



# IGC

## A consumer incentive scheme experiment: Pilot study

---

Brian Dillon (Cornell University) and Twivwe Siwale (IGC)

- Providing consumers with a lottery-based incentive to collect Electronic Fiscal Devices (EFD)-generated invoices led to a 93 percent increase in the total value of such invoices submitted, without changing the total value of all receipts and invoices submitted.
- Most households did not report any difficulty in identifying VAT-registered firms, although treated households were somewhat more likely to report such difficulties, possibly because they worked harder to find such firms.
- Consumers reported only modest pushback from sellers, in the form of refusing to issue receipts or offering handwritten rather than machine-generated receipts.
- The above findings are based on a relatively small sample from a middle- and upper-income neighbourhood of Lusaka, and cannot be taken as representative of other population subgroups. The findings provide reason for cautious optimism about how consumers might react to VAT-related incentives.

POLICY BRIEF ZMB-21087

APRIL 2022

This project was funded  
by IGC Zambia

[www.theigc.org](http://www.theigc.org)

DIRECTED BY



FUNDED BY



## Overview

---

This policy brief describes the goals, activities, and findings of a pilot experiment on the use of consumer incentives to increase VAT compliance for retail sales transactions in Zambia. This pilot experiment is part of a research collaboration between the Research Department of the Zambia Revenue Authority (ZRA), the International Growth Centre (IGC), and Cornell University.

The overall goal of this collaboration is to understand whether consumer incentives and/or some other set of policies can serve as effective complements to Electronic Fiscal Devices (EFDs) to increase VAT compliance in Zambia.

## Background

A VAT consumer incentive scheme involves either a guaranteed or a probabilistic reward for consumers, activated when they request a formal tax invoice from a VAT-eligible purchase. For guaranteed rewards, the value of the reward is typically increasing in the value of VAT remitted in a retail transaction. For probabilistic rewards (i.e., lotteries), the probability of success is typically increasing in the value of VAT remitted in a retail transaction. Our focus is on probabilistic rewards.

A number of countries have used consumer incentive programs to increase VAT remittances. (ZRA ran VAT lotteries as far back as 2008, although those lotteries were short-term events rather than permanent programs). If consumers can be effectively incentivised to request VAT invoices, then they begin to act like tax auditors *en masse*. VAT-registered firms are the targets of the program; consumers are the pathway through which those firms can be induced to fulfil their legal tax obligations. The practical challenges to implementing these programs vary across settings. By testing the design before rolling it out at large scale, we hope to identify and adjust to implementation challenges that are specific to Zambia.

## Research questions

In the first pilot study, we aimed to address the following question: What is the effect of a probabilistic reward (i.e., a lottery) on consumer request for and retention of VAT invoices when making retail purchases?

**What is the effect of a probabilistic reward (i.e., a lottery) on consumer request for and retention of VAT invoices when making retail purchases?**

The specific goals were as follows:

- To gain experience with the implementation of a consumer incentive program in Lusaka.
- To experimentally evaluate the performance of one type of incentive program
- To elicit feedback from consumers on their experience with the program

Because this was a small-scale pilot, we were not focused on identifying an unbiased estimate of the program's impact at larger scale. Instead, the pilot will inform our approach to future experiments that will inform other potential components of a large-scale consumer incentive program.

## Research design

### Study population and recruitment

The target population for the pilot experiment included all households in Kabwata, a middle- and upper-income neighbourhood of Lusaka. We focused on this income bracket because relative to lower-income households, we expected middle- and upper-income households to make more purchases during the relatively short study period, and to make more purchases from VAT-registered businesses.<sup>1</sup> The only other eligibility criteria were that only one participant could be enrolled per household (the project was framed as a household-level project).

Individuals were invited to participate in the study through direct door-to-door recruitment. Groups of enumerators walked systematically through each block of Kabwata, knocked on the door of every residence, and explained the project to an available adult. If the head of household was not there, the enumerator made a plan to return when the head of household was available. The team worked primarily in the afternoon and evening, and on Saturday, to maximise the chance of encountering the head of household.

