How Do Taxpayers Respond to Public Disclosure and Social Recognition Programs? Evidence from Pakistan

Joel Slemrod, Obeid Ur Rehman & Mazhar Waseem*

February 2020

Abstract

We examine two Pakistani programs to see if the public disclosure of tax information and social recognition of top taxpayers promote tax compliance. Pakistan began revealing income tax paid by every taxpayer in the country from 2012. Simultaneously, another program began recognizing and rewarding the top 100 tax paying corporations, partnerships, self-employed individuals, and wage-earners. We find that both programs induced strong compliance responses. The public disclosure caused on average a 9 log-points increase in the tax paid by individuals exposed to the program. The increase was even larger for the social recognition program, around 17 log-points. Our results suggest that such programs can be important policy levers to mobilize resources, especially in weak-enforcement-capacity economies.

Keywords: Tax evasion, income tax, social norms
JEL Classification: H24, H25, H26

*Email Addresses: Joel Slemrod jslemrod@umich.edu, Obeid Ur Rehman obeidr@umich.edu, Mazhar Waseem mazhar.waseem@manchester.ac.uk. Helpful comments on an earlier draft were received from participants at seminars at the Bank of Mexico, the University of Michigan, the University of Manchester, Harvard University, the University of Munich, ETH Zurich, the University of Zurich, the World Bank, the University of Southern California, , the National Tax Association annual meeting, the University of Exeter, the University of Essex, the International Institute of Public Finance annual congress, the University of Oxford, the World Bank, and UNU-WIDER. Usama Jamal provided outstanding research assistance. This paper was earlier circulated with the title "Pecuniary and Nonpecuniary Motivations for Tax Compliance: Evidence from Pakistan".

I Introduction

Tax evasion is a pervasive problem in developing countries and a non-trivial one in developed countries (Slemrod, 2019). Economic theory suggests that tax evasion is deterred by the risk of detection and punishment (Allingham & Sandmo, 1972), and it may be influenced by social and psychological factors, such as guilt or shame from evading, pride from fulfilling one's civic duty, and approval or sanctions from peers (Luttmer & Singhal, 2014). To leverage these motivations, many countries employ policies that disclose tax information, shame tax delinquents, or honor top tax payers. Given that these policies entail little resource costs, they are becoming increasingly common.[1] Yet, there is little evidence, especially from the emerging economies, on how effective they are in promoting tax compliance.

In this paper, we exploit two Pakistani programs to fill this gap in literature. In the first of these programs, the government began revealing the amount of income tax paid by every taxpayer in the country. The public disclosure program was instigated by a series of press reports documenting that the majority of lawmakers of the country had not been fulfilling their tax obligations. It began in tax year 2012 and has continued since then. Each year, two tax directories are published, one for the Members of Parliament (MPs) and one for all taxpayers. The directories are available online in a searchable PDF format and can be downloaded freely by anyone. The directory for general taxpayers reveals the name, a numerical tax identifier, and the tax paid by each taxpayer. The directory for MPs also lists the constituency they serve.

The second program we examine publicly recognizes and rewards top taxpayers of the country. The Taxpayers Privileges and Honour Card (TPHC) program began concurrently with the public disclosure program. It acknowledges the top 100 taxpayers in each of four categories—self-employed individuals, wage-earners, partnerships, and corporations—and grants them certain privileges, such as invitation to a special ceremony hosted by the Prime Minister and eligibility for benefits such as fast-track immigration and gratis passports.

These programs can influence taxcompliance through a number of channels.

[1] Dwenger & Treber (2018), for example, report that one-half of the OECD countries have the legal power to publish the names of tax delinquents and nearly 90% of them used this power in 2015. Similarly, 23 US states run shaming programs, maintaining online lists of tax delinquents with their names and addresses (Perez-Truglia & Troiano, 2015).

