

Evaluation Replication Worksheet

Landlord Rental Unit Registration Letters - Newark, NJ

Use this worksheet to develop your city's Trial Protocol – the document that will guide your evaluation.

Since this is an evaluation replication guide, we've gotten you started by filling in some of the elements of your approach that stay the same. The rest is up to you!

Before we dive in - a few reminders:

- A big part of designing and running an evaluation is **careful planning** - we've designed this guide following our "Trial Protocol" template. A protocol can help you to document the decisions and intentions around your intervention. It's a forward-looking plan that will help you down the road - and also make it easy for people to replicate your work in the future (like we are doing now).
- We provide some guidance within the document for each section, but this guide has been designed for those who are comfortable handling data and understand a power calculation. If that's not you yet - that's ok! We'd recommend you partner with someone in your city to complete this document. Or, you can try [courses](#) on the What Works Cities Academy to get up to speed or join a BIT Sprint for "Evaluation Foundations" to learn some of the key concepts.

To use this document: Please download a copy and fill in the "Your Approach" column

Part 1 - Scoping Your Evaluation

Step & Guidance	EXAMPLE – Newark, NJ	Your Approach [Fill in this column]
<p>Define your problem statement</p> <p>What is your high-level goal?</p>	<p><i>Landlords do not register their rental units</i></p>	<p>Landlords do not register their rental units.</p>
<p>Describe your solution</p> <p>How will you try to reach your goal?</p>	<p><i>We will provide a reminder letter to landlords to encourage them to register their units.</i></p>	<p>We will provide a reminder letter to landlords to encourage them to register their units.</p>
<p>Describe your comparison</p> <p>What will you be comparing your approach to?</p> <p>What does your comparison or control group receive?</p> <p>Is the comparison business-as-usual, or an alternate version of a new process? If not business-as-usual, why?</p>	<p><i>We will randomly assign Newark property owners to receive either the treatment letter or no letter (business-as-usual control group).</i></p>	<p>We will randomly assign property owners in our city to receive either the treatment letter or no letter (business-as-usual control group).</p>
<p>Define your outcome & indicator</p> <p>How will you measure the success of your approach? What is the target outcome and what is the indicator?</p> <ul style="list-style-type: none"> What is the unit of your outcome measure? How does it relate to your unit of assignment? How will you link your outcome measure to treatment assignment? 	<p>Target outcome: <i>Property owners registering their rental units with the Office of Rent Control</i></p> <p>Indicator: <i>The proportion of property owners who register their rental units with the Office of Rent Control within 30 days of the intervention mailing. This is a direct binary outcome measure recorded at the individual level whose exact matching process is described in the comprehensive overview's Outcome Data Preparation section.</i></p>	<p>Target outcome:</p> <p>Indicator:</p>
<p>Describe your research question</p>	<p><i>Does sending a reminder letter increase the number of landlords who register their units?</i></p>	<p>Research Question:</p>
<p>Identify your population</p> <p>Describe who you are trying to target with this intervention</p>	<p>Target Population: <i>Our target population is all landlords in Newark who have not registered their rental units with the Office of Rent Control in 2019.</i></p>	<p>Target Population:</p>
<p>Identify your sample</p> <p>Describe who you will test your intervention on (and if / how that is different than the population)</p> <ul style="list-style-type: none"> Who will you actually be able to reach with your communication? How will you identify and reach your sample? Does this sample represent your population of interest? Do you anticipate any selection bias in who is represented in your sample? 	<p>Trial Sample: <i>All property owners of Newark properties with 2 or more units, per the Newark property tax assessor, who have listed owner addresses that we can verify through SmartyStreets USPS Address Verification. This is not an exact match for our target population, but we do expect that it includes a large number of them.</i></p>	<p>Trial Sample:</p>

DATA QUALITY CHECKPOINT

Before going too much further - it's important check the quality of your data

Before you design your evaluation, you need to understand **what data is available** and **any potential issues** you may have with the data.

Reviewing your data in advance helps you determine how you design your evaluation. For example, you might find out that you have less available data than we thought, which could make your planned evaluation more difficult and may compromise your ability to measure the outcomes you seek to evaluate. Or you may discover your outcome measures are structured differently in your data than you had thought because it may require you to make different decisions about your evaluation design or explore other opportunities to generate data to help you with your evaluation.

