

Technology, Taxation, and Corruption: Evidence from the Introduction of Electronic Tax Filing[†]

By OYEBOLA OKUNOGBE AND VICTOR POULIQUEN*

Many e-government initiatives introduce technology to improve efficiency and avoid potential human bias. Using experimental variation, we examine the impact of electronic tax filing (to replace in-person submission to tax officials) using data from Tajikistan firms. E-filing reduces the time firms spend on taxes by 40 percent. Further, among firms previously more likely to evade, e-filing doubles taxes paid. Conversely, evidence suggests that e-filing reduces tax payments among firms previously less likely to evade. These firms also pay fewer bribes, as e-filing reduces extortion opportunities. These patterns are consistent with differential treatment of firms by tax officials prior to e-filing. (JEL D22, H25, H26, O14, O23)

Technology is transforming the way governments function across various sectors (World Development Report 2016). From electronic public financial management systems to digital delivery of social programs and many other functions, these e-government initiatives typically seek to improve service delivery and efficiency. Often, they also aim to combat corruption by automating systems and reducing officials' discretion.

Tax administration is an important application of e-government in developing countries. Traditional tax systems in these countries are often characterized by high

*Okunogbe: World Bank Development Research Group (email: ookunogbe@worldbank.org); Pouliquen: University of Oxford (email: victorpouliquen@gmail.com). Erzo F.P. Luttmer was coeditor for this article. We are grateful to three anonymous referees for their helpful comments and suggestions that helped to improve this paper. We thank Emil Abdykalykov, Penelope Fidas, Alisher Isaev, Rustam Karimov, Vazha Nadareishvili, Dmitry Pyatachenko, Inomjon Sadulloev, World Bank Group (WBG) Tajikistan country office, and partners from the Tax Committee of Tajikistan for their invaluable cooperation. We also appreciate input provided by Alejandra M. Alcantara, Jean Marie Baland, Anne Brockmeyer, Ana Goicoechea, Marc Gurgand, Elise Huilery, Asim Khwaja, David McKenzie, Christopher D. Miller, Jennifer Murtazashvili, Paul Novosad, Owen Ozier, Dina Pomeranz, Simon Quinn, Gabriel Tourek, Liam Wren Lewis, and several seminar participants. The study protocol received approval from Harvard Institutional Review Board. The study was registered on the AEA RCT registry website AEARCTR-0000914 (Okunogbe and Pouliquen 2017). The research for this paper was funded by the Impact Program managed by the WBG Trade and Competitiveness Global Practice and the Impact Evaluation to Development Impact Trust Fund (i2i) managed under the WBG Development Impact Evaluation Unit. The findings, interpretations, and conclusions expressed in this paper, including any errors, are entirely ours and do not necessarily represent the views of the Tax Committee of Tajikistan, the World Bank Group, its Executive Directors, or the countries they represent.

[†]Go to <https://doi.org/10.1257/pol.20200123> to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

compliance costs and frequent interactions between tax officials and taxpayers that present opportunities for corruption. These features of the tax system may contribute to the observed low fiscal capacity in many developing countries (Gordon and Li 2009; Besley and Persson 2014) that limits their ability to provide services and growth-promoting public goods. A number of governments have responded by adopting electronic filing (henceforth, e-filing) of taxes. E-filing refers broadly to online submission of tax declarations, typically replacing in-person submission to tax officials.¹ As of 2015, 32 percent of developing countries had introduced e-filing, and its prevalence continues to grow rapidly (World Development Report 2016).²

E-filing may lower tax compliance costs by removing the need for time-consuming visits to the tax office. Further, by limiting tax officials' discretion in verifying submission of tax declarations, it may protect taxpayers from tax officials holding them up and extorting them. In addition, by reducing interactions between taxpayers and tax officials, e-filing may reduce collusion between the two parties and curb tax evasion.³ However, it is possible that e-filing may not deliver its expected benefits or may even lead to worse outcomes. When third-party reporting or other means of verifying income are limited, tax officials may have important private information gathered through frequent interactions with taxpayers, which they use in verifying filing submissions. Since e-filing removes this *ex ante* check, it may lead to lower tax revenues.⁴ Given the rising prevalence of e-filing and its ambiguous potential effects, it is important to understand its impact.

In this paper, we examine the impact of e-filing adoption on compliance costs, tax payments, and bribe payments using data from 1,498 small and medium-sized businesses in Dushanbe, the capital city of Tajikistan. We implement a randomized experiment with an encouragement design whereby we provide a randomly selected group of firms in-depth training on e-filing as well as logistical assistance with completing all e-filing registration requirements (Okunogbe and Pouliquen 2017). Firms in this group have to explicitly *opt out* of e-filing if they do not want to use it. Ninety-three percent of firms in this group adopt e-filing, compared with 60 percent in the comparison group where firms have to *opt in* and complete the registration process by themselves. We use this difference in adoption rates to estimate the impact of e-filing adoption on firms.

We find that firms that e-file because of our intervention save almost 5 hours on average every month, about 40 percent of the total amount of time spent on tax-related activities. Overall, we find no significant average effects on the amount of taxes paid or on bribe payments (measured using both a direct survey question and a list experiment). For taxes, the point estimate is positive, and we can rule out reductions of more than 14 percent of tax paid. For bribes, the point estimate is

¹Mail submission is also possible but less common in low-income countries with unreliable postal systems.

²E-filing use increases with national income. It is present in 85 percent of high-income countries and in 65 percent of middle-income countries (World Development Report 2016). The most common feature of tax reforms reported in the 2015 World Bank Doing Business Indicators was the introduction or enhancement of electronic systems for filing and paying taxes (World Bank 2016).

³An additional benefit of e-filing we are unable to examine is that it may produce system-wide efficiency gains in processing and analyzing tax information by removing the need for physical collection of forms and data entry.

⁴In contexts with limited information technology coverage, certain taxpayers may also experience significant costs of adoption rendering e-filing inaccessible to them (Yilmaz and Coolidge 2013).

negative, and we can rule out that the share of taxpayers who reported paying a bribe increased by more than 7 percentage points.

However, the absence of a statistically significant average effect masks heterogeneity across firms by their likelihood of tax evasion at baseline, measured using a risk profile score developed by the tax authority. We find that among firms with a risk score above the median, e-filing doubles tax payments. Conversely, we find suggestive evidence that among firms with a risk score below the median, e-filing reduces their tax payments. These results hold when controlling for observable differences between these two groups of firms and for other potential sources of impact heterogeneity, such as the firm size, sector, and use of technology. We also use machine learning methods from Chernozhukov et al. (2020) to confirm that the risk score is an important variable in explaining the impact heterogeneity of e-filing. Supporting the use of the risk profile as a proxy for baseline evasion, we observe that conditional on observable firm characteristics such as firm size, sector, and turnover, higher risk scores are correlated with lower tax payments.

These divergent impacts of e-filing adoption suggest that the nature of the interaction between firms and tax officials during in-person submission differs for firms with different risk profiles. In-person submission results in lower tax payments for firms with a higher risk score but higher payments for firms with a lower risk score. Using the available quantitative and qualitative evidence, we examine possible interpretations of these results.

Based on insights we gained from interviews, our interpretation is that under paper filing, firms with a higher risk score receive support from tax officials to reduce their tax liabilities. With e-filing, they lose this benefit, and their tax payments increase. This interpretation is supported by striking patterns in the selection of firms into e-filing adoption. Among the control group, controlling for firm characteristics, we find that a one standard deviation increase in a firm's risk score is associated with a 7 percentage point decrease in its likelihood of adopting e-filing. Among the treatment group, for whom adoption was almost universal, firms with a higher risk score are also 13 percentage points more likely to stop using e-filing within 6 months (22 percent against 9 percent among firms with a lower risk score). Over time, the positive impact on tax payment fades as more firms with above-median risk score revert to paper filing. In line with this, there is no significant impact of e-filing on bribes one year after the intervention.

For firms with a lower risk score, there are two potential explanations for the lower payments. First, under paper filing, officials may have used private information to enforce compliance of firms' true liability; with e-filing, firms experience less direct monitoring and evade more. Second, as some firms claim, officials may have required firms to pay beyond their true liability in order to meet their revenue targets; e-filing relieves this pressure. In addition, firms with a lower risk score pay fewer bribes as a result of e-filing, likely due to fewer opportunities for extortion given less frequent interactions with tax officials.

In sum, while the intervention does not meaningfully change the total tax revenue collected by the government, it changes the distribution of tax payments across firms by increasing the share of tax revenue coming from firms flagged as likely evaders.

This paper contributes to different areas of the literature. To the best of our knowledge, this paper provides the first causal evidence on the impact of e-filing (to replace in-person submission) on tax revenue. For the impact of e-filing on compliance costs and bribes, the only other study we are aware of is Kochanova, Hasnain, and Larson (2016), which uses variation in when different countries introduced e-filing to determine that e-filing reduces compliance costs and bribe solicitation. By using firm-level variation in e-filing adoption that arises from a randomized experiment, this paper avoids endogeneity bias that may be otherwise present.

More broadly, this paper adds to a growing body of work on the role of technology in enhancing state capacity. Recent studies have examined the impact of e-government initiatives in increasing transparency, reducing leakages, and monitoring compliance in different government functions ranging from procurement (Lewis-Faupel et al. 2016) to elections (Fujiwara 2015) to social programs (Muralidharan, Niehaus, and Sukhtankar 2016; Banerjee et al. 2020) and many others (Olken and Pande 2012; World Development Report 2016). In the field of taxation, recent papers examine the impact of electronic billing machines and the computerization of VAT invoices (Ali et al. 2015; Bellon et al. 2019; Fan et al. 2020). While much of the existing literature across sectors indicates positive effects, this paper highlights the fact that technology may have heterogeneous effects on users based on their compliance behavior prior to the introduction of the technology. Further, when adoption is voluntary, firms and individuals on whom the technology may have the greatest monitoring impact may be the least likely to adopt.

Lastly, this paper contributes to the rapidly growing literature on taxation in developing countries. It focuses on interactions between tax officials and taxpayers similar to Khan, Khwaja, and Olken (2016) and Amodio et al. (2021), who study the impact of performance pay for tax collectors on tax revenue and bribes. Whereas those interventions focus on the incentives of tax officials, this paper focuses on the level of interactions between officials and taxpayers.

I. Context: Tax Administration and E-filing in Tajikistan

Like many developing countries, Tajikistan⁵ faces significant challenges with improving tax administration and reducing corruption, making it an interesting setting to study the introduction of e-filing.

Firms face high tax compliance costs: small and medium firms must file monthly declarations for income tax withholdings on employee salaries and social insurance tax for employees. They also file quarterly declarations for turnover tax.⁶ In the absence of e-filing, firms submit their tax declarations in person at local tax offices, spending otherwise productive time waiting in line for multiple checks and signatures from tax officials. On average, firms in the study sample report spending six hours each month fulfilling tax obligations (excluding accounting), with about three

⁵With a 2015 GDP per capita (current US\$) of \$919, Tajikistan is classified as a low-income country in Central Asia (World Development Indicators 2015).

⁶Firms in the study sample are drawn from the simplified tax regime and as such are subject to a turnover tax rather than a corporate income tax. The tax rates are 6 percent of turnover, 25 percent of salaries for social insurance tax, and 13 percent of salaries for income tax.

hours going toward visits to the tax office. Furthermore, corruption is a major concern in Tajikistan. At the time this project commenced, the World Bank Enterprise Survey (Enterprise Surveys, The World Bank 2013) indicated that 32 percent of firms expect to give gifts in meetings with tax officials, and 37 percent expect to give gifts to any public officials to “get things done.” Opportunities for corruption in this context may be reinforced by the fact that tax officials are responsible for a portfolio of taxpayers, which means that they interact with the same taxpayers frequently.

