

West Bank and Gaza Local Governance and Infrastructure (LGI) Project Evaluation

Pablo Beramendi, Duke University
Soomin Oh, Duke University
Erik Wibbels, Duke University

September 4, 2017

DISCLAIMER

The authors' views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

Introduction

USAID/West Bank and Gaza (WBG) Local Governance and Infrastructure (LGI) program is a four-year, 104 million dollar program that was in effect between 2012 and 2016. The program worked in a total of 28 municipalities across 11 governorates of WBG for both its governance and infrastructure components of the program.

Amongst a number of goals of the program, this evaluation focuses on the *governance* component of the program. The programming aimed at strengthening local government capacity in Palestinian municipalities by implementing three categories of activities: providing support for institutional development, community development, and participatory governance. This evaluation is designed to assess the impact of these efforts on improving Palestine's local government capacity and services.

The program was planned and executed with the general goal of providing “an enabling environment for good local governance” and to “provide the basic infra necessary for sustainable improvements in the quality of life” for the Palestinians in WBG (LGI Program Description).

The project brought substantial development benefits to the citizens of WBG. There was concern that the local government units (LGUs) lacked the institutional capacity and skills to provide quality services and to attract skilled staff, and lacked sufficient guidance and monitoring from the Ministry of Local Government (MoLG). In light of these concerns, the program was designed to increase institutional capacity and skills of the LGUs.

This evaluation assesses the impact of the three categories of governance activities that were implemented, namely institutional development, community planning, and participatory governance. This document describes the endline analysis plan for LGI evaluation and serves as a guide to conducting the final analysis. The authors will complete and register the plan prior to the analysis of the endline data.

We examine the impact of the activities on both the demand side (citizens) and the supply side (administrators) of service provision. Throughout these efforts, we use two methods. First, given that the 28 targeted localities for the governance component of the program were not chosen at random (i.e. it was not a randomized control trial), we employ a statistical method of matching to find suitable “control” units with which we can compare the programmed LGUs. We measure impact via endline surveys of service users and municipal administrators that we conducted in the West Bank. Second, we draw on household and administrator surveys collected from the universe of municipalities for the forthcoming Local Government Performance Assessment (LGPA) program. In this document we refer to the municipalities that received programming as “programmed municipalities” and refer to the controls as “comparison municipalities.”

Specifically, we seek to answer three questions:

1. Whether there have been improvements in the municipalities' governance practices and citizens' satisfaction levels in programmed municipalities and comparison municipalities ;
2. Whether programming in LGI municipalities has improved strategic and physical planning; and
3. Whether civic engagement mechanisms in LGI municipalities have increased citizens' involvement in local governance.

This document describes the endline analysis plan for our evaluation of LGI. The authors will complete and register the plan prior to the analysis of the endline data.

1. Program Description

Development of good democratic local governance institutions and basic infrastructure play a central role in the social and economic requirements for the Palestinian state. The donor community, including USAID, has recognized that there is a need to improve local government capacity and basic services for citizens. The donor community has run projects to provide basic infrastructure and develop service provision institutions. In spite of those efforts, however, there remains a gap in the *institutional capacity* of the Palestinian state to provide public services and infrastructure.

In order to fill this gap in institutional capacity, USAID awarded Global Communities a six-year Cooperative Agreement to carry out the Local Government and Infrastructure Program (LGI), with the goal of “promot[ing] an enabling environment for good local governance and provide the basic infrastructure necessary for sustainable improvements in the quality of life for Palestinians in the West Bank and Gaza” (LGI Program Description).

The project aimed to achieve three objectives:

Objective 1: Improve living conditions for Palestinians through the provision of sustainable, multi-sector community infrastructure packages.

Objective 2: Enhance the impact of USAID sector-specific programs and priorities through the provision of high priority infrastructure

Objective 3: Strengthen local government capacity to respond effectively and efficiently to community needs by promoting and institutionalizing good democratic governance practices

Objective 3 is the main focus of this impact evaluation. There are specific goals that objective 3 aimed to achieve:

(1) Institutional development packages

The program attempted to strengthen local government capacity by institutionalizing good governance practices, examples of which include sound financial practices, transparent procurement and tendering system, performance based appraisal human resources

management, and responsive service delivery mechanisms. The program also developed an e-municipalities concept and internal audit functions within targeted LGUs.

(2) Skills enhancement/community planning support packages

The program provided support to increase skills in strategic and physical planning. Strategic planning included updating Strategic Development Framework of the municipality and implementing strategic planning processes. Physical planning involved naming of streets and buildings, developing the local registrar, and training the employees in physical planning (e.g. Geographical Information Systems).

(3) Citizen engagement/participatory governance tools and mechanisms

Achieving this goal involved introducing and institutionalizing mechanisms of civic engagement in governance, such as civic engagement committees, enhancing dialogue and communication between LGUs and citizens through e.g. town hall meetings and public hearings, raising public awareness on issues of good governance and community development, and promoting youth inclusion in local governance through Youth Local Councils.

1.4. Hypotheses

The evaluation of the interventions is organized by the outcomes of interest and the target population, which we discuss below. The overarching hypotheses that we test reflect the main goals of LGI.

Outcome 1: Institutional development

As a result of the intervention:

H1: We expect that citizen perception of the quality of governance will be higher in programmed than in comparison municipalities.

H2: We expect citizens in programmed municipalities to be more satisfied with the provision of services than their counterparts in comparison municipalities.

H3: We expect that good governance practices among officials and administration will be more pervasive in programmed than in comparison municipalities.

Outcome 2: Community planning

H4: We expect that municipal government's community planning will be more pervasive in programmed than in comparison municipalities.

Outcome 3: Participatory Governance

H5: We expect that citizen participation in government will be more pervasive in programmed than in comparison municipalities.

In section 3 below, we talk about how we operationalize these hypothesis.

2. Design: Overcoming non-randomized selection of targeted municipalities

2.1. Site selection

28 municipalities (list in Appendix 1) in the West Bank received programming between July 2013 and June 2016. Note that not all parts of the programming were applied equally to all the municipalities. For instance, while all municipalities received training in restructuring and process reengineering under the first objective of institutional development, only 4 out of 28 municipalities received training for e-municipality.

The selection of the targeted municipalities was non-random--that is, municipalities that were chosen for programming were not chosen randomly but were instead based on the following criteria¹:

- Regional representation
- Size of the population
- Level of development and current operational capacity

These criteria resulted in a set of municipalities that is considerably more urban than average. In addition to the 28 municipalities, 6 “Vulnerable Communities” were chosen based on Vulnerability Assessment. Vulnerable Communities, however, are not included in the evaluation for two reasons: first, hardly any programming was done in these communities. Of the small list of programming that was done, half or less from the list of vulnerable communities received programming. For instance, “strategic planning process” was completed in 2/6 communities and “development of land use plans and aerial photo” was done in 1/6 communities. Second, the communities chosen as “vulnerable communities” are the six most vulnerable communities in Palestine, according to the Vulnerability Index--they are the most resource constrained and some are targeted for demolition. This is to say, these communities have very different characteristics from the rest of the municipalities in Palestine.

