whether or not it could be effective in encouraging this kind of behavioral change.

OQP assumes that established CBOs are best at mobilizing local populations to contribute to public goods. In addition, their knowledge of the local context is what makes them well-situated to formulate activities to keep neighborhoods clean. In this sense, OQP is designed to be a light-touch and results-based intervention, where CBOs are given little instruction on how and which activities to implement. The objective is thus to change behaviors and beliefs rather than to provide new tools. The expectation is that this will help sustain potential impacts, as communities can easily continue activities after the program, but also facilitate scale-up and reproducibility due to OQP's low intensity.

OQP functions as follows:

i. A focal CBO is identified in each neighborhood. Criteria for selection of a CBO are fairly broad and include that its activities should be mainly confined to a single neighborhood (to maximize local knowledge and minimize spillovers), that it should have some minimum capacity needed to participate, and that it should have some influence within the community.[6]
ii. CBOs are supported in the development of action plans, giving them guidance on how to innovate and implement cleaning activities in their local neighborhood.
iii. CBOs receive an initial endowment package, consisting of cleaning tools and materials. The contents of the endowment are detailed in Table 1.
iv. CBOs sign letters of engagement with the commune's mayor, affirming their intent to participate in OQP. This is considered the start of the intervention.
v. CBOs are first evaluated at mid-term, six months after the start of the intervention. The evaluation criteria are described in Table 2.[7] CBOs that pass the evaluation threshold are recognized through a public ceremony and receive in-kind low-value goods such as plastic chairs, cooking utensils, and t-shirts.
vi. CBOs are evaluated a second time 12 months after the start of the intervention. Rewards for passing the threshold score included a similar ceremony and token prizes such as t-shirts and cooking utensils. All selected CBOs were aware of the prizes and the cleanliness criteria required to meet the threshold grade.[8]

[5] PROGEP was implemented by the Senegalese Municipal Development Agency (Agence de Développement Municipale).
[6] If a flood management CBO was present in the community, it was automatically selected as the focal CBO. This was the case for 3% of the CBO sample.
[7] A neighborhood had to achieve a grade of 60 or more in order to be awarded the mid-term or the final prize with the same grading scale applied to both evaluations. 68% of neighborhoods received the mid-line prize.
[8] 64% of neighborhoods received the end-line prize.

PROGEP's infrastructural work was rolled out in two phases, with the infrastructure component implemented in Phase 1 areas between 2014 and 2016 and in Phase 2 areas between 2017 and 2019. Community engagement activities, including OQP, were implemented by a different local NGO in each phase. The randomization of OQP was stratified across these two phases of infrastructure construction and was implemented simultaneously in these areas. This means that new infrastructure had already been constructed in Phase 1 areas but had not yet been constructed in Phase 2 areas when OQP was implemented. This allows us to investigate the heterogeneous impact of OQP, depending on the existing level of infrastructure in the area.[9]

3. Conceptual Framework

The aim of OQP is to engage local CBOs to work within their community to improve and maintain the cleanliness of public spaces and drainage infrastructure. Because of the short time frame involved in this experiment, we do not expect social norms around cleanliness to change as a result of the intervention. Instead what we have in mind is a situation where an external stimulus can help communities to coordinate their actions and shift from an inefficient equilibrium to an efficient one. The empirical question is whether OQP can provide this stimulus. Not much is known about how to induce behavioral change, particularly in densely populated urban settings, but we hypothesize that any lasting social change must be rooted in existing community structures. For this reason, OQP is targeted at existing, established CBOs, which have strong social connections and influence within the community.

In what follows we develop a simple model that helps to provide a conceptual framework for thinking about our experiment and the underlying mechanisms at work. We begin with a standard public good game played at the level of the neighborhood.[10] The public good is a clean neighborhood and we assume that there is a cost to each individual of contributing to the public good, e.g., taking their rubbish to a collection point instead of dumping on the street, etc. Each neighborhood is composed of n identical inhabitants, each of whom has an endowment E. Each community member can choose to make a contribution, c_i, to the public good, where $c_i \in [0, E]$. Decisions about how much to contribute are made simultaneously. All participants benefit from the public good, regardless of their contribution.

There are two potential benefits to keeping the neighborhood clean. First, there is the general benefit from having a cleaner neighborhood which is more pleasant to live in and might have additional positive impacts, for example better health or increased private investment. This can be thought of as a standard linear public good.[11] The more an individual exerts effort to keep the neighborhood clean, the better the public good but these benefits are shared with everyone. Second, having a clean neighborhood may impact on flooding, but only when cleanliness passes a certain threshold. A small amount of extra cleaning is not likely to have much impact on flooding – it is necessary to reach a minimum level of cleanliness in order for the infrastructure to function properly.

[9] While this was not part of our original pre-analysis plan, it became apparent after the baseline that the initial level of infrastructure may be an important factor for the effectiveness of the intervention. The difference in the availability of new infrastructure across the two phases allows us to consider heterogeneity along these lines ex-post.

[10] In our theoretical framework, we abstract away from the possibility of spillovers across neighborhoods and assume that the public good is limited to the boundary of the neighborhood. However, we address the possibility of spillovers in the empirical section.

[11] We are specifically interested in the cleanliness of public spaces in the neighborhood, rather than the area immediately around the house of a particular individual, which would not be a pure public good.

Initially, we will focus on the situation where there is no risk of flooding so the only public good is a clean neighborhood. We assume that individuals are risk neutral and all individuals receive a portion a of all contributions to the public good, where $\frac{1}{n} < a < 1$. The individual's utility function is as follows[12]:

$$u_i = E - c_i + a\sum_{j=1}^{n} c_j$$

$$\frac{\partial u_i}{\partial c_i} = -1 + a < 0$$

In equilibrium, the optimal contribution is therefore $c_i = 0$.[13]

We next extend the model to consider the additional potential benefits to individuals from reduced flooding as a result of having a cleaner neighborhood. In addition to receiving the benefit from a cleaner neighborhood, if the sum of contributions in a neighborhood is above a certain threshold, T^F, then flooding in the neighborhood is reduced and inhabitants receive a benefit of R^F (regardless of their individual contribution to the public good). Assuming that $R^F > \frac{T^F}{n}$ and $E < T^F$ (with the latter implying that one individual alone could not contribute enough to reach the threshold), the individual's utility function is:

$$u_i = E - c_i + a\sum_{j=1}^{n} c_j \qquad if \quad \sum_{j=1}^{n} c_j < T^F$$

$$u_i = E + R^F - c_i + a\sum_{j=1}^{n} c_j \qquad if \quad \sum_{j=1}^{n} c_j \geq T^F$$

This will give rise to multiple equilibria. We focus on the symmetric pure strategy equilibria, of which there are two[14]:

$$c_i = 0 \tag{1}$$

$$c_i = \frac{T^F}{n} \tag{2}$$

The former inefficient equilibrium implies that no individual will contribute to cleaning the neighborhood but the latter efficient equilibrium could be achieved through coordination of the beliefs within the community. The second equilibrium would imply that each individual would contribute their share in order to reach the threshold whereby flooding is reduced. If this is the prevailing equilibrium, then the intervention will have no effect as communities are already

[12] We assume here that individuals do not have social preferences, and therefore do not consider the impact of their actions on others and are not influenced by social norms. These kinds of social preferences could be included in the utility function but would not add anything to the analysis, since we assume that the intervention will not change social norms in the short run.

[13] In an infinitely repeated game, it is possible to have equilibria with positive contributions as long as players are sufficiently patient. But the no-contribution equilibrium will still exist. It is possible that the intervention could help shift beliefs about which equilibrium the community is in and move towards the efficient equilibrium. We do not consider this mechanism in this paper, however, given the relatively short time horizon of the intervention.

[14] There is an infinite number of efficient equilibria where the threshold for public good provision is exactly met but with varying levels of contributions by individual members. Only one inefficient equilibrium exists.

contributing optimally to the public good to achieve reduced flooding. However, in cases where communities are initially in the inefficient equilibrium then it may be possible to shift that equilibrium through interventions that help to coordinate the activities of individuals within the community.

In this paper, we test whether providing an incentive for a subset of the community can shift the equilibrium and therefore improve outcomes for the community as a whole. We do this through an intervention that rewards Community Based Organizations (CBOs) in neighborhoods that achieve a level of cleanliness above a certain threshold. This could change the payoff from contributing to the public good for members of the CBO.[15] We assume that a CBO has $m < n$ members. The reward for reaching the flooding threshold is the same as before, but now there is an additional reward for reaching the cleanliness threshold required to get the social recognition. The cleanliness threshold and associated reward are denoted by T^C and R^C, respectively. We assume that $T^C < T^F$. The utility function for CBO members will be:

$$u_i^{CBO} = E - c_i + a\sum_{j=1}^{n} c_j \qquad if \ \sum_{j=1}^{n} c_j < T^C$$
$$u_i^{CBO} = E + R^C - c_i + a\sum_{j=1}^{n} c_j \qquad if \ T^F > \sum_{j=1}^{n} c_j \geq T^C$$
$$u_i^{CBO} = E + R^C + R^F - c_i + a\sum_{j=1}^{n} c_j \qquad if \ \sum_{j=1}^{n} c_j \geq T^F$$

We consider two possible scenarios.[16] First, we consider the case where the CBO is large enough to be able to reach T^C themselves, i.e. $\frac{T^C}{m} < R^C$. Members communicate regularly and observe each other's actions, and so it seems reasonable to expect that they may be able coordinate on the equilibrium where they contribute to the public good. In this equilibrium, they will each contribute $c_i = \frac{T^C}{m}$ where m is the total number of members in the CBO.

It is also possible in this scenario that the contributions of the CBO members are observed by the general community. This can lead to two possible outcomes: 1) this has no impact on the beliefs of other members of the community so while the cleanliness of the community increases the threshold required for reduced flooding is not reached and the intervention has no impact on flooding; or 2) the actions of the CBOs lead to a focal point for equilibrium (2) above where other members of the community contribute to the cleanliness of the community and the flooding threshold is met. In this latter case the cleanliness of the community increases, and flooding is reduced.

[15] One other potential mechanism through which the intervention could work is that it provides information and salience. We do not believe that this is the key mechanism, since an information campaign is ongoing in both treatment and control communities as part of the wider PROGEP project. It is unlikely that an additional information effect associated with our intervention would be significant enough to explain our results.
[16] We do not consider the case where the CBO is large enough to be able to meet the flooding threshold by themselves, as we assume that if this were possible, then they would already have done so, even without this intervention.