These two concerns—improving service delivery by eliminating long wait times for submitting declarations and curtailing corruption by reducing interactions with tax officials—were among the primary reasons the tax authority introduced e-filing. By making it easier for people to file taxes and by closing off avenues for corruption, the government expected e-filing would increase voluntary compliance and thus boost tax revenues. Other motives for introducing e-filing were to improve the quality of tax records by reducing arithmetic mistakes by firms⁷ and data entry mistakes by clerks and to improve the efficiency of tax administration by releasing officials from routine work to focus on higher-value activities.

In order to file taxes online, firms must first register for e-filing by submitting application documents to the tax authority. After the application is approved, firms receive an e-token containing their digital signature, which they must use to verify online declarations. Once a firm registers for e-filing, it can file all taxes online using the e-filing software either on its own computer or at public e-filing terminals located in tax offices.⁸ However, firms are still required to visit the tax office quarterly to confirm that their payments are up to date.

Adoption of e-filing by firms was slower than expected following its introduction in 2012. Initially, firms had to purchase e-tokens for \$40 and pay an additional \$40 registration fee. Even after the elimination of these fees, only about 30 percent of firms registered. The low take-up was puzzling given the anticipated benefits from adoption, prompting significant interest in understanding the constraints to adoption and potential ways to address them. Focus group meetings and interviews with both business owners and tax officials indicated firms were not using e-filing for a variety of reasons, including lack of awareness, lack of trust in the system’s reliability and in the security of information submitted online, difficulties navigating the registration process, and lack of access to computers and Internet.⁹ In addition, some firms indicated that they did not file online because they preferred to deal directly with the same tax official on a regular basis when submitting their tax declarations (perhaps for benign reasons such as having someone trusted cross-check their submissions, or perhaps intending to evade tax obligations).

⁷The online system has inbuilt formulas to perform calculations automatically based on inputted entries.

⁸Online payment of taxes through e-banking is also available to firms but is separate from the e-filing system: firms that e-file do not have to pay online and vice versa. Although this paper does not study the impact of electronic tax payment, we observe that 66 percent of firms that e-file also pay online.

⁹From the baseline survey, 50 percent of firms in our study sample were unaware of the possibility of e-filing, and 37 percent did not have high-speed Internet on their premises.

II. Experiment Design

The experimental treatment arms of this study aimed to address those barriers to e-filing identified by firms. Firms were randomly assigned into two treatment groups and one control group (Okunogbe and Pouliquen 2017). In the intensive treatment arm (Group A), firms received training and information about e-filing intended to increase awareness and trust in the system. In the training session, they learned about e-filing availability, its benefits, and registration procedures and participated in an interactive demonstration of the e-filing system. In addition, these firms received logistical support in registering for e-filing to mitigate the hindrances firms willing to use e-filing face in accessing the system: a representative of the implementing partner helped firms complete all the steps required for registration. Firms not willing to register for e-filing had to explicitly “*opt-out*” of the program. Firms in the second treatment arm (Group B) received an identical e-filing training session, but they did not receive the logistical help for registration. Firms willing to register had to “*opt-in*” and complete the registration process by themselves. In the control group (Group C), firms did not receive any e-filing training but were also free to register on their own.

To hold the delivery format of the treatments constant, all firms in the three groups received a general training on taxation not specific to e-filing. Rather, the training included a review of different tax laws and procedures. Due to a requirement by the Tax Committee, this general tax training included one statement about the availability of e-filing on a slide listing the three modes of filing taxes: “by paper, by mail and electronically.” Firms in Group C would also have become aware of the existence of e-filing from a reference to it in the invitation materials and from some questions in the baseline survey. This observation is important for interpreting our results since it means Group C is not a pure control group but rather a group with a light information treatment.

This design allows us to answer two sets of questions. The first set examines firms’ decisions to adopt e-filing, specifically, the impact of information and training about e-filing, the additional impact of helping firms to register, and other firm characteristics that predict e-filing adoption. The second set of questions focuses on the impact of using e-filing on firms, in particular, their compliance costs, tax payments, and bribe payments.

A. Sampling and Randomization

The study draws from the universe of firms in Dushanbe that are registered in the tax authority database. All legal entities and individual entrepreneurs that are (i) simplified tax regime payers, (ii) have been active in the system for at least two years (i.e., not new enterprises or liquidated ones), and (iii) not currently e-filing were eligible for the study. There were 5,218 firms in the tax database meeting these three criteria.

We randomly selected a list of 2,004 firms from this overall population with stratification on status of the firm (legal entities or individual enterprises) and tax district. Based on discussions with the tax authority and the implementing partner

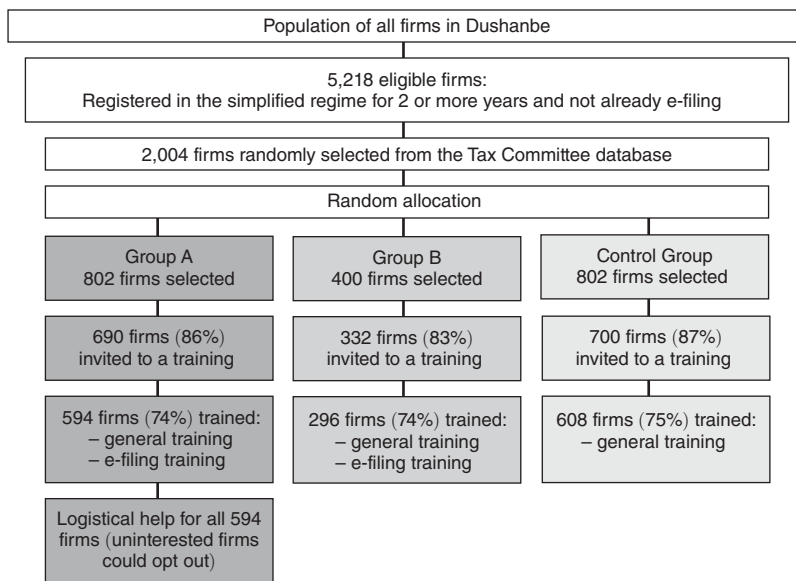


FIGURE 1. STUDY DESIGN AND PROGRAM IMPLEMENTATION

Note: The company implementing the program was contracted to train 1,500 firms and stopped inviting firms once it reached this target.

on expected response rates, we estimated that we needed to contact 2,000 firms to have 1,500 firms attend the training sessions. Since we expected the intervention to be more effective on legal entities, which are usually bigger firms than individual enterprises, we oversampled legal entities to achieve a relative proportion of 75 percent of legal entities to 25 percent of individual enterprises in the study population. We randomly allocated the 2,004 firms into 3 groups: 802 into Group A, 400 into Group B, and 802 into Group C.¹⁰ We conducted the randomization by computer using STATA and stratified using legal status, sector of activity, and tax district. Figure 1 illustrates the experimental design and the sampling strategy.

B. Program Implementation

The training programs and logistical support were delivered by a Dushanbe-based firm with the support of the tax authority from October to December 2014. Firms were invited to attend a general training on taxation through telephone calls by trained operators.¹¹ Although we assigned firms randomly into the treatment and control groups before they were invited (to facilitate scheduling into the different training sessions), all firms received invitations following exactly the same script

¹⁰ More firms were allocated to Groups A and C than to Group B in order to increase our statistical power to measure the impact of e-filing on firms adopting it since we expected the Group A treatment to have the greatest impact on e-filing adoption.

¹¹ The tax authority periodically organizes events and trainings for firms, so this program was not unusual.

(see online Appendix A for the text of the script). Out of the 2,004 firms in the study database, 1,722 (86 percent) received telephone calls and 1,498 (75 percent of the full sample and 87 percent of those invited) attended training. The company implementing the program was contracted to train 1,500 firms and stopped invitations once that number was reached. These response rates were balanced across the three groups. Since the invitations to all firms followed the same protocol, each firm had to decide whether to attend the training before learning its treatment group. Therefore, we are confident that the decision to attend training is orthogonal to treatment status and did not introduce selection bias into the study design. As such, we use the sample of the 1,498 firms that attended training to study the program's impact, with the caveat that this could slightly limit the external validity of our findings.

The training sessions took place either on the tax authority's premises or in the implementing partner's office. Groups A and B received identical training, which included both a general tax presentation and an e-filing presentation and demonstration, with a question-and-answer session at the end. The training for Group C (control group) differed in that it did not include the e-filing presentation and demonstration. On average, training lasted for two hours in Groups A and B and for one hour in Group C. A few days after each training, the implementing partner called back all firms in Group A and assisted firms in registering for e-filing unless the firm opted out. This logistical assistance involved visiting each firm, collecting all necessary documentation, submitting it to the tax authority on the firm's behalf, obtaining the e-token, and installing the software. In contrast, firms in Groups B and C had to *opt in* and register by themselves if they wanted to use e-filing.

III. Conceptual Framework

In this section, we outline potential impacts of e-filing on a firm and how these expected impacts may affect a firm's decision to adopt e-filing. Where relevant, we highlight differences between firms by their baseline tax evasion profile (see Figure 2 for a summary). We expect that compliance costs will be lower for all firms under e-filing since there is no longer any need to spend time traveling to the tax office (except for those using public terminals) and obtaining approval.¹² Further, we expect that the greater the potential savings in compliance costs from e-filing, the more likely a firm will be to adopt e-filing (for example, firms with high paper filing compliance costs).

Under paper filing, tax officials can influence the amount of tax declared by a firm. Each firm is assigned to a tax official who reviews the firms' declarations and signs before the declaration is considered accepted. Given their authority to accept or reject tax declarations, tax officials can insist that firms increase their reported tax liability. They can also help firms minimize their liability. Tax officials face two sets of incentives that may affect the way they exercise their discretion in reviewing

¹²E-filing does not necessarily reduce compliance costs for firms in all contexts. For example, if firms submit both electronic and paper documents due to legal requirements or lack of trust in the system (Yilmaz and Coolidge 2013), they may indeed face higher compliance costs. However, we do not think this should be the case in our context given that it is not technically possible to submit a paper and an electronic declaration simultaneously.

	Expected impact of e-filing on:		
	Compliance costs	Unofficial payments	Tax payments
Collusive firms	Decrease	Decrease (collusion)	Increase
Proxy: Higher risk scores	Taxpayer spends less time on travel and on queues	Official and taxpayer have less opportunity to collude	Official and taxpayer have less opportunity to collude
Noncollusive firms	Decrease	Decrease (coercion)	Decrease
Proxy: Lower risk scores	Taxpayer spends less time on travel and on queues	Official has less opportunity to hold up taxpayer	Official is less able to require taxpayer to increase tax paid

FIGURE 2. CONCEPTUAL FRAMEWORK

firms’ declarations. On the one hand, tax officials face a monthly tax revenue target that is expected from firms assigned to them.¹³ This revenue target provides an incentive to maximize the tax collected from firms. On the other hand, tax officials may have private incentives to receive unofficial payments from taxpayers in return for helping them reduce their tax liability.

Certain other channels are not available for the tax officials who receive declarations to influence firms’ tax payments. For example, audits are centrally determined and are conducted by a different set of tax officials. In addition, tax officials do not directly handle tax payments—after the declaration is submitted to the official, the firm then makes the tax payment at the bank or through electronic transfer.

Given these two opposing incentives faced by tax officials in how to exercise their discretion over firms’ tax submissions under paper filing, we expect that the impact of e-filing on tax payments will be ambiguous, as it depends on firms’ experiences prior to e-filing. On the one hand, e-filing may lead to lower tax payments for firms that tax officials had previously prevented from declaring taxes they considered to be too low. The review during in-person submission may serve as a mini-audit and help to limit evasion, as the tax official may have private information about firms that may be useful in estimating how much taxes the firms should pay. Alternatively, this check could be an opportunity for tax officials to force firms to declare more than their actual tax liability. In interviews and focus group discussions, some firms mentioned that a major consequence of e-filing was the ability to avoid officials forcing these (potentially unjustified) increases in tax payment.¹⁴ In either case, for firms in this category, e-filing may lead to lower taxes paid since tax officials will no longer be able to

¹³Due to data limitations, we do not observe which firms are assigned to the same official.

