Online Appendix

A Example of BIT trial report

THE BEHAVIORAL INSIGHTS TEAM

Reducing errors in business license renewal applications

A Trial Report from the Behavioral Insights Team

What was the context and goal?

The City of [redacted] worked with the Behavioral Insights Team to see if we could reduce the error rate for business license renewal applications. There are approximately 7,200 businesses and individuals that hold one or more [redacted] business licenses. Licenses are renewed annually.

The requirements for renewing a business license are complex, and the existing renewal notice does not effectively help applicants navigate that complexity. Approximately half of license renewal applications have one or more errors. When an error is spotted, city staff call the applicant to resolve the issue (which often include re-submission of paperwork) or even have to mail back the application, which re-starts the process. By reducing the error rate, the city will save time and resources, as will business licensees.

What did we test?

We designed a new license renewal notice aimed at better supporting applicants in navigating the renewal process. There were three primary changes:

1. We developed and included a comprehensive guide to what supporting documentation each licensee needed to provide. Based on their specific licenses up for renewal, the end of each notice include a picture of each required document, guidance on how to avoid common errors, and contact information for the relevant authority should the applicant have questions.
2. We re-organized the information in the notice so that the requirements for all licenses were grouped together. For example, we listed the required documents for all licenses to be renewed in one consolidated list. This makes the task appear simpler and reduces potential duplication of work effort. Similarly, we consolidated all payment requirements into one step.
3. We prompted applicants to set aside a specific date and time for completion of the renewal application.

We tested the re-designed notice to determine whether it reduces the error rate in license renewal applications.

Why did we think it might work?

Our new renewal notice was designed to respond to several key issues that we thought might play a significant role in the overall error rate. These issues were identified by program staff who manage the renewal process and were

supplemented by our prior experience with similar complex administrative tasks that governments require businesses to undertake.

1. Frontline staff indicated that the greatest source of error were issues with supporting documentation (e.g. certificates, affidavits or other licenses that licensees were required to submit to get their business license renewed). We believed that by providing licensees with a guide at the end of the notice, customized to their requirements, it would be easier for them to understand what supporting documentation they needed to provide.
2. To get their licenses renewed, businesses need to complete their renewal applications, provide the required documentation and make a payment of the correct amount. This is a complex administrative task that business owners or responsible employees may seek to avoid or put off. We thought that prompting licensees to set a specific date and time for completing the task would help avoid procrastination. Similar approaches to help people set "implementation intentions" have been effective in other contexts.
3. We thought that the structure of the existing notice, which listed the requirements for each notice separately, made the task seem more complex and burdensome than it was in reality. For example, if three licenses required a copy of the business owners drivers' license, the owner might think they need to provide three photocopies of that license, but one would be enough. By consolidating the requirements of all licenses to be renewed, we sought to reduce the perceived complexity of the task and reduce procrastination.

How did we test it?

We designed a two-armed randomized controlled trial (RCT), sending either the original or re-designed renewal notice to all business license holders who were scheduled to renew business licenses between January 2018 and December 2019. Randomization was clustered on the first letter of the business owner's name.

By randomly assigning some businesses to receive the new notice and others to receive the old notice, we could be confident that differences in error rate between these two groups would be the result of the new notice itself, rather than any other factors. For this reason, RCTs are often considered the "gold-standard" in evaluating the impact of new approaches.

What did we find?

The re-designed notices **reduced the error rate in license renewal applications from about 48% to about 44%, a 7% relative decrease.** The re-designed notices also required staff to mail back about 19% fewer applications (it was 9.3% for the old notice and 7.5% for the new notice). This suggests that in addition to reducing the

proportion of renewal applications that had any error, the new notice reduced the severity of the errors.

In further exploratory analysis, we found that the re-designed notice reduced documentation errors and renewal information errors, but had no effect on payment errors.

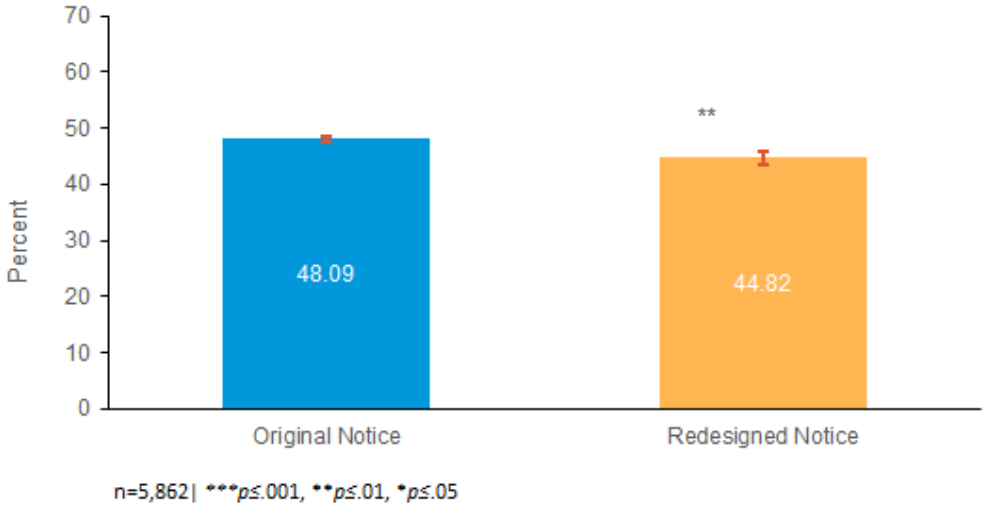
Recommendations

- We recommend that the city switch all licensees to the new format as soon as is convenient to reduce error and mail back rates.
- Even with the new notice, the error rate (44%) and mail back rate (7.5%) are still quite high. We believe that further changes to the renewal notice are unlikely to substantially bring this rate down. As a result, we recommend that the city consider changes to:
    - Policy, with a focus on simplifying renewal requirements;
    - Process, potentially reducing the frequency of renewals (which are currently annual); and
    - Systems, as moving to online renewals could potentially allow the city to pre-populate renewal applications, automatically validate payments or documentation, or take other actions to limit the likelihood and impact of errors.