The non-random selection of targeted units poses a threat to accurate evaluation of the program effects since the municipalities chosen to be programmed are inherently different from those that were not and especially since the targeted municipalities already have higher levels of development and operational capacity (Dunning, 2012; Green, 2012). Likewise, no municipalities were explicitly assigned as control municipalities for evaluation purposes. On top of the non-random selection of targeted units, the non-uniform assignment of programs across

¹ Comes from “Governance Baseline Report - Targeted Location”, p. 6.

the municipalities poses a threat to evaluation since the assignment of the particular programs is not likely to have been random either.

In the absence of an experimental design and in light of the evaluation budget, the evaluation team chose to examine the effects of the interventions using two approaches: first, matching the programmed municipalities to comparison municipalities and conducting surveys (municipal office users and employees) in a total of 56 municipalities and second, using the World Bank/USAID survey of citizens and local government employees of the universe of municipalities.

2.2. Matching

The goal of matching techniques is to reduce imbalance between the programmed municipalities and comparison municipalities on observed pre-treatment confounders (Stuart, 2010). Reducing imbalance between programmed and comparison municipalities and thus increasing their similarity offers several benefits. First, it decreases the risk that municipality level characteristics other than the treatment are responsible for the differences in outcomes between the programmed and comparison municipalities. Differences between the programmed and comparison municipalities that are correlated with the program outcomes can lead to biased, incorrect inferences about program effects. Second, it increases the robustness of subsequent analyses by decreasing the prevalence of extrapolation across observed variables. Third, directly comparing districts that are most similar on pre-treatment municipal-level characteristics that predict the outcomes of interest can decrease noise in the data and thereby lead to improvements in statistical power and our ability to detect effects of the programming.

In order to ensure that the municipalities were as similar as possible, we conducted matching on a number of pre-treatment characteristics at the municipal level that are expected to be highly correlated with the outcomes of interest (Brookhart et al., 2006), namely service provision and local governance capacity. The full list of covariates included in the matching algorithm can be found in Appendix 2.

There are several approaches to statistical matching, ranging from exact matching, nearest neighbor, to genetic matching. Given the non-random selection of programmed municipalities and the multi-dimensionality of the covariates we are matching on, most matching algorithms produce imbalanced pairs of matches. In light of this consideration, we used genetic matching technique that tries to maximize the balance of observed covariates across programmed and comparison units. It searches a range of distance metrics to find the particular measure that optimizes post-matching covariate balance. Each potential distance metric considered corresponds to a particular assignment of weights for all matching variables, and the algorithm weights each variable according to its relative importance for achieving the best overall balance.

The matching algorithm resulted in pairs of programmed and comparison municipalities that are as similar as possible on the observed pre-treatment covariates included. Figure 1 shows the standardized mean differences of the covariates between the programmed and comparison

municipalities. The “adjusted” points are the balance in covariates of the matched pairs and the “unadjusted” points are the balance in the entire universe of municipalities in Palestine. Ideally,

Figure 1 shows the balance in covariates between the treated and control units in terms of standardized mean differences. The "unadjusted" points are the balance in covariates for the entire sample of municipalities in Palestine and the "adjusted" points are the balance in covariates of the matched pair of municipalities (the 56 matched treated and control units). Figure 2 shows the map of treated and matched control units, and the full list of the matched pairs is in Appendix 3.

Ideally, the standardized mean differences should be as close to 0 as possible, meaning that there is no difference in the covariates between the treated and the control units. The absolute standardized mean differences should be less than 0.25 (Rubin 2001) or more conservatively, 0.1 (Stuart et al. 2013) for the adjustment to be trustworthy. Overall, there is a better balance in all the covariates in the adjusted sample than the unadjusted sample, i.e. the standardized mean differences in the adjusted sample are closer to 0 than the unadjusted sample, but there is still large imbalance across many covariates between the treated and control units, including population. The imbalance is to be expected, especially on variables such as population and levels of development, given that the units that were chosen to be treated were chosen on specific characteristics. The imbalance implies that the strong ignorability condition is not met and there would still be bias in estimation of the program effects. Most importantly, the evaluation is not in the position to make strong causal claims about the effect of LGI.

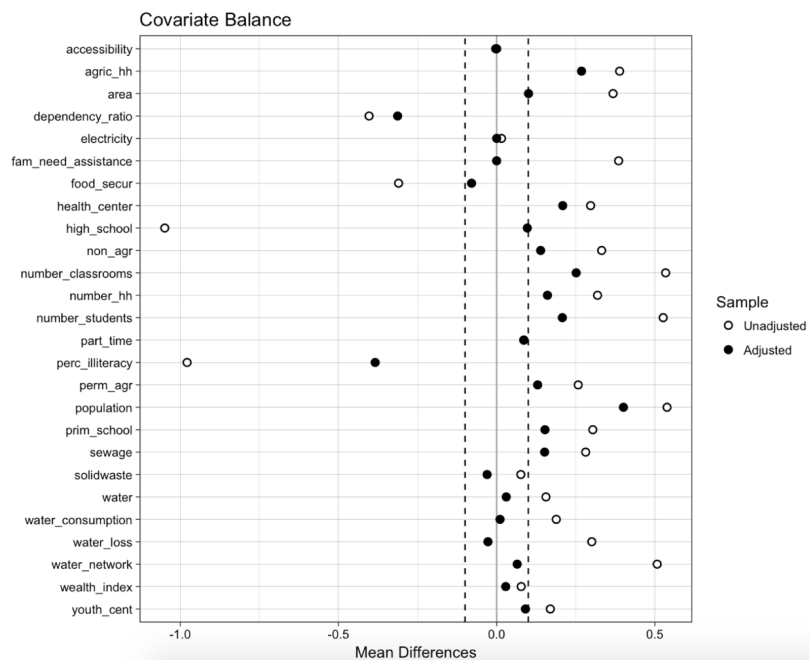


Figure 1. Covariate balance between unadjusted and adjusted samples

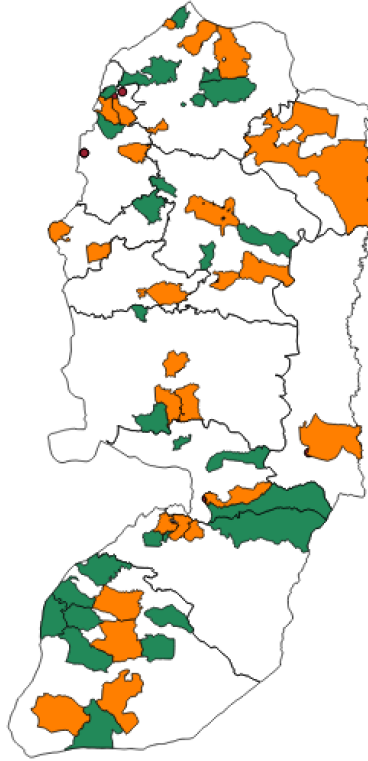


Figure 2. Map of the treated (orange) and control (green) units

2.3. Survey sampling and subjects

Surveys were conducted to evaluate the program effects in the programmed and comparison municipalities. Municipal office users and the municipal officials (administrative and political) were surveyed between March and April 2017. This section explains the details and the logistics of the survey design and sampling procedure.

2.3.1. Users

The goal for the endline survey was to interview 30 users in each of the 58 municipalities. This yielded a target of 1740 surveys; 1751 surveys were ultimately conducted. Within each municipality, respondents were sampled at the municipal office; enumerators were instructed to ask the users that had just completed their task at the municipal office whether they would be willing to participate in a 25-minute survey. When they agreed, the enumerators took the respondents to a room provided by the municipal office to conduct the surveys. The interviews were conducted in private rooms so that the respondents would not be affected by the presence of political or administrative officials when responding to the questions. When they refused, the enumerators filled out the survey as “refused” and asked the respondent why they refused to participate in the survey.