Before you move forward, check your dataset to determine the following:

- ✓ You have access to the data set
- ✓ Your data set contains the information or outcome variable you want to measure
- ✓ You are aware of the ways the data could be inaccurate or unreliable (e.g. self-reported, incomplete, etc.)
- ✓ If new data is being collected or if data is being collected manually, consider could it go wrong?
- ✓ You've reviewed an output of this data to determine accuracy and how difficult it is to produce (if possible)
- ✓ You are handling any personal identifying information and consent processes in an accordance with legal guidelines and ethical best practices

Part 2 - Designing your evaluation

Step / Description / Resources

EXAMPLE – Newark, NJ

Your Approach

Decide randomization strategy - check out our guide for the City of Newark Here

Unit: <i>What is being randomized (e.g. individual, household, school, etc.)?</i>	Unit: We will randomize at the individual property owner level identified by unique owner address	Unit:
Method: <i>How will you conduct the randomization? (E.g., through a random number generator, lottery, coin toss, randomized paper sequence, etc.)</i>	Procedure: We will conduct a complete randomization using Stata.	Procedure:
Verification: <i>How will you make sure that every participant is assigned to one, and only one, treatment group? If there is a risk that participants receive both treatments, will you be able to track this and control for it in analysis?</i>	Verification: We have dropped all duplicate mailing addresses in our sample to ensure that each owner and their linked address are only assigned to one treatment group.	Verification:
Blinding & Masking: <i>Will participants know their own treatment assignment? Will their treatment status be known to others involved in the trial or intervention? (For example, will frontline staff know the difference between people in the control group versus treatment group?)</i>	Blinding and masking: Participants will not know their own treatment assignment, and frontline staff will only know if participants mention their letter.	Blinding and masking:

<p>Spillovers: <i>What are the ways in which someone's treatment status might affect the outcome of someone in a different treatment group? Is it possible that a previous intervention might influence their behavior in this trial? If so, can these be minimized by creating distance between participants?</i></p>	<p>Spillovers: Someone who receives the letter may inform other landlords of the registration requirement, but we do not believe we can minimize this in any meaningful way. And while this may artificially increase the control group's registration rate, it is still a positive social good for Newark.</p>	<p>Spillovers:</p>
--	--	---------------------------

Calculate your sample size and power requirements

<p>Baseline: <i>What is the current average for your indicator? Or, if you don't have historical data, what data do you have available that might give you an indication of what your current average might be? This should be your best guess for the expected outcome for your control group in your trial. If you don't have a precise number, run the power calculations for a range of different baselines based on your assumptions.</i></p> <p>Standard deviation: <i>If your outcome indicator is a continuous measure, what is the baseline standard deviation?</i></p>	<p>Baseline: We do not have a precise estimate for our baseline data. As described previously, Rutgers CLiME estimates that 50% of properties are registered, but this proportion includes all those who have ever registered, not those who have completed their 2019 registration. Based on data provided by the Office of Rent Control, it appears that between 2 and 4 percent of the properties in our dataset have registered in 2019. To account for this, below we include power calculations for a wide range of different baselines.</p>	<p>Baseline:</p>
--	---	-------------------------

<p>Power calculation: <i>Use the power calculator to complete your power calculations using the numbers you listed above.</i></p> <ul style="list-style-type: none"> • <i>Per group sample size:</i> • <i>Minimum detectable effect size:</i> • <i>Significance level: 0.05</i> • <i>Power level: 0.80</i> 	<p>Number of participants available: 15,601 Newark property owners of properties with two or more units.</p> <p>Power calculation: We use a significance level of 0.05 and a power level of 0.80</p> <p>Even at the most conservative baseline estimate of 50%, we are comfortable with the MDES of 2.24 pp.</p>	<p>Number of participants available:</p> <p>Power calculation:</p> <p>Per group sample size:</p> <p>Minimum detectable effect size:</p>
---	--	---

<p>Target effect size: <i>How large of an effect size do you think is reasonable to expect from this intervention (based on prior evidence if available)? Alternatively, what would be a meaningful effect size (based on break-even point if applicable), and why?</i></p>	<p>We hope to increase the amount of landlords who register their properties by 10-15%</p>	
--	--	--

<p>Attrition: <i>What is your best guess of the number/percentage of participants that will leave your sample between randomization and outcome data collection. Adjust your sample to ensure it's large enough for analysis at the end of the trial.</i></p>	<p>There may be some landlords that may sell their properties during this period, but we do not anticipate it to be large enough to significantly decrease our sample size.</p>	
--	---	--

FEASIBILITY CHECKPOINT:

Is the minimum detectable effect size reasonable? If not, can you change the sample size by running the trial for longer or changing the intervention to make it potentially more impactful? If no changes can be made, do you still want to go ahead with the trial as an implementation pilot?