Results

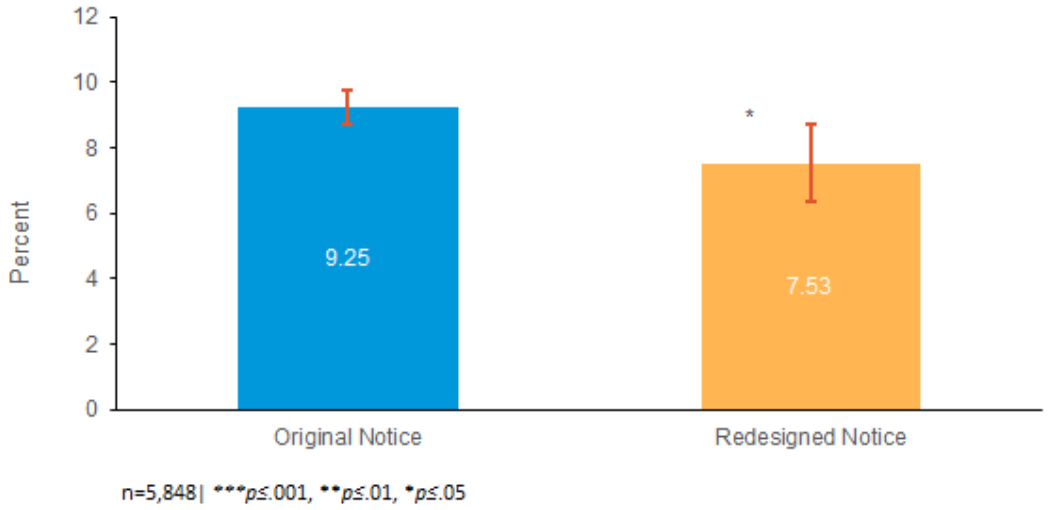
The re-designed notice decreased the likelihood that a renewal application would have a documentation, renewal information, or payment error by 3.27 percentage points (standard error = 1.19) for a relative decrease of 7 percent. This difference is statistically significant at a p-value of less than 0.01, meaning it is very unlikely to have occurred by chance.

Figure 1: Error rates in license renewal applications



The re-designed notice also decreased the likelihood that a renewal application would be mailed back to the applicant by 1.72 percentage points (standard error = 0.78) for a relative decrease of 19 percent. This difference is statistically significant at a p-value of less than 0.05, meaning it is very unlikely to have occurred by chance.

Figure 2: Mailback rates in license renewal applications



B Marginal Cases of Non-Adoption

As mentioned in the text, we defined a city as adopting a trial if the city has used one of the communications in the nudge treatment arms again following the RCT. This includes cases when the city had used the communication after the trial but was not currently doing so, for example, because it was not an election year. When cities have made further changes to the communication since the trial, we counted adoption as incorporating at least 50% of the nudge features as pre-specified in the internal trial protocol or report.
Most cases of (non-)adoption were clear according to this rule, but there were 4 cases of non-adoption for which the post-trial communication seemed to include some nudge features, but did not meet our criteria upon close inspection. We describe each of these marginal non-adoption cases below.

1. This city sent new postcards encouraging local business owners to renew their license online. The control arm used a slogan on convenience, whereas the treatment arm used one with normative language. Both postcards were equally effective. The city no longer sends the postcards, but uses the same exact slogan from the control arm in different letters sent to businesses about their licenses.

2. This city police department sent recruitment postcards to local neighborhoods. The version with a message emphasizing the benefits and salary had the strongest effect. Now on its website, the police department has adopted this type of messaging for recruitment.

3. This city police department used online ads to recruit applicants from Historically Black Colleges and Universities (HBCUs). The control ads, based on a prior pilot, highlighted the relatable background of current police officers, and the treatment ads also offered a "personal concierge" service to guide applicants through the process. The control ads were significantly more effective. The police department still uses online ads to target applicants from HBCUs, but does not use the treatment messaging.

4. In both the control and treatment arms of this trial, the city added a new checkbox to the water utility bill for easy enrollment into a local charity program. The utility bill in the treatment arm also included colorful ASCII art and a message requesting recipients to sign up for the program. There was not a significant difference between the control and treatment arms. The city continues to send the utility bill with only the checkbox from the control arm.

Furthermore, in 11 out of the 20 cases of adoption, the city contacts verbally provided a description of the communication they had used after the trial that matched the

treatment arm, but they could not send us a template of the exact communication for us to independently verify due to bureaucratic or technical issues (e.g., they were no longer using the same email system from which the newsletter had been sent before). As a robustness check in Table A.4, we drop the marginal non-adoption cases and/or the verbal-only adoption cases. We replicate the main specification (Column 4) from Table 1. The key finding, namely on the importance of pre-existence, remains large and significant.

C Broad Cases of Adoption

As introduced in Section 2.3, we also code cases of broad adoption in which the city contacts stated that either the trial with BIT inspired another communication in a separate context or the city adopted the process of experimentation for their own internal trials. A brief description of each case is listed below with the number of related trials counting as broad adoption shown in parentheses.

- In this trial, a police department sent postcards encouraging applications from minority groups. From the trial, the city identified successful language that they added to subsequent phone and email recruitment scripts. (1 trial)
- A city police department used implementation intentions in an email trial targeting inactive applicants. The department did not continue the prompts for implementation intentions, but incorporated emailing inactive applicants in their long-term recruitment process. (1 trial)
- In this trial, the city sent text reminders for show-cause hearings. The department no longer sends these text reminders, but now sends similarly worded texts for citations, a step prior to the show-cause hearings. (1 trial)
- This city ran an email trial to recruit police applicants. After the trial, the city conducted three internal RCTs in other contexts. (1 trial)
- A city used a nascent text messaging system in two trials to remind citizens under a Medicaid waiver program to use their free health check-up. Motivated by these trials, the city began to use text reminders for a variety of purposes. (2 trials)
- This city ran three trials with BIT for charitable giving, police recruitment, and paperless utility billing. These collaborations inspired the city to create its own internal team to experiment with nudge interventions in city communications. (3 trials)

- In this trial, the city sent postcards to encourage applications to the police force. The Tax and License Division adopted nudges and the process of experimentation in their communications. (1 trial)
- This city conducted two trials with BIT for donations to local charities and voter registration. After the trials, the city established an internal nudge team that has run at least four trials, for example, on library fine payment and water conservation. (2 trial)
- This city police department sent postcards for police recruitment in a trial. The postcards were discontinued, but the findings informed other recruiting materials such as bus advertisements and language on the website. (1 trial)
- This city police department ran three trials with BIT to improve recruiting practices by implementing social media advertisements as well as email and text reminders. These trials led to expanded communication efforts for police recruitment through these mediums. (3 trials)

D Forecasting Survey

This section details the 10-minute forecasting administered through the Social Sciences Prediction Platform[8]. In total, 118 forecasters submitted their predictions on the platform over 25 days. The survey first summarized the setting and main result of DellaVigna and Linos (2022), and then introduced the focus for the current paper on the adoption rate of the nudge interventions *after* the RCT collaborations with the cities and on the determinants of adoption. The survey described the sample of trials and highlighted that each trial was co-designed by BIT and the partnering city and that the results were shared with the city in a report after the trial. Next, the survey showed two randomly selected examples of communications used in trials with a brief description of the policy area and targeted outcome.

The forecasters then made their first prediction on the baseline adoption rate. Specifically, we asked, "*What percent of the 73 trials do you think have been adopted by the cities?*" The forecasters provided their answer in percentages (from 0 to 100). We defined adoption as: "*We count a city as "adopting" a trial if one of the nudge treatment arms has been used in city communications after the trial with BIT.*" We gave an example of an adopted trial, showing the nudge communication used in the trial next to the comparable current communication in use by the city. For reference, we provided two statistics: 78% of the trials had at least one nudge intervention arm that led to an

[8] `https://socialscienceprediction.org/`

improvement relative to the control group, and 45% of the trials found a nudge that led to a significant improvement with $p < 0.05$. On the same page, we asked forecasters to write a short list or a couple sentences in an open-ended text box on which determinants of adoption they expect to matter most.