Public disclosure can encourage whistle-blowing, evoke shame and guilt, and inspire pride.[2] Social recognition of top taxpayers can stimulate a sense of pride and self-fulfillment. Some individuals may obtain higher utility from the public appreciation of their affluence (Akerlof & Kranton, 2000; Glazer & Konrad, 1996), while others may monetize the goodwill offered by the programs, translating the social recognition into higher sales and profits. Through these channels, the two programs promote tax compliance. On the other hand, these programs could conceivably backfire, if for example they reveal others to be even less compliant (Schultz *et al.*, 2007) or if they crowd out intrinsic motivation (Benabou & Tirole, 2003).

We use a novel empirical strategy to estimate the impacts of the public disclosure. As noted above, the tax directory published under the program lists the name and a numerical identifier of each taxpayer. The numeric identifier is effectively private information, known primarily to the agent and the tax administration. Thus, the only publicly-disclosed information that can link an observation in the directory to a particular taxpayer is the name. Pakistani names do not follow the standard Western syntax of given name + middle name + surname. Instead, a typical Pakistani name is composed of two or more given names. One of these given names—usually the most-called name of the father or husband—serves as the surname. Surnames in this way are usually not fixed across generations and vary even within the nuclear family. Because of these naming conventions, it is quite common for people to have the same full name. For example, the most frequent name in our data, Muhammad Aslam, appears 15,598 times in four years, with a typical year's directory containing more than 60 pages listing the name Muhammad Aslam alone. On the other hand, about one-third of taxpayers have unique names. This variation in name commonness implies that the intensity of the disclosure varies considerably across individuals depending upon how common their name is. Taxpayers with very frequent names enjoy virtual anonymity in the disclosed records; uniquely-named taxpayers, on the other hand, are exposed perfectly. We exploit this variation in treatment intensity in our empirical strategy, comparing the change in tax payments across taxpayers with frequent and unique names

Of course, names are not randomly assigned. Instead, they are chosen by par-

[2]Public disclosure may evoke shame and guilt if a taxpayer perceives her tax payments too low relative to the consumption or wealth observed by peers. Similarly, it may inspire pride if one is revealed to be a compliant or top taxpayer.

ents and hence may be correlated with parental traits such as income, education, and ethnicity. We always include individual fixed effects in our empirical models, implying that parental traits will influence our estimates only if their effect changes over time, in particular contemporaneously with the program. We provide two sets of tests to rule out this and related concerns. First, we show through both visual and regression-based evidence that the tax payments of the compared groups were trending similarly in the six pre-program periods: the relative difference in the outcome was indistinguishable from zero for virtually all these years. Second, we show that the name of a taxpayer bears no association with the outcome in the sample of taxpayers (MPs) where the disclosure intensity is independent of the name commonness.

The TPHC program applies only to the top 100 taxpayers of each category. We leverage this discontinuity in program eligibility to estimate its impacts. If social recognition and related benefits offered by the program are valued, taxpayers close to the eligibility cutoff will increase their tax payments in order to remain in, or enter into, the top 100 club. We test this by comparing the yearly growth in tax liability reported by agents close to the cutoff with other top taxpayers. To show that our estimates are not driven by factors unrelated to the program such as rising inequality at the top, we run placebo regressions estimating the program effects in pre-intervention periods and on unaffected groups.

We combine the disclosed data of the years 2012-2015 with administrative tax return data from 2006-2012 to create a long panel of tax records from 2006 to 2015. Our populations of interest are the universe of self-employed tax filers for the public disclosure program and the top-1000 taxpayers of each category for the TPHC program. We document three key findings. First, the exposure of tax information induced a substantial response from the treated taxpayers. The tax liability reported by taxpayers with less common names on average increased by around 9 log points as a result of the program. Consistent with our expectations, the estimated effect varies directly with the program intensity. It is strongest in the left-tail of the name-frequency distribution, declines monotonically as we move rightward, and becomes insignificant as the name-frequency approaches 300 (i.e., the name of the taxpayer appears at least 300 times in the four years of disclosed data). Along the extensive margin, the program caused a 1-2 log points increase in tax filing by individuals with less common names relative to others. Second, the TPHC program also had a large impact. The tax liability reported by 70-130 ranked taxpayers grew by nearly 17 log points faster than others as a result of the program. This estimate declines slightly as we widen the treatment window, suggesting that, as hypothesized, the effect is concentrated around the eligibility cutoff of the program. Finally, we document that our estimates are highly robust to alternative specifications and the identification concerns noted above.