In cases where municipal offices (one-stop shops) did not exist, such as in village councils, enumerators visited the villages and conducted surveys at households of the villages. 96 surveys were conducted in either the respondents' homes or businesses. Enumerators asked whether the respondent from the household had visited at the municipal office and whether the respondent would be willing to participate in a 25-minute survey. In case the respondent refused, the enumerators were instructed to go to the next house. If no respondent was present at the household the first time, enumerators were instructed to go back to the household.

2.3.2. Officials

The goal was to interview two administrative officials and two political officials per municipality, totaling in 224 officials, and a total of 225 officials were interviewed. The enumerators were instructed to balance the number of administrative and political officials. For surveys with the political officials such as the mayor, JMCC, the survey company, had set up appointments in advance. For administrative officials, enumerators were instructed to go through the list below to inquire whether they were available for a 25-minute survey. If the official refused, enumerators were told to inquire about the next administrative official on the list. If the official agreed to take the survey, enumerators went through the survey questions in his/her office.

The list of political officials, in the order of priority:

- Mayor
- Deputy mayor
- Municipal council member

The list of administrative officials, in the order of priority:

- Head of tax office
- Manager of the service center
- Director of planning
- Director of waste/water management
- Head of electricity department
- Head of local economic development and investment
- Head of audit department
- Human resources manager
- Head of public relations
- Financial manager
- Accountant

2.4 Survey Practicalities

2.4.1. Survey company and enumerators

The survey team was composed of 17 enumerators employed by Jerusalem Media and Communications Center (JMCC).

2.4.2. Training

Training was provided over the course of three days, from March 20th to March 22nd. There were three parts to training: first, the evaluation team and the enumerators went through each of the questions in both the user and official surveys. This was to ensure that the language of the survey was suited to the local conditions and to ensure that enumerators were clear with the meaning and the intention of the survey questions.

Second part of the training was on the technicalities of the tablet and the survey program (Open Data Kit; ODK). Supervisors were trained in how to load the survey form into the Android tablets and to upload the completed forms to the encrypted server.

The last part of training was practicing the surveys on tablets to become familiarized with ODK. The enumerators were paired with partners to go through all the questions and practice putting the responses into ODK. During this process, technical errors were reported and fixed.

In addition to in-person training by the evaluation team, an interviewer manual was prepared detailing the survey procedures that was translated into Arabic by JMCC.

2.4.3. Pilot and review of the pilot

The pilot of the user survey was launched on March 23rd. The enumerators went to 12 municipalities and conducted a total of 62 user surveys. On March 25th, the review of the pilot survey was conducted. On March 27th, the surveys went into the field and the last survey came in May 6th.

2.4.4. Assessing the quality of data

Supervisors were asked to upload the completed survey forms once every two days for the first two weeks and once a week thereafter. The evaluation team conducted weekly assessment of the quality of the data to address problems. The weekly reviews focused on the following aspects:

- Duration and the number of the surveys by enumerator
- Number of surveys by day of the week
- Survey start hours
- Checks for completion

- Balance check, including male/female balance, randomization balance for survey experiment questions
- Checks for GPS locations

2.5. WB/USAID SURVEY

A second method to evaluate the effects of the program is to use the World Bank/USAID survey on households and administrators. Two surveys will be used for this portion of the evaluation: the Local Government Performance Assessment (LPGA) survey, designed by Fotini Christia, Ruben Enikolopov, and Erin York, that provides data on household level (demand side of services), and the Questionnaire on Municipal Service Provision that provides data on administrator level (supply side of services).

2.5.1. Demand side (LPGA survey)

The LPGA survey was conducted between August 2016 and February 2017. The survey covers more than 380 Palestinian municipalities and village councils, and 11970 households were surveyed. The survey included a wide range of questions with regards to service provision by the municipality, including citizen satisfaction with service reliability, service quality, and service cost. For the purpose of this evaluation, survey questions that are in line with LGI programming have been chosen to be used for the analysis.

2.5.2. Supply side

The supply side survey was conducted in 2016. The survey is intended to depict the supply side perspective of service provision--that is, how the municipal administrators in the West Bank understand their role and capacity to be. The survey covers 116 Palestinian municipalities. From each municipality, the enumerators asked the administrator in charge of the task that the survey was asking about, which meant multiple administrators answered the supply side survey. Questions covered include how the municipal administrators perceived the quality of service provision to be and how the quality had changed. Like the demand side survey, questions that are in line with LGI programming have been chosen to be included in the analysis.

3. Measures

For the evaluation of the program, there will be two types of analyses: primary and secondary. Primary analysis pertains to analyzing survey responses that are directly related to the activities done by the program and it will use the user and official surveys as well as LPGA and supply side surveys. The purpose of LPGA and supply side surveys is to validate the analysis using the matched sample, as the balance in the sample, as explained, is not optimal.

There are two components to secondary analysis. First component of the secondary analysis pertains to analyzing the “branch questions” from user and official surveys. Branch questions

are those that are asked in a survey if and only if the respondent has given a particular answer in the previous question. For instance, the question of how satisfied the respondent (user) is with the process of getting a business permit only will be asked to a respondent who has answered “yes” to a previous question asking whether the respondent has applied for a business permit before. The reasons for separating out branch questions will be discussed in section 5; put simply, we will be imputing the missing values for the surveys but there is no justified way to impute the branch questions. In light of this, we will separate out the branch questions and test each of the questions separately. The second component of the secondary analysis pertains to analyzing survey questions that are not directly on the activities implemented by the program but are questions that can be used to supplement the primary analysis, as a check for robustness of the findings.

Per the hypotheses listed in section 1.4 above, this study examines the effect of programming on three sets of outcome measures: (1) institutional development, (2) community planning, and (3) participatory governance.

As mentioned, there are two approaches to the analysis: first is to use the matched sample of municipalities where user and official surveys were run, and second is to use the universe of municipalities where LPGA and supply side surveys were conducted. For the first approach, dependent and independent variables were constructed from user and official survey data. In several instances, we rely on multiple related survey items bearing on a core concept to reduce the noise of the measurement. In such cases, we follow the standard approach of calculating z-scores (Kling, Liebman, and Katz, 2007) as summary measures.

Table 1 (user survey) and table 2 (official survey) provide a summary of the indicators and details on how we create the indices for the models that will be run using the matched sample of municipalities. Table 3 (LPGA household survey) and table 4 (supply side survey) provide a summary of the indicators and details on how we create the indices for the models that will be run for the universe of municipalities in Palestine.

Table 1: User indicator table here

Table 2: Administrator and political indicator table here

Table 3: Citizen indicator table here

Table 4: Administrator and politician indicator table (supply side survey) here

3.1 Outcome Measures

In this section, we provide primary and secondary measures for each category of objectives pursued by LGI. Primary indicators are indicators that are the most closely related to the activities that were implemented; secondary indicators are alternative dependent variables for

each of the outcome families. Some families of outcomes have both primary and secondary indicators, and some only have primary, depending on the questions that the surveys cover.