We then introduced the determinants of adoption that we consider: statistical significance, effect size, retention of the original city staff collaborator, state capacity (proxied by city population), What Works Cities certification, pre-existence of the communication in the trial, and behavioral mechanism used in the nudge intervention. (At this point, forecasters could not return to the previous page to change their baseline prediction nor their open-ended responses.) Next, the survey asked for the predicted adoption rate for each of these determinants separately page-by-page. The survey randomized the order of the determinants between two different orderings.

For each determinant, the sample of trials was separated into relevant bins with the number of trials in each bin shown, and forecasters predicted the adopted rate within each bin. For example, for statistical significance, we asked what percent the forecasters think have been adopted for trials that found: (i) a statistically significant improvement (i.e., $t \geq 1.96$, covers 45% of all trials, $n = 33$), (ii) a statistically insignificant improvement (i.e., $0 < t < 1.96$, covers 33% of all trials, $n = 24$), and (iii) a zero or negative effect (covers 22% of all trials, $n = 16$).

On every page, we reminded forecasters of their predicted baseline adoption rate from the very first question. For comparison, we displayed the weighted average adoption rate implied by their forecasts for the determinant on the page as a soft "nudge" to help them give forecasts that were consistent with their initial predicted baseline rate. For example, if they predicted that the adoption rates were 50% for statistically significant trials, 30% for statistically *in*significant trials, and 10% for zero or negative trials, then the weighted average we calculated for them would be $(50\% \times 0.45) + (30\% \times 0.33) + (10\% \times 0.22) = 34.6\%$.

Lastly, we asked forecasters to compare our sample of RCTs in U.S. municipal cities with similar representative samples of trials conducted by large multinational firms and by governments of low-income countries. We asked forecasters to rank these three samples by the overall adoption rate and by the responsiveness to evidence in adoption.

E Categorizing trials combining of new and pre-existing communications

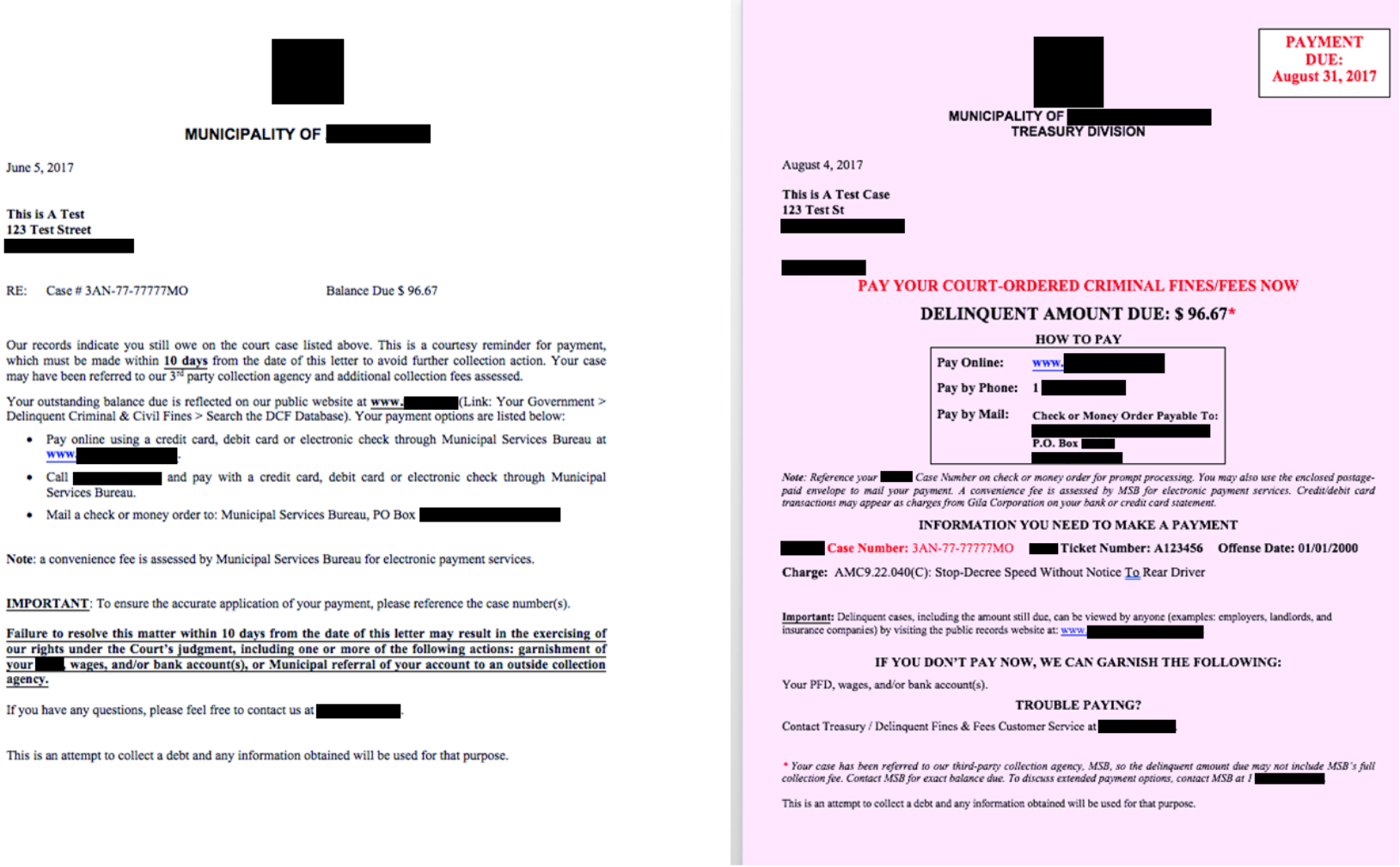
In 6 trials, a new insert or letter was sent in addition to a pre-existing mailer. For example, a new postcard insert was added in the same envelope for a pre-existing routine utility bill. In these cases, we focus on the adoption of the new insert or letter and count them as new communications, since the pre-existing component (i.e., the utility bill in the example) remained unchanged and the nudge was entirely contained in the new inserts. In 2 trials, pre-existing email and letter notices were modified to include

behavioral mechanisms and new text reminders were also used. We count these two cases as pre-existing communications, as the nudge intervention involved changes to the pre-existing email and letter. In one of these trials, the city adopted both the modified nudge email and the new text reminder, and in the other trial, the city adopted neither the nudge version of the letter nor the text reminder.

Figure A.1: Example of adoption of BIT-NA trial

(a) Status-quo control arm communication

(b) Nudge treatment arm communication



[redacted]

MUNICIPALITY OF [redacted]

June 5, 2017

This is A Test
123 Test Street
[redacted]

RE: Case # 3AN-77-77777MO Balance Due $ 96.67

Our records indicate you still owe on the court case listed above. This is a courtesy reminder for payment, which must be made within **10 days** from the date of this letter to avoid further collection action. Your case may have been referred to our 3rd party collection agency and additional collection fees assessed.