Second, it may be the case that the CBO is not large enough to be able to reach T^C themselves, i.e. $\frac{T^C}{m} > R^C$. In this case, the CBO needs to encourage members of the community to also contribute. CBO members can exert effort, e, in order to encourage members of the community to contribute to the cleanliness of the community. Assuming $e \in \{0, e^*\}$, the utility function for CBO members will be:

$$u_i^{CBO} = E - c_i + a\sum_{j=1}^{n} c_j - e \qquad if \quad \sum_{j=1}^{n} c_j < T^C$$
$$u_i^{CBO} = E + R^C - c_i + a\sum_{j=1}^{n} c_j - e \qquad if \quad T^F > \sum_{j=1}^{n} c_j \geq T^C$$
$$u_i^{CBO} = E + R^C + R^F - c_i + a\sum_{j=1}^{n} c_j - e \qquad if \quad \sum_{j=1}^{n} c_j \geq T^F$$

The utility function for non-CBO members will be[17]:

$$u_i = E - c_i + a\sum_{j=1}^{n} c_j \qquad if \quad \sum_{j=1}^{n} c_j < T^F$$
$$u_i = E + R^F - c_i + a\sum_{j=1}^{n} c_j \qquad if \quad \sum_{j=1}^{n} c_j \geq T^F$$

A subgame perfect equilibrium exists, where the strategy of non-CBO members is to contribute $\frac{T^F}{n}$ if they observe $e = e^*$ for all CBO members, and not to contribute otherwise. The strategy of CBO members is to contribute $\frac{T^F}{n}$ and choose $e = e^*$ if $R^C + R^F \geq \frac{T^F}{n} + e$, and to contribute 0 and choose $e = 0$, otherwise. Assuming $R^F < e$ – otherwise, CBOs would have been incentivized to exert this effort even in the absence of the intervention – we expect to see improved cleanliness, reduced flooding and actions by CBOs such as cleaning events, information campaigns or training in relation to flooding.[18]

To summarize, where an inefficient equilibrium exists and individual community members do not contribute to the provision of a public good, this behavior can potentially be shifted by engaging CBOs to motivate their members to contribute to the provision of the public good, which in turn can potentially shift the behavior of non-members thereby reaching a new efficient equilibrium where the public good is provided. It is through this mechanism that we expect the OQP intervention to impact on flood-related outcomes.

It is possible that the effects of the intervention will be heterogeneous across different types of CBOs. Of particular importance is the size of the CBO. As suggested above, larger CBOs may

[17] The effort of CBO members does not directly benefit non-CBO members and therefore does not enter their utility function directly. Instead it acts as a signal observed by non-CBO members, which influences their beliefs about the strategies of others and therefore influences their own choice of strategy.

[18] Another equilibrium exists where the community members do not contribute regardless of the choice of e, and therefore CBO members choose $e = 0$. In this case, there will be no impact of the intervention on cleanliness or flooding.

be more likely to reach the cleanliness threshold for receipt of the reward by themselves and so will not have an incentive to engage community members. If this is the case it may be that the cleanliness of the neighborhood improves as a result of the intervention but the level of engagement of non-members in the community does not so the cleanliness threshold necessary for reducing flooding may not be reached. It is also possible, however, that the cost of reaching the flooding threshold is lower for non-members in a community where there are many contributing CBO members, and so larger CBOs may be more likely to end up in the contribution equilibrium which will result in reduced flooding. Smaller CBOs will be unable to reach the cleanliness threshold without contributions by community members and so will have a greater incentive to both coordinate actions among members and mobilize the community. The extent to which the size of the CBO may matter for reducing flooding is an empirical question.

Underlying differences in the neighborhoods in terms of their vulnerability to flooding could also have an impact on the effectiveness of the intervention. Areas that are more prone to flooding may require a higher threshold of cleanliness (T^F) to be reached in order to have an impact on flooding. In addition, areas that are receiving more drainage infrastructure may receive a lower reward (R^F) from reaching the contribution threshold. Either of these factors, could discourage contributions to the public good (Cadsby and Maynes, 1999). As Phase 1 areas are both more prone to flooding and also received the infrastructure earlier, there are reasons to believe that the intervention might have a different impact in these neighborhoods than the neighborhoods located in Phase 2.

4. Experimental Design

We use a randomized controlled trial to evaluate the impact of OQP on cleanliness and other flood-related outcomes. The unit of randomization is the *quartier*, or neighborhood. Pikine and Guediawaye are comprised of a set of 16 and 5, respectively, communes which are formal political structures with an elected local government. Neighborhoods are primarily informal geographic groupings within those communes; though there are often local governance structures and community leaders, these are informal institutions which vary in nature across neighborhoods. In total, there are 395 neighborhoods in the PROGEP area. For political reasons, to ensure even representation across the area, neighborhoods were stratified into 48 sub-groups on the basis of geographic and social ties. The total number of communes, sub-groups and neighborhoods in each area is provided in Table 3. Of the 395 neighborhoods, 160 were randomly selected for inclusion in our study with proportional representation of neighborhoods in each sub-group. This number was primarily informed by the available budget for OQP, which was sufficient to cover 80 treatment neighborhoods. Randomization of neighborhoods was done within these sub-groups with the number of treatment and control neighborhoods proportional to the size of the group. A total of 80 neighborhoods were assigned to the OQP treatment group and 80 to the control group. CBOs were selected in *all* 160 study neighborhoods, prior to random assignment, and each focal CBO within a particular sub-group was invited to a public meeting where the lottery took place.

Random assignment into treatment and control neighborhoods was carried out through a public lottery held with representatives of all CBOs selected into the study, stratified by grouping, in which the representatives themselves drew their CBO's treatment status. A public lottery was selected for maximum transparency and to avoid allegations of corruption or clientelism to the maximum extent possible. Operationally, this was considered essential, especially as the proximity of neighborhoods and the population density made it highly unlikely that control neighborhoods would not eventually find out about the intervention. We pay close attention to spillovers in our analysis. A map of the intervention area and the selected neighborhoods in Phase 1 and Phase 2 of the larger infrastructure project, PROGEP, is provided in Figure 1. The full time-line of the project is described in Figure 2.

Finally, we note that PROGEP was rolled-out in two phases.[19] PROGEP's infrastructure component (rehabilitation of old water basins, implementation of new drainage pipelines, etc.) was implemented in Phase 1 areas between 2014 and 2016 and in Phase 2 areas between 2017 and 2019. Given that the former benefitted from improved drainage infrastructure before the latter, we might expect to see differences in the effect of the OQP intervention in each Phase. This is something we take into account in our empirical analysis.

5. Data collection and empirical approach

Data collection

Data were collected at the individual, household and CBO levels. In each neighborhood, 15 households were randomly chosen to be surveyed, giving a total sample size at baseline of 2,400 households. The household questionnaire collected information on household demographics, livelihoods and income sources, socioeconomic characteristics, health outcomes, exposure to flooding, knowledge of risk mitigation methods, and attitudes towards community participation and one's general responsibilities vis-à-vis the community (and vice-versa). Within each household, data relating to certain variables (education, health, etc.) were collected for all individual household members, with a total sample size of 28,010

A CBO-level questionnaire was also administered as a group survey. This survey focused on basic group characteristics, their motivations for participating, their attitudes towards civic participation, and the nature of CBO activities. The end-line survey differed slightly for treatment and control CBOs as it aimed to collect data to allow us to gauge the awareness of control CBOs about OQP and whether or not this had influenced their activities, in order for us to construct a measure of potential spillovers.

The baseline survey was carried out in July 2015.[20] The end-line survey was carried out during November 2016. The attrition rate was low. Out of the 2,400 households surveyed at baseline, only 115 could not be included in the end-line survey (an attrition rate of less than 5%).

Baseline characteristics

Table 4 presents the average values for a range of baseline household characteristics for the treatment group and the control group, along with the difference between groups and the p-value of a t-test of whether this difference is statistically significant at conventional levels (p-value < 0.10; this is identifiable by the presence of one or more stars next to the p-value).

The first set of variables relates to the characteristics of the household head. Approximately 70% of household heads are male and the vast majority (96%) are Muslim. Most are married

[19] Phase 1 covers Dalifort-Thiouorour (communes of Wakhinane Nimzatt, Djeddah Thiaroye Kao, Dalifort-Foirail, Hann Bel Air and the Western halves of Yeumbeul Nord and Yeumbeul Sud) and Phase 2 covers Yeumbeul-Mbeubeuss (Keur Massar, Malika and the Eastern halves of Yeumbeul Nord and Yeumbeul Sud).

[20] The baseline survey took place before neighborhoods were assigned to the treatment and control groups, to avoid any anticipation effects from CBOs and households.

(around 80%), either in monogamous or polygamous marriages. The average age of household heads is 55 and the average household size is 10 people. For household heads for whom data are available on salary, the average daily salary in FCFA is around 11,000 (approximately USD 19). There are no statistically significant differences between the mean head of household characteristics across treatment and control households at baseline.

The second set of variables relates to the perceptions of households in terms of the cleanliness of the neighborhood, whether they were victims of flooding,[21] and whether they received any training in relation to flood prevention. Around 30% of households rate their neighborhood as clean and 20% were victims of flooding in the previous year. There are no statistically significant differences between the mean of these variables across treatment and control households at baseline. Households were also asked, at baseline, whether any members received training in relation to flood prevention. We find that 16% of households in the control group and 13% of households in the treatment group received training and that this difference is statistically significant. If households in the control group are more informed about flood prevention measures, then this would likely work against us finding an effect. We nonetheless ensure that all our results are robust to the inclusion of baseline values for training in the empirical specification.

Table 4 also presents balance tests for treatment and control groups within each Phase where we also achieve reasonably good balance. There are some exceptions. In Phase 1, a slightly higher proportion of household heads in the control group are married but this difference is only statistically significant at the 10% level. We achieve balance on all the other characteristics within Phase 1. In Phase 2, we find that households in the treatment group are more likely to have been flood victims in the previous year and are less likely to have received training about flood prevention. Both differences would work against us finding an effect. We also include values at baseline in the empirical specification to correct for any baseline differences.