### Activities and experimental design

The experiment involved two visits to each participating household: an initial visit to recruit the household, conduct the baseline survey, and assign treatment, and a follow-up visit two weeks later to conduct an endline survey and to collect the receipts and invoices retained by the household during prior two weeks.

---

<sup>1</sup> The target populations for our future, larger scale studies, will include a wider range of households.

At the time of initial enrolment, participating households were randomly assigned by the survey software to one of the following treatments, with 50/50 probability:

Control: No VAT-related incentive.

Treatment: During the follow-up visit, treated participants had the opportunity to draw for a prize based on the total value of EFD-generated VAT invoices that they submitted. The prize was 400 Kwacha. To win, the respondent had to select a red card from a hat containing 9 black cards and 1 red card (10% chance of winning). The number of attempts each participant had to win the prize was based on the total value of their submitted invoices, using a formula shown below.

Participants in the Treatment arm were informed about the lottery after agreeing to participate.

At the end of the baseline survey, enumerators read the following script to the Control households:

*“Thank you for answering those questions. Now I will introduce the other part of the study. We would like to track the number and total value of all formal VAT invoices issued to households like yours when you make purchases. To do this, we are asking you and other participants to retain all of the receipts and invoices that you receive over the coming two weeks. Please shop as you normally would, but when you or anyone else in your household receives a receipt or invoice from a seller, store it in this envelope. [Enumerator: hand the respondent their plastic envelope]. Only include receipts or invoices received by current members of this household. We are especially interested in measuring the value of machine-generated (not handwritten) VAT invoices that your household receives, so please be sure to retain all of those. You can begin retaining receipts and invoices from this moment forward.*

*In two weeks, someone from my team will return to collect this envelope and to conduct a brief follow-up survey with you. At that time, we will add up the total value of your household expenditures that are reflected on machine-generated VAT invoices. To participate, I will need to ask for your mobile phone number, so that we can send you reminder texts and schedule our follow-up survey.”*

And this script to the Treated households:

*“Thank you for answering those questions. Now I will introduce the other part of the study. We would like to track the number and total value of all formal VAT*

invoices issued to households like yours when you make purchases. To do this, we are asking you and other participants to retain all of the receipts and invoices that you receive over the coming two weeks. Please shop as you normally would, but when you or anyone else in your household receives a receipt or invoice from a seller, store it in this envelope. [Enumerator: hand the respondent their plastic envelope]. Only include receipts or invoices received by current members of this household. We are especially interested in measuring the value of machine-generated (not handwritten) VAT invoices that your household receives, so please be sure to retain all of those. You can begin retaining receipts and invoices from this moment forward.

In two weeks, someone from my team will return to collect this envelope and to conduct a brief follow-up survey with you. At that time, we will add up the total value of your household expenditures that are reflected on machine-generated VAT invoices. You will then have the opportunity to draw for a prize based on the total value of machine-issued (not handwritten) VAT invoices. The prize is 400 Kwacha. To win, you will have to draw a yellow card from a hat that contains 9 black cards and 1 yellow card, for a 10% probability of winning on each draw. The number of draws that you are allowed to make will depend on the total value of VAT invoices that your household receives over the two week period. The formula for the number of draws is as follows:

[Enumerator: show the information sheet with the table below].

Total kwacha value of purchases reflected in formal tax invoices from machine printers	Number of draws
0	0
1-1000	1
1000-2000	2
2000+	3

After each draw we will return the card to the hat, so that the probability of winning is always 10%. You will only be able to win once, so if you are eligible for 2 draws and pick a yellow card on your first draw, the game will end.

To participate, I will need to ask for your mobile phone number, so that we can send you reminder texts and schedule our follow-up survey.”