¹⁴Some examples of comments are: “[Tax inspectors] do not care whether we make a profit or how much was our turnover during the reporting period, they insist on an amount [they want us to pay] to fulfill their targets. So when filing in person, sometimes they reject a declaration if the amount is too small. Now, by filing online, you don’t depend on tax inspectors.” “After moving to e-filing, we’re more certain of the amount we need to pay. Before, during paper-filing, each time we didn’t know what figures the tax inspector will want us to pay for that filing period.” “I was tired of always arguing with my tax inspectors on the amount of taxes to be paid. They always insist on me filing a certain amount, which did not seem to depend on my turnover, which was usually about half of the figure I was asked to submit.”

influence the amount of taxes firms can declare at the time of submission. We expect that firms in this category will be more likely to adopt e-filing.

On the other hand, e-filing may lead to an increase in firms' tax payments for firms that actively collude with tax officials to receive assistance in completing their declarations to minimize their tax liability. The introduction of e-filing increases the transaction costs for this collusion because forms submitted online are not easily modifiable (unlike paper declarations that can be easily redone). Since e-filing may lead to higher taxes paid, we expect that these firms will be less likely to adopt e-filing. Importantly, these two channels may be self-reinforcing because, given the revenue targets, if a tax official is colluding with some taxpayers to reduce their tax payments, he may need to compensate by collecting higher tax revenue from others.

We hypothesize that, under e-filing, firms will pay fewer bribes to tax officials because the less frequent interactions between taxpayers and tax officials provide fewer opportunities for corruption.¹⁵ Qualitative evidence from focus group discussions and interviews indicate that at least two different types of bribe payments may occur during interactions between taxpayers and tax officials: (i) coercion (tax officials demanding bribes or creating administrative hurdles to put firms in a position such that they would voluntarily offer a gift) and (ii) collusion (tax officials and taxpayers working together to help taxpayers evade taxes, usually with a bribe for tax officials). Bribe coercion will be more difficult under e-filing since officials will be less able to hold up taxpayers, causing them to submit their declarations late and incur fines. In addition, collusion may be more difficult to coordinate if submissions are now done electronically. For firms that voluntarily make bribe payments to reduce their tax liability, their likelihood of e-filing will depend on the combination of savings from both bribe payments and tax payments. Since e-filing reduces but does not eliminate interactions with tax officials, these impacts may be muted if tax officials and/or taxpayers adjust their behavior to compensate for fewer opportunities to interact.

Since we are unable to directly observe tax evasion and bribe payments, it is difficult to know which firms are colluding with tax officials to evade taxes. In subsequent analyses of the heterogeneous impact of e-filing, we distinguish firms using a risk index developed by the tax authority to identify firms by their likelihood of tax evasion. Section IV (Data) provides a description of this index. As such, we expect that firms in the first category (noncollusive firms) will have lower risk scores, as they will be less likely to be evading taxes, whereas firms in the second category (collusive firms) will have higher risk scores.

IV. Data

We rely on three main sources of quantitative data. First, we use administrative data from the tax authority on firm characteristics (such as number of employees, legal status, industry, and district), e-filing use,¹⁶ and tax behavior (such as monthly

¹⁵Sequeira and Djankov (2014) find significantly higher levels of customs corruption in Maputo ports relative to Durban ports. Online processing of clearance documentation in Durban results in minimal interactions between clearing agents and customs agents compared to higher levels of interaction with in-person submissions in Maputo.

¹⁶We have monthly data on e-filing use from January to June 2015 but only a one-time measure from July to December 2015.

tax payments). These data include all 1,498 firms in the study sample as well as the additional 506 firms originally selected but that did not participate in the training. The other two sources are the baseline survey of firms conducted prior to the trainings and the endline survey conducted one year later (Okunogbe and Pouliquen 2022). The surveys include information on firm characteristics and economic behavior as well as experiences of firms with taxation (such as compliance costs).

The baseline survey was relatively short and self-administered by firm representatives at the beginning of training (before revealing treatment status) with detailed instructions and examples provided by implementing partner trainers. All 1,498 firms in the study sample completed the baseline survey. The endline survey had more questions and was administered by enumerators in person at the firm's premises. Eighty-four percent of the study sample (1,263 firms) completed the endline survey. Twelve percent of the sample did not complete the survey because the firm was liquidated, while the remaining 4 percent proved untraceable or had moved outside Dushanbe. There were no significant differences in attrition across treatment groups at the endline survey (online Appendix Table A1). In addition to these data, the study relies on extensive interviews and focus groups with tax officials and firms at different stages of the project to design the intervention, understand potential channels of impact, and interpret findings.

Table 1 shows summary statistics for some variables from the administrative data and the baseline survey. Firms have on average three employees (including the entrepreneur). In 40 percent of firms (the modal frequency), the entrepreneur is the only employee. About 42 percent each are in the trade and services sector. Women own 8 percent of firms. Table 1 also shows that randomization achieved balance across the different treatment groups for most variables.¹⁷ The *p*-values of joint orthogonality tests obtained from regressing the treatment variable on firm characteristics also indicate that the variables are not systematically correlated with the treatment. Since the analysis will emphasize heterogeneous results by firms' baseline risk score, Table 1 also includes randomization checks for the two subgroups (firms above and below the median risk score). As Section VI describes, the information and training only treatment has no significant impact on e-filing adoption; therefore, Table 1 combines Groups B and C as the control and compares with Group A. Online Appendix Table A2 provides randomization balance tables for the original three groups. Analogous to Table 1, online Appendix Table A3 shows that randomization balance persists in the sample surveyed at endline, consistent with the nondifferential attrition across groups.

We measure firms' tax compliance costs using the number of visits to the tax office in the last six months, the amount of time spent on different activities during visits to the tax office,¹⁸ and the amount of time spent calculating taxes and completing tax forms.

¹⁷ Firms' compliance costs are a notable exception. Firms in Group A visited the tax office 6.53 times compared to 6.4 times in the control group. Results from the endline survey show a reduction in the number of visits in Group A, so, if anything, the results may underestimate the reduction in compliance costs due to e-filing.

¹⁸ We ask firms separately about each of the following: Time to travel (both directions); Staying in line; Meeting with tax inspector to check reports and submission; Staying in line to sign tax reports by a head of tax office; Obtaining tax payments reconciliation act; Calculation of tax due and filling in tax reports at tax office; and Other.

TABLE 1—DESCRIPTIVE STATISTICS AND BALANCE CHECKS

	All sample		Below median of 2014 risk profile score		Above median of 2014 risk profile score	
	Mean [SD] control group (B&C) (1)	Diff. in Group A (2)	Mean [SD] control group (B&C) (3)	Diff. in Group A (4)	Mean [SD] control group (B&C) (5)	Diff. in Group A (6)
<i>Panel A. Administrative data from Tax Committee (2014)</i>						
Legal entities ^β	0.729 [0.445]	—	1 [0]	—	1 [0]	—
Sector of activity is Trade ^β	0.415 [0.493]	—	0.374 [0.485]	—	0.248 [0.432]	—
Sector of activity is Services ^β	0.42 [0.494]	—	0.495 [0.501]	—	0.582 [0.494]	—
Sector of activity is Manufacturing	0.116 [0.321]	0.012 (0.011)	0.1 [0.301]	−0.003 (0.017)	0.113 [0.317]	0.012 (0.021)
Female owner	0.076 [0.266]	0.023 (0.015)	0.07 [0.255]	0.007 (0.023)	0.048 [0.215]	0.032 (0.024)
No employee	0.406 [0.491]	−0.01 (0.025)	0.195 [0.396]	0.003 (0.035)	0.469 [0.5]	−0.025 (0.045)
Number of employees	3.287 [3.443]	−0.127 (0.171)	4.271 [3.71]	−0.57 (0.306)	3.257 [3.781]	0.108 (0.331)
Risk profile score in 2014 ^λ	62.6 [28.9]	0.1 (1.7)	41.3 [10.5]	1.6 (0.9)	85.1 [24.7]	−0.3 (2.1)
Risk profile score in 2014 above median ^λ	0.486 [0.5]	−0.015 (0.031)	0 [0]	—	1 [0]	—
<i>Panel B. Baseline survey data (2014)</i>						
Firm has an accountant	0.73 [0.444]	0.028 (0.023)	0.799 [0.401]	−0.032 (0.037)	0.762 [0.427]	0.041 (0.036)
Share of technological practices implemented	0.538 [0.437]	0.024 (0.021)	0.639 [0.422]	−0.011 (0.038)	0.665 [0.402]	0.074 (0.036)
Number of visits to tax authority office in Jan.–June 2014	6.382 [0.948]	0.13 (0.056)	6.514 [1.027]	0.109 (0.101)	6.444 [1.111]	0.158 (0.106)
Time spent on tax-related activities during a typical month (hours)	6.185 [2.926]	0.131 (0.162)	6.26 [2.852]	−0.005 (0.241)	6.657 [3.195]	0.188 (0.359)
Number of times tax inspectors visited the company in Jan.–June 2014	1.324 [0.948]	−0.027 (0.054)	1.271 [0.871]	−0.044 (0.107)	1.199 [0.879]	0.022 (0.08)
Ever used e-filing (with another company)	0.14 [0.348]	−0.013 (0.018)	0.198 [0.399]	−0.037 (0.034)	0.177 [0.382]	0.000 (0.035)
Think that it is common for firms to make informal payments	0.18 [0.385]	0.004 (0.02)	0.255 [0.437]	−0.046 (0.035)	0.174 [0.379]	0.019 (0.035)
<i>p</i> -values of joint orthogonality tests ^α	0.585		0.483		0.568	
Observations	1,498		547		520	

Notes: Columns 1, 3, 5: Standard deviations presented in brackets. Columns 2, 4, 6: Coefficients and standard errors (in parentheses) from an OLS regression of the firm owner/firm characteristic on a treatment variable (Group A), controlling for strata dummies. β : Variables used for stratification. λ : Risk profile scores are only calculated for legal entities (Observations = 1,067). Column 2 assigns the mean value for legal entities to individual enterprises. α : *p*-values of joint orthogonality tests obtained from regressions of the treatment dummy on all the variables in the right-hand column (not used for stratification) and testing for all coefficients are jointly equal to zero.

To measure tax payments, we aggregate monthly payments from administrative records from the tax authority to obtain the total amount of tax paid over the course of the year. As an alternative measure, we use firms' self-reports of their

tax payments in two focal months of the year (June and December). Given the large number of zeros in the distribution of taxes paid, we also use an indicator variable for whether a firm makes any positive tax payments during the year, the number of positive payments made as well as the median tax amount paid in that year. Since firms are subject to a turnover tax, paying zero taxes throughout the year is suggestive of evasion.¹⁹

Our main measure of bribe payments in the endline survey is a list experiment (Kuklinski et al. 1997). In response to the same question about *how many* actions on a given list a company had taken in the last year to “solve or prevent problems with the Tax Administration,” half of the respondents receive a list of four options that are neutral and legal.²⁰ The other half receive a list of five options that includes “made unofficial payments” as an additional option. The difference between the average numbers selected by firms presented with the longer list and by those presented with the shorter list gives a measure of how common it is for firms to pay bribes. In addition, we use a survey question (asked at both baseline and endline) asking how often it is true that firms “in their line of business” typically have to make unofficial payments to tax officials and create a dummy variable for firms saying it is “always true” or “often true.”²¹

Risk Score.—We use the firm’s risk score at baseline as a proxy for its likelihood of evading taxes. Tajikistan’s Tax Committee uses a risk-scoring system to assess firm risk as part of its audit strategy. Risk-based audits are widely used by tax authorities across the world, as they help allocate limited resources to cases likely to yield significant revenue while reducing the monitoring burden on compliant firms, as they will be less likely targets for an audit. Khwaja, Awasthi, and Loeprick (2011) provide examples from dozens of countries’ experiences with risk-based audits. In Tajikistan, a firm’s risk score is the product of a proprietary algorithm incorporating observed firm characteristics and results of prior audits on other firms. Similar to most countries (both developed and developing), Tajikistan’s tax authority keeps confidential the details of the risk-scoring process to prevent taxpayers from engaging in strategic behavior in order to avoid an audit. That said, the risk criteria incorporated into basic models for small firms in different countries typically include, *inter alia*, the sector(s) of activity and business structure, quality of record keeping, past compliance behavior, cost ratios and deductions from gross revenue, and comparisons with industry and sector averages (Loeprick and Engelschalk 2011).

