Training Policymakers in Econometrics II:

Tax Officers and Fiscal State Capacity

By Sultan Mehmood, Shaheen Naseer, and Daniel L. Chen[1]

May 2024

A field experiment with tax collectors in Pakistan aimed to assess whether econometrics education could influence tax policy and societal revenue generation. This experiment utilizes administrative tax data and reveals that econometrics education impacts tax policy, such as sending tax reminder letters, and increasing tax collection. The results from Mehmood, Naseer, and Chen (2023) and this separate field experiment provides evidence that econometrics education shapes beliefs, policy choices, and real-world outcomes. This impact spans from fostering a greater recognition of policies with causal evidence to enhancing the effectiveness of tax collection efforts. (JEL D72, D78, O17)

Keywords: randomized clinical trials, policy, credibility revolution, paradigm shift

[1] New Economic School (E-mail: smehmood@nes.ru). Lahore School of Economics (E-mail: ranashaheen133@gmail.com) and Toulouse School of Economics (Email: daniel.chen@iast.fr). An IRB is obtained from the Lahore School of Economics with IRB Number RERC-022021-01. We would like to thank the training Director and staff at the Civil Service Academy of Pakistan and National School of Public Policy. Sajwaar Khalid, Bakhtawar Ali, Hafeez-ul-Hassan and Sameen Tariq provided excellent research assistance. Daniel Chen acknowledges IAST funding from the French National Research Agency (ANR) under the Investments for the Future (Investissements d'Avenir) program, grant ANR-17-EUR-0010. This research has also benefited from financial support of the research foundation TSE-Partnership. Pre-registered at AEA RCT Registry with ID AEARCTR-0006655.

1. Introduction

Many commentators have recognized civil service as a crucial engine of state capacity and public welfare (e.g., Weber, 1922). Civil servants implement policies and provide services in hopes of achieving efficient and effective governance, introducing innovative approaches to public service, and improving governmental processes (e.g., Lipsky, 1980). Accordingly, governments have invested heavily in training programs to increase their 'stock' of skilled bureaucrats (World Bank, 2017). They spend significant amounts annually on civil servant training (Credibility Engine, 2021). Unlocking the 'production function' of high-quality civil servants is pivotal to addressing numerous governance challenges around the funding and provision of public goods (Besley and Persson 2009). Public agencies, accounting for a substantial portion of employment worldwide, are important contributors to policy implementation and societal development. They represent a significant share of the workforce and account for a major part of public expenditure in emerging economies (Lindauer, 1988; Finan et al., 2015). Government Workers in Developing Countries.. Civil service forms the backbone of effective governance, representing a major sector of employment and playing a critical role in policy implementation, many of whom are frontline workers and administrators. The stakes to produce high-quality civil servants are even higher considering the overwhelming majority of government employees are career bureaucrats who lack alternative career paths in public administration (Perry and Hondeghem, 2008). To assess the extent to which successful civil servants (Checchi et al., 2021; Khan et al., 2019; Muralidharan and Sundararaman, 2011; Bloom et al., 2015b) can be 'made' or trained, we designed and implemented a scaleable intervention on problem-solving and numeracy (Deming, 2021), namely econometrics training of causal inference.

The nexus between bureaucratic training and the efficacy of public service delivery, and by extension, state capacity, remains a topic of considerable debate and inquiry. A burgeoning corpus of research underscores the pivotal role of bureaucrats. However, the focus of these studies predominantly lies in the optimal assignment of bureaucrats rather than their training. Our paper shifts this perspective, honing in on the training of front-line civil servants who are the linchpins of policy implementation, as highlighted in studies like Banerjee et al. (2021). Central to our exploration is a compelling question: Can tailored training, specifically designed to

address cognitive biases through econometrics, effectively bolster fiscal state capacity? This inquiry not only studies how to improve bureaucratic efficiency but also probes the potential of such training to reshape and enhance the very framework of public service delivery.

In the past fifty years, empirical economics has experienced a significant transformation (Angrist and Pischke, 2014). This change, known as the credibility revolution, emphasizes causality and empirical accuracy, aiming to enhance the reliability of econometric methods (Leamer, 1983). However, the influence of this shift on policy-making remains unclear. While policymakers are generally open to research, they often face a dilemma: choosing between innovative empirical discoveries and established policy doctrines. This predicament is often compounded by a strong inclination towards existing beliefs, potentially hindering the adoption of policies vital for long-term economic growth (Hjort et al., 2021; Baekgaard et al., 2019; Banuri et al., 2019; Metzger et al., 2020; Vivalt and Coville, 2021; Lu and Chen, 2021). The critical questions that arise are: How can we encourage policymakers to be more accepting of evidence-based findings? Could educating them about the principles of the credibility revolution lead to a shift in their beliefs and influence their policy decisions? Furthermore, does the advantage of learning econometrics extend beyond top-level bureaucrats (Mehmood, Naseer, and Chen 2023) to include those working at the grassroots level? To address these queries, we conduct a randomized control trial and utilize administrative data, examining the interplay between empirical results, policy-making, and their societal outcomes, to understand the extent to which the credibility revolution can shape policy choices and produce broader economic impacts.

The experiment uses a well-known summary of the credibility revolution, Mastering 'Metrics: The Path from Cause to Effect (Angrist and Pischke 2014), as an instrument. It essentially involved trying to influence state and fiscal capacity by changing tax policy – the sending of tax reminders in order to increase tax revenue. Across 310 frontline tax officers, we examined how the philosophical foundation associated with the credibility revolution influences their beliefs and policy decisions in a real-world experimental setup. A similar and yet a more intensive study involving Deputy Ministers in Pakistan was also conducted at around the same time (Mehmood, Naseer, and Chen 2023). The earlier study was not pre-registered thus may be interpreted as a pilot. All outcomes and hypotheses of the tax officer experiment are

pre-specified and the pre-analysis plan can be found at the pre-registry (RCT ID: AEARCTR-0010583). We find our results consistent across both the studies, indicating the effectiveness of econometrics training in improving policy decisions in potentially any bureaucratic context.

Specifically, our experiment randomly assigned an econometrics education workshop relative to a placebo macroeconomics workshop. This workshop consisted of a total of 10 hours: Two three-hour lectures were administered, accompanied by two two-hour self-persuasion exercises. These exercises involved group discussions within each treatment group, all geared towards aiding the integration and application of the lecture's key messages and insights into the participants' roles as policymakers. A similar group discussion was also employed for the macroeconomics placebo workshop as well. The tax officers were then cross-randomized to receive either a causal signal or a correlational signal regarding the effectiveness of a tax reminder policy. These signals were selected based on the emerging consensus in the literature (as summarized in Abdul Latif Jameel Poverty Action Lab, 2022) that sending tax reminders leads to an increase in tax revenue. Our results reveal a significant impact of econometrics education on tax policy and revenue collection. Specifically, relative to the control, econometrics education training led to a 20% increase in the use of tax reminder policy, with a notable 40% increase in tax collection. To put these findings in perspective, the use of tax reminders increased by 0.25 standard deviations and tax revenue increased by 0.19 standard deviations. Notably, we observe that the impact of econometrics education on the decision to send tax reminders or improve tax collection is evident solely among those who received the causal signal. These findings emphasize how econometrics education can effectively shape tax policy, bolster fiscal capacity, and enhance the generation of tax revenue. A number of robustness tests support these observed effects are valid. First, the distribution of tax collection shows a remarkable shift to the right for the econometrics-trained tax officers relative to the placebo-trained group, suggesting that a large number of tax officers were affected by the econometrics educational workshop. Second, the findings are robust to the inclusion of controls such as the demographic characteristics of the tax officers, consistent with randomization of our treatment assignments. Third, we can trace the pre-treatment collection of tax revenue and pre-treatment sending of tax reminders which we also find to be uncorrelated with the treatment assignments.