Table 5 presents descriptive statistics on the characteristics of the CBOs at baseline, focusing again on the balance between treatment and control groups.[22] The first set of characteristics relates to the flood prevention and cleaning activities of the CBO. A number of CBOs (approximately 70%) undertook some action in the fight against flooding in the 12 months prior to the baseline survey. Many are also engaged in activities to counteract flooding including raising awareness and providing financial assistance for victims of flooding. Around 19% of CBOs have as their objective raising awareness on flood-related issues or organizing cleaning events.[23] The average number of days per year spent on flood communication or cleaning activities is 12 and approximately 49 members engage in these activities.

We achieve balance on almost all these measures with the exception of whether the CBO engaged in cleaning activities, where we find a marginally statistically significant difference in the proportion of CBOs engaged in cleaning activities at baseline, with control CBOs more

[21] While in the theoretical framework we considered the level of flooding in the neighborhood to be a public good and therefore the same for all members of the community, in practice, the actual amount of flooding experienced will vary by household even within the same community. We can therefore think of the public good as a reduction in the probability of experiencing flooding within your community, with the actual level of flooding experienced measured at the household level.

[22] We also perform balance tests within Phase 1 and Phase 2. The results are presented in Tables A1 and A2 of the Appendix, respectively.

[23] If a flood prevention CBO was present in a community, then this was automatically chosen as the focal CBO.

likely to do so. This difference would likely work against us finding an effect of the treatment. Moreover, all the results presented in later sections are robust to the inclusion of these baseline CBO controls.

Table 5 also includes baseline statistics for other characteristics of the CBO. The average number of members is around 177 with around 39 members present at the previous meeting. On average, around 94 of their members engage in the activities of the CBO. It should be noted, that the variance in the size of the CBOs is large with a few very large CBOs driving up the mean. The median number of members is 67 with a median of 50 members taking part in CBO activities. On average members engage in CBO activities for 12 hours per week. This suggests that individuals in our study area are indeed very actively engaged with CBOs. The majority of members of CBOs are women (around 60%). The proportion of members that are youths is small at around 5%. Around two-thirds of CBOs vote in their leaders by election.

There are some differences across the treatment and control groups in these characteristics at baseline. In particular, CBOs in the control group report more active engagement of their members. We also find that the CBOs in treatment neighborhoods are less likely to vote in their leaders by election, but are more likely to be a CBO head office and are more likely to be a flood prevention committee. All our results presented in subsequent sections are robust to the inclusion of baseline controls.

The final set of CBO characteristics we consider relates to the engagement of CBOs with CBOs in other neighborhoods. This is important given the close physical distance between neighborhoods in our sample and the likelihood that the activities of CBOs in the treatment area might affect the activities of CBOs in the control areas, potentially contaminating our experiment. We find that while there is some degree of collaboration between CBOs within the neighborhood and with CBOs in neighboring neighborhoods, there is no statistically significant difference between the treatment and the control groups on these characteristics. However, given that 63% of CBOs in the treatment group collaborate with outside CBOs spillovers are likely. The possibility of spillovers is explored further in the empirical analysis.

Empirical approach

Because our treatment was randomly assigned, a direct comparison of outcomes between our treatment and control groups will give us a causal estimate of the impact of the program on those outcomes. The basic specification that we use is as follows:

$$Y_{ij1} = \alpha + \beta_T T_j + \beta_P Phase1_j + \delta Y_{ij0} + \theta X_{ij} + \phi_S S_j + \varepsilon_{ijs} \tag{3}$$

Y_{ij1} represents our outcome of interest for household i, in neighborhood j, at endline. T_j is a dummy variable indicating treatment at the neighborhood level. $Phase1_j$ is a dummy variable equal to 1 if the neighborhood is in the Phase 1 area. Y_{ij0} is a measure of the outcome variable of interest at baseline, which is included when available. X_{ij} is a set of baseline control variables. S_j are dummy variables for the strata or sub-groupings of neighborhoods.

Phase 1 neighborhoods began to benefit from PROGEP drainage infrastructure before OQP was implemented whereas Phase 2 had yet to receive infrastructure at the end of the intervention. In addition, the OQP intervention was implemented by different local NGOs in each of the two phases. Hence, there are a number of reasons why we might expect the impact of the treatment to be different in the Phase 1 and Phase 2 areas. Therefore, we also include an interaction term between the Phase 1 dummy indicator and the treatment indicator in a number of specifications:

$$Y_{ij1} = \alpha + \beta_T T_j + \beta_P Phase1_j + \gamma_T Phase1_j * T_j + \delta Y_{ij0} + \theta X_{ij} + \phi_S S_j + \varepsilon_{ij} \quad (4)$$

6. Results

As described in Section 3, the aim of our intervention is to motivate CBOs to engage in cleaning activities and to encourage others in the neighborhood to do the same. We would expect to see increased levels of cleanliness in treatment neighborhoods as CBO members attempt to reach the threshold level of cleanliness for the non-monetary reward. Where the intervention also manages to shift the behavior of non-members so that they too contribute to the public good, we would also expect to see reduced levels of flooding. For this mechanism to work it is important that there is a direct link between the cleanliness of the neighborhood and the probability of experiencing flooding. Before beginning our main analysis, we explore the correlation between cleanliness and flooding at baseline. To measure cleanliness, households were asked to rate the cleanliness of their neighborhood on a 5-point scale ranging from 'Very Clean' to 'Very Dirty'. A 'Clean' dummy was created which is equal to 1 if the household rated their neighborhood as 'Clean' or 'Very Clean'. We run a simple regression of the probability of experiencing flooding in the previous year on this measure of cleanliness and find a statistically significant negative correlation suggesting that households that report having cleaner neighborhoods are 5% less likely to have experienced flooding in the previous winter season.[24] While this evidence is not causal, it does provide us with assurance of a positive relationship between increased cleanliness and decreased flooding.

We structure the presentation of our main results in line with our conceptual framework. First, we examine the extent to which the OQP led to increased cleaning activities by the CBOs and households. Second, we examine whether the OQP led to increased levels of cleanliness and reduced flooding. Third, we explore heterogeneity in the impact of the intervention on flooding across the size distribution of CBOs. Finally, given the close proximity of neighborhoods and the fact that selection into treatment was done by public lottery, we examine whether there is evidence of spillovers of the OQP into bordering neighborhoods.

Impact of OQP on CBO behavior

Since the OQP program was implemented via CBOs, we begin by looking at the impact that treatment had on their behavior. The key outcome that we are interested in is whether or not the CBO carried out cleaning events as part of its activities.[25]

The results are presented in Table 6. Excluding baseline controls, we can see in column 1 that CBOs in treatment neighborhoods are 18.8 percentage points more likely to include cleaning events in activities compared with those in control neighborhoods. Once baseline controls are included (column 2), this increases to 23.2 percentage points. This provides evidence that

[24] The results are presented in Table A3 of the Appendix.

[25] It should be noted that power calculations were conducted on the basis of household-level outcomes and not CBO-level outcomes. We cannot rule out that null effects in relation to CBO outcomes are due to lack of power. The number of neighborhoods and, with this, the number of CBOs was fixed due to the budget of our implementing partner.

CBOs in treatment neighborhoods were more likely to engage in keeping the neighborhood clean. We also explore heterogeneity across the two infrastructure phases but do not find a significant difference in the treatment effect between Phase 1 and Phase 2 areas. This suggests that visible investment in the community by an external authority does not affect the likelihood that CBOs will engage in an OQP-type intervention.

Impact of OQP on household knowledge

We next consider the impact of treatment on household knowledge of OQP and flood-risk prevention for variables for four different specifications. First, a simple bivariate regression. Second, a regression including the baseline value of the outcome variable (where available), group (randomization strata) fixed effects; baseline household controls (reported level of cleanliness of the community and whether or not any household member is a member of the CBO); and baseline CBO characteristics (size of the CBO, the number of hours a week members on average engage in CBO activities, type of CBO, an indicator for whether an objective of the CBO is to reduce flooding in the community, and a dummy indicator for whether the CBO engaged in flood cleaning activities at baseline). Third, we also control for whether the neighborhood is in Phase 1 or Phase 2. Fourth, we consider separately the differential effect of the treatment in each phase. Given that outcomes are self-reported by households we include enumerator fixed effects to avoid any specific experimenter demand effects in specifications 2 to 4. Tests of statistical significance are based on standard errors that are clustered at the neighborhood level which is the unit of randomization. We also present the randomization inference p-values which account for the sample stratification and correct for the cluster randomization.[26]

Table 7 presents the results for a regression investigating the impact of the OQP program on whether or not households had heard of OQP. This variable is important for two reasons: first, it provides a measure of whether or not treatment CBOs were active in executing OQP; second, households may be more willing to contribute to reaching the flood-prevention threshold of cleanliness if they are aware that others in their community are working actively towards this. We find a statistically significant difference in the awareness of the program between treated and control households, suggesting that CBOs did engage households in treatment neighborhoods.[27] The coefficient on the treatment indicator in column 2 shows that households in treatment neighborhoods were almost 20 percentage points more likely to have heard of OQP than households in control neighborhoods.

In column 3, the coefficient on the Phase 1 dummy is also statistically significant and positive, suggesting that households in Phase 1 areas are generally more likely to have heard of the program, which is not surprising given that they are receiving infrastructure and issues around PROGEP are likely to be more visible to them regardless of whether they are in OQP treatment

[26] To perform randomization inference, we randomly assign neighborhoods to treatment and control groups within each strata and estimate each specification. We repeat this exercise 1,000 times for each outcome and construct the p-value as the proportion of times that the absolute value of the randomization inference coefficient is greater than the absolute value of the actual coefficients from our sample (see Young, 2017). We use the STATA command developed by Heß (2017) for most specifications and our own code for specifications including interaction terms.

[27] The neighborhoods selected for OQP were chosen via public lottery, so we would expect that some households in control areas would have heard of the OQP. However, we would be concerned if there was no difference between control and treatment areas in terms of their awareness of the program, as this would suggest either that it was not implemented effectively or that spillovers were very high. Our results suggest that this is not the case.

or control neighborhoods. In column 4, we consider the difference in knowledge of OQP in the treatment group in Phase 1 and Phase 2 neighborhoods by including an interaction effect between the treatment indicator and the indicator for which phase the neighborhood is in. The effect for Phase 2 neighborhoods is now stronger as treated households are 25.2 percentage points more likely to have heard of OQP than control households. However, the coefficient on the interaction term with the Phase 1 dummy is negative, significant and quite large in magnitude (13.2 percentage points) implying that the effect of the treatment was smaller in Phase 1 areas.[28] A possible interpretation of this is that the construction of flood prevention infrastructure was underway in Phase 1 neighborhoods while in Phase 2 there were no such activities relating to flood prevention making the OQP more salient for treatment households. It is still the case that households in the treatment neighborhoods in Phase 1 areas are more likely to be aware of OQP than those in the control neighborhoods, but the magnitude of the effect is greater in Phase 2 areas.