Your outstanding balance due is reflected on our public website at **www.**[redacted](Link: Your Government > Delinquent Criminal & Civil Fines > Search the DCF Database). Your payment options are listed below:

- Pay online using a credit card, debit card or electronic check through Municipal Services Bureau at www.[redacted].
- Call [redacted] and pay with a credit card, debit card or electronic check through Municipal Services Bureau.
- Mail a check or money order to: Municipal Services Bureau, PO Box [redacted]

Note: a convenience fee is assessed by Municipal Services Bureau for electronic payment services.

IMPORTANT: To ensure the accurate application of your payment, please reference the case number(s).

Failure to resolve this matter within 10 days from the date of this letter may result in the exercising of our rights under the Court's judgment, including one or more of the following actions: garnishment of your [redacted] wages, and/or bank account(s), or Municipal referral of your account to an outside collection agency.

If you have any questions, please feel free to contact us at [redacted].

This is an attempt to collect a debt and any information obtained will be used for that purpose.

[redacted]

PAYMENT DUE: August 31, 2017

MUNICIPALITY OF [redacted]
TREASURY DIVISION

August 4, 2017

This is A Test Case
123 Test St
[redacted]

[redacted]

PAY YOUR COURT-ORDERED CRIMINAL FINES/FEES NOW

DELINQUENT AMOUNT DUE: $ 96.67*

HOW TO PAY

Pay Online:	www.[redacted]
Pay by Phone:	1[redacted]
Pay by Mail:	**Check or Money Order Payable To:** [redacted] **P.O. Box** [redacted]

Note: Reference your [redacted] Case Number on check or money order for prompt processing. You may also use the enclosed postage-paid envelope to mail your payment. A convenience fee is assessed by MSB for electronic payment services. Credit/debit card transactions may appear as charges from Gila Corporation on your bank or credit card statement.

INFORMATION YOU NEED TO MAKE A PAYMENT

[redacted] **Case Number:** 3AN-77-77777MO [redacted] **Ticket Number: A123456** **Offense Date: 01/01/2000**

Charge: AMC9.22.040(C): Stop-Decree Speed Without Notice To Rear Driver

Important: Delinquent cases, including the amount still due, can be viewed by anyone (examples: employers, landlords, and insurance companies) by visiting the public records website at: www.[redacted]

IF YOU DON'T PAY NOW, WE CAN GARNISH THE FOLLOWING:

Your PFD, wages, and/or bank account(s).

TROUBLE PAYING?

Contact Treasury / Delinquent Fines & Fees Customer Service at [redacted]

* *Your case has been referred to our third-party collection agency, MSB, so the delinquent amount due may not include MSB's full collection fee. Contact MSB for exact balance due. To discuss extended payment options, contact MSB at 1*[redacted]

This is an attempt to collect a debt and any information obtained will be used for that purpose.

Figure A.1: Example of adoption of BIT-NA trial

(c) Current communication

PAYMENT DUE: COLUMN A

[redacted]

MUNICIPALITY OF [redacted]
TREASURY DIVISION
FINAL DEMAND

COLUMN B

COLUMN C COLUMN D COLUMN E COLUMN F
COLUMN G
COLUMN H, COLUMN I COLUMN J

PAY YOUR COURT-ORDERED TRAFFIC FINES/FEES NOW

DELINQUENT AMOUNT DUE: $ COLUMN K*

HOW TO PAY

Online:	**www.[redacted].com – Local Payments, Jurisdiction [redacted], Type: DCF Payments**
Phone:	**1 [redacted] (Press 3, then Jurisdiction [redacted])**
Mail:	**Check or Money Order Payable To:** **Municipality of [redacted]** **P.O. Box [redacted]** **[redacted]**
In Person:	**City Hall, [redacted]** **[redacted]**

***Note**: Reference your [redacted] Case Number on check or money order for prompt processing. A convenience fee of 2.55% is assessed by ACI for electronic payment services. Credit/debit card transactions may appear as charges from ACI Payments Inc. on your bank or credit card statement.*

INFORMATION YOU NEED TO MAKE A PAYMENT

[redacted] Case Number: COLUMN L **[redacted] Ticket Number: M** **Offense Date: N**

Charge: O

Important: Delinquent cases, including the amount still due, can be viewed by anyone (examples: employers, landlords, and insurance companies) by visiting the public records website at: http://www.[redacted].

IF YOU DON'T PAY NOW, WE CAN GARNISH THE FOLLOWING:

Your [redacted], your wages, and/or your bank account(s).

IF YOU HAVE ANY QUESTIONS

Contact Treasury / Delinquent Fines & Fees Customer Service at [redacted].

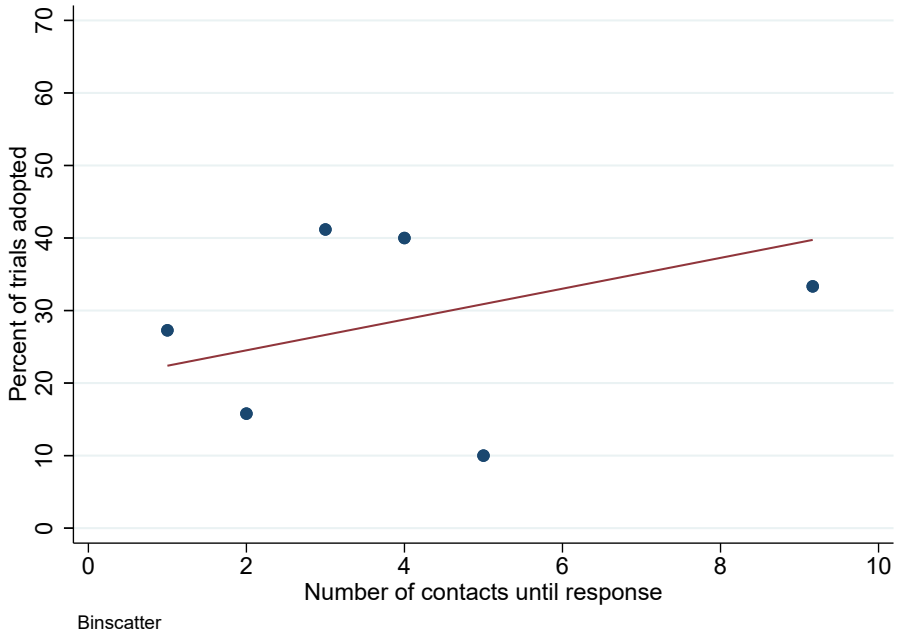
* *If you do not fully pay the delinquent amount due by the payment due date above, the [redacted] may refer your case to our third-party collection agency, Professional Credit, who will add a collection fee of 35.14% to your balance. Your new amount due would then become $P.*

To discuss extended payment options after your case has been referred, contact Professional Credit at 1 [redacted].