The rest of the paper is organized as follows. In Section II, we present the background information that contextualizes the experiments. Section III delves into the specifics of the experimental setup tailored for the study. Section IV discusses the results of our analysis in detail while Section V introduces a sequence of sensitivity tests. A final section concludes.

II. Background

In our field experiment, we work with tax inspectors in Punjab, who are pivotal provincial government employees primarily affiliated with the Punjab Revenue Authority (PRA). These front-line bureaucrats are at the forefront of tax collection, a critical function for the country's revenue system. Their responsibilities include the assessment of provincial taxes, with a particular focus on evaluating tax returns for sales tax on services and verifying the accuracy of income declarations and tax calculations. They are instrumental in ensuring tax compliance, auditing financial records of businesses and individuals to confirm conformity with provincial tax laws, and investigating instances of suspected tax evasion or fraud. Beyond mere assessment, these tax inspectors are actively involved in the actual collection of taxes, diligently working to amass revenues within the State of Punjab and follow up on outstanding tax arrears such as by sending reminders or "notifications" to individuals and firms. Due to their significant involvement in the revenue collection process, these professionals are frequently referred to as tax "collectors," underscoring their essential contribution to maintaining the fiscal health of the province.

We enlisted around 10% of the total tax officers (inspectors) in Punjab, which is Pakistan's largest province, resulting in a cohort of 310 participants. These tax officers at the forefront play a vital role in tax collection and compliance, ensuring timely payments, penalizing defaulters, and aiding in the rollout of tax policies. We arranged in-person training sessions for them at Punjab's Tax and Customs Academy. During these sessions, these tax officers were (1) randomly designated for econometrics training or a placebo macroeconomic workshop (2) further cross-randomized to either causal or correlational studies on the impact of tax policy. The tax officers, similar to the self-persuasion writing exercises of the deputy ministers, also engaged in 2 two-hour-long structured discussions after lectures. During these discussions, they were prompted to succinctly encapsulate the key takeaways from the lecture and contemplate how

they could effectively implement the workshop's insights into their policy practices. Additional specifics, including the discussion of the precise training content, can be found in the subsequent sections.

III. Empirical Design

In this section, we describe the tax officers experiment that investigates (i) whether econometrics training impacts Pakistani tax officers' tax reminder "notification" policy, (ii) whether the training affects tax revenue collection, and (iii) whether the kind of information tax officers receives (causal or correlational) impacts their behavior. All outcomes and hypotheses of the second experiment are pre-specified and the pre-analysis plan can be found at the pre-registry (RCT ID: AEARCTR-0010583).

A. *Treatment Design Details: Tax Officers*

The workshop for tax officers consisted of two three-hour lectures and two two-hour structured discussions that effectively condensed the material covered by the deputy ministers into a 10-hour session. The workshop on econometrics for tax officers was similar except for three major differences. First, the participants were not given the option to choose one of two workshops and then randomly assigned the workshop, instead, they were randomly assigned into two workshops directly, each comprising 2 three-hour lectures of metrics or macroeconomics. Following the lectures, a self-persuasion exercise involved a group discussion within each treatment arm and case study to apply lessons of the workshop to their career. Second, in addition to the random assignment of the workshop, each group was further randomly assigned to receive either a correlational or causal handout–information on the efficacy of tax reminder policy on tax collection. Third, the experiment with tax officers had pre-specified exact specifications and two pre-defined outcomes in the pre-analysis plan: sending tax reminders and tax revenue collected.

Assignment of treatment.–The econometrics workshop delivered to tax officers is also rooted in the principles outlined in the 'Mastering Metrics' book, with a particular emphasis on causal identification. It provides a summary of key concepts related to the credibility revolution. One of the central areas of focus was the concept of random assignment, its significance in

causal inference, and how experimental studies diverge from "naive" correlational comparisons when interpreting causal effects. This workshop distilled the essence of the deputy minister workshop into two three-hour lectures and two two-hour structured discussions, effectively compacting the material that the deputy ministers had covered in a 10-hour workshop. Specifically, tax officers were individually randomly assigned to in-person classes at the Tax and Customs' Academy in Punjab. The macroeconomics workshop was of similar length and medium, where the lecturer discussed some macroeconomic facts and trends in Pakistan and held similar structured discussion on this content. A snapshot of the slides for the two lectures, along with the complete course material, is provided in Figure A1 of Appendix. Screenshots of these in-person lectures are presented as Figure A2 of Appendix.

B. *Data, Outcomes and Empirical Specification*

Data.—We were granted access to tax officers as they entered their training program. The date of training and cities these officers are stationed at have been anonymized to ensure their anonymity. Our strong partnership with the institute enabled us to effectively match them with administrative data pertaining to individual traits and pre-treatment policy outcomes, such as previous year's policy and tax collection figures. We conducted a balance test on these characteristics and pre-treatment outcome variables, as demonstrated in Table 1, and find the treated and control officers as statistically similar.

Outcome Variable on Tax Reminders and Revenue.— We pre-registered two key outcomes for the experiment with tax miniders (AEA RCT Registry ID: AEARCTR-0010583). The first outcome variable is a dummy variable indicating whether the officer dispatched a "tax notification" to prompt citizens and businesses to settle their tax obligations. This information is sourced from the training institute. The second variable is the individual-level tax collected during the fiscal year following the treatment. This data is acquired from the tax office and cross-referenced with information from the Federal Bureau of Statistics (FBR).

Causal vs Correlational Signal.— The “causal signal” is a binary variable that becomes active when tax officers are assigned a handout summarizing the contents of the article titled "Behavioral Interventions in Tax Compliance: Evidence from Guatemala," authored by Kettle, S., Hernandez, M., Ruda, S., and Sanderson, M.A. (2016). In contrast, the binary variable takes

the value of zero if the tax officer is assigned the correlational study on tax reminder policy authored by Budiman, I. and Inayati, I. (2021), titled "Effect of Notice of Tax Warning, Notice of Tax Collection, and Tax Education Programs on Tax Compliance in West Sumatera and Jambi," published in Publik (Jurnal Ilmu Administrasi), volume 10, issue 1, pages 45-63. For precise details of the handouts assigned to the tax officers, refer to Figure A3 and A4 in the Appendix. The causal signal is randomly administered across treatment and control groups. This is evidenced from the balance across demographic traits of participants who were assigned to treatment and received the causal signal in Table A5.