<p>Consider experimental threats & risks</p> <ul style="list-style-type: none"> • <i>What things can you do to make sure people receive the intervention as it was intended?</i> 	<p>Risk: Printing process messes with letter format / content</p> <ul style="list-style-type: none"> • Likelihood: Low • Impact: High 	<p>Risk:</p> <ul style="list-style-type: none"> • Likelihood: • Impact: • Mitigation Approach:
--	--	---

<ul style="list-style-type: none"> • <i>Is there a way you can double check that your randomization has been implemented correctly and the participants are receiving the treatment they were assigned?</i> • <i>Are there things you can do to make sure that staff is prepared and ready to implement the trial (e.g., training on data collection, a pilot period to adjust to new workflows, extra time to adjust to a new call volume)?</i> 	<ul style="list-style-type: none"> • Mitigation Approach: BIT will check with printers before the final printing is complete <p>Risk: Mailing does not go smoothly and reach addresses</p> <ul style="list-style-type: none"> • Likelihood: Low • Impact: High • Mitigation Approach: BIT will include the Newark’s Rent Control Office address and BIT’s Brooklyn office address in the mailing list to see if the mailing arrives <p>Risk: Newark isn’t ready for an influx of registrations, isn’t able to process them in time, and the outcome data is underreported</p> <ul style="list-style-type: none"> • Likelihood: Medium • Impact: Medium • Mitigation Approach: Newark’s Rent Control Office is fully aware of the trial launch so that they can prepare well in advance • Newark is submitting a RFA grant request to contract a data processing specialist who will handle the registration processing • BIT will check in with Newark during the trial to assess their workload and how they’re processing new registrations • BIT will create a secondary outcome of registrations within some longer time period to account for any outcome delays 	
<p>Consider ethical risks</p> <ul style="list-style-type: none"> • <i>How might participating in the trial harm people in your sample or others?</i> • <i>How will you monitor the trial to ensure you can detect the harm early and change your implementation if needed?</i> 	<ul style="list-style-type: none"> • We did not identify any risks in this trial. • We will monitor calls into 311 and potential landlord or tenant complaints to ensure there is no harm done. 	<p>Your answer:</p>
<p>Consider race & equity implications</p> <ul style="list-style-type: none"> • <i>If not already included above, how will you check to see if your intervention had differential effects by race? Disaggregating your data by race is often a good first step here, but consider descriptive analyses that might help explain any trends you are seeing. In cases where you do not have race demographics, can you use proxy variables (e.g. Census tract information matched to zip codes?)</i> 	<ul style="list-style-type: none"> • Note: We were not able to include any race & equity data within this trial given - but if you have access to this data we’d recommend you build it into your analysis. 	<p>Your answer:</p>

Plan for Data Analysis

Understand and Specify Your Variable for Analysis

There are two parts to planning for your analysis. First, you need to think through your different variables and make sure to document how you plan to use them for your trial analysis. This step helps you to ensure you have the data you need and sets you up for success for your analysis plan. Additionally, it's good to specify what you will do with bad data or data that doesn't match.

Here are some questions to guide your data checks:

- Where are you getting the data you need to complete the trial analysis (e.g., treatment assignment, outcome indicator data, other participant characteristics)?
- Are there any data security procedures that need to be followed?
- How will you assemble the data? How will your variables be constructed (e.g., units, interpretation of values, etc.)?
- How will you check your data for accuracy (e.g., to make sure that any data merges were done correctly, or that missing values have been identified and dealt with as needed)?

Specify Your Analysis Plan

Next, you need to specify your analysis plan. You should have a "hypothesis" you are testing - that your intervention will work -- but specifying ahead of time will allow you to say what statistical test & analysis you will use to determine if it worked. See below for an example of Newark's analysis plan. If you have questions on how to choose what test to run, check out [resources here](#).

Questions to guide your analysis plan:

- *For your Primary outcome:*
 - **What statistical test will you run for your outcome indicator?** For example, will you run a t-test to see if the mean outcome for your treatment group is statistically different than the mean outcome for your control group, a regression to control for other factors, or more complex analysis?
- *For your Secondary analysis:*
 - **Are there any other statistical tests you would like to run** (e.g., other outcome indicators, or looking at sub-groups)?
 - **Will you want to do any cost-benefit analysis?**

Sample Data

Variable name	Type	Source	Measurement
Standardized Owner Address	Unique identifier	Tax Assessor (and then BIT cleaning and SmartyStreets)	Unique Identifier in USPS address format (string variable)
Owner Address	Unique identifier	Tax Assessor	Identifier as recorded in the Tax Assessor database (string variable)
Property Address	Unique Identifier	Tax Assessor	Identifier as recorded in the Tax Assessor database (string variable)
Owner Name	Unique Identifier	Tax Assessor	Name of owner (string variable)
Treatment	Independent	BIT	Binary (0 if Control Group, 1 if Treatment Group)
Number of Properties Owned	Covariate	Tax Assessor (plus BIT)	Continuous number (>0)
Number of Units Owned	Covariate	Tax Assessor (plus BIT)	Categorical variable (some discrete numbers along with other categories (e.g. ">7" as described above)
Registered	Dependent	Newark Office of Rent Control (and then BIT cleaning and SmartyStreets)	Binary (0 if not registered, 1 if registered)
Registration Date	Dependent	Newark Office of Rent Control	Date-Time