¹⁹ It is possible for a firm to legitimately have zero tax liability over the 12 months of our study if it has no employees (and thus has no salary withholdings or social insurance tax) and no revenue. Firms can have long periods of inactivity, as firms that wish to cease operations must continue to submit tax declarations until the tax authority conducts a final audit of the business, in many cases with a delay of several months. In our sample, 17 percent of firms make 0 positive payments. Of these, 92 percent report having no employees, and 24 percent are either liquidated or in the process.

²⁰ The four items on the short list are 1. Received help from trade associations, 2. Had detailed discussions with tax officials, 3. Provided additional documents, 4. Pursued court action.

²¹ The question is worded in this manner to avoid asking firms about their own behavior, as corruption is a sensitive topic.

TABLE 2—BASELINE CORRELATES OF RISK PROFILE SCORE

Dependent variables:	Risk Profile Score in 2014 (1)	Above median of Risk Profile Score in 2014 (2)
<i>Panel A. Administrative data</i>		
<i>Sector of activity: omitted = manufacturing</i>		
Trade	−6.161 (2.783)	−0.115 (0.048)
Services	−0.607 (2.565)	−0.016 (0.044)
Female owner	−1.924 (3.396)	−0.078 (0.061)
No employee	20.330 (2.147)	0.299 (0.036)
Number of employees	0.120 (0.240)	0.006 (0.005)
<i>Panel B. Baseline survey data</i>		
Firm has an accountant	3.353 (2.089)	0.041 (0.037)
Share of technological practices implemented	3.144 (2.030)	0.029 (0.038)
Number of visits to tax authority office in Jan.–June 2014	−1.029 (0.586)	−0.011 (0.013)
Time spent on tax-related activities during a typical month (hrs)	0.288 (0.271)	0.008 (0.004)
Number of times tax inspectors visited the company in Jan.–June 2014	0.514 (0.739)	0.003 (0.015)
Ever used e-filing (with another company)	2.063 (2.270)	−0.000 (0.041)
Think that it is common for firms to make informal payments	−2.594 (2.052)	−0.083 (0.036)
Observations	1,057	1,057
R^2	0.203	0.141
Mean dependent variable	63.069	0.487

Note: OLS regressions with standard errors presented in parentheses controlling for dummies for the tax district.

To investigate empirically correlates of the risk score in our setting, we regress firms' risk scores at baseline against different firm characteristics from administrative and survey data. Table 2 indicates that risk score is negatively correlated with operating in the trade sector and positively correlated with claiming to have no other employees besides the entrepreneur. In addition, online Appendix Table A4 shows that, controlling for observable firm characteristics such as industry, gender of owner, number of employees, and (survey data report of) turnover, firms with a higher risk score at baseline were 12 percentage points less likely to make any positive tax payments in 2015 (column 4), suggesting that the risk score is predictive of noncompliance.

The tax authority only calculates risk scores for legal entities (75 percent of the sample), so it is unavailable for individual entities.²² Nevertheless, from Table 1 we find that, within the two subgroups, most variables are balanced, with a few

²²Since randomization was stratified on legal status, this characteristic is perfectly balanced among the different treatment groups.

exceptions.²³ As with the full sample, p -values of the joint F -test indicate that the variables are not jointly correlated with the treatment group.

V. Empirical Specification

We use equation (1) below to examine the relative impact of the two treatments in promoting e-filing adoption as well as the firm characteristics associated with adoption:

$$(1) \quad D_i = \beta_0 + \beta_1 T_{A,i} + \beta_2 T_{B,i} + \alpha \mathbf{X}_i + \lambda \mathbf{S}_i + \epsilon_i,$$

where D_i is an indicator variable for whether a firm registers for and uses e-filing, $T_{A,i}$ and $T_{B,i}$ are indicators for the training with logistical help and the training alone treatments, respectively. Parameters β_1 and β_2 estimate the causal effects of receiving the two treatments respectively on adoption, and the difference between them, $\beta_1 - \beta_2$, estimates the differential impact of the provision of logistical support in addition to the training; \mathbf{S}_i is a vector of strata dummies (one dummy for each triplet of the following business characteristics): legal status (individual entrepreneur or legal entity), sector of activity (manufacturing, trade, or services), and tax district (four tax districts in Dushanbe) (Bruhn and McKenzie 2009). To examine firm characteristics associated with adoption, we include in equation (1) a range of firm-level variables \mathbf{X}_i , in particular, the measures of firms' compliance costs, bribe payments, and the proxy for their likelihood of evading taxes (as discussed above).

To assess the impact of e-filing, we use a variant of equation (1) above in which we replace D_i with Y_i , the outcome variables of interest, to estimate the effect of assignment into either of the two treatment groups (the intent-to-treat estimate). We control for baseline measures of outcome variables when available. In addition, we use equation (2) below to estimate the effect of e-filing on firms that adopted it as a result of the program by using assignment to Group A (the intensive treatment arm) only as an instrumental variable (IV) for adopting e-filing (Local Average Treatment Effect estimate) since assignment to Group B has no meaningful effect on e-filing adoption. As such, in the IV estimates, the effective control group consists of Groups B and C (and not Group C alone).²⁴ Formally,

$$(2) \quad Y_i = \beta_0 + \gamma \widehat{Efile} + \pi Y_{i,t=0} + \lambda \mathbf{S}_i + \epsilon_i,$$

where Y_i are outcome measures such as compliance costs, bribe payments, and tax payments. $Y_{i,t=0}$ are baseline measures of the outcome variable.

We assess the differential impact of e-filing by firms' risk profiles (above and below the median of the tax authority risk score) using equation (3) below, which includes an interaction of e-filing and an indicator for a firm having below-median

²³ Among the low-risk score group, firms in Group A are smaller and have slightly higher risk scores (both at 10 percent level). We control for these variables in robustness checks and find no significant difference on our main results.

²⁴ The results are similar, but less precise, when we compare Group A to Group C alone (see online Appendix Table A5).

risk score (*LowRisk*) and an interaction of e-filing and an indicator for a firm having above-median risk score (*HighRisk*). Here, the instrumental variables are the interaction of assignment to Group A and the indicator for the respective risk profile category. Now,

$$(3) \quad Y_i = \beta_0 + \gamma \widehat{Efile \times LowRisk} + \alpha \widehat{Efile \times HighRisk} \\ + \rho HighRisk + \pi Y_{i,t=0} + \lambda S_i + \epsilon_i.$$

The terms γ and α are the coefficients of interest and give the impact of adopting e-filing among firms with a lower risk score and those with a higher risk score, respectively. For assignment to Group A to be a valid instrument in equations (2) and (3), it must affect outcomes only by inducing firms to e-file and not have a direct effect. For example, it must not be the case that receiving a visit by the implementing firm (as part of the logistical help) caused firms to think that their tax declarations were under greater scrutiny. While we are unable to explicitly rule out a direct effect, we think this is unlikely because firms were accustomed to receiving visits from the tax authority (over two times a year on average). Further, given the opposite pattern of results observed for high- and low-risk score firms, the logistical help would need to have affected these firms in opposite ways.

Since we examine the average impact of e-filing as well as the heterogeneous impacts by risk category, to reduce the risk of over-rejection of the null hypothesis due to multiple inference, we correct p -values for multiple hypothesis testing. We do that using the False Discovery Rate control method introduced by Benjamini, Krieger, and Yekutieli (2006) and described in Anderson (2008). This method provides sharpened two-stage q -values adjusted to account for the expected proportion of all rejections that are type I errors. We correct p -values obtained from the estimation of equation (3).

VI. Results and Discussion

A. Adoption of Electronic Tax Filing

By December 2015, about one year after program implementation, 93 percent of firms in Group A had registered for e-filing and used it at least once, relative to 63 percent in Group B and 59 percent in the control group. The combination of training and logistical support for registration was successful in increasing e-filing adoption by 34 percentage points relative to the control group (panel A of Table 3). On the other hand, there was no significant difference between the adoption rates for firms in Groups B and C, indicating that e-filing training and demonstration did not meaningfully promote e-filing adoption compared to the brief mentions of e-filing the firms in the control group encountered.

The large (30 percentage point) difference between the impact of the treatments in Groups A and B indicates that the logistical help with registration addressed an important constraint to e-filing. This may be due to a number of reasons, including reducing the hassle costs of navigating a complex registration process (Currie 2006)

TABLE 3—IMPACT ON E-FILING ADOPTION

	Mean [SD] control group	Difference between control group and [...]		Observations	<i>p</i> -values of the test Group A = Group B
		Group A	Group B		
<i>Panel A. Administrative data from TC (August 2014–December 2015)</i>					
Used e-filing	0.59 [0.492]	0.337 (0.023)	0.038 (0.033)	1,498	0
Used e-filing conditional on survival	0.641 [0.48]	0.34 (0.022)	0.042 (0.035)	1,275	0
Still using e-filing in Dec. 2015	0.548 [0.498]	0.238 (0.027)	0.046 (0.034)	1,498	0
<i>Panel B. Endline survey data (February 2016)</i>					
Firm used electronic filing to submit tax reports in 2015	0.564 [0.496]	0.439 (0.022)	0.038 (0.037)	1,263	0
Found out about e-filing during intervention training	0.796 [0.403]	0.202 (0.018)	0.054 (0.029)	1,263	0
Found out about e-filing from business network	0.17 [0.376]	−0.169 (0.017)	−0.047 (0.027)	1,263	0
Found out from another source (other training, Tax Committee publication ...)	0.033 [0.18]	−0.033 (0.008)	−0.007 (0.013)	1,263	0.011

Notes: Column 1: Standard deviations presented in brackets. Columns 2–3: Coefficients and standard errors (in parentheses) from an OLS regression of the firm owner/firm characteristic on treatment dummies, controlling for strata dummies.

or helping firms overcome procrastination. One additional possibility is that since all the logistical steps for registration were covered for firms in Group A, they had to opt out of the system if they did not wish to use e-filing. Firms may have assumed that opting out would be perceived as a signal that the firm was deriving (unofficial) benefits from paper filing.

The lack of a significant difference between Group B and the control group could be because neither of the two interventions had any impact on firms or because control group firms’ limited exposure to information on e-filing had effects as strong as the e-filing training. We find evidence that the second possibility is more likely. Indeed, 80 percent of firms in the control group indicated that they had found out about e-filing from the general tax training session they attended,²⁵ only 5 percentage points lower than the share of Group B firms that found out about e-filing through the detailed training sessions (panel B of Table 3).²⁶

We examine whether firms’ compliance costs, experiences with bribes, and their propensity to evade taxes affect their likelihood of adopting e-filing as described in the conceptual framework. Since almost all firms in the treatment group adopt

²⁵The other main source of information about e-filing for the control group was business networks (17 percent).