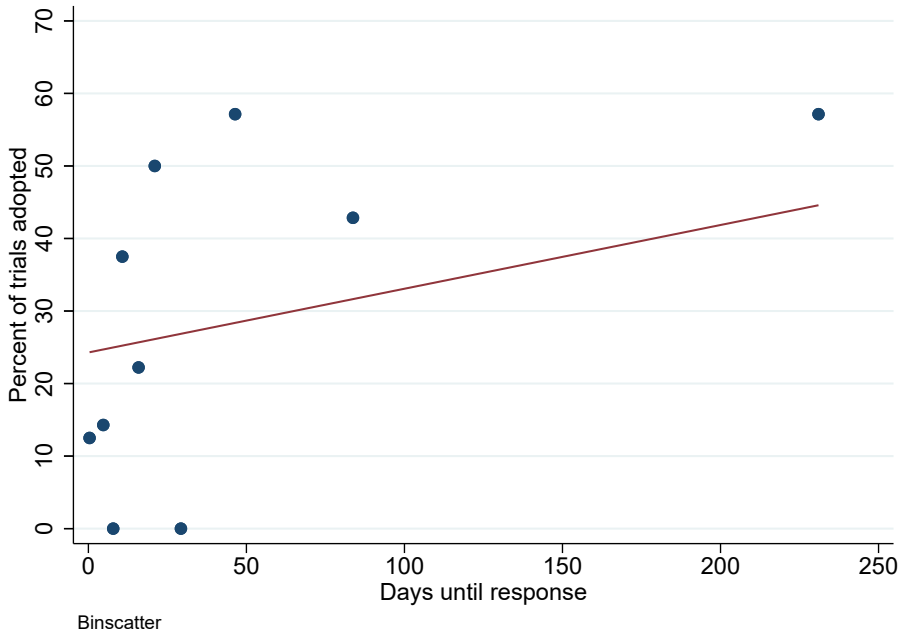
This is an attempt to collect a debt and any information obtained will be used for that purpose.

Figure A.2: Adoption by response times (bin scatters)

(a) Number of times contacted until final response

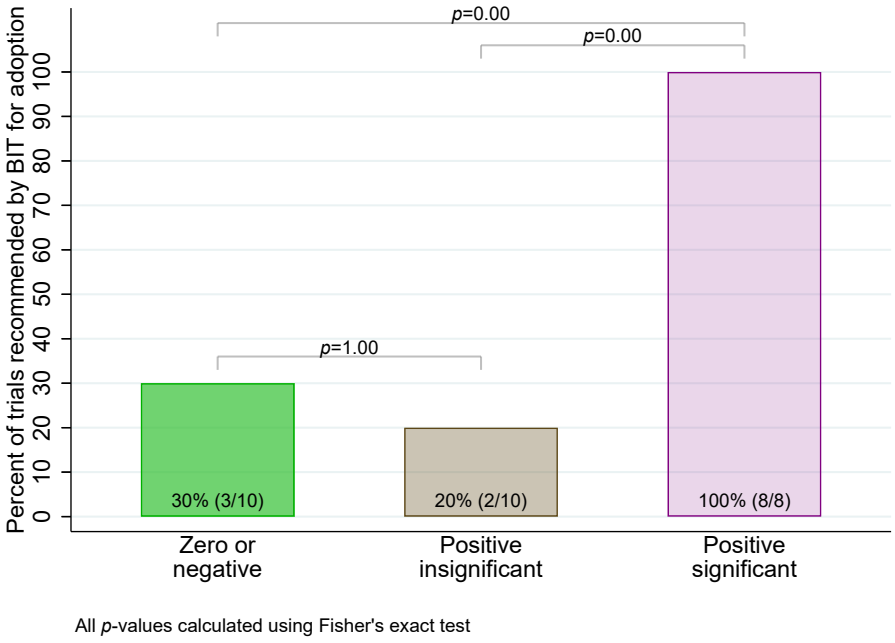


(b) Number of days from first request to final response



Figures A.2a and A.2b show binscatters relating the adoption rate to (a) the number of exchanges with the city contact (e.g., emails and phone calls) and (b) the days from first contact to final response, until all the information on the trial was collected.

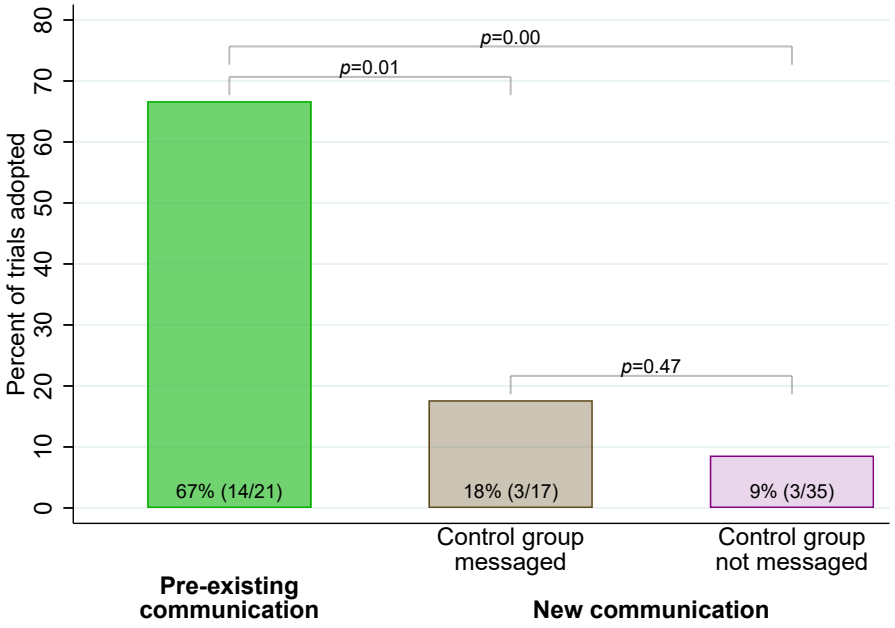
Figure A.3: BIT recommendations for adoption by statistical significance



This figure shows the percent of trials that BIT recommended for adoption within three groups separated by the sign and significance of the best-performing nudge treatment effect. BIT began documenting recommendations in their trial reports in mid-2017. 28 trials in the sample have these recommendations.