Our empirical strategy implicitly assumes that the public disclosure did not affect the tax payments of more-common-named taxpayers. However, we have noted above that such programs can backfire and decrease compliance, for example if they cause a perception that others are even less compliant. To rule out such a possibility in our setting, we compare the tax payments of wage-earners and self-employed taxpayers in a difference-in-differences research design. Under the assumption that the public disclosure had no effect on wage-earners given that their earnings are third-party-reported, we estimate the average effect of the program on the more-common-named self-employed, finding it to be positive. This implies that the public disclosure had an overall positive effect on the self-employed and a much stronger effect on the less-common-named individuals amongst these, for whom the exposure to the program was more intense. To this extent, our estimates reported above provide a lower bound on the true effect of the program.

As we note above, programs similar to the ones we study are becoming increasingly common. Public disclosure of taxes with varying degrees of coverage and access is now in place in a diverse group of countries including Norway, Finland, Sweden, Iceland, Australia, Japan, and Pakistan. Of these, Norway's program is closest to the Pakistan's. Exploiting a unique feature of the Norwegian program, Bø *et al.* (2015) estimate that it caused at least 3 percent increase in income reported by the self-employed. Unsurprisingly, the effect we find is stronger given that the baseline noncompliance in our setting is expected to be larger (see Hasegawa *et al.*, 2012 and Hoopes *et al.*, 2018 for analyses of the Japanese and Australian programs).

Shaming programs, which although not identical to the programs we study, rely on similar behavioral factors are even more common. For example, as we note above some version of the shaming program was in use in one-half of the OECD countries in 2015. Dwenger & Treber (2018) study one such program from Slovenia finding that taxpayers reduce their debt by 8.5% to avoid shaming, particularly in sectors where reputational concerns are more important. Similarly, 23 US states implement some type of shaming program via maintaining online lists of tax delinquents with their names and addresses. Using a randomized intervention, Perez-Truglia & Troiano (2015) find that increasing the visibility of tax delinquency status increases compliance by individuals, a result qualitatively very similar to ours. Most of the above programs and the related studies have developed-country settings. In developing countries, tax enforcement capacity is limited and evasion is pervasive. In such settings, the programs we study have a particular appeal, offering potentially cost-effective options to mobilize resources. Of course, any such policy needs to balance the revenue gains against concerns such as privacy and security.[3] Our estimates provide a basis for such an evaluation.

Our paper is also related to another strand of literature that studies social motivations in tax compliance, mostly through lab and field experiments (see Mascagni, 2018 for a survey). Del Carpio (2013), for example, randomizes deterrence messages to study the role of social norms in property tax compliance in Peru. Castro & Scartascini (2015) run a similar experiment in Argentina and Kettle *et al.* (2016) in Guatemala (please see Slemrod *et al.*, 2001, Fellner *et al.*, 2013, and Dwenger *et al.*, 2016 for three similar studies from developed countries). Relative to these studies, we provide evidence on the impacts of two national programs that appeal, among other things, to social motivations of taxpayers.

The Pakistani public tax disclosure program has been studied in one recent political science paper. Malik (2019) investigates the impact of the program on the tax reporting behavior of MPs. She uses two years' publicly available data to assess if MPs in more competitive races respond more aggressively to the program than others and similar political economy questions. As we note above, the primary focus of our paper is the universe of tax filers and not MPs.

II Context

In this section, we describe features of the Pakistani environment that are important for our empirical analysis.

[3]Please see Lenter *et al.* (2003); Blank (2014); Perez-Truglia (2019) for the non-tax effects of public disclosure.