²⁶In addition, we also compare the study treatment groups to two groups of firms not included in the study: firms not contacted at all after reaching the required number of firms and contacted firms that declined to participate in the training (these two groups are not necessarily random). In both groups, the e-filing adoption rate is about two-thirds of that of the control group at 39 percent and 44 percent, respectively, suggesting that the brief mentions of the availability of e-filing in the control group had a significant effect not dissimilar to the effect of detailed training on e-filing procedures and demonstration.

TABLE 4—DETERMINANTS OF E-FILING ADOPTION IN GROUPS B AND C

Dependent variable:	Firm used E-Filing		
	Group B & C all sample		Group B & C only legal entities
	(1)	(2)	
<i>Treatment variables</i>			
Group B (Training alone)	0.026 (0.033)	0.033 (0.033)	0.015 (0.040)
<i>Administrative data (baseline)</i>			
Female owner	-0.008 (0.063)	-0.008 (0.063)	0.010 (0.081)
No employee	-0.133 (0.040)	-0.092 (0.041)	-0.147 (0.052)
Number of employees	0.018 (0.005)	0.018 (0.005)	0.017 (0.005)
Standardized risk profile score in 2014 ^λ		-0.066 (0.018)	-0.056 (0.018)
<i>Survey data (baseline)</i>			
Firm has an accountant	0.059 (0.038)	0.062 (0.038)	0.039 (0.048)
Share of technological practices implemented	0.023 (0.043)	0.031 (0.043)	0.013 (0.049)
Number of visits to tax authority office in Jan.–June 2014	0.012 (0.017)	0.008 (0.017)	0.005 (0.016)
Time spent on tax-related activities during a typical month (hours)	0.010 (0.005)	0.010 (0.005)	0.011 (0.006)
Number of times tax inspectors visited the company in Jan.–June 2014	-0.016 (0.017)	-0.017 (0.017)	-0.039 (0.022)
Ever used e-filing (with another company)	-0.017 (0.050)	-0.014 (0.050)	-0.029 (0.052)
Think that it is common for firms to make informal payments	-0.007 (0.043)	-0.010 (0.043)	0.036 (0.047)
Observations	894	894	634
R ²	0.173	0.188	0.161
Mean dependent variable in control	0.120	0.134	0.115

Notes: Robust standard errors are in parentheses. Omitted group are firms in the control group that received a general tax training. Regressions include fixed effects for strata. λ : For individual entities, there are no risk profile scores, so missing values were replaced by the mean of the variable.

e-filing, we run equation (1) using only Groups B and C and present the results in Table 4 to show which firm characteristics are associated with e-filing adoption. The first two columns of Table 4 provide results for all firms, while the third column covers only legal entities. We find a positive association between e-filing adoption and number of hours spent on tax preparation, supporting the hypothesis that firms with higher compliance burdens under paper filing would be more willing to switch. In addition, we find that firm size is strongly associated with e-filing adoption: one-(wo)man businesses are significantly less likely to e-file, and, among firms with employees, e-filing increases with the number of employees, suggesting that firms with more complex taxes are more likely to adopt.

Importantly, Table 4 shows that firms with higher risk scores are significantly less likely to adopt e-filing. One standard deviation increase in the risk score is associated with a 6.6 percentage point decrease in a firm's likelihood of e-filing. This result is consistent with firms with a higher risk score preferring to deal directly with tax inspectors with whom they have ongoing relationships and are able to collude to pay less in taxes.

We analyze the differential selection into e-filing use by firm risk score by assessing complier characteristics and calculating the relative likelihood that compliers have a higher risk score (Angrist and Pischke 2008). This statistic, given by the ratio of the "first stage" among firms with above-median risk score to the overall "first stage" ($0.39/0.349 = 1.12$), shows that compliers are 12 percent more likely to have above-median risk score.²⁷ This selection pattern is consistent with the notion that the treatment for Group A firms resulted in some firms registering for e-filing that would otherwise not have registered because Group A treatment replaced a situation where firms could opt in to e-filing with a system where they had to opt out to avoid e-filing (and opting out may have sent a negative signal that they were evading).

Lastly, we find that other firm characteristics, such as having a female owner, prior exposure to e-filing, a firm's level of comfort with technology (measured by an index of whether a firm has high-speed Internet, uses e-mail for business communications, and maintains accounting records electronically), and having an accountant, do not predict e-filing adoption.

B. *Impact of Electronic Tax Filing*

This section examines the impact of e-filing on compliance costs, tax payments, and bribe payments. The top panel of Table 5 presents results for the overall sample, showing both intent-to-treat (ITT) and instrumental variable estimates, where assignment to Group A is an instrumental variable for e-filing use. The middle panel shows IV results for legal entities alone, while the bottom panel shows IV heterogeneous results for firms above and below the median risk score. This last panel includes the sharpened two-stage q-values adjusted for multiple hypothesis testing described previously.

We find a significant impact of e-filing on compliance costs (Table 5, columns 1 and 2). Firms that e-file as a result of the program visit the tax authority 1.4 times fewer each month. In all, e-filing reduces the time firms spend on tax-related activities by 4.7 hours a month. As such, e-filing does fulfill the intended goal of reducing tax compliance costs of firms. This impact is large in magnitude, as the time savings corresponds to 40 percent of the overall 12 hours firms in the control group report spending on tax-related activities. A "naïve" cost-effectiveness analysis comparing program costs to the amount of money firms save from lower compliance costs²⁸

²⁷ The "first stage" is the share of compliers and is given by the adoption rate in Group A minus the adoption rate in the control group. Online Appendix Table A6 shows the first stage for the two risk score categories. In our calculation of complier characteristics, for comparability, we use the overall first stage for only legal entities (0.349).

²⁸ A full cost-effectiveness analysis would account for other benefits such as savings in tax administration costs.

TABLE 5—IMPACT OF E-FILING ADOPTION ON MAIN OUTCOMES

Dependent variables:	# visits per month to tax authority office in 2015 (1)	Time spent monthly on tax-related activities in 2015 (hours) (2)	Tax paid in 2015 ^α (admin. data) (in TJS) (3)	At least one positive tax payment in 2015 (4)	Median of tax paid in 2015 ^α (admin. data) (5)	Tax paid in 2015 ^α (survey data) (in TJS) (6)	Think that it is common for firms to pay bribes (7)	Ever paid a bribe to a tax official (list exp.) (8)
<i>Overall sample</i>								
(ITT) Impact of assignment to Group A	-0.47 (0.03)	-1.58 (0.33)	1,859 (2,925)	-0.002 (0.019)	-50 (153)	1,701 (3,195)	0.023 -0.028	-0.039 (0.057)
(IV) E-filing impact (all firms)	-1.39 (0.06)	-4.71 (0.99)	5,721 (8,934)	-0.005 (0.058)	-155 (474)	5,049 (9,445)	0.07 -0.084	-0.117 (0.168)
Observations	1,263	1,252	1,498	1,498	1,498	1,263	1,263	1,263
Mean outcome control group (B&C)	0.79	11.7	28,245	0.847	1,605	25,305	0.613	0.075
Mean outcome Group A	0.33	10.08	30,152	0.843	1,551	26,928	0.632	0.02
<i>E-filing impact, legal entities sample</i>								
(IV) E-filing impact (legal entities)	-1.47 (0.08)	-5.07 (1.11)	7,043 (11,192)	-0.017 (0.071)	-216 (591)	5,059 (11,848)	-0.025 -0.093	-0.200 (0.188)
Observations	934	925	1,096	1,096	1,096	934	934	934
Mean outcome control group (B&C)	0.84	11.8	31,222	0.806	1,660	28,049	0.617	0.123
Mean outcome Group A	0.32	10	33,745	0.801	1,586	29,860	0.608	0.044
<i>Impact by baseline risk profile score</i>								
(IV) E-filing impact for below-median risk score	-1.64 (0.12)	-6.30 (1.72)	-15,930 (19,412)	-0.147 (0.077)	-1,896 (1,023)	-20,386 (19,076)	-0.236 -0.144	-0.587 (0.296)
Sharpened two-stage <i>q</i> -values	0.001	0.001	0.26	0.124	0.097	0.171	0.214	0.104
(IV) E-filing impact for above-median risk score	-1.33 (0.08)	-4.05 (1.44)	25,737 (12,970)	0.045 (0.102)	1,145 (671)	26,403 (14,721)	0.161 -0.119	0.154 (0.235)
Sharpened two-stage <i>q</i> -values	0.001	0.001	0.104	0.495	0.097	0.171	0.214	0.345
<i>p</i> -value diff. (low versus high risk score)	0.031	0.319	0.076	0.131	0.013	0.052	0.035	0.050
Observations	934	925	1,067	1,067	1,067	934	934	934
<i>Mean outcome in control group (B&C) for</i>								
Firms below median risk score	0.84	12.05	40,087	0.951	2,295	34,140	0.637	0.143
Firms above median risk score	0.84	11.52	23,751	0.701	1,090	21,529	0.596	0.098
<i>Mean outcome in Group A for</i>								
Firms below median risk score	0.33	10.03	36,314	0.908	1,785	29,072	0.569	-0.039
Firms above median risk score	0.32	9.97	32,671	0.722	1,455	30,733	0.652	0.139

Notes: The first two parts of the table present results from 2SLS regressions measuring the impact of e-filing adoption (instrumented by assignment to Group A) on outcomes presented in columns. The last part of the table presents results from 2SLS regressions measuring the heterogeneous impact of e-filing adoption (instrumented by assignment to Group A) by risk profile scores. For the heterogeneous analysis, samples are restricted to legal entities because risk scores are not calculated for individual enterprises. Column 8 presents results from a list experiment measuring bribe payments (see main text for a description). Robust standard errors are in parentheses. All regressions include control dummies for strata. α : Winsorized at the ninety-ninth percentile. Sharpened two-stage *q*-values corrected for multiple hypothesis testing as described in Anderson (2008).

(using the average wage of an accountant in our sample) shows that potential benefits from reducing compliance costs more than compensate program costs after seven months (online Appendix C and online Appendix Table A7 detail program costs and calculations).²⁹

²⁹ While it is possible for lower compliance costs to increase voluntary compliance (by increasing tax morale) or to increase profits (by reducing business costs), which could both lead to higher taxes paid, we find no evidence for improved attitudes on taxation issues, and the lower compliance costs are small relative to the tax amounts; hence, we do not emphasize these channels.

We do not find a significant impact of e-filing on tax payments but observe opposite effects for firms above and below the median risk score. The top and middle panels of Table 5 (columns 3–6) show no statistically significant results for the four different measures of taxes paid (total taxes for the year from administrative records, at least one positive tax payment, median tax paid, as well as self-reported tax paid). Standard errors are large, and we can only rule out reductions larger than 14 percent and increases greater than 27 percent of tax payments (column 3, ITT estimates).

As described in the conceptual framework, we expect that the impact of e-filing on firms will differ based on firms' experiences at baseline: firms for which meeting with a tax inspector served as a constraint on the minimum payments they could make would reduce their tax payments, whereas firms that received support from tax officials to reduce their tax liability would increase their tax payments. Our heterogeneous analysis of tax payments by firm risk of evasion at baseline provides evidence for this hypothesis, as shown in the bottom panel of Table 5.

All four measures of tax payments show a similar pattern of impact from e-filing: there is a decrease in tax payments among firms with a lower risk score and an increase in taxes paid among firms with a higher risk score. These effects are statistically significant for the median and making any positive tax payments among firms with a lower risk score (measures that are sensitive to having many zeros in the distribution) and for the mean, median, and self-reported survey mean among firms with a higher risk score.