Empirical Specification.— The evaluation of the impact of metrics training can be carried out within a straightforward regression framework. For each individual-level tax officers' outcome, the estimation equation takes the form:

$$Y_i = \alpha + \theta\, Econometrics\ Training_i + X_i'\mu + \epsilon_i \quad (2)$$

where Y_i represents either tax revenue in millions of Pakistani Rupees (PKR) or a dummy variable that switches on if officer *i* sends a tax reminder notification. $Econometrics\ Training_i$ is a dummy equal to one if the tax officer is assigned to metrics training and zero if assigned the placebo macroeconomics workshop. X_i is a vector of individual-level controls provided in balance check in Table 1. We cluster standard errors at the tax region level. Figure 1 depicts the 78 tax regions in our study. θ is our primary coefficient of interest, in this second experiment, estimating the causal impact of metrics training on tax reminder policy and tax collection. We also cross-randomized the tax officers to receive causal versus correlational signals within the metrics training and placebo workshops. These signals are depicted in Figure A3 and Figure A4 of the Appendix report, illustrating experimental versus observational studies showing tax reminder policy increases tax collection.[2] These handouts were randomly assigned within each

[2] Causal Handout Distributed among Tax Officers uses the following study "Behavioral Interventions in Tax Compliance: Evidence from Guatemala" by Kettle, S., Hernandez, M., Ruda, S., and Sanderson, M.A. (2016). This paper shows that sending tax reminders increases tax collections in a field experiment. Correlational Handout Distributed among Tax Officers uses the following study: "Effect of Notice of Tax Warning, Notice of Tax Collection, and Tax Education Programs on Tax Compliance in West Sumatera and Jambi" by Budiman, I., and

treatment arm for the tax officers. We have pre-specified the following two hypotheses that we aim to test through this cross-randomization:

"*H1: Econometrics training will lead to a higher increase in the adoption of sending tax reminders for those officers that received the signal study to send tax reminders with RCT evidence.*

H2: Econometrics training will not impact or decrease the adoption of sending tax reminder for those officers that received the signal study to send tax reminders with correlational evidence."

Figure 1 offers an illustrative map depicting the fraction of metrics-assigned tax officers in each district alongside the corresponding tax collection. This visualization reveals the geographic correlation between tax officer econometrics workshop assignments and tax collection levels in their respective districts. The assignment pattern of tax officers into the econometrics workshop and the placebo workshop also appears to be randomly distributed throughout Punjab, the setting of the tax officers' experiment. Furthermore, relative to placebo assigned officers, regions where metrics-trained tax officers were stationed seem to exhibit a more pronounced, darker shade, indicative of higher tax collections. Examining the results visually, the metrics-trained tax officers' revenue distribution experiences a shift toward higher values compared to placebo tax officers (refer to Figure 2 for an estimate of probability density functions).[3]

IV. Results

Treatment effects on Tax Reminders and Tax Collection. —Table 2 presents the outcomes of metrics training. As indicated in Columns 1 and 2 of Table 2, tax officers assigned to the metrics group are approximately 12 percentage-points more likely to send tax reminder notifications to citizens. This increase corresponds to a 20% surge in policy adoption compared to the placebo macroeconomic assigned group. Similarly, Columns 3 and 4 of Table 2 correspond

Inayati, I. (2021). This paper shows that sending tax reminders increases tax collections in an observational panel data study.

[3] Values larger than PKR 80 million are dropped for readability. Keeping these values yields a similar conclusion. Results are available on request.

with a substantial body of existing evidence summarized in the Abdul Latif Jameel Poverty Action Lab (2022) policy paper that concludes that tax reminders increase tax revenue generation. Our findings, significant at the 10% level, suggest a substantial increase in tax revenue collection attributed to metrics training—approximately PKR 25 million (USD 85,000) or a 40% elevation in revenue over the mean dependent variable. The enhanced performance seen in deputy ministers who underwent metrics training, resulting in a notable improvement on public policy exam scores, is similarly reflected in tax officers' outcomes. Tax officers who received the metrics training displayed a 0.25 standard deviation increase in sending tax reminders and a 0.19 standard deviation improvement in the amount of tax revenue collected.[4] The MDE for the policy exam effects, sending of tax reminders, and amount of tax collected are 0.47, 0.31, and 0.30 respectively.[5] To put these MDEs in perspective, other interventions in training of civil servants have MDEs of 0.11 to 0.26 for the Socratic method, a 5 minute-video, on course outcomes and a behavioral measure of motivated reasoning (Chen, Ramos-Maqueda, Silveira 2023) or 0.4 for both teacher vaccination and absenteeism in a role model vaccination message campaign (Mehmood, Naseer, Chen 2023). Examining the results visually, the metrics-trained tax officers' revenue distribution experiences a shift toward higher values compared to placebo tax officers (refer to Figure 5 for an estimate of probability density functions).[6] Next, we further unpack the mechanism that explain the increase in tax collection.

Treatment effects by Causal vs Correlational Signal. — Table 3 examines whether the nature of the signal—causal vs. correlational—regarding the impact of sending tax reminders to enhance tax collection affects tax reminder policy and tax collection. The objective here is to understand the mechanism at play—how metrics training influences causal thinking—and how it subsequently shapes policy and fiscal capacity. Table 3 reveals that the influence of metrics training is exclusively attributed to those tax officers who receive the causal signal (as illustrated in Figure A3), while metrics trained officers exposed to correlational signal (Figure A4 show

[4] The MDE for these effects are 0.47, 0.31, and 0.30 respectively. To put these MDEs in perspective, other interventions in training of civil servants have MDEs of 0.11 to 0.26 for the Socratic method, a 5 minute-video, on course outcomes and a behavioral measure of motivated reasoning (Chen, Ramos-Maqueda, Silveira 2023) or 0.4 for both teacher vaccination and absenteeism in a role model vaccination message campaign (Mehmood, Naseer, Chen 2023).

[5] The MDE is calculated as 2.8 times the standard error from the regression coefficient (Ioannidis, et al., 2017).

[6] Values larger than PKR 80 million are dropped for readability. Keeping these values yields a similar conclusion. Results are available on request.

minimal impact from the training. These findings suggest that the metrics training is only effective if the information about the policy's causal impact is known. The collective evidence indicates that policymakers, when confronted with policy decisions that carry tangible risks to their reputation, present implementation difficulties, and are constrained by public budgeting, tend to opt for policies that are backed by causal evidence.

V. Conclusion

This study demonstrated how econometric education, specifically focusing on the principles of the credibility revolution, can profoundly influence policy decisions and outcomes in the domain of tax collection. Through a field experiment involving tax officers in Punjab, Pakistan, our research provides evidence that training in causal inference leads to more effective and evidence-based tax policy implementation. Tax officers who underwent this specialized training not only increased their use of tax reminder policies but also significantly boosted tax revenue collection.

The marked increase in the adoption of tax reminder policies and the subsequent rise in tax revenue among those who received econometrics training underlines the transformative potential of this educational intervention. This shift is particularly notable as it occurs within a bureaucratic context often characterized by resistance to change due to reputational risks, implementation challenges, and budgetary constraints. Our findings align with prior research, including the comprehensive experiment with Deputy Ministers in Pakistan (Mehmood, Naseer, and Chen 2023), reinforcing the notion that even 10 hours of econometric training can transcend hierarchical levels, benefiting frontline officers and top bureaucrats alike.

Furthermore, our study show that the type of information—causal versus correlational—presented to the tax officers matters. The results indicate that the effectiveness of the econometrics training is contingent upon the officers receiving a causal signal about the policy's impact. This suggests that an understanding of causality is critical in shifting policy choices and enhancing fiscal capacity.