Next, we investigate whether households in OQP treated areas were more likely to receive interventions related to cleanliness and flood reduction which would provide further evidence that CBOs did in fact engage with households in the community as part of OQP. We ask households whether they received training related to managing the risks of flooding in the last year. The results are presented in columns 5 to 8 of Table 7. We find a positive and significant effect of the treatment. Treated households are over 5 percentage points more likely to have received training than control households once we have included all relevant control variables. This is a large effect relative to a baseline mean for the control group of 17 percent. We do not find any statistically significant difference across Phases (column 7) or in in the impact of the treatment on training in Phase 1 and Phase 2 areas (column 8).

Impact of OQP on household outcomes related to cleanliness and flooding

We now examine whether the OQP had an impact on perceived levels of cleanliness and whether this translated into actual reduced flooding in the community. The first set of results examines the impact of treatment on our measure of cleanliness described above.[29] The results are presented in columns 1 to 4 of Table 8. Households in treatment areas are 12.3 percentage points more likely to give their neighborhood one of the higher cleanliness ratings than households in control areas and this difference is statistically significant (column 2). We find no differential impact of the treatment across Phase 1 and Phase 2 neighborhoods. Together with the results from above showing that there is no difference across Phases in the OQP related interventions conducted by CBOs, this suggests that the mechanics of the implementation of OQP were similar in both Phases.

To measure flooding, households were asked questions about their experience of flooding in the most recent rainy season and in the previous season. Both of these seasons occurred between our baseline and endline surveys, while OQP was running, although we might expect any effects to be stronger in the most recent season as OQP had been running for longer and

[28] It should be noted that the associated p-value using randomization inference is 0.112, and so is just outside conventional statistical significance levels.

[29] We find a positive and statistically significant correlation of 0.35 (0.41 including baseline control variables) between the household reported level of cleanliness of the neighborhood and the community-level score awarded by the OQP implementing NGOs to determine if the CBO won the prize after the end-line. This gives us some confidence that our measure of cleanliness is not picking up experimenter demand effects.

the first round of incentives had been delivered to the successful CBOs thereby giving credibility to the program.

The results for whether the household was a victim of flooding in the current year are presented in columns 5 to 8 of Table 8 while columns 9 to 12 present the results for whether the household was a victim of flooding in the previous year. The coefficients on the treatment indicator are negative in all specifications but are only statistically significant in the case of flooding in the most recent year. We find in column 6 that households in treatment neighborhoods are 2 percentage points less likely to have experienced flooding than those in control neighborhoods. When the interaction term with Phase 1 is included, the magnitude of the coefficient on the treatment dummy increases suggesting that the impact of OQP on reducing flooding was larger in Phase 2 neighborhoods. The probability of experiencing flooding in the most recent year declines by 3.9 percentage points in Phase 2; given that the average probability of experiencing flooding in control households in Phase 2 areas in this time period is 16.6 percent, the magnitude of this effect is meaningful. The coefficient on the interaction term is positive for all flooding outcomes and is statistically significant in the case of flooding in the most recent year, although it falls below the 10% cut-off when randomization inference p-values are used. Nevertheless, the result provides some tentative evidence OQP may not have been as effective at reducing flooding in Phase 1 areas. There are two possible explanations for this. On the one hand, given that Phase 1 had just received new flood prevention infrastructure the scope for OQP to reduce flooding may have been less than Phase 2. On the other hand, neighborhoods in the Phase 1 area are in general more vulnerable to flooding. This could mean that the threshold of cleanliness required to reduce flooding in these areas is higher than in Phase 2 neighborhoods.

Heterogeneous effects

As hypothesized in Section 2, we might expect the effectiveness of the OQP to depend on the type of CBO. In particular, the size of the CBO is likely to matter. Larger CBOs are more likely to reach the cleanliness threshold by themselves and so while the OQP might lead to higher levels of cleanliness in these neighborhoods it may not induce participation by non-CBO members and so may have less of an impact on flooding. It could also be, however, that where there are many CBO members engaged in cleaning activities the marginal cost for non-members to participate is lower, in which case they may be more likely to contribute. On the other hand, smaller CBOs will be unable to reach the cleanliness threshold by themselves and so may more actively engage non-members. We explore these possibilities empirically by introducing heterogeneity along the size distribution of CBOs. We include the number of members of the CBO at baseline (scaled by 100 for easier interpretation of coefficients) and its interaction with the treatment indicator. We also trim the sample to exclude six neighborhoods with very large CBOs with over 500 members (three treatment neighborhoods and three control neighborhoods). The results are presented in Table 9.[30]

In column 1 we see that the level of awareness of OQP among households in treatment neighborhoods is decreasing in the size of the focal CBO.[31] This suggests that larger CBOs

[30] We also considered whether the heterogeneous impact of the size of the CBO on outcomes differs across phases. The results are presented in Table A4 of the Appendix. The only statistically significant difference of note is that in Phase 1 areas quartiers with larger CBOs are perceived to be cleaner.

[31] It should be noted, however, that the randomization inference p-value in this case is above the standard 10% threshold.

may be less dependent on non-members for achieving the cleaning thresholds and may be less likely to engage with them. Indeed, in columns 2 and 3, we do not find any differential effect in the perceived cleanliness of the neighborhood or the probability of experiencing flooding. This suggests that larger CBOs may be more likely to reach the cleanliness threshold for receipt of the non-monetary reward and reduced flooding through their own actions rather than through engaging with other members of the community.

Secondary outcomes

We also consider whether there is any evidence for positive impacts on health, education and work-related outcomes for individuals living in OQP neighborhoods.[32] As mentioned in Section 2, because of the short time frame of the experiment, we do not expect there to be very large impacts on such secondary outcomes but it is possible that they are affected through both improved cleanliness and reduced flooding.[33] In Table 10, we find that individuals in treated neighborhoods are around 2 percentage points less likely to report that they have been sick in the previous 30 days, are less likely to have been ill in the most recent rainy season and, in particular, are less likely to have been ill due to malaria in the most recent rainy season. Given that at baseline, 22.4% of households in the control group reported that they were sick in the previous 30 days, the magnitude of the effect is relatively large. This suggests that the OQP, through its impact on cleanliness and reduced flooding, also impacts on the health of individuals living in the neighborhood. In Table 11 we explore the extent to which there are effects on the number of missed days at work and school and the number of days that schools were closed. While most coefficients are not statistically significant, the signs in all cases are negative and are negative and statistically significant for work days in Phase 2 (column 4).

Spillovers

Our study is set in a densely populated urban area. While communities have clear boundaries, they are in very close proximity to each other and so there is potential for spillover effects along a number of dimensions that could work in either direction. This is exacerbated by the fact that the selection into treatment was done by public lottery and so CBOs in both treatment and control neighborhoods know about OQP. It is therefore possible that households in untreated neighborhoods are also informed about OQP and change their behavior in response. On the one hand, they could observe the behavior of households in treated neighborhoods and copy this behavior in their own neighborhood. This would work against us finding an effect of the treatment. On the other hand, if households in untreated neighborhoods know that CBOs in treated neighborhoods are being incentivized to engage in cleaning behavior (albeit through non-monetary incentives) they may be discouraged from engaging in any cleaning activities. This would lead to an over-estimation of the impact of OQP on behavior. It is also possible that the proximity of neighborhoods means that poorly maintained public infrastructure in one neighborhood leads to flooding in a neighboring neighborhood, even if they perform well at

[32] Baseline balance tests for whether the household was ill in the last 30 days and the number of work days missed are presented in Table A5 of the Appendix. While the sample is balanced across treatment and control group in Phase 1, there are baseline imbalances in Phase 2. In phase 2, individuals are more likely to have been sick in the control group but miss fewer days of work. This should be borne in mind in interpreting these results.

[33] At baseline we find a statistically significant correlation between flooding and illness in the previous 30 days, with households that experienced flooding 5% more likely to report having been sick. There is also a small positive correlation between cleanliness and health, but this is not statistically significant.

keeping their community clean. If the latter are treated, then this will lead to a downward bias on our estimate of the impact of OQP on flooding. It might also impact on the behavioral response in treated neighborhoods if they feel that their efforts to keep the neighborhood clean are pointless.

We explore the extent to which there are spillovers of this kind by taking into account the treatment status of bordering neighborhoods. We first consider whether the level of awareness of OQP and the behavior of households in control neighborhoods is impacted by whether they share a border with a treated neighborhood. We estimate the following regression for the control group for our four primary household-level outcomes:

$$Y_{ij} = \alpha + \beta_T border\_T_j + \beta_P Phase1_j + \beta_B nr\_border_j + \delta Y_{ij0} + \theta X_{ij} + \phi_S S_j + \varepsilon_{ij} \qquad (5)$$

where Y_{ij}, $Phase1_j$, Y_{ij0}, X_{ij}, S_j and ε_{ij} are as for equation (3), $border\_T_j$ is a dummy indicator for whether neighborhood j shares a border with a treated neighborhood and nr_border_j is a control for the total number of bordering neighborhoods. Around 75% of control neighborhoods share a border with a treatment neighborhood, 86% in Phase 1 and 71% in Phase 2. Where relevant to the outcome of interest, we also include the same set of baseline control variables as included in the main regressions. We also separately include in each specification an interaction term between the border-treatment indicator and the dummy indicator for the Phase that the neighborhood is in.

Table 12 presents the results for how bordering a treatment neighborhood affects each of the key outcome variables for the control group. Bordering a treated neighborhood has no impact on the level of awareness of OQP of households (columns 1). When we include an interaction with the Phase 1 dummy indicator, however, we find that control households in Phase 1 that border a treatment neighborhood are more likely to have heard of OQP. This is likely due to the fact that infrastructure was being constructed in Phase 1 at that time, making PROGEP related activities, including OQP, more salient.