3.1.1. Outcome family 1: Institutional Development

USER

Outcome family 1, institutional development, is primarily focused on a series of variables pertaining to good governance practices--whether they have been implemented, and how the citizens perceive the quality of governance.

To reiterate the hypotheses for the outcome of institutional development from user and household surveys: The interventions will increase citizen perception of the quality of governance (H1) and increase the citizen satisfaction with the provision of services (H2)

(1) *Primary indicators (user and LGPA survey):*

Variable	Hypothesis	Survey source
Index of user perception of the quality of financial management	H1	User
Index of user experience with e-municipality	H1	User
Ordinal measure of opinion on employee selection process	H1	User
Index of user perception of fairness of service delivery for a range of services	H1	User
Index of citizen satisfaction with the quality of municipal services provided to the community	H2	User/LGPA
Index of citizen satisfaction with the quality of municipal services provided at the municipal office	H2	User/LGPA
Binary indicator of the presence of services (water, electricity, sewage, and waste collection) in the municipality	H2	LGPA
Binary indicator of whether the services are provided by the municipality/village council within which they live	H2	LGPA
Index of the change in quality of services provided over the last 3 years	H2	LGPA
Index of citizen satisfaction with problem resolution mechanisms on public services	H2	LGPA
Binary indicator of whether the household needs to resort to personal connections for services	H1	LGPA

(2) *Secondary indicators (user survey, branch questions):*

Variable	Hypothesis	Survey source
Binary indicator of whether the user ever paid side payments to obtain permits	H1	User
Ordinal measure of the fairness of permit provision by the municipality	H1	User
Ordinal measure of the frequency of municipal staff misplacing/losing documents	H1	User
Ordinal measure of user satisfaction with paying citizen taxes at the municipal office	H2	User
Ordinal measure of user satisfaction with paying business taxes at the municipal office	H2	User
Ordinal measure of user satisfaction with getting services for the disabled at the municipal office	H2	User

(3) *Secondary indicators (user survey, supplemental questions):*

Variable	Hypothesis	Survey source
Index of employee responsiveness to citizens	H1	User
Index of user perception on employee competency in service delivery	H2	User
Continuous measure of the length of time it took to get the service provided in the office	H2	User

ADMINISTRATOR

To reiterate the hypotheses for the outcome of institutional development from administrator surveys: Within municipal governments (i.e. among administrators and political officials), the interventions will increase the number of good governance practices practiced by the municipality (H3).

(1) *Primary indicators (official and supply side survey):*

Variable	Hypothesis	Survey source
Index of the implementation of good governance practices, including e-municipality, internal audit process, and human resource management	H3	Official
Index of political independence from external parties	H3	Official
Index of political responsiveness to citizens	H3	Official
Index of the presence of corruption (in form of side payments) in public service provision	H3	Official

Index of public disclosure of municipal documents	H3	Official/supply side
Index of the extent of skills training and the presence of manuals	H3	Official
Index of the presence of manuals	H3	Supply side
Index of whether the municipality provides services (water, electricity, physical planning and maintenance, piped sewage, solid waste collection) to citizens in the municipality	H3	Supply side
Index of officials' judgement on the quality of services (electricity, piped water, spatial and road planning, piped sewage, solid waste collection)	H3	Supply side
Index of the change in quality of service provision in the past 3 years as judged by the officials	H3	Supply side

(2) *Secondary indicators (official and supply side survey, branch questions):*

Variable	Hypothesis	Survey source
Binary indicator of whether the Citizen Service Center staff received additional training in the past 3 years	H3	Official
Continuous measure of the number of years electronic billing has been in place	H3	Official
Binary indicator of whether the internal audit staff received additional training in the past 3 years	H3	Official
Binary indicator of whether the human resources staff received additional training in the past 3 years	H3	Official
Continuous measure of the number of hours of training per year the human resources staff received	H3	Official
Binary indicator of whether the procurement staff received additional training in the past 3 years	H3	Official
Continuous measure of the total number of hours of training the procurement staff received	H3	Official
Ordinal measure of the frequency of procurements being subject to competitive bids	H3	Official
Ordinal measure of the frequency of procurement exceeding budgeted amounts	H3	Official

(3) *Secondary indicators (official and supply side survey, supplementary questions):*

Variable	Hypothesis	Survey source
Index of employee responsiveness to citizens	H3	Official
Continuous measure of the number of minutes it takes on average to provide services to citizens at the office	H3	Official

3.1.2. Outcome family 2: Community planning

The interventions will improve the municipal governments' community planning (H4).

Primary indicators (official and supply side survey):

Variable	Hypothesis	Survey source
Index of the extent of physical planning implementation, including naming of streets and buildings	H4	Official
Index of physical planning training, including GIS training and urban planning	H4	Official

3.1.3. Outcome family 3: Participatory governance

The interventions will increase citizen participation through various mechanisms including feedback and local councils (H5).

(1) Primary indicators (user and LGPA survey):

Variable	Hypothesis	Survey source
Index of community participation through feedback to the municipal employees and meetings organized by the municipality	H5	User
Index of whether the citizens have ways to communicate with the municipal officials on feedback, complaints, or information	H5	LGPA

(2) Primary indicators (official and supply side survey):

Variable	Hypothesis	Survey source
Index of youth participation through Youth Local Council	H5	Official/supply side
Binary indicator of whether the municipality actively involves citizens in local government planning	H5	Official/Supply side

Index of the presence of public meetings/campaigns to educate citizens about their rights and good governance practices	H5	Supply side
---	----	-------------

(3) *Secondary indicators (administrator and politician survey):*

Variable	Hypothesis	Survey source
Continuous measure of the number of hours in a typical week the municipal employees spend dealing with citizens and the community	H5	Official

Indicator tables 1 - 4 provide further details on how each of these measures are produced.

3.2 Treatment Indicators and Control Variables

3.2.1. Independent variables: Treatment indicators

We test for the effect of LGI programming, using an indicator variable that equals 1 if a district was assigned to receive LGI programming, and 0 if a district was not assigned to receive the programming. If we are able to decipher the extent of individual activities in the municipalities with further information, we will incorporate it into the analysis by assigning 1 for the municipality that was treated with the specific activity and 0 for the non-treated municipality.

3.2.2. Independent variables: Control variables

The models will include a set of control variables. In the user and citizen surveys, these include:

- Key respondent characteristics, including gender, age, employment status, and whether the respondent has connections in the office.
- A respondent asset index, calculated as the first component in principal component analysis conducted on a series of asset questions.²
- Matched indicators, including the municipality's size of the population, area, and the number of households in the locality.

The user- and citizen-level measures account for standard features of respondents likely correlated with attitudes toward and experiences with the public sector. The matched indicators address pre-treatment municipality-level features that may be correlated with the outcomes of interest.

In the analysis of the administrator/official data, we introduce a measure for the length of time respondents have been in their current positions. While not a perfect measure, it attempts to address the potential for newly rotated administrators into treatment municipalities with little

² The asset questions include whether or not the user has a car, motorcycle, tractor, animal-drawn car, fridge, washing machine, television, DVD, satellite, and a radio.

experience with the programming, or newly rotated administrators into control municipalities who were previously in treatment municipalities. To the extent that much of the programming was done on municipality level, the rotation of the administrators may mute the capacity to identify a treatment effect.