In conclusion, we find robust evidence that econometric training, rooted in the principles of the credibility revolution, can significantly influence policy decisions and outcomes in the

public sector. By fostering an appreciation for causal evidence among policymakers and front-line implementers, such training can lead to more informed, effective, and efficient policy implementation. This paradigm shift in academic thought holds the promise of enhancing public service delivery and, ultimately, contributing to more effective governance and improved societal outcomes.

Illustration 1: Summary statistics and variable description for Tax Officers Experiment

Panel A: Dependent Variables

Variable	Obs	Mean	Std. Dev.	Min	Max
Tax Reminders	310	.584	.494	0	1
Tax Revenue	310	55.315	128.198	.01	774.374

Note: The table contains pre-registered dependent variables used in the regressions in the tax officers' experiment.

Panel B: Baseline Characteristics

Variable	Obs	Mean	Std. Dev.	Min	Max
Gender (Male)	310	.61	.489	0	1
Birth capital	310	.297	.458	0	1
Land	310	.3	.459	0	1
Income	310	55203.226	8531.722	40000	70000
Age	310	32.865	2.733	28	42
Religion	310	.948	.222	0	1
Experience	310	4.906	2.572	1	12
Pre-Tax Revenue	310	53.422	128.735	.005	879.622
Pre-Tax Reminders	310	.539	.499	0	1

Note: The table contains all baseline or control variables used in the regressions with controls in the tax officers experiment.

Illustration 2. Baseline Characteristics

Variable	Description
Gender (Male)	Dummy Variable equals 1 if the person is male
Birth capital	Dummy Variable equals 1 if the birth city is the provincial capital
Land	Dummy Variable equals 1 if the person owns any land
Income	Monthly income in PKRs
Age	Age in Years
Religion	Dummy Variable equals 1 if the person is Muslim
Experience	Work experience in Years

Note: The table describes baseline or control variables used in the regressions in the tax officers experiment.

Table 1: Tax Officers - Balance over individual characteristics and pre-treatment outcomes

	(1) Gender	(2) Age	(3) Land	(4) Income	(5) Birth capital	(6) Religion	(7) Experience	(8) Pre-treatment Revenue	(9) Pre-treatment Reminder Letter	(10) Metrics Assigned
Metrics Assigned	-0.0263	-0.369	0.0497	1,300	-0.00211	-0.0364	0.00459	-10.76	-0.0321	
	[0.0558]	[0.268]	[0.0527]	[964.7]	[0.0531]	[0.0258]	[0.257]	[14.79]	[0.0580]	
Gender										-0.0283
										[0.0601]
Birth capital										-0.00255
										[0.0639]
Land										0.0609
										[0.0646]
Income										0.0000046
										[0.00000339]
Age										-0.0170
										[0.0121]
Religion										-0.187
										[0.125]
Experience										0.000233
										[0.0131]
Pre-treatment Revenue										-0.000165
										[0.000213]
Pre-treatment Notification										-0.0319
Observations	310	310	310	310	310	310	310	310	310	310
R-squared	0.049	0.291	0.054	0.050	0.031	0.029	0.273	0.036	0.013	0.026
Mean	0.610	32.86	0.300	55203	0.297	0.948	4.906	53.42	0.539	0.500
F Statistics	0.222	1.898	0.891	1.817	0.00158	2.001	0.000318	0.529	0.306	0.965
p-values	0.638	0.169	0.346	0.179	0.968	0.158	0.986	0.468	0.580	0.469

Robust standard errors appear in brackets (clustered at the tax region level). Metrics assigned is a dummy variable that switches on when the causal inference workshop is randomly assigned to participants relative to the placebo macroeconomics workshop. The causal inference book is randomly assigned conditional on the book being chosen. The controls include other available individual characteristics obtained from administrative data (i.e. all remaining column dependent variable except the dependent variable used in the respective column). *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Impact of Metrics Training on Tax Reminder Letters and Tax Revenue

	(1) Tax Reminders	(2) Tax Reminders	(3) Tax Revenues	(4) Tax Revenues
Metrics Assigned	0.123**	0.123**	28.34*	24.53*
	[0.0545]	[0.0543]	[14.81]	[13.65]
Controls	No	Yes	No	Yes
Observations	310	310	310	310
R-squared	0.015	0.053	0.012	0.041
Mean	0.584	0.584	55.32	55.32

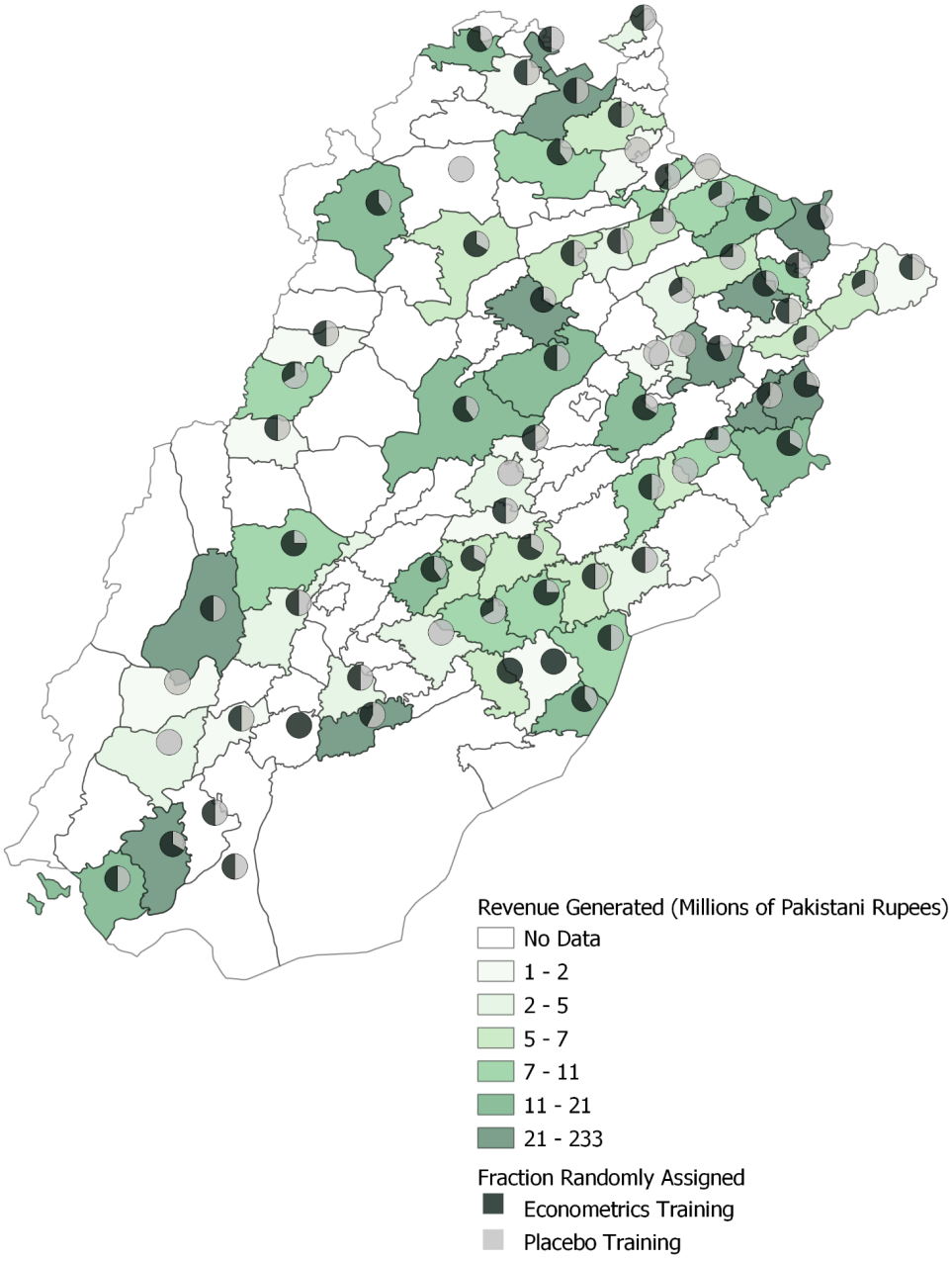
Robust standard errors appear in brackets (clustered at the tax region level). Metrics assigned is a dummy variable that switches on when a causal inference workshop is randomly assigned to participants relative to the placebo macroeconomics workshop. The controls include available individual characteristics in the balance table. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Impact of Metrics Training and Signals on Tax Reminder Letters and Tax Revenue