If implemented at scale, a consumer incentive scheme would likely be an individual- rather than household-level program. We assigned treatment at the household level for two reasons. First, it would be natural for participants in the treatment group to aggregate invoices within the household, to increase their chance of winning the lottery. Simply asking them not to do that would be effective for some, but not for all. By encouraging both Treatment and Control households to bundle all invoices received by household members, we aimed to minimise differences in measurement error between the two groups. The second rationale for household-level treatments is that shopping patterns within a household are not independent of each other. If one person does the grocery shopping today, others in the household do not need to. The total effect of the incentive on purchases is best represented by its effect on the entire household unit, rather than on any specific individual within the household. This does not detract from the central goals of the study, which were to determine whether consumers react to these incentives and to learn about implementation challenges, but is one reason that the quantitative estimates of this small-scale pilot should not be treated as reliable estimates of the effect of the program at scale.

The enrolment script emphasised collection of machine-generated rather than hand-written VAT invoices. ZRA has expressed an interest in moving toward a single-accounting VAT system based on EFD-generated invoices. Future administration of any lottery scheme at scale would likely focus on machine-generated invoices, because they record lottery entries automatically, whereas paper invoices would have to be submitted and processed by hand.

Payments to lottery winners were made via mobile money within 1-2 days of winning.

### **Data collection and timeline**

The baseline survey and enrolment took place from December 1-8, 2021. The baseline survey took 10-15 minutes to complete, and included modules on household demographics, education, occupation, mobile money use, shopping patterns, and experience receiving receipts and invoices. For households assigned to treatment, it also included an explanation of the lottery game that consumers would be able to play, the rules of the game and the prize they were eligible to win.

The endline survey began two weeks after the baseline survey, and ran from December 15-24. The survey covered retained receipts and invoices, recent shopping experiences, and any challenges experienced while trying to collect invoices or identify VAT-registered firms. Treated households played the lottery

game during the endline visit. Enumerators also collected the retained receipts and invoices from respondents at this time.

From December 15-16 the endline survey was conducted in person. The omicron variant of the coronavirus then began to spread rapidly through Lusaka, forcing the team to switch to phone surveying. Survey team members returned to Kabwata on pre-arranged days for brief visits with the endline respondents who were interviewed by phone, to collect the receipts and play the lottery game (if the household was treated).

In January 2022, one of the survey team members received training from ZRA on how to identify a valid VAT invoice. This person then examined the receipts and invoices submitted by each study household, and recorded the total values of valid VAT invoices and other submitted receipts.

### **Empirical design and evaluation method**

To measure the impact of the incentive treatment on the request for and retention of VAT invoices, we will compare the average total value of invoices submitted by the Control and Treatment participants. Specifically, we will estimate OLS regressions of the following form:

$$Outcome_i = \alpha + \beta Treatment_i + Controls_i + \varepsilon_i \quad (1)$$

where  $Outcome_i$  is one of the four endline outcomes described below,  $Treatment_i$  is a binary variable that takes a value of 1 for members of the Treatment group and 0 for members of the Control group,  $Controls_i$  includes as control variables the type of endline survey (phone vs. in-person), household income, and shopping expenditure over the past week (at baseline), and  $\varepsilon_i$  is a statistical error term. With random treatment assignment, the estimated coefficient  $\hat{\beta}$  is an unbiased estimate of the Average Treatment Effect (ATE) of the incentive on the outcome variable. Inference will be based on robust standard errors.

We estimate equation (1) for four outcomes:

- i. The value of EFD-generated VAT invoices, verified by a trained enumerator
- ii. The value of invoices submitted by the respondent as valid VAT invoices
- iii. The difference between the value of verified EFD invoices and the value indicated by the respondent (we call this “respondent accuracy”)

The first outcome measures the incentive effect on actual retention of valid, EFD-generated VAT invoices. The second measures the incentive effect on the respondent's stated belief about the value of the valid VAT invoices that they collected for the project. The third outcome is a proxy for the accuracy or attention paid by the respondent to the validity of the submitted invoices.