It is also possible that the behavior of households in control neighborhoods is impacted as a result of bordering treated neighborhoods, even if they have not heard of OQP. As suggested above, they may observe persons or groups in treated neighborhoods engaging in cleaning and flood prevention activities and decide to also engage. On the other hand, households that share a border with a treated neighborhood may be dis-incentivized and clean less or perceive their neighborhood to be less clean. As revealed in columns 5 and 6, we find no evidence that the cleanliness perceptions of households in control neighborhoods that share a border with treated neighborhoods are different to those that do not. We also do not find any evidence that these households are differentially affected by flooding. In sum, we find some evidence for spillover effects in terms of awareness of and activities relating OQP but only in Phase 1, which is not surprising given that they have already been exposed to other components of the larger infrastructure program. There do not appear to be any spillovers in relation to the main outcomes of interest, cleanliness and flooding.

As mentioned above, it is also possible that treated neighborhoods experience flooding as a result of poorly maintained public infrastructure in neighboring control neighborhoods. This could have two effects: 1) it could demotivate households in treated neighborhoods from keeping their neighborhood clean; and 2) it could lead to a downward bias in our estimate of OQP on flooding. To explore this, we estimate equation (5) for the treatment group and include an indicator of whether they border a control neighborhood. The results are presented in Table

13. We find some evidence in columns 1 and 2 to suggest that households in treatment neighborhoods that share a border with a control neighborhood are more likely to report their area as clean. This appears to be the case in both Phases but to a greater extent in Phase 1. This suggests that households are not demotivated by being in close proximity to control neighborhoods, although it may be that they perceive their neighborhoods as being relatively cleaner. Flood related outcomes in treated neighborhoods, however, are not affected by bordering control neighborhoods (columns 3 and 4), suggesting that there are no spillover effects from flooding in either direction.

7. Conclusions

Understanding how individuals in large, urban communities can coordinate to overcome public goods problems is extremely important. Using a field experiment, this paper tests the effectiveness of an intervention targeting a key subset of the community to motivate them to keep their local area clean, in order to increase the effectiveness of drainage infrastructure and therefore reduce the risk of flooding in that area. This was a light-touch, bottom up intervention: OQP engaged existing community-based organizations and empowered them to use their local knowledge and networks to work towards improved community cleanliness, while providing minimal guidance, basic materials, and relatively inexpensive incentives. We find that the program had a significant positive effect on households' likelihood of receiving training, their perception of the cleanliness of their neighborhood and their vulnerability to flooding. Our results demonstrate that providing incentives for a subset of the community to contribute to a threshold public good can lead to a shift to a more efficient equilibrium for the community as a whole.

This intervention took place in the context of a larger infrastructure program that was being rolled out in this area in two phases. In addition to the overall impact of OQP, we also find important differences in the results for Phase 1 and Phase 2 areas. This is relevant to note because the Phase 1 areas had received upgraded infrastructure by the start of the OQP intervention, whereas the Phase 2 areas had not. This provides important insight into the interaction between this type of community-based intervention and the provision of infrastructure by the state. The intervention led to greater awareness of the program in treatment areas and more significant impacts on flooding in Phase 2 areas, although the behavior of the CBOs seems to be similar in both areas. This suggests that the motivation of CBOs may not be affected by the investment by the state in public infrastructure. However, the impact of their actions on overall flooding levels may be. This raises the possibility that this type of intervention could act as a substitute for investment in infrastructure, at least in the short term. It is important to note that this study only evaluated the impact of this program after a 12-month period. It would be interesting to see if this CBO activity can be sustained over a longer time-period and, if so, then what impact this has on the sustainability of these infrastructure investments in these areas.

More broadly, the OQP impact evaluation provides lessons for other projects and initiatives aiming to engage communities in the upkeep or provision of public goods. First, the intervention relies almost wholly on local knowledge, and gives participating CBOs full autonomy in determining the types of activities to implement. Second, CBO rewards are not based exclusively on the activities they conduct (this is only one of six evaluation criteria), but primarily on an external assessment of the cleanliness of their neighborhood. This is therefore an example of a "results-based" intervention at the very local level. Third, such light-touch, relatively low-cost, non-interventionist approaches can be effective at shifting community-level behaviors, which is fundamental to achieving sustainability and returns on other types of community investments.

References

Ansink, E. and Bouma, J. (2013) "Effective support for community resource management." *Forest Policy and Economics*, 37:C, 94-103.

Ariely, D., Bracha, A., and Meier, S. (2013) "Doing good or doing well? Image motivation and monetary incentives in behaving prosocially." Working paper No. 07-09, Federal Reserve Bank of Boston.

Ashraf, N., Bandiera, O. and Jack, J. (2013) "No margin, no mission? A field experiment on incentives for public service delivery." *Journal of Public Economics*, 120, 1-17.

Baland, J.M. and Platteau, J.P. (1996) *Halting degradation of natural resources. Is there a role for rural communities?* Rome: Food and Agricultural Organization of the United Nations.

Banerjee, A., Banerji, R., Duflo, E., Glennerster, R. and Khemani, S. (2010) "Pitfalls of participatory programs: Evidence from a randomized evaluation in education in India." *American Economic Journal: Economic Policy*, 2:1, 1-30.

Beath, A., Christia, F. and Enikolopov, R. (2013) "The National Solidarity Programme: Assessing the effects of community-driven development in Afghanistan." UNU-WIDER Working Paper No. 2013/112.

Besley, T. and Ghatak, M. (2006) "Public goods and economic development." In: *Understanding Poverty* (ed.) Banerjee, A., Benabou, R. and Mookherjee, D. (eds.), New York: Oxford University Press.

Bosetti, V., Heugues, M. and Tavoni, A. (2017) "Luring others into climate action: coalition formation games with threshold spillover effects." *Oxford Economic Papers*, 69:2, 410-431.

Cadsby, C.B. and Maynes E. (1999) "Voluntary provision of threshold public goods with continuous contributions: experimental evidence," *Journal of Public Economics*, 71, 53–73.

Cappelen, A., Reme, B.A., Sorenson, E.O. and Tungodden, B. (2013) "Leadership and incentives." Working Paper Norwegian School of Economics, ISSN0804-6834.

Carlsson, F., Johansson-Stenman, O. and Nam, P. (2015) "Funding a new bridge in rural Vietnam: a field experiment on social influence and default contributions." *Oxford Economic Papers*, 67:4, 984-1014.

Croson, R. and Marks, M. (2000) "Step returns in threshold public good: A meta- and experimental analysis." *Experimental Economics*, 2:3, 239-259.

Jack, K. and Recalde, M.P. (2015) "Leadership and the voluntary provision of public goods: Field evidence from Bolivia." *Journal of Public Economics*, 122, 80-93.

Fafchamps, M. (2006) "Development and Social Capital." *The Journal of Development Studies*, 42:7, 1180-1198.

Feltovich, N. and Grossman, P. (2015) “How does the effect of pre-play suggestions vary with group size? Experimental evidence from a threshold public-good game.” *European Economic Review*, 79:C, 263-280.

Gächter, S. and Renner, E. (2018). “Leaders as role models and ‘belief managers’ in social dilemmas.” *Journal of Economic Behavior and Organization*, 154:C, 321-334.

Guiteras, R., Levinsohn, J. and Mobarak, A. (2015) “Encouraging sanitation investment in the developing world: a cluster-randomized trial.” *Science*, 348:6237, 903-906.

Heß, S. (2017) “Randomization inference with Stata: A guide and software.” *Stata Journal*, 17:3, 630-651.

Kremer, M., Leino, J., Miguel, E. and Peterson Zwane, A. (2011) “Spring cleaning: rural water impacts, valuation and property rights institutions.” *Quarterly Journal of Economics*, 126, pp. 145-205.

Kremer, M., Miguel, E., Mullainathan, S., Null,C. and Peterson Zwane, A. (2009) "Making Water Safe: Price, Persuasion, Peers, Promoters, or Product Design." Working Paper, Harvard University, August 2009.

Mansuri, G. and Rao, V. (2004) “Community-based and –driven development: a critical review.” *World Bank Research Obse*rver, 19:1, 1-39.

Mansuri, G., and Rao, V. (2013) “Can Participation Be Induced? Some Evidence from Developing Countries.” *Critical Review of International Social and Political Philosophy*, 16:2, 284-304.

Moxnes, E. and van der Heijden, E. (2003) “The effect of leadership in a public bad experiment.” *Journal of Conflict Resolution,* 47:6, 773-795.

Ostrom, E. (1990) *Governing the commons: the evolution of institutions for collective action*. New York: Cambridge University Press.

Ostrom, E. (1999) “Coping with tragedies of the commons.” *Annual Review of Political Science*, 2, 493-535.

Ostrom, E. (2008) “Tragedy of the Commons.” In: Durlauf, S. and Blume, L (eds.) *The New Palgrave Dictionary of Economics*, Second Edition.

Ostrom, E. (2009) “A polycentric approach for coping with climate change.” World Bank Policy Research Working Paper No. 5095.

Ostrom, E., Gardner, R. and Walker, J. (1994) *Rules, games and common-pool resources.* Ann Arbor: University of Michigan Press.

Ruaniyar, G., Orbeta, A. and Sugiyarto, G. (2011) “Impact of water supply and sanitation assistance on human welfare in rural Pakistan.” *Journal of Development Effectiveness*, 3:1, 62-102.

Sheely, R. (2013) “Maintaining local public goods: Evidence from rural Kenya.” Harvard Kennedy School, Faculty Research Working Paper Series, RWP13-051.

UNICEF (2012) UNICEF *Water, Sanitation and Hygiene Annual Report 2011*. UNICEF: New York.

World Bank (2004) *World Development Report: Making services work for poor people*. World Bank and Oxford University Press: Washington.

World Bank. (2010). *Cites and Climate Change: An Urgent Agenda*. Washington, DC: World Bank.

Young, A., 2017. “Channeling Fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results.” Mimeo.