	(1) Tax Reminders	(2) Tax Reminders	(3) Tax Revenues	(4) Tax Revenues
Metrics Assigned	-0.00616	-0.00544	15.63	11.44
	[0.0770]	[0.0756]	[13.79]	[13.23]
Causal Signal	-0.00616	0.0159	-1.849	-7.233
	[0.0812]	[0.0799]	[16.15]	[14.59]
Metrics Assigned x Causal	0.256**	0.259**	25.29**	26.26**
Signal	[0.108]	[0.107]	[12.36]	[10.61]
Controls	No	Yes	No	Yes
Observations	310	310	310	310
R-squared	0.042	0.080	0.024	0.042
Mean of Dep. Variable	0.587	0.587	55.32	55.32

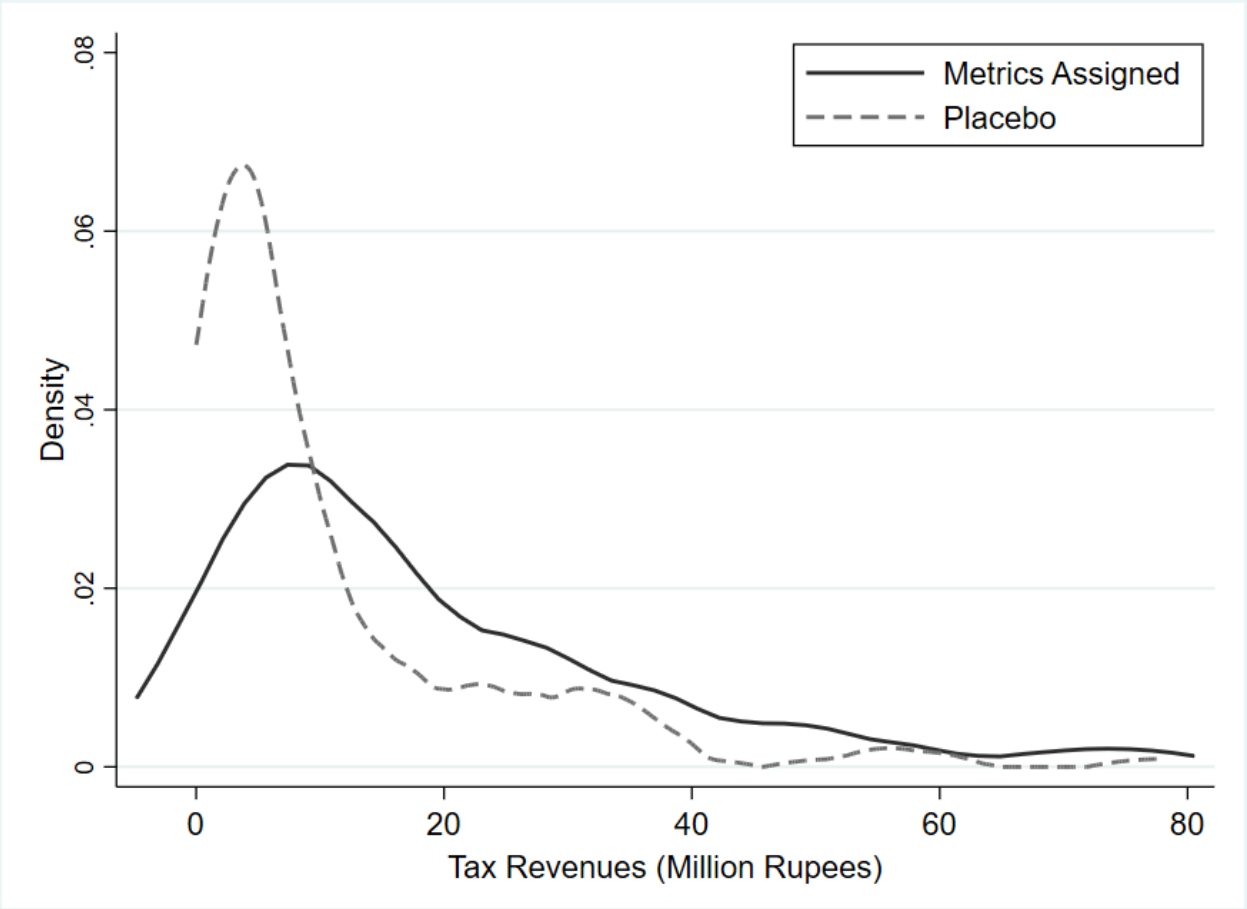
Robust standard errors appear in brackets (clustered at the tax region level). Metrics assigned is a dummy variable that switches on when a causal inference workshop is randomly assigned to participants relative to the placebo macroeconomics workshop.The "causal signal" is a binary variable activated when tax officers receive a summary handout of the article "Behavioral Interventions in Tax Compliance: Evidence from Guatemala" by Kettle et al. (2016). Alternatively, the variable takes zero if officers are assigned the correlational study "Effect of Notice of Tax Warning, Notice of Tax Collection, and Tax Education Programs on Tax Compliance in West Sumatera and Jambi" by Budiman and Inayati (2021). The controls include available individual characteristics in the balance table. *** p<0.01, ** p<0.05, * p<0.1.

Figure 1: Treated versus Control Regions by Tax Revenue Units in Punjab



Note: The figure illustrates tax units in Punjab, with darker colors representing higher tax collections in the region and lighter regions representing lower tax revenues. The units are in millions of rupees. The pie charts on the regions indicate the proportion of metrics-trained tax officers versus placebo-trained tax officers. Within each region, tax officers were either treated with a metrics workshop or a placebo macroeconomics workshop. Full dark circles represent all tax officers as treated, while full light pie charts represent no one in the region being treated with metrics training.

Figure 2: Tax Revenue distributions of Metrics Trained versus Placebo Trained Tax Officers



Note: The figure presents estimates of probability density functions for tax revenue collected by metrics-trained and placebo-trained tax officers. The units are in millions of Pakistani rupees.