II.A The Public Disclosure Program

In the first of two programs we study, the Pakistani government started publishing a tax directory each year, revealing income tax paid by every taxpayer in the country.[4] The policy change (in large part) was instigated by a string of investigative reports that began appearing in the Pakistani press in the latter half of 2012. The reports focused primarily on the tax affairs of lawmakers of the country, documenting that a majority of them had apparently not been fulfilling their tax obligations. Combining data leaked by whistle-blowers with the official data obtained through the Election Commission of Pakistan, the reports painted quite a bleak picture of tax compliance among the MPs of the country. It was reported that around 66% of them—including 34 out of 55 federal ministers—had not filed their tax return for the latest year; in fact, about 20% of them had not even obtained the National Tax Number, which is the first requirement for tax filing (Center for Investigative Reporting in Pakistan, 2012). These revelations, compiled into two papers published by the Center of Investigative Reporting in Pakistan (CIRP), generated strong reaction. The Federal Tax Ombudsman, upon a representation filed by a citizen, ordered the government to begin disclosing the tax remitted by every public office holder in the country. The leading opposition party at the time went even further, pledging to publish the amount of tax remitted by all taxpayers in the country if elected to power. This party won the next elections and formed the federal government in May 2013. It fulfilled its election promise and began publishing the tax records for the tax year 2012 onward, which were due to be filed by December 15, 2013.[5]

Since the institution of the program in 2012, two tax directories are published each year, one for MPs and the other for all taxpayers. These directories are posted online on the Federal Board of Revenue (FBR)'s website in a searchable PDF format.[6] They can also be downloaded freely by anyone. The directory for general taxpayers reveals the name, tax identifier, and tax liability of each taxpayer. This

[4] Tax paid here refers to the self-assessed tax liability reported by a taxpayer in their annual income tax return, which includes any tax withheld at source. The Pakistani tax code requires that this self-assessed tax liability should be deposited into the treasury at the time of filing of return. For this reason, we use the terms tax paid and tax liability interchangeably in this paper.

[5] The Pakistani tax year runs from July to June. Any year t in this paper denotes the tax year from July t to June $t+1$.

[6] In fact, the title page of the directory contains the following direction in a very salient yellow box: "Please press CTRL + F Key to Search the Record".

information—sorted alphabetically on the full name—is provided separately for corporations, partnerships, and individuals. The tax identifier is either the nine-digit National Tax Number (NTN), disclosed with the tax year 2012 data, or the 13-digit Computerized National Identity Card Number (CNIC), disclosed with the 2013 tax year data and thereafter, both of which are effectively private information of agents.[7] Therefore, the only information through which an observation in the directory can be readily linked to a taxpayer is the name.[8] In contrast, the directory of parliamentarians also contains the constituency number an MP serves, and therefore the disclosed information can be linked to them fairly easily.

Table A.I lists important events in the public disclosure program. The timing of these events is important for our empirical analysis, in particular in deciding from which period the program would begin affecting behavior. As we note above, the political party committed to the full public disclosure had come into power in May 2013. The last date for filing the 2012 tax return was December 15, 2013.[9] Thus, by the time the 2012 returns were filed, it was clear that the tax remitted through them would be made public. We accordingly treat tax year 2012 (which covers July 2012 - June 2013) as the first post-program year in our analysis. Although the exact format of the disclosure was not known at the time, it was clear that it would, at a minimum, include the name of the taxpayer. The name is a primary, and to some extent the only, information through which the public can link a tax return to a taxpayer, and therefore there could be no meaningful disclosure without it.[10]

[7]The NTN is used exclusively for tax filing. It was issued sequentially beginning in 1995, so the number reveals some information about how long a taxpayer has been in the tax net. The CNIC is the primary identification and proof of citizenship document in Pakistan. It is required for most official services including obtaining a passport, driving license, utility connection, opening and operating bank accounts. The first few digits of the CNIC indicate the district (of 128 in Pakistan) where the individual resided at the time of initial registration.