Tables

Table 1: Initial endowment package

	Materials	Amount per *neighborhood*
1	Boots	15
2	Shovel	15
3	Wheel Barrow	05
4	Pitchfork	10
5	Broom	25
6	Professional Gloves	15
7	Sifter	05
8	Rake	10
9	Machete	05
10	Shears	05

Table 2: Assessment Criteria

Criteria	Rationale	Weight
1. Non-obstruction of natural drainage structures and waterways	Neighborhoods are often flooded due to waterway obstruction from solid waste.	30%
2. Lack of waste dumping in the streets	Household often dump their waste in public spaces.	20%
3. Cleanliness of abandoned houses and lands	Many houses are abandoned in the flood zones and they are often transformed into dumping grounds for waste.	20%
4. Cleanliness of public spaces	The cleanliness of social and collective spaces is an indicator of the level of community awareness of sanitation.	10%
5. Lack of water on streets and house fronts	Households often discharge domestic wastewater into the streets.	10%
6. Innovative initiatives taken by the CBO	OQP aimed to encourage CBOs to be creative in addressing community cleanliness.	10%

Table 3: Communes, neighborhoods, and groups in the study area

	Communes	Neighborhoods	Groups
Pikine	6	349	42
Guediawaye	1	47	5
Dakar	1	2	1
Total	8	398	48

Table 4: Baseline household characteristics

	n	Control	Treatment	p-value
Characteristics of household head				
Male	2,285	0.70	0.68	0.29
Muslim	2,285	0.96	0.95	0.23
Married	2,285	0.80	0.78	0.10
Age	2,285	55.45	55.36	0.86
Salary (FCFA per day)	1,062	12,656.05	9,113.88	0.12
Household size	2,285	10.14	10.19	0.82
Cleaning and flooding related characteristics				
Cleanliness of neighborhood	2,285	0.29	0.30	0.51
Flood victim in last year	2,285	0.21	0.23	0.36
Training about flood prevention	2,285	0.16	0.13	0.04**

Phase 1	n	Control	Treatment	p-value
Characteristics of household head				
Male	922	0.69	0.64	0.13
Muslim	922	0.97	0.95	0.22
Married	922	0.77	0.73	0.07*
Age	922	57.14	56.14	0.26
Salary (FCFA per day)	405	9,535.17	8,222.07	0.61
Household size	922	11.10	10.88	0.59
Cleaning and flooding related characteristics				
Cleanliness of neighborhood	922	0.26	0.25	0.68
Flood victim in last year	922	0.28	0.25	0.28
Training about flood prevention	922	0.15	0.13	0.47

Phase 2	n	Control	Treatment	p-value
Characteristics of household head				
Male	1,363	0.71	0.71	0.92
Muslim	1,363	0.95	0.94	0.55
Married	1,363	0.82	0.81	0.54
Age	1,363	54.31	54.83	0.44
Salary (FCFA per day)	657	14,588	9,661	0.13
Household size	1,363	9.48	9.73	0.37
Cleaning and flooding related characteristics				
Cleanliness of neighborhood	1,363	0.32	0.34	0.26
Flood victim in last year	1,363	0.16	0.21	0.02**
Training about flood prevention	1,363	0.17	0.13	0.04**

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

Table 5: Baseline CBO characteristics

	Control	Treatment	p-value
CBO Flood prevention and cleaning activities			
CBO has taken action in relation to flooding in last year	0.72	0.66	0.42
Flood activities: raising awareness	0.30	0.30	1.00
Flood activities: financial assistance for victims	0.28	0.20	0.27
Flood activities: cleaning of canals and lakes	0.11	0.11	1.00
Flood activities: surveillance of canals and lakes	0.04	0.05	0.70
Flood activities: cooperation with other stakeholders	0.25	0.28	0.72
Flood activities: small works to avoid flooding	0.15	0.14	0.82
Objective of CBO: raise awareness on flood-related issues/organize cleaning activities	0.19	0.19	1.00
CBO engaged in flood communication campaign	0.24	0.20	0.57
CBO engaged in cleaning activities	0.64	0.50	0.08*
Number of days spent on flood communication or cleaning activities	10.32	14.30	0.38
Number of members engaged in flood communication or cleaning activities	59.38	38.76	0.17
CBO Characteristics			
Number of members	200.01	155.22	0.54
Number of members present at the last meeting	38.48	39.67	0.83
Number of members engaged in activities of CBO	117.55	71.36	0.07*
Average hours per week members engage in CBO activities	11.93	12.56	0.83
Proportion of female members	0.59	0.62	0.63
Proportion of youth members	0.06	0.04	0.46
CBO represents a head office	0.85	0.95	0.04**
CBO votes in leaders by election	0.69	0.53	0.04**
Most important benefit of membership of CBO are benefits for the community	0.14	0.07	0.20
CBO type: Sports and culture association	0.16	0.09	0.15
CBO type: Economic interest group	0.10	0.16	0.24
CBO type:	0.20	0.15	0.41
CBO type: Development association	0.49	0.44	0.53
CBO type: Flood prevention committee	0.00	0.06	0.02**
CBO type: Other association	0.05	0.10	0.23
Collaboration between different CBOs			
CBO only intervenes in this neighborhood	0.65	0.65	1.00
CBO collaborates with other CBOs in the neighborhood	0.65	0.71	0.40
CBO collaborates with other CBOs in other neighborhoods	0.69	0.63	0.41

N=160

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

Table 6: CBO listed cleaning events among top 3 main activities

	(1)	(2)	(3)
Treatment	0.188**	0.232***	0.210*
	(0.078)	(0.082)	(0.108)
Phase 1		0.037	0.011
		(0.083)	(0.108)
Treatment x Phase 1			0.053
			(0.164)
Baseline outcome	No	Yes	Yes
Baseline controls	No	Yes	Yes
Constant	0.362***	0.429**	0.445**
	(0.054)	(0.178)	(0.184)
Observations	160	160	160
R-squared	0.035	0.079	0.080

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 7: Household Awareness of OQP and impact of OQP on flood risk training

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent variable:	Awareness of OQP				Received training on flood prevention			
Treatment	0.168***	0.196***	0.199***	0.252***	0.041**	0.053***	0.053***	0.042**
	(0.040)	(0.029)	(0.028)	(0.036)	(0.019)	(0.014)	(0.014)	(0.017)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.055]	[0.009]	[0.014]	[0.023]
Phase 1			0.465***	0.466***			0.079	0.079
			(0.084)	(0.075)			(0.050)	(0.050)
			[0.028]	[0.075]			[0.111]	[0.211]
Phase 1*Treatment				-0.132***				0.027
				(0.042)				(0.032)
				[0.091]				[0.562]
Constant	0.354***	0.648***	0.206	0.250*	0.100***	0.227	0.152	0.143
	(0.030)	(0.125)	(0.139)	(0.136)	(0.011)	(0.147)	(0.153)	(0.148)
Baseline outcome	No	No	No	No	No	Yes	Yes	Yes
Baseline controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Group fixed effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Enumerator fixed effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Observations	2,400	2,271	2,271	2,271	2,285	2,271	2,271	2,271
R-squared	0.028	0.334	0.338	0.341	0.004	0.182	0.182	0.182

Robust standard errors clustered at the neighborhood level in parentheses. Randomization inference p-values (generated using the STATA command developed by Heß (2017)) are presented in brackets (Young, 2017). Note that there is no baseline value of the outcome variable 'awareness of OQP' in columns (1) to (4). *** p<0.01, ** p<0.05, * p<0.1 for the t-test based on the clustered standard errors.

Table 8: Impact of OQP on perceptions of cleanliness and flood related outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Cleanliness of neighborhood				Flood victim this year				Flood victim last year			
Treatment	0.075*	0.123***	0.123***	0.120***	-0.014	-0.022**	-0.021**	-0.039***	-0.015	-0.019	-0.016	-0.021
	(0.038)	(0.024)	(0.024)	(0.033)	(0.019)	(0.010)	(0.010)	(0.013)	(0.025)	(0.015)	(0.015)	(0.018)
	[0.028]	[0.000]	[0.000]	[0.005]	[0.337]	[0.150]	[0.166]	[0.018]	[0.415]	[0.403]	[0.439]	[0.397]
Phase 1			-0.101	-0.102			0.175***	0.174***			0.368***	0.368***
			(0.063)	(0.063)			(0.047)	(0.047)			(0.066)	(0.065)
			[0.965]	[0.873]			[0.827]	[0.756]			[0.715]	[0.809]
Phase 1*Treatment				0.008				0.045**				0.012
				(0.049)				(0.018)				(0.031)
				[0.908]				[0.170]				[0.813]
Constant	0.358***	0.978***	1.075***	1.072***	0.095***	0.167***	0.001	-0.014	0.194***	0.328***	-0.022	-0.026
	(0.024)	(0.074)	(0.094)	(0.099)	(0.013)	(0.052)	(0.067)	(0.068)	(0.018)	(0.073)	(0.085)	(0.085)
Baseline outcome	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Baseline controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Group fixed effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Enumerator fixed effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Observations	2,285	2,271	2,271	2,271	2,285	2,271	2,271	2,271	2,285	2,271	2,271	2,271
R-squared	0.006	0.226	0.226	0.226	0.001	0.122	0.123	0.124	0.000	0.160	0.165	0.165

Robust standard errors clustered at the neighborhood level in parentheses. Randomization inference p-values (generated using the STATA command developed by Heß (2017)) are presented in brackets (Young, 2017). *** p<0.01, ** p<0.05, * p<0.1 for the t-test based on the clustered standard errors.

Table 9: Heterogeneity by size of CBO

	(1) Awareness of OQP	(2) Cleanliness	(3) Flood victim this year
Treatment	0.299***	0.092**	-0.017
	(0.051)	(0.039)	(0.020)
	[0.000]	[0.228]	[0.532]
Treatment x CBO Size	-0.107**	0.030	-0.004
	(0.051)	(0.038)	(0.021)
	[0.146]	[0.705]	[0.905]
CBO Size	0.071**	-0.048*	-0.021
	(0.033)	(0.028)	(0.015)
	[0.093]	[0.773]	[0.595]
Constant	0.192	1.150***	-0.006
	(0.142)	(0.083)	(0.067)
Baseline outcome	No	Yes	Yes
Baseline controls	Yes	Yes	Yes
Group fixed effects	Yes	Yes	Yes
Enumerator fixed effects	Yes	Yes	Yes
Observations	2,126	2,126	2,126
R-squared	0.349	0.230	0.130

Robust standard errors clustered at the neighborhood level in parentheses. Randomization inference p-values (generated using the STATA command developed by Heß (2017)) are presented in brackets (Young, 2017). *** p<0.01, ** p<0.05, * p<0.1 for the t-test based on the clustered standard errors.