## Findings

---

In this section we present three sets of findings. First, we provide a descriptive analysis of the study participants, based on data collected in the baseline survey. Then we present experimental estimates of the impact of the incentive on our three outcomes. Finally, we present descriptive statistics regarding shopping and VAT identification, from the endline survey.

### Baseline sample characteristics

A total of 337 households were successfully enrolled at baseline, with 168 assigned to Treatment, and 169 to Control. During the period between baseline and endline, the omicron variant of the coronavirus began to spread rapidly in Lusaka. The study team was forced to switch to a phone survey after two days of conducting the endline in-person. The team successfully reinterviewed 185 of the original 337 participants; of these, 77 were interviewed in person, and 108 were interviewed by phone. Attrition was slightly higher in the control group, with the team reinterviewing 101 of the 168 treated participants, and 84 of the 169 control participants. Multiple factors contributed to the high attrition rate, including concerns about the coronavirus, discomfort with phone surveys, the approaching Christmas holidays, and the absence of survey team members who became infected with the coronavirus or needed to take time off for family responsibilities.



**Table 1. Baseline characteristics and Treatment/Control balance**

	Control (N=169)	Treatment (N=168)	p-value of the difference
Respondent is female (=1)	0.62	0.63	0.94
Age of household head	43.69	42.68	0.50
Household size	4.94	4.85	0.71
Typical monthly household income			
Less than K1000 (=1)	0.09	0.07	0.44
K1000-K2000 (=1)	0.12	0.15	0.42
K2000-K3000 (=1)	0.14	0.2	0.14
K3,000 - K4,000 (=1)	0.14	0.21	0.08*
K4,000- K10,000 (=1)	0.37	0.25	0.02**
K10,000 - K20,0000 (=1)	0.12	0.11	0.75
More than K20,000 (=1)	0.02	0.01	0.42
Education of household head			
No formal education (=1)	0.00	0.01	0.16
Some Primary school (=1)	0.01	0.01	1.00
Completed primary (=1)	0.04	0.03	0.56
Some Secondary (=1)	0.09	0.12	0.36
Completed Secondary (=1)	0.15	0.14	0.78
Some university (=1)	0.39	0.40	0.79
Completed university or above (=1)	0.31	0.28	0.50

*Notes:* Authors' calculations from baseline survey data with 337 household heads.

Table 1 presents descriptive characteristics from the baseline survey. 62% of respondents were female, including both the heads of female-headed households and wives from male-headed households where the head was not available for the interview. The average household head is about 43 years old in both the treatment and control groups, and has income and education levels reflective of the middle- and upper-income status of the Kabwata neighbourhood. The average household has just under five members. The sample is generally balanced, with the only statistically significant difference between the treatment and control households arising from a slight compositional difference in the middle of the income distribution.

Table 2 describes participants' shopping experiences in the week before the baseline interview. Control households had shopped slightly more than treated households at formal and semi-formal businesses, making 3.85 purchases to

the treated groups 3.27 on average, and spending 1511 kwacha compared to 1182. Those differences are not statistically significant at conventional values, although the p-value for the Number of Purchases is just slightly over 0.1. Both groups of households made more purchases at informal than formal businesses over the last week, but spent substantially less, with average expenditure of 293 kwacha in both groups. The Control group reported a higher propensity to receive receipts or invoices from informal purchases, 17% to 2%, and that difference is statistically significant with 96% confidence.