4. Empirical Approach/Models

4.1. Primary analysis: Main effects for user and administrator surveys

We present the results of models that include an indicator variable for LGI and a set of control variables. There is debate as to the costs and benefits of introducing covariates into the analysis of randomized control trials; because LGI was not designed as an RCT and the selection of programmed municipalities was intentional, it is particularly important to include a set of covariates to account for differences across municipalities that could have an impact on the outcomes of interest independent of LGI programming.

We use various specifications, mainly OLS including linear probability models, and logistic regressions to estimate the effect of each treatment arm of interest using the following baseline specification:

$$[1] Y_{ij} = \beta_0 + \beta_1 T_{ij} + \psi_j + \mu_{ij}$$

where Y_{ij} is the outcome measure of respondent i in municipality j . As described in Section 3, most outcome indicators are constructed from survey data and are continuous, dichotomous and ordinal. We also include matched pair indicators, ψ_j per Bruhn and McKenzie (2009) to account for the matching design. We rely on OLS (rather than a combination of linear and non-linear models) across all these variable types in light of evidence on the robustness of OLS in a wide range of settings. T_{ij} is the treatment dummy for each of the two treatment arms of interest, and μ_{ij} are robust standard errors clustered at the municipal level using Huber-White sandwiched standard errors (Lin et al., 2013).

4.2. Primary analysis II: Main effects for citizen and administrator surveys

In a second set of models, using the citizen and administrator surveys, we include all municipalities in the sample, not just the matched pairs, and hence exclude the matched indicators. Thus, we estimate:

$$[2] Y_{ij} = \beta_0 + \beta_1 T_{ij} + \mu_{ij}$$

4.3. Secondary analysis I: User and administrator surveys

The secondary analysis for the evaluation

5. Missingness, and Multiple Hypothesis Testing

Treatment of attrition and missing values

A number of tests will be run to determine whether there is attrition bias; these tests include t-tests and ANOVA tests to answer the following:

- Is the magnitude of attrition different between programmed and comparison municipalities?
- Are there baseline characteristics of attrited users/citizens in the comparison municipalities significantly different than the baseline characteristics of the attrited users/citizens in the programmed municipalities?

If statistical tests reveal that attrition happened at random and is rare enough to not strongly affect the power of the design, attrited users/citizens will be dropped from the analysis. However, if statistic tests reveal non-random differences in magnitude or baseline characteristics of the attrited in the treated and control areas, the analysis will adjust the sample through the use of imputation, weighting or Lee bounds so the share of observed individuals is equal for both treatment and control groups.

Missingness is likely to exist in the survey data that underpins our outcome measures. We drop from the analysis questions for which more than 30 percent of responses are missing. When missingness is lower, we use multiple imputation (Rubin 2004) to fill in the missing values. We use level of education, asset index, region, their political linkages, as well as outcome measures as predictors. We remove variables that have 90 percent of observations with the same value for a response category from the analysis. These variables are not included as covariates or outcome indicators. Note, as mentioned in section 3, the “branch questions” will not be imputed for the lack of a justified method of imputation.

Multiple hypothesis testing

Given the large number of hypotheses that we will test in our statistical analyses and our reliance on null hypothesis significance testing, we expect our inferences to lead to a non-negligible count of Type I errors (“false positives”) and Type II errors (“false negatives”). Our evaluation will report both raw p-values that are unadjusted for the multiple comparisons we make throughout our analysis and also more conservative p-values adjusted for multiple comparisons using the Benjamini & Hochberg (1995) False Discovery Rate Correction. Our main findings and summary sections will rely on the uncorrected values, because we are analyzing a number of closely related interdependent outcomes and, therefore, the standard corrections for the false discovery rate are likely too conservative (Gelman, Hill, and Yajima, 2012).³ Nevertheless, we will clearly note in the body of the main report when Benjamini and

³ Gelman and co-authors note that for most social science studies, where the effects may be small but are unlikely to be exactly zero, the corrections are likely too conservative.

Hochberg-adjusted p-values are substantively different and suggest a conservative interpretation.

7. Threats to Inference

There are a number of threats to inference that should be taken into account.

Balance

Although imbalance on pre-treatment variables is often a threat to impact evaluations, the use of the matching procedure for sampling of control municipalities from the total population of municipalities led to a sample of municipalities more balanced across a wide range of observed confounders, although there is still imbalance across a number of variables, including the size of the population, since the selection was based on these variables. To that end, the citizen and administrator and politician (supply side) surveys with the entire sample of municipalities will also be used for the analysis.

Heterogeneity in implementation across municipalities

According to the Global Communities document on the implementation of activities across municipalities, we can see that not all activities were implemented across all municipalities. For instance, only 4 out of 28 municipalities received e-municipality treatment, and 26 out of 28 municipalities received human resources management training. Furthermore, many of the activities are still marked as “ongoing”, hampering the judgement with regards to the extent of programming in the 28 municipalities. If we are able to get enough information on the intensity of treatment for each of the treatment arms, we will make use of the information to analyze the intention-to-treat (ITT).

Spillovers

There is a potential source of spillover from treated to control municipalities and from control to treated municipalities via the rotation of administrators from one municipality to another. To the extent that a lot of the programming was inside the municipal administration, rotation of the administrators mutes our capacity to identify a program effect: administrators from the treated municipalities could rotate into the control municipalities and hence have had exposure to the training that were implemented in treated municipalities, or the administrators from control municipalities could rotate into the treated municipalities and be less exposed to the activities even if they are surveyed in the treated municipalities. Absent a registry for all administrative rotations, we rely on a measure of the length of time an administrative respondent has been in their current position as a control variable in the primary analysis.

Another source of spillover is from related donor programming in other municipalities by World Bank and other agencies. Similar to the first point, related donor programming on administrators could have potential spillover through rotations of staff across municipalities. Again, absent a

registry for all administrative rotations as well as a list of programs that were implemented by other donors, we rely on the measure of the length of time an administrative respondent has been in their current position as a control.

External factors

Major factors impacting local services, such as electricity, water, and road access, are outside the control of local governments. Depending on the frequency of service cuts or the reliability of services provided by the municipality, the positive effect of LGI could be muted. We control for this by adding as control variable the frequency of service cuts to the analysis of the user and citizen surveys.

7. Ethical Considerations

Participation in the study was voluntary, and all respondents were required to give their informed consent at the beginning of the survey process. The evaluation team received Institutional Review Board approval for the survey instruments from Duke on March 22, 2017; minor modifications were made to question wording as a result of training sessions and piloting. Informed consent was received from each participant after reading a statement about the purpose of the research, the content of the survey, any risks or benefits, and the time commitment. Participants were assured their participation was voluntary and could be withdrawn at any point and that their answers would be kept confidential.

As described in section 2, quantitative data was collected through the ODK platform on Android tablets. Tablets were password protected, and data was uploaded to an encrypted server when network connectivity was available. Data is stored on password encrypted computers, with PII removed.