References:

1. **Mehmood S., Nasser S. and Chen C.** Training policymakers in econometrics: Evidence from Pakistan. https://users.nber.org/~dlchen/papers/Training_Policy_Makers_in_Econometrics.pdf
2. **Angrist, Joshua D., and Jörn-Steffen Pischke.** 2010. "The credibility revolution in empirical economics: How better research design is taking the con out of econometrics." *Journal of economic perspectives* 24 (2): 3-30.
3. **Weber, M**. 1922. *Economy and Society: An Outline of Interpretive Sociology*.
4. **Lipsky, M.** 1980. *Street Level Bureaucracy: Dilemmas of the Individual in Public Services*. Russell Sage Foundation. http://www.jstor.org/stable/10.7758/9781610447713
5. World Bank Annual Report. 2017. https://thedocs.worldbank.org/en/doc/908481507403754670-0330212017/original/AnnualReport2017WBG.pdf
6. **Besley, T, and Persson,** T. 2009. "The Origins of State Capacity: Property Rights, Taxation, and Politics." American Economic Review, 99 (4): 1218-44.
7. **Lindauer, D.L**. 1988. Government Pay and Employment Policies and Government Performance in Developing Economies.
8. **Finan, F., Olken, B.A., & Pande, R.P.** 2015. Government Workers in Developing Countries.
9. **Perry, J. L., & Hondeghem, A.** 2008. Building Theory and Empirical Evidence about Public Service Motivation. International Public Management Journal, 11(1), 3–12. https://doi.org/10.1080/10967490801887673
10. **Checchi D., Fenizia A. and Lucifora C.** 2021. Public Sector Jobs: Working in the Public Sector in Europe and the Us. IZA Discussion Paper No. 14514, Available at SSRN: https://ssrn.com/abstract=3879362 or http://dx.doi.org/10.2139/ssrn.3879362
11. **Khan A.Q., Khwaja A.L., Olken B.A.** 2019. Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors , The Quarterly Journal of Economics, Volume 131, Issue 1, February 2016, Pages 219–271, https://doi.org/10.1093/qje/qjv042
12. **Muralidharan K. and Sundararaman V.** 2011. Teacher Performance Pay:Experimental Evidence from India. Journal of Political Economy 119 (1): 39-77
13. **Bloom N., Propper C., Seiler S., Van Reenen J**. 2015b. The Impact of Competition on Management Quality: Evidence from Public Hospitals, The Review of Economic Studies, Volume 82, Issue 2, April, Pages 457–489
14. **David J. Deming. and Weidmann, Ben.** 2021. *Team Players: How Social Skills Improve Team Performance*. Forthcoming Econometrica.
15. **Banerjee A., Chattopadhyay R., Duflo E., Keniston D., and Singh N.** 2021. "Improving Police Performance in Rajasthan, India: Experimental Evidence on Incentives, Managerial Autonomy, and Training." American Economic Journal: Economic Policy, 13 (1): 36-66.
16. **Leamer, Edward E.** 1983. "Let's take the con out of econometrics." *The American Economic Review* 73 (1): 31-43.

17. **Hjort, Jonas, Diana Moreira, Gautam Rao, and Juan Francisco Santini.** 2021. "How research affects policy: Experimental evidence from 2,150 brazilian municipalities." *American Economic Review* 111 (5): 1442-80.
18. **Baekgaard, Martin, Julian Christensen, Casper Mondrup Dahlmann, Asbjørn Mathiasen, and Niels Bjørn Grund Petersen.** 2019. "The role of evidence in politics: Motivated reasoning and persuasion among politicians." *British Journal of Political Science* 49 (3): 1117-1140.
19. **Banuri, Sheheryar, Stefan Dercon, and Varun Gauri.** 2019. "Biased policy professionals." *The World Bank Economic Review* 33 (2): 310-327.
20. **Metzger, L., Svoronos, T. and Khan, A., 2020.** Policy decisions and evidence use among civil servants: A group decision experiment in Pakistan. CID Working Paper Series.
21. **Vivalt, Eva, and Aidan Coville.** 2021. "How Do Policymakers Update?" *Mimeo*.
22. **Lu, Wei, Elliott Ash, and Daniel L. Chen.** 2021." Motivated Reasoning in the Field: Polarization of Precedent, Prose, and Policy in U.S. Circuit Courts, 1930-2013." *Mimeo*.
23. **Kettle S., Hernandez M., Ruda S., and Sanderson, Michael A.** Behavioral Interventions in Tax Compliance: Evidence from Guatemala (June 1, 2016). World Bank Policy Research Working Paper No. 7690, Available at SSRN: https://ssrn.com/abstract=2811337

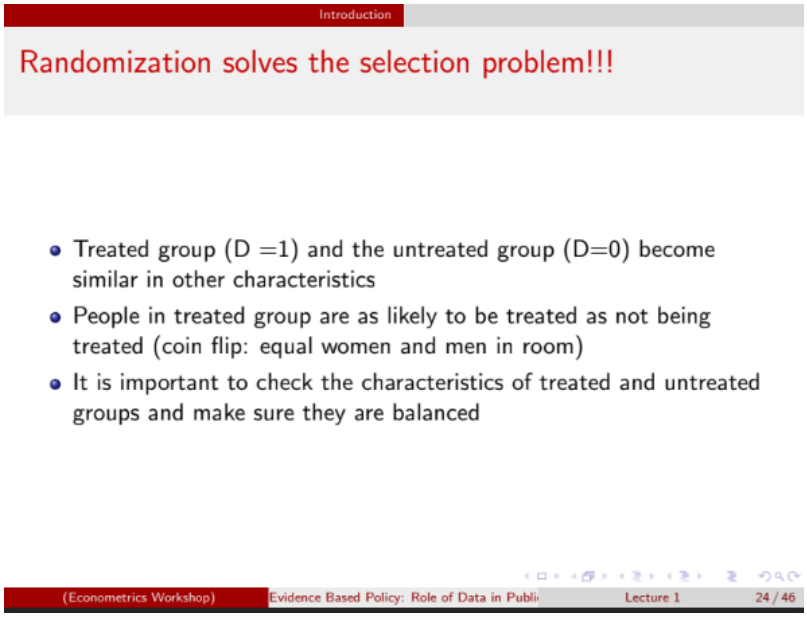
Online Appendix to:

Training Policymakers in Econometrics II:

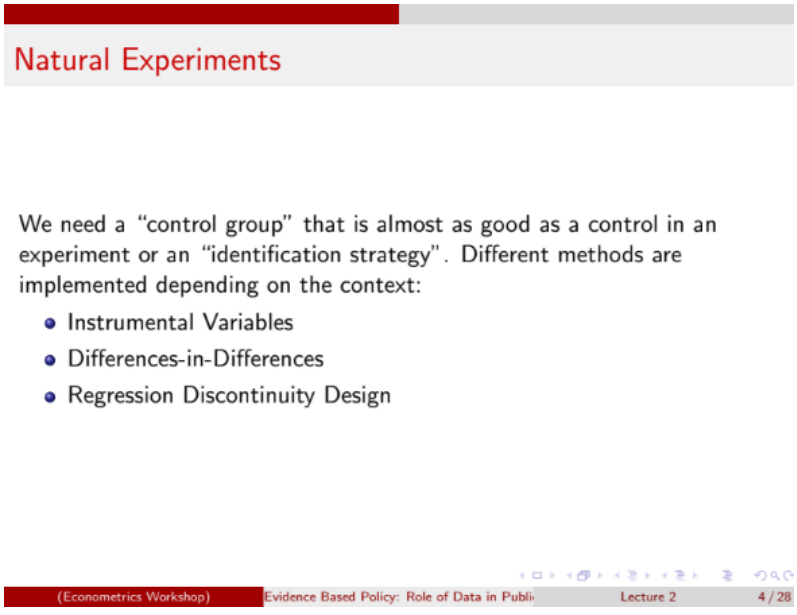
Tax Officers and Fiscal State Capacity

Figure A1: Slides of Tax Officers Workshop

Panel A: A slide from Lecture 1



Panel B: A slide from Lecture 2



Note: Two representative slides from Lecture 1 and 2 in Panel A and B, respectively, are presented in the figure above. The link to access the complete slide deck for both Lecture 1 and Lecture 2 can be found HERE. Placebo lecture on macroeconomics can be found HERE.

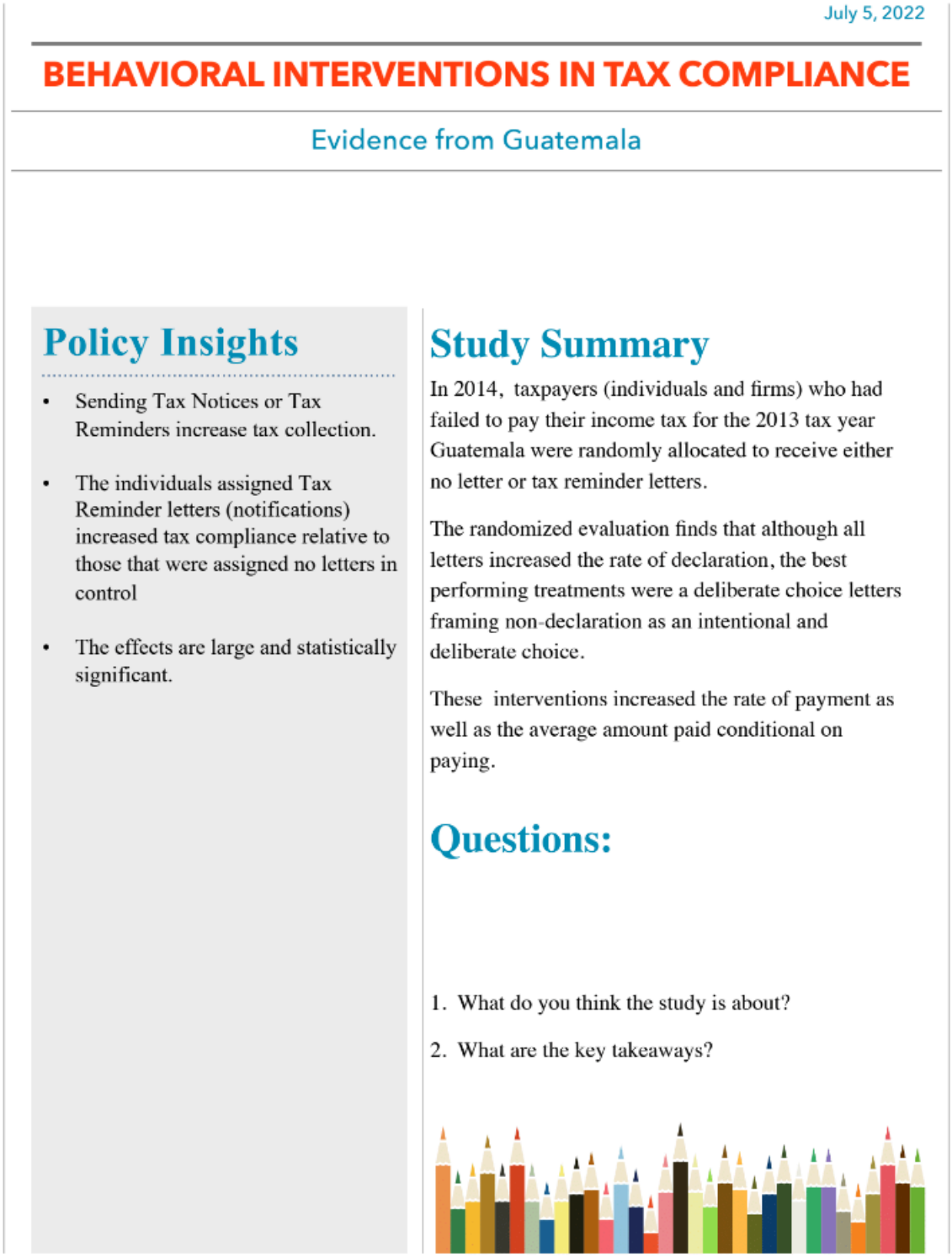
Figure A2: Tax Officers Econometrics Workshop

Panel A: Illustration 1

Panel B: Illustration 2

Note: Two images from the tax officers' workshop are displayed in this figure.

Figure A3: Causal Signal Distributed to Tax Officer



July 5, 2022

BEHAVIORAL INTERVENTIONS IN TAX COMPLIANCE

Evidence from Guatemala

Policy Insights

- Sending Tax Notices or Tax Reminders increase tax collection.
- The individuals assigned Tax Reminder letters (notifications) increased tax compliance relative to those that were assigned no letters in control
- The effects are large and statistically significant.

Study Summary

In 2014, taxpayers (individuals and firms) who had failed to pay their income tax for the 2013 tax year Guatemala were randomly allocated to receive either no letter or tax reminder letters.

The randomized evaluation finds that although all letters increased the rate of declaration, the best performing treatments were a deliberate choice letters framing non-declaration as an intentional and deliberate choice.

These interventions increased the rate of payment as well as the average amount paid conditional on paying.