Figure A.4: Adoption of nudges by pre-existence: Additional results

(a) By control group communication



(b) By effect size (three bins)

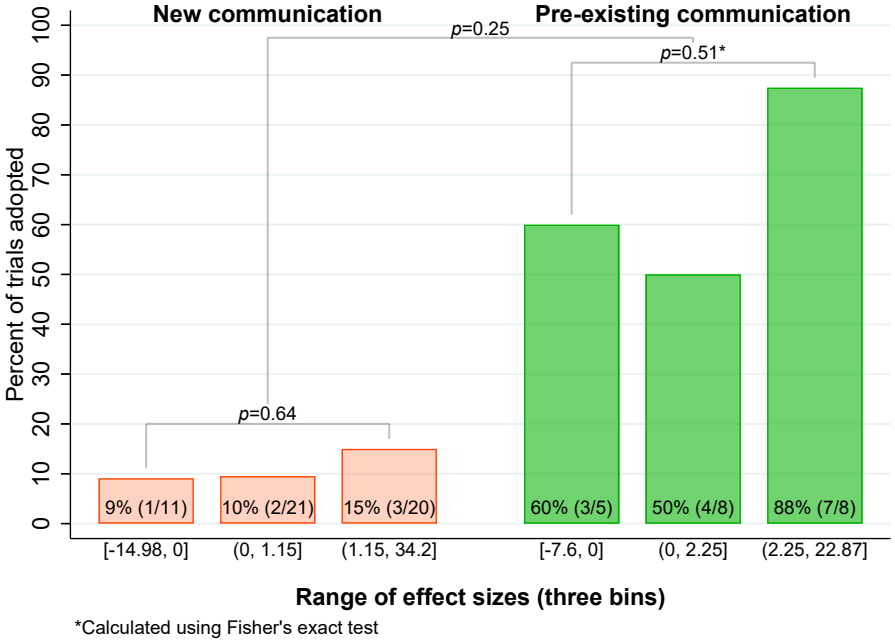
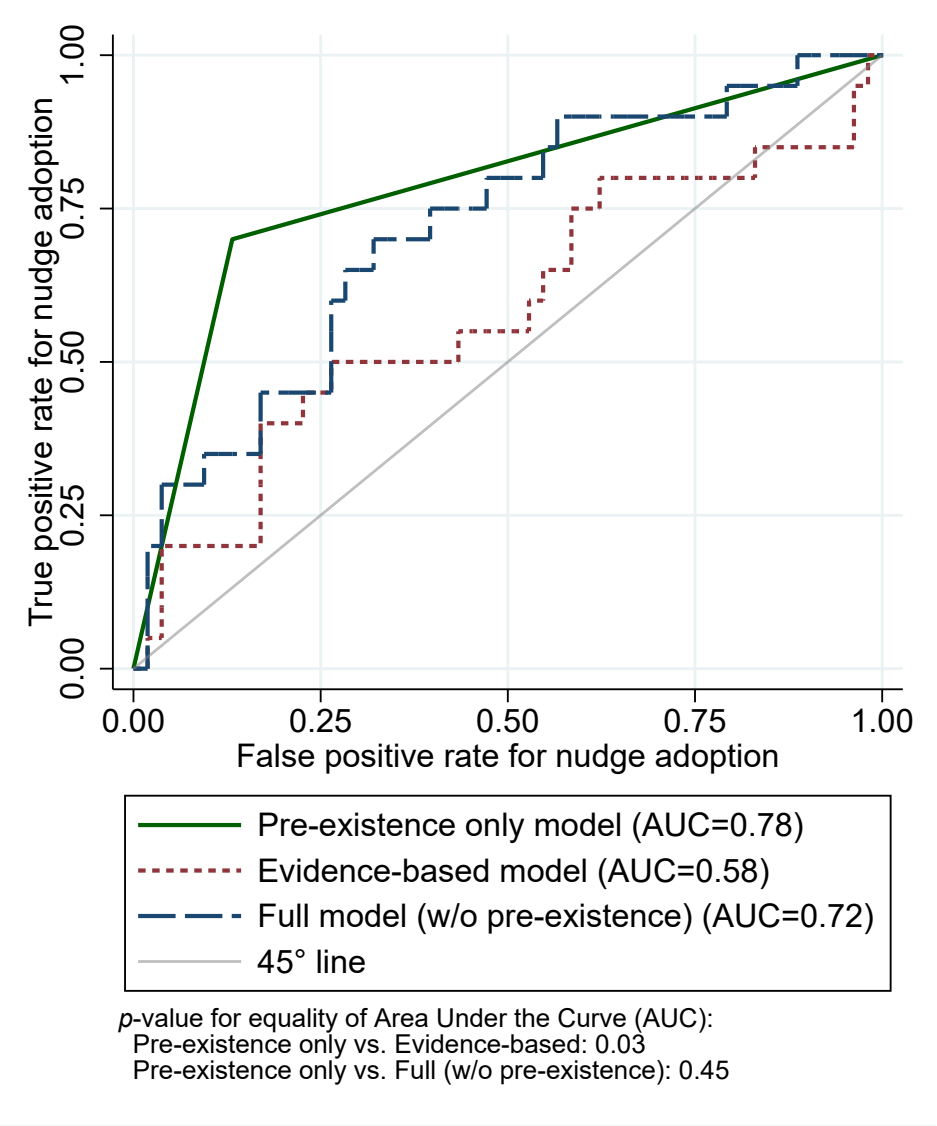


Figure A.4a shows the adoption rate for trials on new communications conditional on whether the control group received any communication in the trial. Figure A.4b bins together trials that found a negative or zero effect, and those that are below or above the median among the trials with a positive effect. It compares the adoption rates within each bin for new and pre-existing communication trials separately.

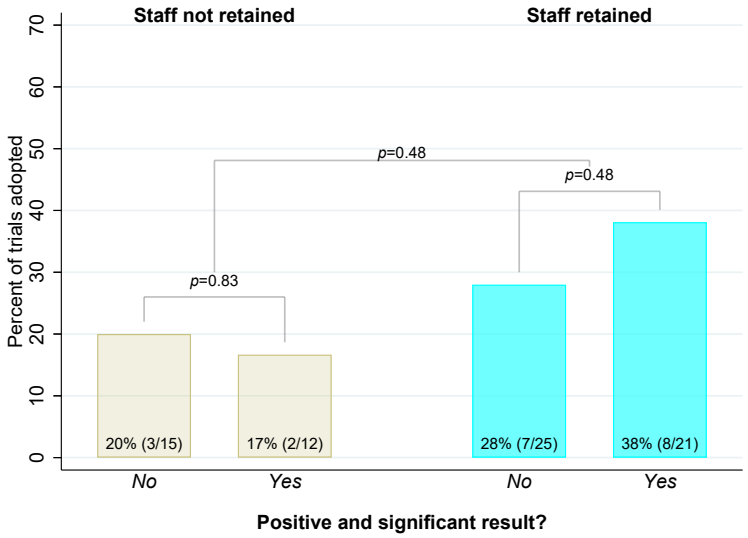
Figure A.5: Receiver operating characteristic (ROC) curves



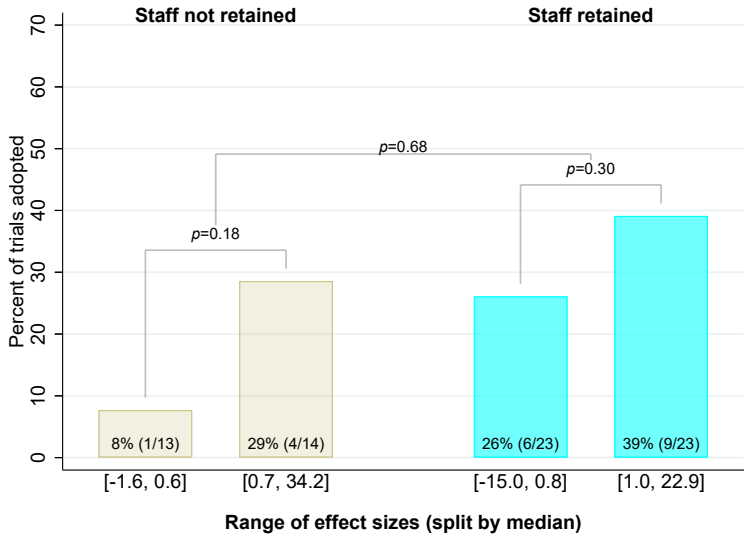
This figure shows the receiver operating characteristic (ROC) curves and the computes the area under the curve (AUC) for three models estimated by logistic regression. Each model includes a constant and different sets of explanatory variables.
The *Pre-existence only* model includes an indicator for whether the communication in the trial was pre-existing.
The *Evidence-based* model includes an indicator for the whether the most effective treatment arm in the trial was positive and significant (i.e. max $t \geq 1.96$) as well as the percentage-point treatment effect.
The *Full without pre-existence* model includes all the controls in Column 4 of Table 2 except for the pre-existing communication indicator.