Tables

Table 1: User indicator table

Table 2: Administrator and politician indicator table

Table 3: Citizen indicator table

Table 4: Administrator and politician indicator table (supply side survey)

Table 1: User indicator table

Description	Operationalization	Items
INSTITUTIONAL DEVELOPMENT		
Financial management	z-score(E2, E3, E6)	E2. What share of municipal government resources do you think your municipality wastes due to LOCAL POLITICIANS using the money for PERSONAL USE? E3. What share of municipal government resources do you think your municipality wastes due to LOCAL ADMINISTRATORS using the money for PERSONAL USE? E6. How likely is your MUNICIPAL COUNCIL REPRESENTATIVE to help members of their own party at the expense of the people of your municipality?
E-municipality	1 if either/both = 'yes', 0 if none = 'yes'	C19. could you apply for services through the website ? C19_2. could you pay for services through the website?
HR management and information systems	F3	F3. (how much do you agree with the statement) Selection of municipal employees reflects a competitive process based on merit
Accountability	Z-score(C25a, C25b, C25c, C25d, C25e, C25f, C25g, F5a, F5b)	C25a. Have you ever had to pay side payments to the municipal office for water provision? electricity provision/tax payments/roads and basic infrastructure/sewage maintenance/solid waste collection/permits? (y/n)
		C25b. Have you ever had to pay side payments to the municipal office for electricity provision?
		C25c. Have you ever had to pay side payments to the municipal office for tax payments?
		C25d. Have you ever had to pay side payments to the municipal office for roads and basic infrastructure?
		C25e. Have you ever had to pay side payments to the municipal office for sewage maintenance?
		C25f. Have you ever had to pay side payments to the municipal office for solid waste collection?
	C26	F5a. How much do you agree with the statement: Municipal employees take side payments from the citizens to help them bend the rules to pay lower taxes (y/n)
		F5b. How much do you agree with the statement: Municipal employees take side payments from local businesses to help them bend the rules to pay lower taxes (y/n)
	Z-score(E4, E6)	C26. Compared to two years ago, has the need to pay side payments increased or decreased? (inc/dec/remians the same)
	If either F6a or F6b has citizens, code as 1. If none of the two has citizens, code as	E6. How likely are your municipal council representatives to help members of their own party at the expense of the people of your municipality?
F6a. In your opinion, who are municipal employees MOST responsive to? (responses: citizens/incumbent party/elected officials)		
Accountability - Branch Q	C25g	F6b. In your opinion, who are municipal employees SECOND MOST responsive to? C25g. Have you ever had to pay side payments to the municipal office for permits?

Attitude	mean(C28a, C28b, C28d, C28e, C28f)	C28a. How fairly does the municipality provide water?
		C28b. How fairly does the municipality provide electricity?
		C28d. How fairly does the municipality provide roads and basic infrastructure maintenance?
		C28e. How fairly does the municipality provide sewage maintenance?
		C28f. How fairly does the municipality provide solid waste collection?
	Z-score(C26, C28c)	C26. In your view, do municipal employees treat everyone fairly? C28c. How fairly does the municipality provide tax collection services?
Attitude - Branch Q	C28g	C28g. How fairly does the municipality provide permits?
Service delivery	C14 (yes = 1, no = 0)	C14. Did you receive guidance on the type of documentation/evidence needed to apply to process your request and if so, how? Select all that apply.
	C22	C22. How knowledgeable do you think the municipal employees are?
	Z score(C2c, C2f, C2g, C2h, C2i)	C2f. Please position your level of satisfaction with water provision of the municipality
		C2g. Please position your level of satisfaction with electricity provision of the municipality
		C2h. Please position your level of satisfaction with solid waste collection
C6	C6. In this visit, how many MINUTES did you need to wait to get the task completed?	
Service delivery - Branch Q	C17	C17. In the last 12 months, how often has the municipality misplaced or lost a document you provided them?
	C2a	C2a. Please position your level of satisfaction with paying taxes at the municipal office
	C2d	C2d. Please position your level of satisfaction with paying taxes for your business at the municipal office
	C2b	C2b. Please position your level of satisfaction with getting services for disabled citizens at the municipal office
STRATEGIC PLANNING		
PARTICIPATORY GOVERNANCE		
Enhancing dialogue b/w LGUs and citizens	C23, C24	C23. In the last 12 months, were you asked to provide feedback to the municipal employees about their service?
		C24. In the last 12 months, have you participated in or heard of any meetings organized by the municipality on how to improve services of the municipality?

Table 2: Administrator and politician indicator table			
Description	Operationalization	Items	
INSTITUTIONAL DEVELOPMENT			
Citizen service center	E12	E12. Does the municipality have a Citizen Service Center?	
Citizen service center - Branch Q	E13	E13. Have your Citizen Service Center staff received additional training in the last 3 years?	
E-municipality	Z-score(I1, I1y, I2)	I1. Does your municipality have electronic billing for any of its municipal services? I2. Does the municipality have a digitized system for information management?	
E-municipality - Branch Q	I1y	I1y. How many years has electronic billing been in place?	
internal audit process	D4	D1. Do you have an internal audit manual? D5. Did you conduct an internal audit in the last 12 months?	
		D4. In the past 3 years, how has the internal auditing skills of the municipal employees changed? (coded as: Increased a lot/increased a little/did not change/decreased a little/decreased a lot)	
		D3. Has your internal audit staff received additional training in the last 3 years?	
Internal audit - Branch Q	D3	D3. Has your internal audit staff received additional training in the last 3 years?	
	D8	D8. In your opinion, how independent is the internal audit staff from the rest of the municipal government?	
HR management and information systems	Z-score(G7, G10)	G2. Is human resource management digitized? G3. Is payroll digitized? G4. Does the municipality have a digitized system for hiring municipal employees? G5. Does the municipality have a digitized system for performance assessment of	
		G7. In your opinion, what percentage of employees in your municipality are hired on merit? G10. Does your municipality have a clear set of criteria for salary increases and promotions?	
		Z-score(G9, G11)	G11. In the past 3 years, how has the importance of MERIT in the PROMOTION process changed? G9. In the past 3 years, how has the importance of MERIT in SELECTION process changed?
	HR management and informaiton systems - Branch Qs	G1	G1. Have your human resources employees received additional training in the last 3 years?
		G1a	G1a. Approximately how many hours of training per year did they receive?
		E6. In the past 3 years, how have the staff's physical planning skills changed?	

Training	Z-score(E6,G6)	G6. In the past 3 years, how has the efficiency of human resource management of the municipal staff changed?
Accountability	Z-score(C4,C5,C6)	C4. (How much do you agree with the statement:) Municipal employees are independent of municipal council when performing their job.
		C5. (How much do you agree with the statement:) Municipal employees are independent of the pressure by the Palestinian National Government.
		C6.(How much do you agree with the statement:) Municipal employees are independent of the pressure by the israeli Government.
	If either C19 or C20 has "citizens", code as 1. If neither has "citizens", code as 0	C19. Who are the municipal employees the MOST responsive towards?
		C20. Who are the municipal employees the SECOND MOST responsive towards?
	Z-score(C7, C8, C9, C10)	C7. (How much do you agree with the statement:) Citizens who provide side payments would get faster services.
	C8. (How much do you agree with the statement:) businesses that provide side payments would get faster services.	
	C10. In the past 2 years, has the share of citizens or businesses giving side payments to process their requests increased or decreased?	
Public disclosure		E2_1. Does the municipality publish municipal balance sheet?
		E2_2. Does the municipality publish summary of approved annual budgets?
		E2_3. Does the municipality publish detailed approved annual budgets?
		E2_4. Does the municipality publish actual executed budgets?
		E2_5. Does the municipality publish municipal investments?
		E2_6. Does the municipality publish SDIP Execution?
		E2_7. Does the municipality publish external audit reports?
		E2_8. Does the municipality publish municipal Council decisions?
Procurement process and manuals	F1	F1. Is your procurement system digitized?
Procurement process and manuals - Branch Qs	F2	F2. Has your procurement staff received additional training in the last 3 years?
	F2a	F2a. Approximately how many hours of training in total did they receive?
	recode F3: reverse t	F3. How often are municipal procurements subject to competitive bids?
	recode F4: reverse t	F4. How often over the last 12 months have procurement costs exceeded budgeted amounts?
		E14a. On average, in minutes, how long would you say it takes a citizen to pay taxes