These impacts are quantitatively large. Taking the total taxes for the year from the administrative data as an example, column 3 reveals an increase of TJS25,737 (US\$3,680) as a result of e-filing among firms with above-median risk scores ($p = 0.046$). It represents a 108 percent increase relative to the control group. In contrast, among firms with below-median risk scores, e-filing adoption results in firms paying TJS15,930 (US\$2,280) less in taxes (a 40 percent reduction relative to the control group), but this amount is not statistically significant (for this variable), and we can only rule out an increase in excess of 55 percent. The estimated impact on firms with above-median risk scores is significantly different from that for firms with below-median risk scores at the 10 percent level ($p = 0.076$). Considering whether the firm pays any taxes during the year, Table 5 (column 4) suggests that, among firms with a lower risk score, those that e-file due to the intervention are 15 percentage points less likely to pay any taxes on average ($p = 0.07$). For firms with a higher risk score, the coefficient is positive but not statistically different. These results are (broadly) robust to False Discovery Rate control adjustment for multiple hypothesis testing.³⁰ We discuss in Section VID the possibility that other differences between firms in the two risk score categories may be driving these results. We show that the results hold when we control for observable baseline differences between the two groups, suggesting this is not the case.

This result implies that while the intervention does not meaningfully change the total level of taxes paid, it changes the distribution of tax payments across firms.

³⁰For the positive impact among firms with a higher risk score, q -values are equal to 0.104 for the outcome tax paid in 2015 and to 0.097 for the median of tax paid. For the negative impact among firms with a lower risk score, q -values = 0.124 for making a positive tax payment and 0.097 for the median tax paid.

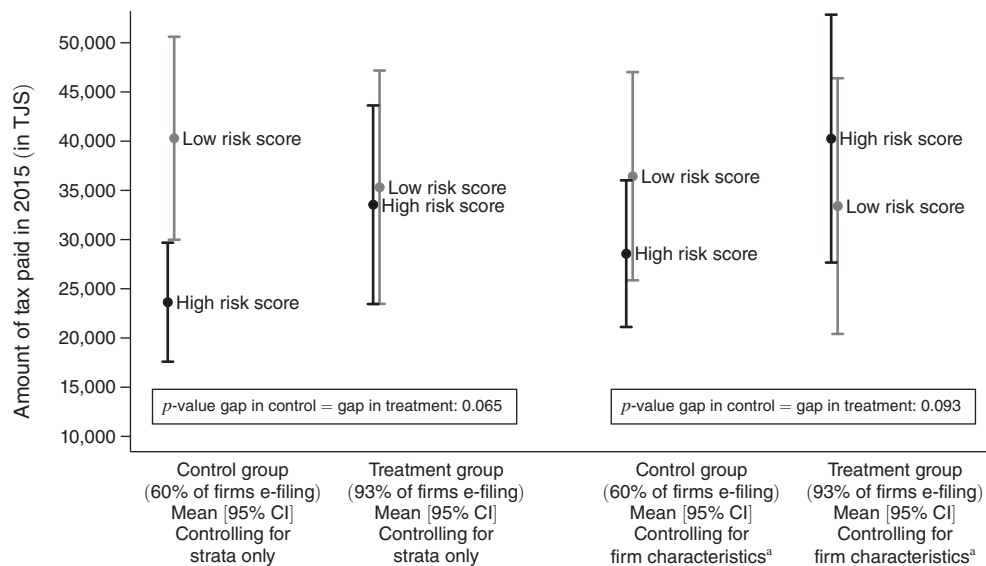


FIGURE 3. TREATMENT IMPACT ON TAX PAID BY RISK PROFILE SCORE

Notes: Number of observations: 1,067. Low (high) risk score refers to firms with a baseline risk profile score below (above) the median. 1 TJS = US\$0.14.

^aControlling for sector, no employee, technological practices, and time spent on tax activity.

Source: Administrative data from tax authority

Figure 3 displays graphically the overall impact of the intervention (ITT results). The first two columns of Figure 3 show that in the control group, firms with lower risk scores pay more tax than firms with higher risk scores but this gap is closed in the treatment group (p -value of the difference in the gaps = 0.065). This suggests that the intervention promotes horizontal equity by removing the tax gap between the two groups. However, in the third and fourth columns that control for firm baseline characteristics correlated with the risk score and across which the effect of e-filing could be heterogeneous,³¹ the increase in tax payments for firms with higher risk scores is large enough that the tax gap observed in the control group is reversed in the treatment group and low-risk score firms pay lower taxes than high-risk score firms (p -value of the difference in the gaps = 0.093).³² In all, the intervention results in an increased share of tax revenue coming from firms with higher risk scores at baseline.

Using our two measures of bribe payments, we find no overall impact of e-filing but observe different patterns by firm risk profile. For the survey question about how widespread bribe payments are, we find no overall impact either in the full sample

³¹ Section VID describes the rationale for the control variables included (sector of activity, no employee, share of technological practices implemented, and time spent on tax-related activities).

³² In online Appendix Figure A1, we show that this result is consistent when we use our other tax outcomes. We also conduct a simultaneous (seemingly unrelated) estimation that captures any correlation among our four different tax outcomes and obtain p -values for the test that the difference in tax gap are jointly zero of 0.069 without controls and 0.115 with controls).

or among legal entities (Table 5, column 7). We can rule out effects below a 3 percentage point reduction and above an 8 percentage point increase (ITT estimates), relative to a mean of 61 percent in the control group. However, we find a statistically different effect for firms with a higher risk score compared to firms with a lower risk score ($p = 0.035$)—we find a decrease of 23.6 percentage points (only marginally statistically significant with $p = 0.103$) among firms with a lower risk score and a positive impact of 16 percentage points (not statistically significant) among firms with a higher risk score (we can rule out effects larger than a 7 percentage point reduction).

Similarly, using the list experiment, we find opposite results for firms in the two risk score categories (p -value of difference = 0.05). With the list experiment, we measure bribe payments by comparing firms presented with the longer list (that includes making unofficial payments to prevent or settle problems with the tax authority) to firms presented with the shorter list. To compare treatment to control group in our regression framework, we interact our treatment variables with a dummy equal to one if the firm was presented with the longer list. Results are presented in the last column of Table 5 (more details on the list experiment are available in online Appendix Table A8). We find a large and statistically significant reduction in bribes (59 percentage points) for firms with below-median risk scores. However, for firms with above-median risk scores, standard errors are large, and we cannot rule out large positive or negative effects (the point estimate is a 15 percentage point increase).

C. Mechanisms

This section provides additional evidence to understand the divergent results by risk group.

For firms with below-median risk score, the lower tax payments under e-filing suggests that tax officials prevent firms from submitting tax declarations they consider too low. Switching to e-filing removes the scrutiny by a tax inspector at the point of submission of a tax declaration, and therefore firms become more likely to pay zero taxes. In the control group, firms with below-median risk score, on average, make positive payments more often than firms with above-median risk score (8.3 months a year in contrast to 5.1 times for firms with above-median risk score). Table 6 (column 1) shows that, among firms with a lower risk score, those that e-file pay zero taxes two more months in a year on average, but this effect is absent for firms with a higher risk score. There are at least two possible scenarios for this to occur. On the one hand, it could be that tax officials compel these firms to pay more taxes (including possibly more than their true liability) in order to meet their tax collection targets set by the tax authority. On the other hand, it could be that tax officials have valuable information about tax liabilities of these firms and can use this information to ensure higher tax payments.

While we are unable to rule out either of these channels given the available data, there is suggestive evidence that some of the effect is coming from the “forced payments” channel alongside the “information-enforcement” channel. First, as detailed in the quotes provided in the conceptual framework (footnote 14), during the interviews we conducted, a number of firms claim that officials were forcing them to

TABLE 6—IMPACT OF E-FILING ADOPTION: MECHANISMS

Dependent variables:	Administrative data					Survey data		
	# months with a positive tax payment in 2015 (1)	Tax paid ^a		At least one positive payment		Monthly time spent by type of task: (hours)		
		January–June 2015 (2)	July–December 2015 (3)	January–June 2015 (4)	July–December 2015 (5)	Submit tax returns and get reconcil. act (6)	Prepare doc. used for tax purposes (7)	Collate records (8)
<i>Overall sample</i>								
(ITT) Impact of assignment to Group A	−0.052 (0.234)	1,173 (1,830)	27 (1,568)	0.009 (0.02)	−0.021 (0.021)	−1.12 (0.06)	−0.45 (0.33)	−0.01 (0.14)
(IV) E-filing impact (all firms)	−0.159 (0.725)	3,608 (5,601)	84 (4,823)	0.028 (0.062)	−0.064 (0.065)	−3.33 (0.14)	−1.35 (0.97)	−0.04 (0.4)
Observations	1,498	1,498	1,498	1,498	1,498	1,252	1,252	1,252
Mean outcome control group	7.377	14,522	14,612	0.809	0.814	1.69	9.49	21.78
Mean outcome Group A	7.315	15,754	14,741	0.816	0.79	0.59	8.99	21.78
<i>E-filing impact, legal entities sample</i>								
(IV) E-filing impact (legal entities)	−0.421 (0.844)	4,237 (7,041)	225 (6,041)	0.017 (0.075)	−0.103 (0.08)	−3.54 (0.18)	−1.51 (1.09)	0.29 (0.47)
Observations	1,096	1,096	1,096	1,096	1,096	925	925	925
Mean outcome control group	6.545	16,011	16,432	0.76	0.769	1.84	9.44	22.26
Mean outcome Group A	6.442	17,564	16,647	0.769	0.732	0.6	8.89	22.36
<i>Impact by baseline risk profile score</i>								
(IV) E-filing impact for below-median risk score	−2.319 (1.218)	−14,200 (12,383)	−10,177 (9,969)	−0.163 (0.089)	−0.200 (0.095)	−3.9 (0.3)	−2.4 (1.7)	−0.7 (0.7)
Sharpened two-stage <i>q</i> -values	0.129	0.144	1	0.158	0.078	0.001	0.441	0.182
(IV) E-filing impact for above-median risk score	0.856 (1.069)	19,555 (8,181)	8,545 (6,966)	0.137 (0.103)	−0.073 (0.113)	−3.2 (0.2)	−0.8 (1.4)	1.2 (0.6)
Sharpened two-stage <i>q</i> -values	0.269	0.036	1	0.158	0.352	0.001	0.441	0.119
<i>p</i> -value diff (low vs. high risk score)	0.051	0.024	0.127	0.028	0.389	0.025	0.479	0.042
Observations	1,067	1,067	1,067	1,067	1,067	925	925	925
<i>Mean outcome in control group (B&C) for</i>								
Firms below median risk score	8.279	21,766	20,400	0.921	0.921	1.85	9.67	22.37
Firms above median risk score	5.109	10,900	13,238	0.637	0.656	1.82	9.18	22.13
<i>Mean outcome in Group A for</i>								
Firms below median risk score	7.665	18,233	17,893	0.876	0.858	0.61	8.9	22.14
Firms above median risk score	5.469	17,707	16,134	0.6940	0.632	0.58	8.870	22.61

Notes: The first two parts of the table present results from 2SLS regressions measuring the impact of e-filing adoption (instrumented by assignment to Group A) on outcomes presented in columns. The third part of the table presents results from 2SLS regressions measuring the heterogeneous impact of e-filing adoption (instrumented by assignment to Group A) by risk profile scores. For the heterogeneous analysis, samples are restricted to legal entities because risk scores are not calculated for individual enterprises. Robust standard errors are in parentheses. All regressions include control dummies for strata. α : Winsorized at the ninety-ninth percentile. Sharpened two-stage *q*-values corrected for multiple hypothesis testing as described in Anderson (2008).

overpay under paper filing. Second, the fact that some of the impact of e-filing on tax payment for firms with below-median risk score is driven by the extensive margin (panel B of Figure 4) suggests that the information channel is also likely to play a role. Indeed, since the simplified tax regime is based on turnover, tax inspectors may be able to easily verify whether a firm was economically active or not during the period, to determine whether its true liability is zero.³³

³³ It is also possible that even without relying on firm-specific information, tax inspectors may, in general, try to prevent zero declarations given the high probability of evasion.