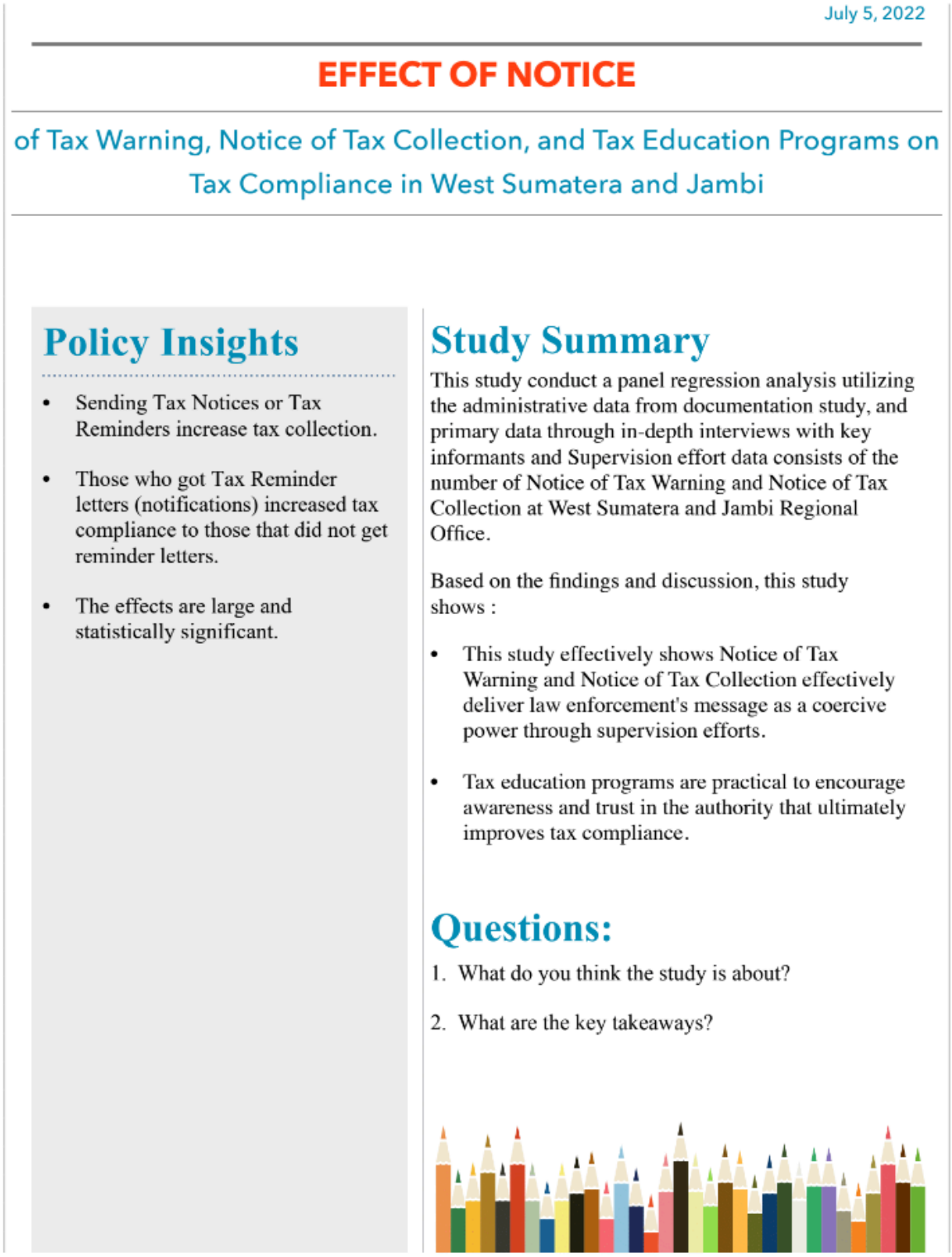
Questions:

1. What do you think the study is about?
2. What are the key takeaways?

Note: The figure is an illustration of a causal handout seen by the tax officers. The study used to prepare this document causal handout is the following: Kettle, S., Hernandez, M., Ruda, S. and

Sanderson, M.A., 2016. Behavioral interventions in tax compliance: Evidence from Guatemala. World Bank Policy Research Working Paper, (7690).

Figure A4: Correlational Signal Assigned to Tax Officer



July 5, 2022

EFFECT OF NOTICE

of Tax Warning, Notice of Tax Collection, and Tax Education Programs on Tax Compliance in West Sumatera and Jambi

Policy Insights

- Sending Tax Notices or Tax Reminders increase tax collection.
- Those who got Tax Reminder letters (notifications) increased tax compliance to those that did not get reminder letters.
- The effects are large and statistically significant.

Study Summary

This study conduct a panel regression analysis utilizing the administrative data from documentation study, and primary data through in-depth interviews with key informants and Supervision effort data consists of the number of Notice of Tax Warning and Notice of Tax Collection at West Sumatera and Jambi Regional Office.

Based on the findings and discussion, this study shows :

- This study effectively shows Notice of Tax Warning and Notice of Tax Collection effectively deliver law enforcement's message as a coercive power through supervision efforts.
- Tax education programs are practical to encourage awareness and trust in the authority that ultimately improves tax compliance.

Questions:

1. What do you think the study is about?
2. What are the key takeaways?

Note: The figure is an illustration of the correlational handout seen by the tax officers. The study used to prepare this document is the following: Budiman, I. and Inayati, I.,

2021. Effect of Notice of Tax Warning, Notice of Tax Collection, and Tax Education Programs on Tax Compliance in West Sumatera and Jambi. Publik (Jurnal Ilmu Administrasi), 10(1), pp.45-63.

Table A5: Balance with interactions

	(1) Gender	(2) Age	(3) Land	(4) Income	(5) Birth capital	(6) Religion	(7) Experience	(8) Pre-treatment Revenue	(9) Pre-treatment Notification	(10) Metrics Assigned
Metrics Assigned	0.0313	-0.415	0.0746	1,634	0.0534	-0.0342	-0.106	-2.060	-0.0176	
	[0.0763]	[0.387]	[0.0723]	[1,360]	[0.0761]	[0.0291]	[0.366]	[19.36]	[0.0817]	
Causal Signal	0.0145	0.0290	0.0335	516.3	-0.0328	-0.0390	-0.739**	9.084	0.0260	
	[0.0790]	[0.406]	[0.0724]	[1,328]	[0.0734]	[0.0302]	[0.372]	[24.34]	[0.0816]	
Metrics Assigned x Causal Signal	-0.114	0.0917	-0.0498	-667.4	-0.110	-0.00347	0.225	-17.33	-0.0290	
	[0.110]	[0.535]	[0.103]	[1,927]	[0.103]	[0.0520]	[0.504]	[28.69]	[0.115]	
Gender										-0.0314
										[0.0302]
Birth capital										-0.0341
										[0.0319]
Land										-0.0155
										[0.0320]
Income										-0.000000598
										[0.00000173]
Age										0.00107
										[0.00626]
Religion										-0.00455
										[0.0682]
Experience										0.00296
										[0.00666]
Pre-treatment Revenue										-0.0000673
										[0.000109]
Pre-treatment Notification										-0.00731
Observations	310	310	310	310	310	310	310	310	310	310
R-squared	0.054	0.291	0.055	0.050	0.043	0.037	0.288	0.037	0.013	0.675
Mean	0.610	32.86	0.300	55203	0.297	0.948	4.906	53.42	0.539	0.255
F Statistic	1.080	0.0294	0.233	0.120	1.135	0.00445	0.198	0.365	0.0632	0.407
p-values	0.299	0.864	0.629	0.729	0.288	0.947	0.656	0.546	0.802	0.931

Robust standard errors appear in brackets (clustered at the tax region level). Metrics assigned is a dummy variable that switches on when a causal inference workshop is randomly assigned to participants. The "causal signal" is a binary variable activated when tax officers receive a summary handout of the article "Behavioral Interventions in Tax Compliance: Evidence from Guatemala" by Kettle et al. (2016). Alternatively, the variable takes zero if officers are assigned the correlational study "Effect of Notice of Tax Warning, Notice of Tax Collection, and Tax Education Programs on Tax Compliance in West Sumatera and Jambi" by Budiman and Inayati (2021). The controls include other available individual characteristics obtained from administrative data (i.e. all remaining column dependent variable except the dependent variable used in the respective column). *** $p<0.01$, ** $p<0.05$, * $p<0.1$.