Service delivery	Z-score	E14b. On average, in minutes, how long would you say it takes a citizen to pay business taxes
		E14c. On average, in days, how long would you say it takes a citizen to get a building permit?
		E14d. On average, in days, how long would you say it takes a citizen to get a business permit?
STRATEGIC PLANNING		
Physical Planning	Z-score(H1, H2, H3y)	H1. in percentage terms how complete is the naming of the streets?
		H2. In percentage terms how complete is the numbering of the buildings?
		H3y. In percentage terms, how complete is the local registrar?
		H4n. Does the municipality have a detailed map of landmarks and municipal services and if so, how complete is it?
	code as 1 if either =	E4. Have the municipal employees been trained in GIS in the past 3 years?
		E5. Have the municipal employees been trained in urban/municipal planning in the last 3 years?
PARTICIPATORY GOVERNANCE		
Youth shadow local councils	Z-score(E10,E11)	E10. Does the municipality have a Youth Local Committee?
		E11. How important would you say the youth input is to the decisions of the municipal council?
Enhancing dialogue b/w LGUs and citizens	E9	E9. Does the municipality have a system to allow public input on proposed policies or actions?
	sum(B12d, B12e)	B12d. In a TYPICAL working week, how many HOURS do you spend on each of the following tasks? Time spent on official meetings with citizens
		B12e. In a TYPICAL working week, how many HOURS do you spend on each of the following tasks? Time spent on with civil society or community groups

Table 3: Citizen indicator table		
Description	Operationalization	Items
INSTITUTIONAL DEVELOPMENT		
Citizen service center	Q10_7_1	Q10_7_1. Does your community have CSC available?
Service delivery: Availability	Q2_03_1, Q2_03_3, Q2_03_4, Q2_03_5	Q2_03_1. Does your household have access to electricity grid?
		Q2_03_3. Does your household have access to piped water?
		Q2_03_4. Does your household have access to piped sewage?
		Q2_03_5. Does your household have access to waste collection?
Service delivery: Satisfaction	mean(Q2_05_1, Q2_05_3, Q2_05_4, Q2_05_5)	Q2_05_1. Overall, how satisfied are you with the reliability of electricity grid?
		Q2_05_3. Overall, how satisfied are you with the reliability of piped water?
		Q2_05_4. Overall, how satisfied are you with the reliability of piped sewage?
		Q2_05_5. Overall, how satisfied are you with the reliability of waste collection?
	mean(Q2_06_1, Q2_06_3, Q2_06_4, Q2_06_5, Q9_10)	Q2_06_1. How satisfied are you with the overall quality of electricity grid?
		Q2_06_3. How satisfied are you with the overall quality of piped water?
		Q2_06_4. How satisfied are you with the overall quality of piped sewage?
		Q2_06_5. How satisfied are you with the overall quality of waste collection?
	Q9_10. Overall, how satisfied are you about the physical condition of the road outside?	
	Q2_07_1, Q2_07_2, Q2_07_4, Q2_07_5	Q2_07_1. How has overall quality of electricity grid changed during the last 3 years?
		Q2_07_2. How has overall quality of piped water changed during the last 3 years?
		Q2_07_4. How has overall quality of piped sewage changed during the last 3 years?
Q2_07_5. How has overall quality of waste collection changed during the last 3 years?		
Service delivery: Responsible actor	Q2_11_3, Q2_11_4, Q2_11_5	Q2_11_3. Is piped water provided by the same [municipality/village council] in which you live?
		Q2_11_4. Is piped sewage provided by the same [municipality/village council] in which you live?
		Q2_11_5. Is waste collection provided by the same [municipality/village council] in which you live?
Service delivery:	Z-score(Q10_13, Q10_10_1, Q10_10_2, Q10_10_3)	Q10_13. DO you agree with the following statement: "My municipality/village council is very responsive to citizen concerns and complaints"
		Q10_10_1. How satisfied are you with CSC to resolve problems with public services?
		Q10_10_2. How satisfied are you with feedback/complaints system to resolve problems with public services?

responsiveness	Q10_10_2, Q10_10_3, Q10_10_4, Q10_10_5)	Q10_10_3. How satisfied are you with grievance redress mechanism to resolve problems with public services?
		Q10_10_4. How satisfied are you with website of municipality/village council to resolve problems with public services?
		Q10_10_5. How satisfied are you with service booklets to resolve problems with public services?
Service delivery: Personal connections	Recode 'no' (currently 2) as 0	Q10_6_1. Does your household ever need to resort to personal connections to obtain Electricity services?
		Q10_6_2. Does your household ever need to resort to personal connections to obtain Piped water services?
		Q10_6_3. Does your household ever need to resort to personal connections to obtain Piped sewage services?
		Q10_6_4. Does your household ever need to resort to personal connections to obtain Waste collection services?
STRATEGIC PLANNING		
PARTICIPATORY GOVERNANCE		
Enhancing dialogue b/w LGUs and citizens	Q10_7_2, Q10_7_3, Q10_7_4	Q10_7_2. Does your community have feedback/complaints systems available?
		Q10_7_3. Does your community have grievance redress mechanism available?
		Q10_7_4. Does your community have website of municipality/village council available?
		Q10_14. Have you ever visited the website / Facebook page of your Municipality / Village Council?

Table 4: Administrator and politician indicator table (supply side survey)

Description	Operationalization	Items
INSTITUTIONAL DEVELOPMENT		
Citizen service center	If 'yes' or 'planned/in process', code as 1. if 'no', code as 0	26.10. Does your municipality have a Citizen Service Center?
Financial management (*NOTE: As per the "Governance activities per location" document, no municipalities received financial management packages! uh oh)	75, 79, 86, 80	75. Does the municipality have an approved integrated computerized financial management system tracking budgeting, collection and billing? (respon: Yes, no)
		79. Does the municipality have a prepared budget for each of the last three years?
		86. Does the municipality employ a professional accountant?
		80. Has the municipality submitted the budget to the Ministry of Local Government for approval?
E-municipality	77	77. Does the municipality have e-billing for any of its municipal services? (yes/no)
Public disclosure	88, 89.10, 89.20, 89.30, 89.40, 89.50, 89.60, 89.70	88. Does the municipality post financial statements to the public?
		89.10. Does the municipality publish summary of approved annual budgets?
		89.20. Does the municipality publish detailed approved annual budgets?
		89.30. Does the municipality publish detailed actual executed budgets?
		89.40. Does the municipality publish municipal investments?
		89.50. Does the municipality publish SDIP execution?
		89.60. Does the municipality publish external audit reports?
		89.70. Does the municipality publish municipal council decisions?
Records and manuals	78.00.1, 78.00.2, 78.00.3, 78.00.5	78.00.1. Does the municipality have Fixed Assets Registry? (yes, updated/yes, not updated/no)
		78.00.2. Does the municipality have Operations and maintenance plan?
		78.00.3. Does the municipality have Accounting and financial procedures manual?
		78.00.5. Does the municipality have Procurement records?
Service delivery: responsible actor	17_1, 17_2, 17_3, 17_4, 17_5, 17_7	17_1. Who provides the electricity service for citizens within the municipality borders? "Entirely the municipality (yes/no)"
		17_2. Who provides water service for citizens within the municipality borders? "Entirely the municipality (yes/no)"
		17_3. Who provides spatial and road planning and maintenance service for citizens within the municipality borders? "Entirely the municipality (yes/no)"
		17_5. Who provides piped sewage for citizens within the municipality borders? "Entirely the municipality (yes/no)"