**Table 2. Shopping experience over the last week (baseline)**

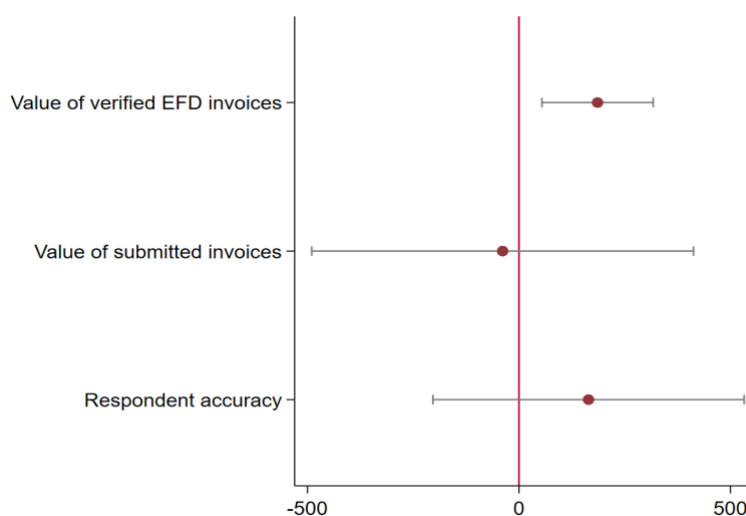
	Control (N=169)	Treatment (N=168)	p-value of the difference
Formal and semi-formal businesses			
Number of purchases	3.85	3.27	0.10
Number of receipts and invoices received	3.79	3.55	0.46
Total value of purchases	1511	1182	0.30
Informal businesses and vendors			
Number of purchases	5.07	5.6	0.46
Number of receipts and invoices received	0.17	0.02	0.04
Total value of purchases	293	293	1.00

The baseline differences in the recent shopping experience of control and treated households are the result of pure chance and the modest sample size. It is not possible to say whether these differences represent fundamental differences in the composition of the two groups. However, if the Control group has a tendency to shop a bit more than the Treated group, and is more likely to receive receipts or invoices in the absence of treatment (i.e., at baseline), then that would make it more difficult for us to detect a statistically significant impact of the incentive treatment. In that sense, our estimates can be thought of as lower bounds on the true treatment effects for this group. In the analysis, we control for baseline shopping expenditure and household income.

### Estimated treatment effects

The estimated treatment effects, with 95% confidence intervals, are plotted in Figure 1. The most important takeaway from Figure 1 is that the incentive was successful in inducing participants to retain and submit more EFD-generated invoices. The treatment effect on the variable “Value of verified EFD invoices” is 186 kwacha. The mean value of this variable in the control group is 200 kwacha, indicating that the incentive increased the submission of EFD-generated invoices by  $100 \times (186/200) = 93\%$ . Consumers seem to be able to successfully identify VAT-registered firms and EFD-issued invoices, at least to some extent.

Figure 1. Average Treatment Effects of the 2-week Incentive Program



The incentive had no effect on the total value of submitted invoices that the respondent believed to be machine-generated EFD invoices. The treatment effect on “value of submitted invoices” is close to zero, and the confidence interval includes a wide range of both positive and negative values. There is also no statistically significant effect on “Respondent accuracy”, which is the difference of the other two variables in Figure 1, although the point estimate is positive (indicating that the treated households overestimated the total value of the EFD invoices that they submitted by less than the control households did).

This pattern of findings is consistent with treated respondents increasing their purchases from businesses with EFDs by substituting away from other types of businesses. The fact that total expenditure does not increase in the treated group may also indicate that participants did not try to cheat at the lottery game by submitting invoices gathered from other households.

These findings are subject to some caveats. The first is that the rate of attrition was substantial in both the treatment and control groups. We do not know how the estimated effects might change if we were able to include all baseline households in the analysis. Second, the treatment effect on the value of EFD invoices is smaller in magnitude (though still positive) and not statistically significant if we do not control for baseline expenditure and income. While we believe it is appropriate to control for these variables in the analysis, as they represent a potential source of bias due to the relatively small sample size, in an ideal setting we would be able to rely on the randomisation and estimate treatment effects without involving additional control variables. Third, while the invoice verification process substantially reduced measurement error for our

main outcome variable, some receipts and invoices were faded or damaged by the time we were able to review them. If the value of faded and illegible invoices was different for the treatment and control groups—which it could be, if the treated respondents substituted away from businesses with low quality point-of-sale machines toward those using higher quality EFD invoices—then the treatment effect on “value of submitted invoices” may be biased. Finally, the subgroup targeted for this study consists of primarily middle and upper income households in Kabwata. Average treatment effects among this group may be different from those for the population at large.