[8]FBR provides an online taxpayer verification service through which tax identifiers can be used to obtain additional taxpayer information, namely address (at the time of registration), registration date and regional tax office. This additional information may improve the chances of linking an observation in the directory to a taxpayer but may still not be sufficient. A taxpayer's address may have changed since they first registered for an NTN or it may not be public information. Additionally, there is a significant effort cost of obtaining the information and it is increasing in the commonness of the taxpayer's name. The tax identifiers of all taxpayers with a particular name would have to be manually entered one at a time to obtain the additional information and online security features prevent the process from being automated. The effective disclosure intensity therefore is still linked primarily to the commonness of the taxpayer's name.

[9]Generally, a majority of tax returns are filed in the last few weeks before the due date. Consistent with this trend, more than 90% of the 2012 returns in our data were filed in or after October 2013.

[10]The CIRP reports that precipitated the full public disclosure program always used the name as

As we note above, the MPs' directory also contains the constituency number they serve. Table A.II reports the composition of the Pakistani legislature. Because the country has a limited number of MPs, their identities are well known, especially in their electoral constituencies. Their exposure to the program therefore must be independent of how common their name is. We use this feature of the program as a specification check on our empirical strategy.

Both sets of directories receive wide coverage in the Pakistani media, especially at the time they are released. Figure A.I plots the time line of Google searches in Pakistan for the phrases "FBR Tax Directory" and "Tax Directory". Clearly, searches for these phrases peak at the time the tax directories are published. In addition, simple Google searches of "FBR Tax Directory" and "Tax Directory" looking for the occurrence of these words as exact phrases return 1,010 and 32,800 results.[11] This indicates that there are at least 1,010 (and potentially many more)[12] active web pages that discuss the Pakistani tax directories. This profusion of information creates a strong first stage in our setting in the sense that many Pakistani taxpayers are aware that their disclosed tax data would remain available online for the foreseeable future and could be accessed anytime by their peer networks. Note that the income tax exemption threshold in Pakistan, like other developing countries, is quite high, set at around the 80th percentile of the income distribution (Waseem, 2019). Income taxpayers in the country are a richer segment of the population and therefore they and their peer networks are extremely likely to be exposed to the disclosed information, be it online or in other formats.

II.B The Taxpayer Privileges and Honour Card Program

The second program we examine is the Taxpayer Privileges and Honour Card (TPHC) scheme. The program was announced at the beginning of the tax year 2012, in July 2012. It acknowledges and grants special privileges to the top 100 taxpayers in each of the following four categories: (a) wage-earners, (b) self-employed individuals, (c) partnerships, and (d) corporations. The special privileges granted by the pro-

the primary identifier of a taxpayer.

[11]This data was accessed on May 28, 2019 in Manchester, UK.

[12]Similar Google searches looking for the occurrence of "FBR Tax Directory" and "Tax Directory" not as exact phrases return 169,000 and 867,000,000 results, suggesting that there are potentially many more active web pages that discuss the two sets of directories.

gram include: (1) automatic invitation to the Annual Excellence Awards hosted by the Prime Minister; (2) automatic invitation to the state dinners held on Pakistan Day (23rd March) and Independence Day (14th August); (3) fast-track immigration through special counters (Figure A.II provides a photograph of such an immigration counter at the Lahore airport); (4) issuance of gratis passports; (5) access to VIP lounges at Pakistani airports; and (6) an increased baggage allowance. These privileges last one complete year, until the new set of recipients are announced. The personal benefits of the program are conferred on the partner with the highest capital contribution in the case of partnerships, and on the CEO in the case of corporations.

Two features of the program need emphasizing. First, while the principal element of the program is the social recognition of top taxpayers,[13] it provides some material benefits as well. To the extent that these benefits are valued, the response to the program would also reflect the willingness to pay of top taxpayers for these benefits. Second, the program has some overlap with the public disclosure, as the latter also identifies top taxpayers, albeit indirectly. In fact, most of the news items that report on the public disclosure program also focus on who are the top taxpayers in the disclosed data. This media recognition, however, is indirect, usually limited to the very top taxpayers (say top 10), and is not as salient or meaningful as one offered by the TPHC program. But to the extent that the two programs overlap, our estimates will capture the combined effects of the two.