Table 10: Impact of OQP on individual level health outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Ill in the last 30 days				Ill in the most recent rainy season				Ill due to malaria in the most recent rainy season			
Treatment	-0.010	-0.022***	-0.020***	-0.027***	-0.013	-0.027***	-0.025***	-0.026**	-0.008	-0.013***	-0.012**	-0.011*
	(0.019)	(0.007)	(0.007)	(0.010)	(0.023)	(0.009)	(0.008)	(0.011)	(0.010)	(0.005)	(0.005)	(0.006)
	[0.332]	[0.135	[0.161]	[0.025]	[0.303]	[0.064]	[0.083]	[0.051]	[0.228]	[0.108]	[0.091]	[0.105]
Phase 1			0.158***	0.157***			0.230***	0.230***			0.149***	0.149***
			(0.028)	(0.028)			(0.029)	(0.029)			(0.019)	(0.019)
			[0.804]	[0.734]			[0.818]	[0.737]			[0.825]	[0.741]
Phase 1*Treatment				0.016				0.003				-0.002
				(0.014)				(0.014)				(0.007)
				[0.559]				[0.935]				[0.912]
Constant	0.240***	0.126*	-0.023	-0.028	0.280***	0.152***	-0.066	-0.067	0.077***	0.047	-0.094**	-0.094**
	(0.012)	(0.068)	(0.072)	(0.069)	(0.016)	(0.046)	(0.047)	(0.048)	(0.008)	(0.033)	(0.037)	(0.037)
Baseline outcome	No	Yes	Yes	Yes	No	No	No	No	No	No	No	No
Baseline controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Group fixed effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Enumerator fixed effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Observations	20,970	20,389	20,389	20,389	20,981	20,399	20,399	20,399	23,227	22,606	22,606	22,606
R-squared	0.000	0.118	0.119	0.119	0.000	0.120	0.121	0.121	0.000	0.036	0.038	0.038

Note: In columns (5) to (12) baseline values of the outcome variable are not available. The baseline level of reported illness in the previous 30 days is included as an additional control in each specification. Robust standard errors clustered at the neighborhood level in parentheses. Randomization inference p-values (generated using the STATA command developed by Heß (2017)) are presented in brackets (Young, 2017). *** p<0.01, ** p<0.05, * p<0.1 for the t-test based on the clustered standard errors.

Table 11: Impact of OQP on individual work and education outcomes

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Number of work days missed				Number of school days missed				School year began on time			
Treatment	-0.049	-0.221	-0.212	-0.842**	-0.032	-0.049	-0.053	-0.167	-0.003	-0.003	-0.003	0.002
	(0.342)	(0.226)	(0.225)	(0.335)	(0.093)	(0.074)	(0.074)	(0.120)	(0.005)	(0.005)	(0.005)	(0.006)
	[0.840]	[0.510]	[0.504]	[0.044]	[0.531]	[0.619]	[0.599]	[0.183]	[0.772]	[0.710]	[0.721]	[0.765]
Phase 1			1.433**	1.402**			-0.467***	-0.466***			-0.000	-0.001
			(0.560)	(0.583)			(0.149)	(0.162)			(0.009)	(0.010)
			[0.652]	[0.627]			[0.335]	[0.451]			[0.863]	[0.915]
Phase 1*Treatment				1.397***				0.277**				-0.013
				(0.412)				(0.124)				(0.012)
				[0.020]				[0.103]				[0.436]
Constant	2.561***	0.657	-0.683	-1.220	0.375***	0.612	1.056*	0.960*	0.982***	0.990***	0.991***	0.994***
	(0.243)	(0.903)	(0.827)	(0.786)	(0.072)	(0.562)	(0.530)	(0.479)	(0.004)	(0.018)	(0.019)	(0.020)
Baseline outcome	No	Yes	Yes	Yes	No	No	No	No	No	No	No	No
Baseline controls	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Group fixed effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Enumerator fixed effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Observations	5,296	5,253	5,253	5,253	5,513	5,476	5,476	5,476	6,027	5,990	5,990	5,990
R-squared	0.000	0.134	0.134	0.136	0.000	0.089	0.089	0.090	0.000	0.054	0.054	0.055

Note: Sample sizes refer to the number of working adults in columns 1 to 4 and the number of children attending school in columns 5 to 12. Robust standard errors clustered at the neighborhood level in parentheses. Randomization inference p-values (generated using the STATA command developed by Heß (2017)) are presented in brackets (Young, 2017). *** p<0.01, ** p<0.05, * p<0.1 for the t-test based on the clustered standard errors.

Table 12: Outcomes in control group and spillovers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Awareness of OQP		Received training in flood prevention		Cleanliness of neighborhood		Flood victim this year	
Treat Border	0.067	0.067	0.070**	0.070**	0.053	0.053	0.038	0.038
	(0.048)	(0.048)	(0.032)	(0.032)	(0.075)	(0.075)	(0.026)	(0.026)
Phase 1	0.515***	0.515***	-0.143*	-0.143*	-0.041	-0.041	0.023	0.023
	(0.127)	(0.127)	(0.081)	(0.081)	(0.178)	(0.178)	(0.067)	(0.067)
Phase 1*Treat Border		0.247**		0.017		-0.212		0.004
		(0.102)		(0.081)		(0.230)		(0.045)
Nr Border	-0.003	-0.003	0.000	0.000	0.009	0.009	-0.006	-0.006
	(0.007)	(0.007)	(0.004)	(0.004)	(0.011)	(0.011)	(0.004)	(0.004)
Constant	-0.733***	-0.733***	-0.045	-0.045	1.164***	1.164***	-0.114	-0.114
	(0.172)	(0.172)	(0.127)	(0.127)	(0.407)	(0.407)	(0.081)	(0.081)
Baseline outcome	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Baseline controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Group fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Enumerator fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,060	1,060	1,060	1,060	1,060	1,060	1,060	1,060
R-squared	0.393	0.393	0.242	0.242	0.278	0.278	0.176	0.176

Robust standard errors clustered at the neighborhood level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 13: Contamination from control group

	(1)	(2)	(3)	(4)
	Cleanliness of neighborhood		Flood victim this year	
Control Border	0.306***	0.244***	0.032	0.028
	(0.058)	(0.060)	(0.026)	(0.031)
Nr Border	-0.010	-0.009	-0.004	-0.004
	(0.007)	(0.008)	(0.003)	(0.003)
Phase 1	0.287***	-0.082	0.018	-0.007
	(0.079)	(0.110)	(0.034)	(0.040)
Phase 1*Treat Border		0.340***		0.024
		(0.082)		(0.036)
Constant	0.492***	0.542***	0.242***	0.245***
	(0.142)	(0.139)	(0.070)	(0.072)
Baseline outcome	Yes	Yes	Yes	Yes
Baseline controls	Yes	Yes	Yes	Yes
Group fixed effects	Yes	Yes	Yes	Yes
Enumerator fixed effects	Yes	Yes	Yes	Yes
Observations	1,098	1,098	1,098	1,098
R-squared	0.295	0.297	0.155	0.155

Robust standard errors clustered at the neighborhood level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Figures

Figure 1: Map of intervention area

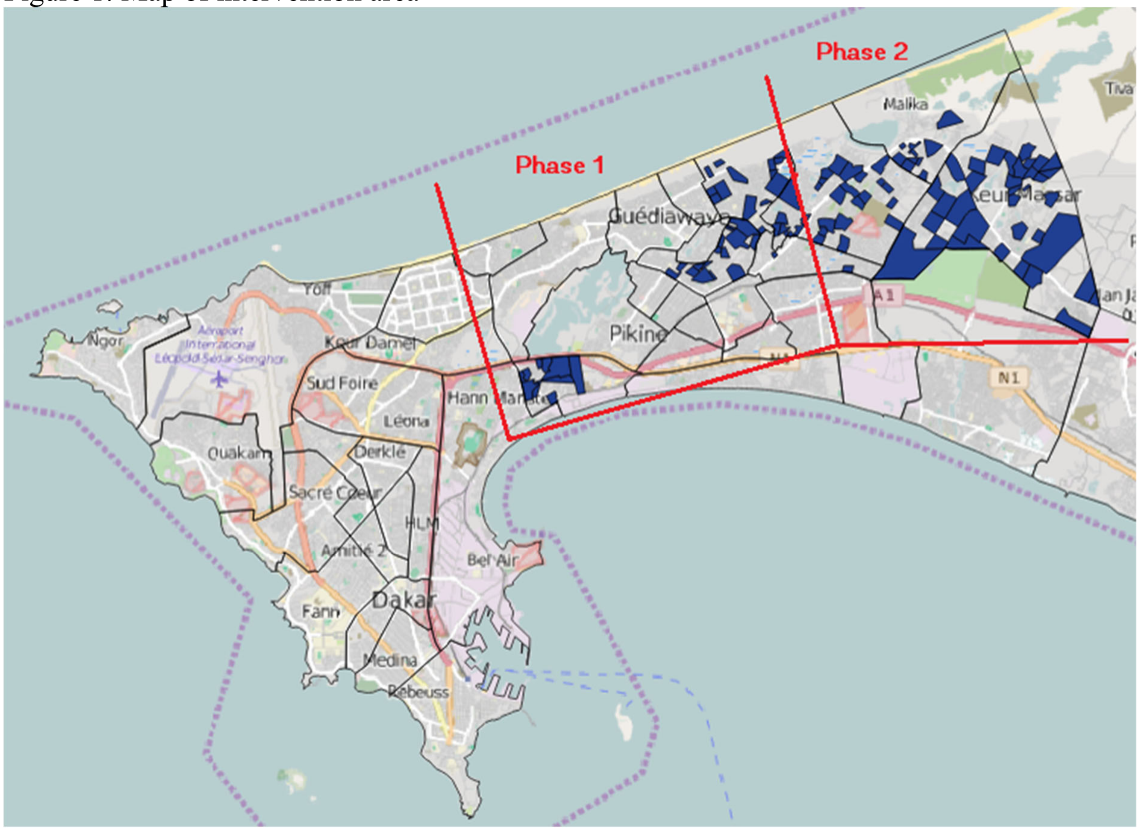
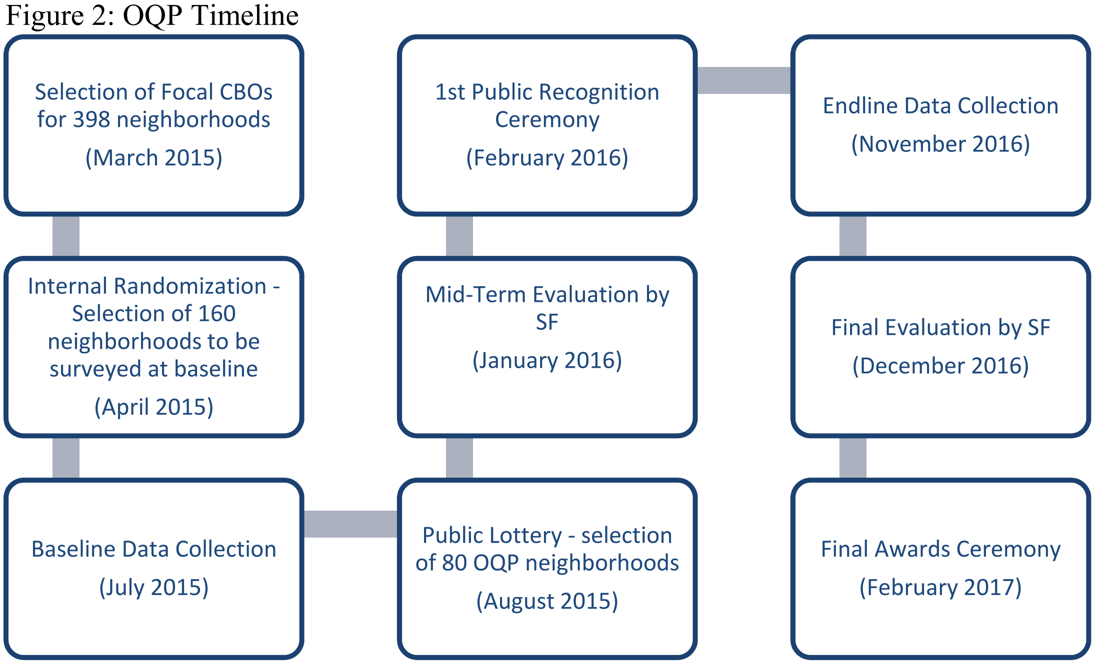