Figure A.6: Staff retention interaction effects

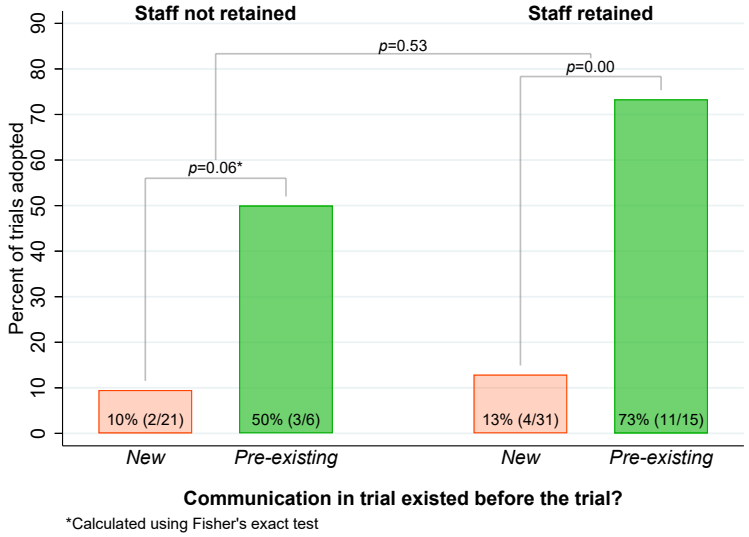
(a) By staff retention and statistical significance



(b) By staff retention and effect size



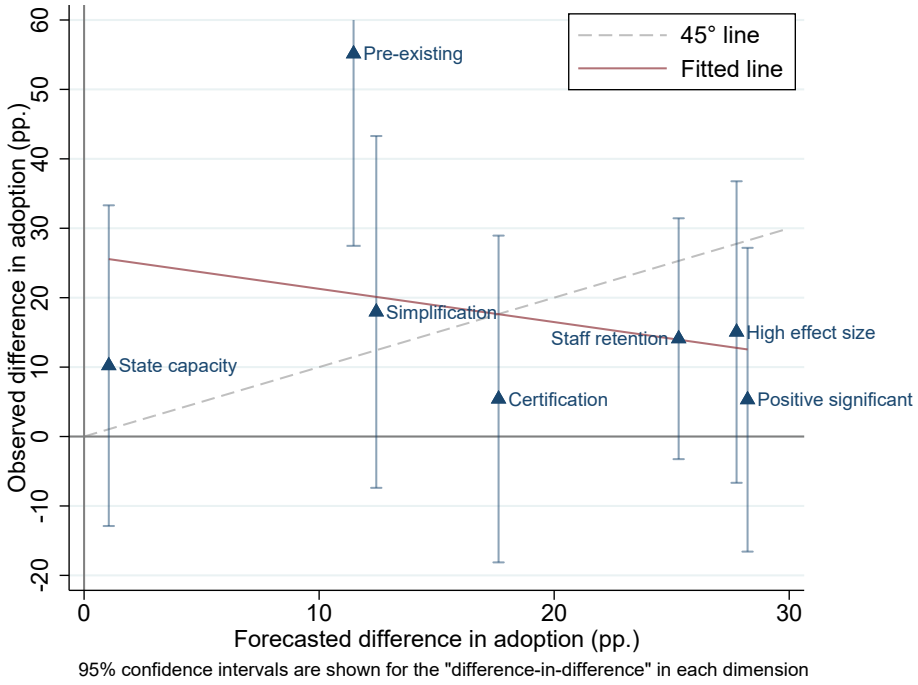
(c) By staff retention and pre-existence



Figures A.6a and A.6b show the adoption rates conditional on finding an effect that is (a) positive and significant and (b) above or below the median effect size, separately for trials where the original city collaborator has left or has been retained by the city department. Figure A.6c compares the adoption rate of interventions in new (orange) versus pre-existing (green) communications separately for cases when the original city collaborator has left or has been retained.

Figure A.7: Comparisons of forecasts and observed adoption

(a) By dimension



(b) Ranking of adoption compared to firms and governments of low-income countries

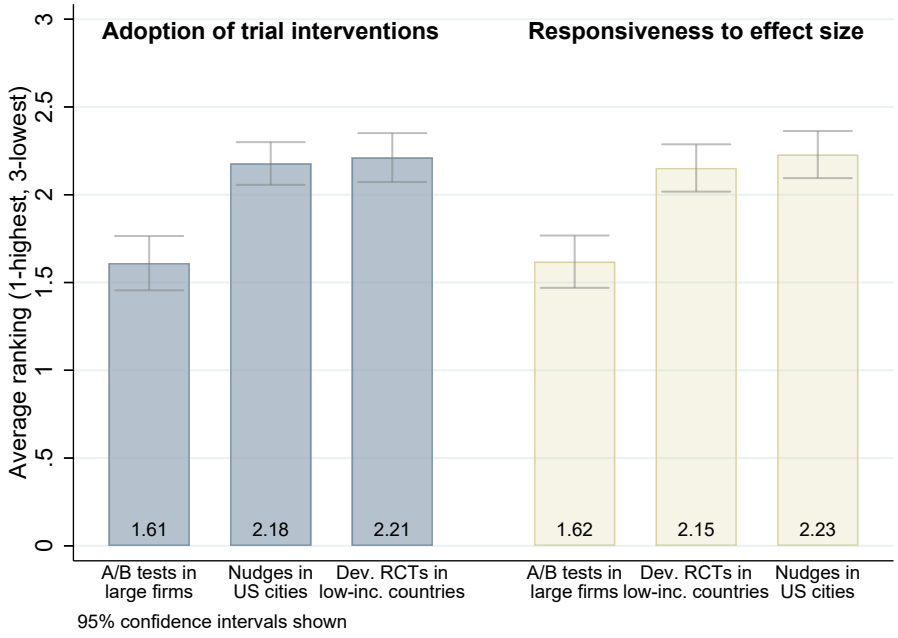


Figure A.7a evaluates the accuracy of the expert forecasts across determinants and plots the predicted impact of each determinant (horizontal axis) against the observed impact (vertical axis). Figure A.7b shows the average of the forecasters' ranking of adoption rates for three samples: (1) nudges in U.S. cities (this paper), and similar hypothetical representative samples of trials conducted by (2) large multinational firms and (3) governments of low-income countries. The left side shows the average rankings for the overall adoption rate (where 1 corresponds to the highest adoption and 3 to the lowest), and the right side shows the ranking for each sample's responsiveness to evidence in adoption.