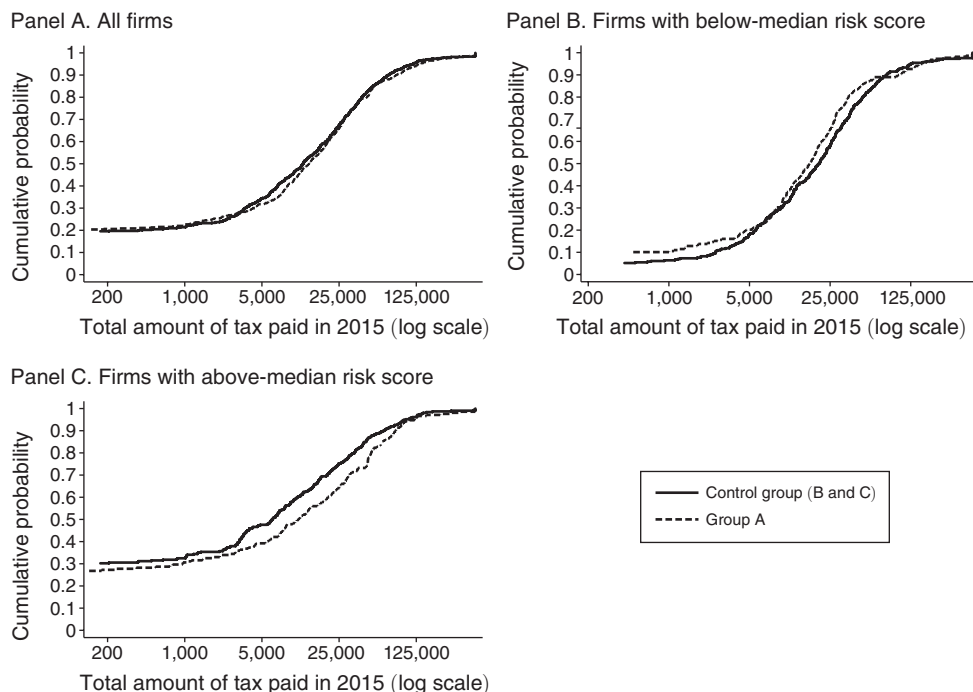


FIGURE 4. CDF OF TAX PAID IN 2015 FOR THE WHOLE SAMPLE AND BY RISK PROFILE SCORE

Notes: Panel A: Observations = 1,096. Panel B: Observations = 547. Panel C: Observations = 520. To avoid long left tail, the graph starts at the lowest positive value.

Source: Administrative data from tax authority

Turning now to the results for firms with a higher risk score, the increase in tax payments occurs on the intensive margin, that is, the amount of taxes paid, conditional on paying in a given month. Similarly, the cumulative distribution function shows that, among firms with a higher risk score, the higher tax payment among Group A firms relative to the combined control group occurs mostly on the middle to upper end of the distribution (panel C of Figure 4).

Our interpretation of this increase in tax payment is that firms with a higher risk score lose the opportunity to collude with tax officials to reduce their tax liabilities. From anecdotal evidence, one of the deterrents to evasion under e-filing is higher transaction costs for collusion because after a firm submits a declaration online, it is difficult to change it (unlike declarations filed on paper, which are easy to change). From the survey data, we find some additional evidence consistent with high-risk score firms becoming less able to rely on assistance from tax officials. We find that the amount of time firms spent monthly collating records *increases* by 1.2 hours among firms with a higher risk score, while among firms with a lower risk score there is no significant impact ($p = 0.042$ for the difference between the two coefficients) (Table 6, column 8), suggesting that firms with a higher risk score may now need to pay more attention to preparing their tax declaration submissions themselves.

The e-filing enrollment behavior of firms with above-median risk score provides additional insight into the differential impact on tax payment. As noted in the discussion of e-filing adoption, controlling for factors such as firm size and compliance costs, risk score is negatively correlated with e-filing adoption in the combined control group, suggesting that firms with high risk scores anticipated a negative impact from e-filing. In addition, we find evidence that firms in the treatment group are more likely to stop e-filing during the course of the year, an effect concentrated among firms with above-median risk score: controlling for observable firm characteristics, the dropout rate for firms with above-median risk score is 22 percent, relative to 9 percent for firms with below-median risk score (online Appendix Table A9).³⁴ In contrast, in the control group, the dropout rates for firms in both risk categories are similar at 3 percent. Given that e-filing led to an increase in tax payment for firms with above-median risk score, it is not surprising that they were more likely to stop using it. This finding is consistent with the notion that firms in Group A that anticipated negative impacts from e-filing may have avoided opting out explicitly when offered e-filing registration but instead dropped out (less conspicuously) after a few months.

In line with the decline in e-filing among Group A firms with above-median risk score, we examine whether the impact of e-filing on tax payments fades over time. It does. Whereas there is a strong and statistically significant increase in tax payments for Group A firms with above-median risk score from January to June, the analogous result for July to December falls to less than half and loses statistical significance (Table 6, columns 2 and 3). In contrast, given the more modest drop in e-filing among Group A firms with below-median risk score, the impact on making a positive tax payment from January to June is sustained, with similar magnitude to the impact from July to December (Table 6, columns 4 and 5).

Next, we examine the results on bribes. Our interpretation of the simultaneous reduction in both bribes and tax payments for firms with a lower risk score is that the unofficial payments were more likely extorted payments and not voluntary payments for collusion to pay lower taxes. If the unofficial payments were voluntary for collusive purposes, we would expect tax payments to increase when unofficial payments fall. Less frequent interaction with tax officials due to e-filing therefore reduces the opportunity for firms with a lower risk score to be forced to give bribes.

The abandonment of e-filing by Group A firms with a higher risk score also provides a potential explanation for the outstanding puzzle that we do not observe any reduction in bribe payments for high-risk score firms, unlike for low-risk score firms. Since the survey measuring this outcome took place in January/February 2016, one year after program implementation, we have only a one-time measure of this outcome, which occurs after a sizable proportion of the firms with a higher risk score have dropped out of e-filing. Since the impact on tax payments has dissipated by then, it is likely that any collusion arrangements have resumed, such that we see no negative impact of assignment to the treated group on the prevalence of bribes.

³⁴ Online Appendix Table A9 shows the differences in the dropout rate only among firms that registered for e-filing, and, as such, the coefficients are descriptive, not causal. The analysis continues only until June 2015 because the data from July to December do not report e-filing use for each month.

Consistent with this, e-filing has a significantly lower impact on number of visits to the tax authority for firms with a higher risk score compared to firms with a lower risk score (Table 5, column 1). One approach to examine this possibility is to conduct the same analysis on only firms that consistently e-filed. However, with our available data, we can only identify firms that dropped out in the first six months, so excluding only these firms would leave both firms that filed consistently and those that dropped out in the second half of the year.³⁵ Another approach is to examine the bribe behavior of those firms that dropped out. Controlling for firm characteristics such as size, sector, and location, the bribe rate among firms that stopped using e-filing is 29 percentage points higher ($p = 0.103$) than those that e-filed consistently (online Appendix Table A10). This pattern, combined with the higher dropout rate among treated firms with a higher risk score, provides some support for the hypothesis that, for firms with a higher risk score, dropouts are in part responsible for the lack of a significant impact of e-filing on bribes.

D. Robustness Checks

This section discusses potential concerns to a causal interpretation of the heterogeneous results obtained by baseline risk profile of firms.

Controlling for Other Sources of Impact Heterogeneity.—One potential concern is that, since the baseline risk profile score of firms is correlated with other firm characteristics (Table 2), if the impact of e-filing adoption is heterogeneous across these characteristics, then initial differences in these characteristics between firms in the two risk score categories may be driving our heterogeneous analysis by risk score.

To examine this possibility, we allow the effect of e-filing adoption to vary with other firm characteristics, following the approach in de Mel, McKenzie, and Woodruff (2009). To do so, we add additional covariates to equation (3) as well as interaction terms of these covariates with the e-filing indicator (instrumented by treatment assignment). We use covariates that are significantly correlated with the risk scores and across which the effects of e-filing adoption could plausibly be heterogeneous: *operating in the trade sector* (this sector of activity may benefit more from e-filing adoption), *declaring no employees* (one-(wo)man businesses may have simpler taxes and benefit less from e-filing), *tax district* (interactions with officials may vary across different tax offices), and *monthly amount of time spent on tax related activities* (firms with high tax compliance costs may benefit more from e-filing).³⁶

Results of these estimations are presented in Table 7 and show that our main results are robust to the introduction of these additional controls. The coefficient on high-risk score firms for total tax paid increases with all controls added (column 5).

³⁵We conduct this analysis and find a slightly smaller but not significantly different coefficient—from 0.052 ($p = 0.56$) to 0.019 ($p = 0.84$).

³⁶All variables are demeaned so we can still interpret our coefficients of interests as the impact of e-filing on low- and high-risk score firms, respectively, measured at the mean of the other variables.

TABLE 7—HETEROGENEOUS IMPACT BY BASELINE RISK PROFILE SCORE INCLUDING ADDITIONAL CONTROLS

	(1)	(2)	(3)	(4)	(5)
<i>Part I: Dependent variable: Tax paid in 2015^α</i>					
(IV) E-filing impact for below-median risk score	-17,753 (22,213)	-13,281 (17,939)	-16,184 (20,853)	-19,126 (21,287)	-17,295 (22,242)
(IV) E-filing impact for above-median risk score	23,658 (13,165)	37,912 (19,642)	32,905 (15,431)	26,595 (13,334)	47,456 (23,351)
<i>p</i> -value diff. (low versus high risk score)	0.106	0.060	0.061	0.072	0.049
Controlling for heterogeneity of e-filing impact with variables:	Trader	No employee	Tech. practices implem.	Time spent on tax	All variables together
Observations	1,067	1,067	1,067	1,057	1,057
<i>R</i> ²	0.040	0.067	0.044	0.046	0.025
<i>Part II: Dependent variable: At least one positive tax payment in 2015</i>					
(IV) E-filing impact for below-median risk score	-0.122 (0.079)	-0.142 (0.079)	-0.140 (0.078)	-0.167 (0.096)	-0.140 (0.100)
(IV) E-filing impact for above-median risk score	0.041 (0.103)	0.077 (0.096)	0.042 (0.106)	0.033 (0.104)	0.076 (0.096)
<i>p</i> -value diff. (low versus high risk score)	0.204	0.075	0.165	0.157	0.115
Controlling for heterogeneity of e-filing impact with variables:	Trader	No employee	Tech. practices implem.	Time spent on tax	All variables together
Observations	1,067	1,067	1,067	1,057	1,057
<i>R</i> ²	0.110	0.215	0.115	0.099	0.190

Notes: 2SLS regressions measuring the heterogeneous impact of e-filing adoption (instrumented by assignment to Group A) by risk profile scores. Samples restricted to legal entities because risk scores are not calculated for individual enterprises. Robust standard errors are in parentheses. All regressions include control dummies for strata. α : Winsorized at the ninety-ninth percentile.

The corresponding coefficient on low-risk score firms remains of similar magnitude as in Table 5, but, with larger standard errors, it loses statistical significance (column 5 of Table 7); hence, we describe the results on firms with lower risk scores as suggestive.

Alternative Definition of Firm Risk Score Categories.—We replicate our main analyses to examine the impact of e-filing across the quartiles of the baseline risk score. Online Appendix Table A11 shows that the negative impact on tax payment is concentrated in the first quartile (lowest risk scores) and the coefficient increases in the other quartiles. The coefficients in the second, third, and fourth quartiles are not statistically different from one another for the total amount of tax paid. Online Appendix Figure A2 replicates Figure 3 for our main tax outcomes but keeps the first quartile alone as the low-risk score group, while the other quartiles make up the high-risk score group. The heterogeneous impacts by risk profile are even more striking, as the negative impact on tax payment for firms with a lower risk score is more pronounced.