Figure 2: OQP Timeline



Appendix

Baseline CBO characteristics within Phases

Table A1: Baseline CBO characteristics (Phase 1)

	Control	Treatment	p-value
CBO Flood prevention and cleaning activities			
CBO has taken action in relation to flooding in last year	0.75	0.78	0.77
Flood activities: raising awareness	0.34	0.38	0.80
Flood activities: financial assistance for victims	0.28	0.13	0.12
Flood activities: cleaning of canals and lakes	0.13	0.16	0.72
Flood activities: surveillance of canals and lakes	0.06	0.13	0.40
Flood activities: cooperation with other stakeholders	0.28	0.38	0.43
Flood activities: small works to avoid flood	0.09	0.19	0.29
Objective of CBO: raise awareness on flood-related issues/organize cleaning activities	0.22	0.25	0.77
CBO engaged in flood communication campaign	0.38	0.22	0.18
CBO engaged in cleaning activities	0.66	0.56	0.45
Number of days spent on flood communication or cleaning activities	13.44	13.16	0.97
Number of members engaged in flood communication or cleaning activities	64.78	40.19	0.09*
CBO Characteristics			
Number of members	177.53	192.91	0.90
Number of members present at the last meeting	35.72	30.41	0.46
Number of members engaged in activities of CBO	143.66	77.03	0.08*
Average hours per week members engage in CBO activities	11.50	14.53	0.53
Proportion of members that are women	0.55	0.53	0.81
Proportion of members that are young	0.05	0.08	0.30
CBO represents a head office	0.84	0.97	0.09*
CBO votes in leaders by election	0.69	0.63	0.61
Most important benefit of membership of CBO are benefits for the community	0.13	0.09	0.69
CBO type: Sports and culture association	0.13	0.16	0.72
CBO type: Economic interest group	0.09	0.13	0.69
CBO type: Women's involvement group	0.19	0.03	0.05**
CBO type: Development association	0.53	0.41	0.32
CBO type: Flood prevention committee	0.00	0.03	0.32
CBO type: Other association	0.06	0.25	0.04*
Collaboration between different CBOs			
CBO only intervenes in this neighborhood	0.59	0.53	0.62
CBO collaborates with other CBOs in the neighborhood	0.56	0.75	0.12
CBO collaborates with other CBOs in other neighborhoods	0.69	0.78	0.40

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

Table A2: Baseline CBO characteristics (Phase 2)

	Control	Treatment	p-value
<u>CBO Flood prevention and cleaning activities</u>			
CBO has taken action in relation to flooding in last year	0.70	0.58	0.23
Flood activities: raising awareness	0.27	0.25	0.82
Flood activities: financial assistance for victims	0.27	0.25	0.82
Flood activities: cleaning of canals and lakes	0.10	0.08	0.73
Flood activities: surveillance of canals and lakes	0.02	0.00	0.32
Flood activities: cooperation with other stakeholders	0.23	0.21	0.81
Flood activities: small works to avoid flood	0.19	0.10	0.25
Objective of CBO: raise awareness on flood-related issues/organize cleaning activities	0.17	0.15	0.78
CBO engaged in flood communication campaign	0.15	0.19	0.59
CBO engaged in cleaning activities	0.63	0.46	0.10
Number of days spent on flood communication or cleaning activities	8.25	15.06	0.26
Number of members engaged in flood communication or cleaning activities	55.77	37.81	0.45
<u>CBO Characteristics</u>			
Number of members	215.00	130.10	0.34
Number of members present at the last meeting	40.31	45.85	0.49
Number of members engaged in activities of CBO	100.15	67.58	0.33
Average hours per week members engage in CBO activities	12.21	11.25	0.80
Proportion of members that are women	0.62	0.67	0.45
Proportion of members that are young	0.06	0.02	0.02**
CBO represents a head office	0.85	0.94	0.19
CBO votes in leaders by election	0.69	0.46	0.02**
Most important benefit of membership of CBO are benefits for the community	0.15	0.06	0.19
CBO type: Sports and culture association	0.19	0.04	0.02**
CBO type: Economic interest group	0.10	0.19	0.25
CBO type: Women's involvement group	0.21	0.23	0.81
CBO type: Development association	0.46	0.46	1.00
CBO type: Flood prevention committee	0.00	0.08	0.04**
CBO type: Other association	0.04	0.00	0.16
<u>Collaboration between different CBOs</u>			
CBO only intervenes in this neighborhood	0.69	0.73	0.66
CBO collaborates with other CBOs in the neighborhood	0.71	0.69	0.83
CBO collaborates with other CBOs in other neighborhoods	0.69	0.52	0.10*

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

Table A3: Baseline correlation between flooding and cleanliness

Outcome: Experienced flooding in the previous rainy season	(1)	(2)
Cleanliness	-0.054**	-0.053***
	(0.020)	(0.020)
Constant	0.234***	-0.237
	(0.017)	(0.165)
Baseline controls	No	Yes
Group fixed effects	No	Yes
Enumerator fixed effects	No	Yes
Observations	2,285	2,271
R-squared	0.004	0.129

Robust standard errors clustered at the neighborhood level in parentheses. **** p<0.01, ** p<0.05, * p<0.1.

Table A4: Heterogeneity by size of CBO and Phase

	(1)	(2)	(3)	(4)	(5)	(6)
	Awareness of OQP		Cleanliness		Flood victim this year	
Treatment	0.299***	0.314***	0.092**	0.088	-0.017	-0.024
	(0.051)	(0.063)	(0.039)	(0.053)	(0.020)	(0.026)
	[0.000]	[0.000]	[0.228]	[0.296]	[0.532]	[0.495]
Treatment x CBO Size	-0.107**	-0.064	0.030	0.036	-0.004	-0.010
	(0.051)	(0.061)	(0.038)	(0.047)	(0.021)	(0.025)
	[0.146]	[0.552]	[0.705]	[0.639]	[0.905]	[0.753]
Treatment x CBO Size x Phase 1		-0.170		0.005		0.017
		(0.127)		(0.058)		(0.052)
		[0.342]		[0.978]		[0.777]
Treatment x Phase 1		-0.049		-0.071		-0.013
		(0.091)		(0.084)		(0.048)
		[0.937]		[0.915]		[0.875]
CBO Size	0.071**	0.031	-0.048*	-0.081***	-0.021	-0.010
	(0.033)	(0.037)	(0.028)	(0.026)	(0.015)	(0.019)
	[0.093]	[0.559]	[0.773]	[0.400]	[0.595]	[0.691]
Phase 1	0.452***	0.337**	-0.115	-0.219***	0.154***	0.186***
	(0.094)	(0.126)	(0.070)	(0.072)	(0.045)	(0.054)
	[0.197]	[0.909]	[0.915]	[0.193]	[0.800]	[0.585
CBO Size x Phase 1		0.100		0.094**		-0.028
		(0.080)		(0.044)		(0.033)
		[0.359]		[0.148]		[0.570]
Constant	0.192	0.200	1.150***	1.136***	-0.006	-0.010
	(0.142)	(0.152)	(0.083)	(0.078)	(0.067)	(0.075)
Baseline outcome	No	No	No	Yes	No	Yes
Baseline controls	No	Yes	No	Yes	No	Yes
Group fixed effects	No	Yes	No	Yes	No	Yes
Enumerator fixed effects	No	Yes	No	Yes	No	Yes
Observations	2,126	2,126	2,126	2,126	2,126	2,126
R-squared	0.349	0.353	0.230	0.232	0.130	0.131

Robust standard errors clustered at the neighborhood level in parentheses. Randomization inference p-values (generated using the STATA command developed by Heß (2017)) are presented in brackets (Young, 2017). **** p<0.01, ** p<0.05, * p<0.1 for the t-test based on the clustered standard errors.

Table A5: Baseline Individual Characteristics

	n	Control	Treatment	p-value
Ill in the last 30 days	22,749	0.224	0.214	0.069*
Number of work days missed	23,227	0.260	0.318	0.057*
Phase 1				
Ill in the last 30 days	9,946	0.236	0.248	0.166
Number of work days missed	10,135	0.348	0.409	0.253
Phase 2				
Ill in the last 30 days	12,803	0.213	0.187	0.000***
Number of work days missed	13,092	0.190	0.250	0.089*

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