Table A.1: Average nudge treatment effects

	Nudge Units*	Updated BIT-NA
	(1)	(2)
Average treatment effect (pp.)	1.390	1.906
	(0.304)	(0.587)
Nudges	241	116
Trials	126	73
Observations	23,556,095	1,800,382
Average control group take-up (%)	17.33	15.07
Distribution of treatment effects		
25th percentile	0.06	0.01
50th percentile	0.50	0.40
75th percentile	1.40	1.72

This table shows the average treatment effect of nudges. Standard errors clustered by trial are shown in parentheses. pp. refers to percentage point.
*Column 1 replicates Column 2 of Table III in DellaVigna and Linos (2022).

Table A.2: Forecasts summary

	Observed	Forecasts (Mean % [SD])			
Category	(%)	(1) Overall ($N = 118$)	(2) Nudge unit staff ($N = 19$)	(3) Reseachers ($N = 67$)	(4) Government workers ($N = 14$)
Baseline adoption rate	27.40	32.47 [19.06]	37.16 [20.64]	31.91 [19.16]	32.00 [20.40]
By sign and significance:					
Positive & significant	30.30	46.58 [23.66]	48.00 [23.91]	46.49 [23.13]	44.29 [28.01]
Positive & insignificant	25.00	23.22 [20.27]	34.84 [23.51]	21.31 [18.59]	23.29 [22.61]
Zero or negative	25.00	11.06 [16.21]	17.47 [19.25]	10.43 [16.35]	12.93 [17.40]
By effect size:					
High third	37.50	49.31 [24.81]	54.26 [24.80]	48.85 [24.15]	45.43 [30.89]
Middle third	28.00	31.61 [19.60]	40.32 [23.51]	29.57 [16.95]	32.29 [25.08]
Low third	16.67	12.94 [15.42]	18.16 [18.39]	12.60 [15.80]	11.57 [11.88]
By staff retention:					
With original staff	32.61	43.75 [22.64]	48.26 [26.59]	42.13 [20.74]	44.07 [28.02]
Without original staff	18.52	18.45 [16.31]	24.32 [15.14]	17.51 [16.33]	19.71 [18.59]
By state capacity (proxied by 2020 city population size):					
Above median	31.82	33.26 [19.62]	40.16 [22.93]	32.07 [17.76]	31.79 [24.12]
Below median	20.69	32.21 [18.67]	35.84 [19.25]	32.15 [18.89]	29.50 [20.26]
By What Works Cities certification:					
Certified	29.55	41.69 [21.23]	45.26 [22.86]	40.06 [20.37]	42.93 [24.94]
Not certified	24.14	24.06 [17.90]	32.42 [20.61]	22.58 [17.21]	22.14 [17.16]
By pre-existing or new communication:					
New	11.54	29.79 [20.42]	36.32 [22.87]	29.48 [19.48]	25.50 [19.00]
Pre-existing	66.67	41.25 [25.43]	45.63 [26.68]	37.78 [23.11]	47.71 [28.97]
By behavioral mechanism:					
Simplification	33.33	42.42 [22.06]	51.16 [24.34]	40.21 [20.56]	41.00 [25.45]
Personal motivaton	19.05	30.14 [19.30]	35.84 [20.55]	28.37 [18.35]	26.79 [19.42]
Social cues	24.39	29.83 [20.97]	34.74 [23.05]	28.81 [19.58]	30.36 [22.01]

Table A.3: BIT recommendations for trial adoption

Dep. Var.:	BIT recommended adopt	Trial adopted
(OLS)	(1)	(2)
Max $t \geq 1.96$	0.71	
	(0.12)	
Max treatment effect (10pp.)	0.05	
	(0.06)	
BIT recommended for adoption		0.25
		(0.17)
BIT did not recommend for adoption		-0.16
		(0.11)
Communication pre-existed		0.45
		(0.14)
Constant	0.25	0.13
	(0.10)	(0.05)
Average rate	0.46	0.18
Trials with recommendations only	✓	
Number of trials	28	73
Number of cities	16	30
R^2	0.46	0.38

Standard errors clustered by city are shown in parentheses. BIT has included recommendations for or against adoption in their trial reports since mid-2017. In Column 3, the omitted group are the earlier trials without BIT recommendations in the trial reports.

Table A.4: Determinants of nudge adoption (robustness)

	Baseline	Robust SEs	Robustness to marginal cases		
Dep. Var.: Nudge adopted (0/1, OLS)	(1)	(2)	(3)	(4)	(5)
Max $t \geq 1.96$	-0.03	-0.03	-0.06	0.06	0.03
	(0.08)	(0.10)	(0.09)	(0.10)	(0.10)
Max treatment effect (10pp.)	0.10	0.10	0.11	0.03	0.03
	(0.08)	(0.07)	(0.08)	(0.06)	(0.06)
City staff retained	0.07	0.07	0.04	-0.00	-0.04
	(0.08)	(0.09)	(0.09)	(0.07)	(0.08)
Above-median city population	0.08	0.08	0.06	0.13	0.11
	(0.09)	(0.09)	(0.11)	(0.08)	(0.08)
What Works Cities certified	0.12	0.12	0.17	0.04	0.09
	(0.11)	(0.11)	(0.13)	(0.12)	(0.12)
Communication pre-existed	0.52	0.52	0.56	0.40	0.45
	(0.13)	(0.12)	(0.13)	(0.17)	(0.17)
Mechanism					
Simplification & information	0.03	0.03	0.05	0.10	0.11
	(0.10)	(0.09)	(0.09)	(0.06)	(0.06)
Personal motivation	-0.12	-0.12	-0.16	-0.13	-0.16
	(0.12)	(0.10)	(0.11)	(0.07)	(0.07)
Social cues	-0.07	-0.07	-0.09	-0.08	-0.10
	(0.08)	(0.10)	(0.09)	(0.09)	(0.09)
Constant	0.04	0.04	0.08	-0.01	0.03
	(0.16)	(0.15)	(0.17)	(0.12)	(0.11)
Average adoption rate	0.27	0.27	0.29	0.15	0.16
Dropping marginal non-adopts			✓		✓
Dropping verbal-only adopts				✓	✓
Number of trials	73	73	69	62	58
Number of cities	30	30	29	29	27
R^2	0.38	0.38	0.41	0.39	0.44

Standard errors (shown in parentheses) are clustered at the city level except in Column 2, which provides heteroskedastic-robust standard errors. "Baseline" replicates Column 4 of Table 2. See Online Appendix Section B for details on marginal non-adoption and verbal-only adoption cases.