Heterogeneity Analysis Using Machine Learning.—The fact that we did not pre-specify our heterogeneous analysis by risk score in a pre-analysis plan raises the concern for type I error and overfitting. To address this concern, we follow the machine learning methodology described in Chernozhukov et al. (2020) for studying heterogeneity of impact of randomized experiments, with the caveat

that our sample size of 1,057 firms is relatively small, so some of the results lack precision.

First, to verify that there is any heterogeneity, we use random forest to compute the best linear predictor (BLP) of impact heterogeneity.³⁷ Second, we perform classification analysis and compute the average baseline characteristics of the 50 percent most affected and least affected firms defined in terms of the random forest predicted effect. This analysis is presented in online Appendix Table A14. The results show that among the 50 percent most affected firms (that is, firms with the greatest predicted increase in tax paid), the share of firms with risk score above the median is 12 percentage points higher than among the 50 percent least affected firms (which, in this context, would also include firms for whom e-filing had a negative impact on taxes paid).

Finally, similar to Carter, Tjernström, and Toledo (2019), we use random forest to examine the relative “importance” of each baseline covariate in explaining heterogeneity of impact by looking at the frequency with which they are used as a splitting variable. Frequencies are then adjusted to give more weights to splits that occur earlier in the tree. The risk profile score is the most “important” variable for tax paid according to this measure (see Figure A3 in the online Appendix). We interpret these results as additional evidence that the risk profile score is important to explain heterogeneity in the impact of e-filing.

VII. Conclusion and Policy Implications

This paper presents, to the best of our knowledge, the first experimental evidence on the impact of electronic tax filing. First, in a context characterized by high compliance costs and frequent (monthly) interactions between taxpayers and tax officials, e-filing adoption reduces the amount of time firms spend on fulfilling tax obligations by 40 percent (5 fewer hours each month). Second, we are unable to detect any average effect on the total amount of tax collected by the administration or on bribe payments. Third, firms’ baseline likelihood of tax evasion (as measured by a risk profile score developed by the tax authority) is an important determinant of their willingness to use e-filing as well as of the impact e-filing has on them.

For firms with above-median risk score, e-filing causes them to pay significantly higher taxes. These firms are also less likely to sign up for e-filing voluntarily, and, when they do (potentially to avoid conspicuously opting out of automatic registration), they are more likely to stop e-filing, indicating a preference to continue interacting directly with tax officials. This evidence suggests that, under paper filing, they may have been obtaining support from tax inspectors to minimize their tax liability.

On the other hand, for firms with below-median risk score, we find some indication (significant at the 10 percent level) that e-filing results in lower tax payments. This effect is concentrated in the first quartile of risk scores. The result suggests

³⁷First, we compare among different machine learning methods to compute the best linear predictor of impact heterogeneity, and we focus on the random forest, as it gives the most precise results (see online Appendix Tables A12 and A13).

that under paper filing, officials responsible for accepting tax declarations prevented these firms from submitting taxes they considered too low. At the same time, as a result of e-filing, firms with below-median risk score pay fewer bribes, likely due to becoming less exposed to coercive demands. These opposing patterns of impact result in a significant change in the distribution of tax payments across firms.

Over time, e-filing use has continued to expand in Tajikistan. The overall impact on tax revenue from the introduction of e-filing (beyond the encouragement design considered in this study) would depend on the distribution of firms' likelihood of evasion among the total population that adopts e-filing and the relative magnitudes of the respective increase and decrease in tax payment across firms in different categories of evasion risk. Beyond the impact on tax payments, the tax authority stands to derive significant system-wide efficiency gains. Back-of-the-envelope calculations indicate that, at current e-filing rates, the tax authority frees up over 5,800 tax-official-hours each month that can be reallocated to other activities likely higher in marginal productivity than receiving tax declarations.³⁸

These findings provide lessons extending beyond e-filing to other technologies designed to increase efficiency and reduce corruption. First, our results highlight the importance of selection among potential users. Those who anticipate negative effects from the new technology due to expected constraints on their behavior may be the least likely to adopt. However, this may be the population segment of greatest interest. As a result, when deciding whether to make new technologies mandatory, governments must weigh carefully the capacity constraints of potential users in adopting a new technology, with the potential of those with high risk of noncompliance opting out. In Tajikistan, while e-filing was not mandatory, updated policies require that once firms start e-filing, they may not revert to paper filing.

Second, our results underscore the possibility of heterogeneous effects of technology. The impact of technology introduced to replace human discretion will depend on the way that discretion was previously exercised. If discretion produces worse outcomes, technology may improve outcomes (as we find with increased tax payments among firms with higher risk scores and lower bribe payments among firms with lower risk scores). However, if discretion produces desirable outcomes (such as monitoring firms), technology may have an unanticipated effect of producing less desirable outcomes. Lastly, on a related note, our results on the distributional consequences of e-filing emphasize the potential of technology to correct possible patterns of bias resulting from human discretion.

REFERENCES

- Ali, Merima, Abdulaziz Shifa, Abebe Shimeles, and Firew Woldeyes. 2015. "Information Technology and Fiscal Capacity in a Developing Country: Evidence from Ethiopia." International Centre for Tax and Development Working Paper 31.
- Amodio, Francesco, Jieun Choi, Giacomo De Giorgi, and Aminur Rahman. 2021. "Bribes vs. taxes: Market structure and incentives." *Journal of Comparative Economics*. <https://doi.org/10.1016/j.jce.2021.09.006>.

³⁸This calculation assumes that a tax official spends six minutes on each tax declaration. This is a conservative estimate, as the firms in our baseline survey report tax officials spending an average of 18 minutes checking their reports.

- Anderson, Michael L.** 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103 (484): 1481–95.
- Angrist, Joshua D., and Jörn-Steffen Pischke.** 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.
- Athey, Susan, Julie Tibshirani, and Stefan Wager.** 2019. "Generalized Random Forests." *Annals of Statistics* 47 (2): 1148–78.
- Banerjee, Abhijit, Esther Duflo, Clément Imbert, Santhosh Mathew, and Rohini Pande.** 2020. "E-governance, Accountability, and Leakage in Public Programs: Experimental Evidence from a Financial Management Reform in India." *American Economic Journal: Applied Economics* 12 (4): 39–72.
- Bellon, Matthieu, Jillie Chang, Era Dabla-Norris, Salma Khalid, Frederico Lima, Enrique Rojas, and Pilar Villena.** 2019. "Digitalization to Improve Tax Compliance: Evidence from VAT e-Invoicing in Peru." IMF Working Paper 19/231.
- Benjamini, Yoav, Abba M. Krieger, and Daniel Yekutieli.** 2006. "Adaptive Linear Step-up Procedures That Control the False Discovery Rate." *Biometrika* 93 (3): 491–507.
- Besley, Timothy, and Torsten Persson.** 2014. "Why Do Developing Countries Tax So Little?" *Journal of Economic Perspectives* 28 (4): 99–120.
- Bruhn, Miriam, and David McKenzie.** 2009. "In Pursuit of Balance: Randomization in Practice in Development Field Experiments." *American Economic Journal: Applied Economics* 1 (4): 200–232.
- Carter, Michael R., Emilia Tjernström, and Patricia Toledo.** 2019. "Heterogeneous Impact Dynamics of a Rural Business Development Program in Nicaragua." *Journal of Development Economics* 138: 77–98.
- Chernozhukov, Victor, Mert Demirer, Esther Duflo, and Iván Fernández-Val.** 2020. "Generic Machine Learning Inference on Heterogenous Treatment Effects in Randomized Experiments." <https://arxiv.org/pdf/1712.04802.pdf>.
- Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins.** 2018. "Double/Debiased Machine Learning for Treatment and Structural Parameters." *Econometrics Journal* 21 (1): C1–68.
- Currie, Janet.** 2006. "The Take Up of Social Benefits." In *Poverty, the Distribution of Income, and Public Policy*, edited by Alan Auerbach, David Card, and John Quigley, 80–148. New York: Russell Sage.
- de Mel, Suresh, David McKenzie, and Christopher Woodruff.** 2009. "Are Women More Credit Constrained? Experimental Evidence on Gender and Microenterprise Returns." *American Economic Journal: Applied Economics* 1 (3): 1–32.
- Enterprise Surveys, The World Bank.** 2013. The World Bank Enterprise Survey. <http://www.enterpris-surveys.org>.
- Fan, Haichao, Yu Liu, Nancy Qian, and Jaya Wen.** 2020. "Computerizing VAT Invoices in China." NBER Working Paper 24414.
- Fujiwara, Thomas.** 2015. "Voting Technology, Political Responsiveness, and Infant Health: Evidence from Brazil." *Econometrica* 83 (2): 423–64.
- Gordon, Roger, and Wei Li.** 2009. "Tax Structures in Developing Countries: Many Puzzles and a Possible Explanation." *Journal of Public Economics* 93 (7–8): 855–66.
- Khan, Adnan Q., Asim I. Khwaja, and Benjamin A. Olken.** 2016. "Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors." *Quarterly Journal of Economics* 131 (1): 219–71.
- Khwaja, Muwer Sultan, Rajul Awasthi, and Jan Loeprick.** 2011. *Risk-Based Tax Audits: Approaches and Country Experiences*. Washington, DC: World Bank.
- Kochanova, Anna, Zahid Hasnain, and Bradley Robert Larson.** 2016. "Does E-government Improve Government Capacity? Evidence from Tax Administration and Public Procurement." World Bank Policy Research Working Paper 7657.
- Kuklinski, James H., Paul M. Sniderman, Kathleen Knight, Thomas Piazza, Philip E. Tetlock, Gordon R. Lawrence, and Barbara Mellers.** 1997. "Racial Prejudice and Attitudes toward Affirmative Action." *American Journal of Political Science* 41 (2): 402–19.
- Lewis-Faupel, Sean, Yusuf Neggers, Benjamin A. Olken, and Rohini Pande.** 2016. "Can Electronic Procurement Improve Infrastructure Provision? Evidence from Public Works in India and Indonesia." *American Economic Journal: Economic Policy* 8 (3): 258–83.
- Loeprick, Jan, and Michael Engelschalk.** 2011. "Simplified Risk Scoring for SMEs." In *Risk-Based Tax Audits: Approaches and Country Experiences*, edited by Muwer Sultan Khwaja, Rajul Awasthi, and Jan Loeprick, 45–54. Washington, DC: World Bank.

- Muralidharan, Karthik, Paul Niehaus, and Sandip Sukhtankar.** 2016. "Building State Capacity: Evidence from Biometric Smartcards in India." *American Economic Review* 106 (10): 2895–2929.
- Olken, Benjamin A., and Rohini Pande.** 2012. "Corruption in Developing Countries." *Annual Review of Economics* 4: 479–509.
- Okunogbe, Oyebola, and Victor Pouliquen.** 2017. "Technology and Corruption: Experimental Evidence from the Introduction of Tax E-filing in Tajikistan." AEA RCT Registry. November 17. <https://doi.org/10.1257/rct.914-2.0>.
- Okunogbe, Oyebola, and Victor Pouliquen.** 2022. "Replication data for: Technology, Taxation, and Corruption: Evidence from the Introduction of Electronic Tax Filing." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. <https://doi.org/10.3886/E131281V1>.
- Sequeira, Sandra, and Simeon Djankov.** 2014. "Corruption and Firm Behavior: Evidence from African Ports." *Journal of International Economics* 94 (2): 277–94.
- World Bank.** 2016. *Doing Business 2016: Measuring Regulatory Quality and Efficiency*. Washington DC: World Bank.
- World Bank Group.** 2016. *World Development Report 2016: Digital Dividends*. Washington, DC: World Bank.
- Yilmaz, Fatih, and Jacqueline Coolidge.** 2013. "Can E-filing Reduce Tax Compliance Costs in Developing Countries?" World Bank Policy Research Working Paper 6647.