Sample Analysis Plan

Our basic estimating equation is:

$$Y_i = \beta_0 + \beta_1 Treatment_i + \beta_i X_i + \varepsilon_i$$

Where i represents the property owner (identified by their unique mailing address) and Y_i represents our outcome of interest for owner – property registration within 30 days of the mailing. $Treatment$ represents an indicator variable equal to 1 if owner was randomly assigned to the treatment group and 0 otherwise. This is our main coefficient of interest and represents the average effect of receiving the intervention letter on property registration. It tests the hypothesis that the difference between the treatment group and the control group is statistically significantly different from zero (i.e. a two-tailed test). X_i represents a vector of covariates, including Number of Properties Owned and Number of Units Owned. We will use heteroskedastic-robust standard errors in our estimations.

We will estimate a primary version of this model where Y represents whether or not an owner registered at least one property by December 31, 2019.

Data Cleaning Notes:

For registered properties whose owner address does not match with an owner address listed in the pre-trial sample, we will clean their property address and perform the following checks to identify if the property address is listed in the pre-trial sample:

- 1) If a registered property's address is listed on the pre-trial sample, we will manually inspect if their owner name is identical.
 - a) If owner names match, we will assume that the owner address changed and assign a positive outcome for that owner.
 - b) If owner names do not match, we will assume ownership transfer and drop the previous owner from analysis as they are no longer responsible for property registration.
- 2) For registered properties whose property address and owner address are both not listed in the pre-trial sample data, we will exclude them from analysis as they were not included in the trial sample.

Your Approach: Please describe how you will analyze your data as well as what steps you will take to ensure the data is clean and ready for analysis.

Part 3 - Implementing your evaluation

Plan for Implementation

Create a project implementation plan and timeline

At this point, an evaluation starts to feel just like any other project -- with a few extra checks! You will need to carefully plan for implementation and ensure randomization is able to occur and data is able to be collected at the correct intervals. Be sure to build in time for getting any necessary approvals and for double-checking your work. We've included a sample timeline here - but we find that implementation works best when it follows a city's normal project management process.

Sample Implementation Plan

Task	Deadline
Finalize draft intervention letter	October 1
QA draft intervention letter	October 4
Receive Newark approval of intervention letter <ul style="list-style-type: none">Confirm that we can remove Cover sheet and Parking page from form	October 9
Finalize trial sample (mailing list)	October 1
Conduct randomization	October 1
Decide on a printing vendor	October 4
Complete the RfA grants request	October 11
Share Newark mailing list and trial mailing materials (Letter, FAQ, 3 page form) with the printing vendor and secure an agreement	November 19
Monitor mailing and printing as possible	November 19 - December 3
Mail letters to treatment group	December 5
Process registrations and enter registration information into a database	December 5 - February 2
Export and share data on all landlord registrations occurring between December 3 and December 31, 2019	February
Analyze data and report	March

Your Approach: Please fill in a chart of your implementation plan for this project - feel free to copy in from above!

Monitor your evaluation

While your evaluation is in the field, be sure to check in and ensure everything is happening to plan. This can take the form of automated data checks, regularly scheduled check-ins with your partners or observations of the process. It's better to catch anything before the end of your evaluation so you can adapt as needed!

Part 4 - Analyzing your results

Analyzing your data

You've already specified your analysis plan before you launched the trial - so now is the time to use it!

If possible, have someone who is not involved with the project review your work. It can help you eliminate any blind spots or highlight any assumptions you've made about the data.

Once you've analyzed your results you can ask yourself a few questions to help you reflect your results:

- Descriptive analysis:
 - What is the average and standard deviation for your full sample?
 - What is it for each of your treatment groups?
 - Are there any characteristics of the sample to describe (demographics, location, time, etc.)?
- Primary outcome:
 - Statistically significant: Yes/No
 - Effect size: What is the effect of your treatment, on average? (e.g., an increase of 2 percentage points or a decrease of \$100, on average)
 - Distribution: What is the range of outcomes – was the confidence interval quite large, or narrow?
 - Interpretation: What does this difference mean practically? Is it large enough to make a meaningful difference? If there was a wide range of outcomes, is it acceptable to implement something with that range of results?

Congrats on running your evaluation!

Now that you have your results, you can decide how you want to use them. If your result was positive, you may consider scaling your solution. To see how Newark scaled up their letter, head back to the Replication Guide.