		17_7. Who provides solid waste collection and treatment for citizens within the municipality borders? "Entirely the municipality (yes/no)"
Service delivery: quality	Z-score	How would you rate the quality of electricity provision? (doesn't apply, poor, fair, adequate, good, very good)
		How would you rate the quality of piped water provision?
		How would you rate the quality of spatial and road planning and maintenance?
		How would you rate the quality of piped sewage service provision?
		How would you rate the quality of solid waste collection and treatment service provision?
Service delivery: change in quality	Z-score	How has the quality of electricity provision changed in the past 3 years? (deteriorated, improved, no change)
		How has the quality of piped water provision changed in the past 3 years?
		How has the quality of spatial and road planning and maintenance changed in the past 3 years?
		How has the quality of piped sewage provision changed in the past 3 years?
		How has the quality of solid waste collection and treatment changed in the past 3 years?
STRATEGIC PLANNING		
Physical Planning	78.00.4	78.00.4. Does the municipality have Strategic Development and Investment Plan?
	9	9.00. Does the municipality have a physical plan?
PARTICIPATORY GOVERNANCE		
Youth shadow local councils	Z-score(11, 13)	11. Does your municipality have a Youth Local Council?
		13. Please indicate how active the Youth Local Council is. (Holds activities... 5 or more times a year, at least once a month, at least once every three months, at least once every six months, at least once a year)
Enhancing dialogue b/w LGUs and citizens	95	95.00 Does the municipality actively involve citizens in local government planning? (yes/no)
Provision of information to the public	If any of the four = yes, code as 1. If none = yes, code as 0.	96.20 Does the municipality hold public campaigns to educate citizens about service provision and associated costs?
		96.30 Does the municipality hold public campaigns to educate citizens about bill payment and collection provisions?
		96.40 Does the municipality hold public campaigns to educate citizens about citizens rights and responsibilities?
		96.50 Does the municipality hold public campaigns to educate citizens about benefits of civic participation?

Appendix

Appendix 1. List of municipalities

- Tubas
- Qalqiliya
- Tulkarm
- Illar
- Kafr Thulth
- Attil
- Aqraba
- Al Yamun
- Qabalan
- Anabta
- Nablus
- Jenin
- Ar Rama
- Al Bireh
- Bir Zeit
- Jericho
- Salfit
- Abu Dis
- Ramallah
- At Tayba
- Hebron
- Yatta
- Beit Jala
- Beit Sahur
- Halhul
- Adh Dhahiriya
- Beit Fajjar
- Bethlehem

Appendix 2. List of covariates included in the matching algorithm

- Population
- Area (km²)
- Availability of water services (binary; yes or no)
- Availability of electricity network (binary; yes or no)
- Availability of sewage and wastewater systems (binary; yes or no)
- Availability of solid waste collection system (binary; yes or no)
- Accessibility to road network (1km of service area is used and thus the accessibility to road network was assessed)
- Water network (% with access)
- Water loss (% of freshwater lost)
- Water consumption (liter per capita per day)
- Number of students
- Number of classrooms
- Dependency ratio - A measure of the portion of a population which is composed of dependents. The dependency ratio is equal to the # of individuals aged below 15 and above 64 divided by the # of individuals aged 15 to 64
- illiteracy (%)
- Presence of a youth center
- Families in need of assistance
- Food security index - Households with income and consumption below \$4.7/capita (adult)/day and household showing decrease in total food and non-food expenditures including HHs unable to further decrease their expenditure patterns
- Wealth index - Is the value of all natural, physical and financial assets owned by a HH, reduced by its liabilities? The wealth index is a composite index composed of key asset ownership variables; it is used as a proxy indicator for HH level of wealth. The criteria reflect % of HHs that are measured within poorest wealth index quintile.
- permanent agriculture
- Primary school
- High school
- Health center - number of health centers per locality
- Part time (%) - % of working force working part time?
- Non-agriculture households
- Number of households
- Agricultural households
- Families in need of assistance - percentage of families that need and did not receive assistance

Appendix 3. Matched pairs of municipalities

Treated	Control
Tubas	Ash Shuyukh
Qalqiliya	Ya'bad
Tulkarm	Beit Ummar
'Illar	Baqa ash Sharqiya
Kafr Thulth	Kafr Qaddum
'Attil	Deir al Ghusun
Aqraba	Beit Furik
Al Yamun	Beit Ula
Qabalan	Huwwara
'Anabta	Ad Doha
Nablus	Al 'Ubeidiya
Jenin	Fahma al Jadida
Ar Rama	Mirka
Al Bireh	Qabatiya
Bir Zeit	Qarawat Bani Zeid
Jericho (Ariha)	Rummana
Salfit	Bir Nabala
Abu Dis	Beituniya
Ramallah	Tarqumiya
At Tayba	Ash Shuhada
Hebron (Al Khalil)	As Sawahira ash Sharqiya
Yatta	As Samu'
Beit Jala	'Anata
Beit Sahur	Al Khadr
Halhul	Idhna
Adh Dhahiriya	Dura

Beit Fajjar	Deir Sharaf
Bethlehem (Beit Lahm)	Bani Na'im

References

- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)*, 289-300.
- Brookhart, M. A., Schneeweiss, S., Rothman, K. J., Glynn, R. J., Avorn, J., & Stürmer, T. (2006). Variable selection for propensity score models. *American journal of epidemiology*, 163(12), 1149-1156.
- Dunning, T. (2012). *Natural experiments in the social sciences: a design-based approach*. Cambridge University Press.
- Gelman, A., Hill, J., & Yajima, M. (2012). Why we (usually) don't have to worry about multiple comparisons. *Journal of Research on Educational Effectiveness*, 5(2), 189-211.
- Gerber, A. S., & Green, D. P. (2012). *Field experiments: Design, analysis, and interpretation*. WW Norton.
- Kling, J. R., Liebman, J. B., & Katz, L. F. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1), 83-119.
- Lin, W. (2013). Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique. *The Annals of Applied Statistics*, 7(1), 295-318.
- Rosenbaum PR, Rubin DB. 1985. "Constructing a control group using multivariate matched sampling methods that incorporate the propensity score." *The American Statistician* 29:33-38.
- Rubin, D. B. (2001). Using propensity scores to help design observational studies: application to the tobacco litigation. *Health Services and Outcomes Research Methodology*, 2(3), 169-188.
- Rubin, D. B. (2004). *Multiple imputation for nonresponse in surveys* (Vol. 81). John Wiley & Sons.
- Stuart, E.A., Lee, B.K. and Leacy, F.P., 2013. "Prognostic score-based balance measures can be a useful diagnostic for propensity score methods in comparative effectiveness research." *Journal of clinical epidemiology*, 66(8), pp.S84-S90.