These caveats notwithstanding, the finding of a statistically and economically significant effect on the value of submitted EFD invoices is a promising start for this line of inquiry.

### Endline descriptive statistics

One of the pilot goals was to identify any challenges encountered by participants when trying to shop at businesses that issue VAT invoices. In the endline survey we asked about this and related issues, such as whether consumers changed their shopping patterns, and whether they encountered resistance from sellers when they requested VAT invoices.

**Table 3. Shopping experiences during the study (endline)**

	Control	Treatment	p-value of the difference
How easy was it to identify businesses that issue VAT invoices?			
Very easy	0.75	0.71	0.62
Somewhat easy	0.16	0.13	0.59
Somewhat difficult	0.01	0.06	0.16
Very difficult	0.00	0.02	0.20
I did not try	0.07	0.06	0.73
Not applicable (no purchases made)	0.00	0.01	0.37
Changed timing, location, or types of purchases (=1)	0.09	0.08	0.92
Seller refused to give any receipt (=1)	0.09	0.05	0.33
Received handwritten receipt instead of requested VAT invoice (=1)	0.10	0.07	0.51
Seller offered to reduce price to deter VAT invoice request (=1)	0.00	0.00	
Has taxpayer ID number (=1)	0.78	0.77	0.90

Table 3 contains summary statistics for these endline variables, as well as the p-values from tests for statistically significant differences between the treatment and control groups. In Table 3 there are no statistically significant differences

between the two groups. However, treated households were more likely to report that it was “Somewhat difficult” or “Very difficult” to identify businesses that issue VAT invoices. Approximately 8% of treated households provided one of these responses, compared to only 1% for control households. When responses are aggregated this way, the difference between the groups is statistically significant (p-value of 0.06 is not reported in the table). This difference could indicate that treated households exerted more effort or paid greater attention to finding VAT-registered businesses.

Overall, the findings in Table 3 suggest that neither the treatment nor the control group encountered resistance from sellers or found it difficult to shop as desired. Across the two groups, 8-9% of respondents changed how they shop in pursuit of machine-generated invoices, 5-9% met a seller who refused to issue any form of receipt, and 7-10% received a handwritten receipt instead of the requested formal VAT invoice. It is not possible to tell whether these failures to issue invoices are indications of active deterrence or fraud on the part of sellers, or are simply reflective of the logistical challenges of having a working EFD or an official receipt book available at all times. No respondents in either group reported that a seller offered to lower the price in exchange for not issuing a VAT invoice. Consumers in prior work have reported that they encountered this form of bargaining from sellers, albeit infrequently. The difference may be due to the relatively small sample and narrow shopping window in the current study, which make it difficult to detect relatively rare events.

Finally, just over three quarters of endline survey participants report having a taxpayer identification number (TPIN). This bodes well for future programs that might require input of a TPIN to the EFD in order to enrol a buyer in a lottery, although we expect much lower rates of TPIN registration in lower-income neighbourhoods and in other parts of the country.

## Conclusion

---

The above findings are based on a relatively small sample from a middle- and upper-income neighbourhood of Lusaka, and cannot be taken as representative of other population subgroups. The coronavirus affected the study in various ways, contributing to a high rate of attrition and forcing us to switch to phone interviews during the middle of the endline survey. Nonetheless, the findings provide reason for cautious optimism about how consumers might react to VAT-related incentives.

The next step for this project are (i) to conduct a larger-scale pilot of the incentive program, involving a wider range of consumer types and varying the level of the incentive, and (ii) to explore the possibility of an EFD software update that would allow a subset of businesses to participate by entering TPIN numbers directly at the time of purchase.