II.C Pakistani Naming Conventions

Pakistani names generally do not conform to the standard Western syntax of given name + middle name + surname. Instead, a typical Pakistani name consists of one or more given names and a surname. The given names are usually derived from Persian, Arabic, or Turkish, and it is quite common for people to have more than one given name. If a person has two or more given names, the less common one serves as the *most-called* name (the person is informally referred to by this given name). For example, if Muhammad is one of the multiple given names, it is usually not the person's most-called name, as being so common it does not serve as a

[13] Addressing the first batch of the Honour Card recipients, the Prime Minister said that the "ceremony has been convened to acknowledge your services for the nation."

useful identifier. Unlike the Western practice, surnames in Pakistan are usually not fixed across generations. The most popular convention is to adopt the most-called given name of father (husband) as the child's (married woman's) surname. As a result, surnames vary even within the nuclear family (father/husband has a different surname). In cases where the surname does not vary within the family, it is rarely unique. For example, virtually all people of Pashtun origin use Khan as their surname.

Because of these conventions, many full names are widely shared in Pakistan. Figure I illustrates this formally. We plot the distribution of full names contained in the public disclosure data for the tax years 2012-2015. To construct the diagram, we treat all English variants of an Urdu name as one. For example, Muhammad spelled as Mohammad, Muhammed or Mohammed is treated as one name (to an Urdu speaker, they would be indistinguishable). To show that adjusting these spelling variations does not change our results materially, we provide the corresponding raw distributions in Figure A.III (the details of our cleaning algorithm are presented in Appendix A.1). A total of 526,425 unique names appear in the publicly disclosed data during the four years. Of these, Muhammad Aslam is the most frequent, appearing 15,598 times. Because a single page of the directory on average consists of 60 rows, a given year's directory contains about 65 (15,598/(4*60)) pages listing the name Muhammad Aslam alone. There are other such very frequent names. In fact, nearly one-third of taxpayers share their full name with at least 500 others. The distribution has a thick tail at the other end as well. Approximately 35% of taxpayers have names that appear fewer than ten times in the four years of data; about 4% appear only once, while 24% of names appear between 2-5 times.

As we note above, the directory carries no publicly-known identifier other than the name. The wide variation in name frequency thus translates into a wide variation in the effective intensity of disclosure. Note that we do not expect, and do not assume, that taxpayers know precisely how common their name is. However, persons with very frequent names such as Muhammad Aslam would very likely have come across numerous other people of the same name in their lives and would have—through a conscious or subconscious process—formed a belief that their name grants virtual anonymity to them. On the other hand, unique-named individuals would likely have a sense that any information with their name on it can be linked to them directly. Once the public disclosure lists became available, it was straightforward to acquire more concrete information about how common one's name is.

III Conceptual Framework

III.A Social and Psychological Motivations in Tax Compliance

Economists have traditionally modeled tax evasion as if it were a choice under uncertainty (Allingham & Sandmo, 1972). Successful evasion provides additional disposable income, but evasion also entails the risk that the evaded amount will be recovered along with penalty in case of detection. Assume a taxpayer earns real income z but reports $\underline{z} \leq z$ with $e \equiv z - \underline{z}$, paying a tax $T \equiv \tau(z - e)$. The taxpayer perceives that evasion will be detected with probability p, triggering a proportional penalty of θ applied to the evaded income upon detection. The taxpayer chooses e to maximize the expected utility of the gamble denoted by

$$\max_{e} \quad (1-p)\,.u\left[(1-\tau)z+\tau e\right]+p.u\left[(1-\tau)z-\theta e\right]. \tag{1}$$

In this model evasion is deterred solely by the fear of penalty. A risk-averse taxpayer balances the disutility of income loss in the detected and penalized state against the utility of extra income in the undetected state.

$$\frac{u'(c_A)}{u'(c_{NA})}=\frac{(1-p)\,\tau}{p\theta}, \tag{2}$$

where c_A and c_{NA} denote consumption in the detected and undetected states.

The deterrence model captures the first-order pattern of tax evasion quite well. For example, cross-matching of third-party information reports means that the detection probability faced by taxpayers (such as wage-earners) on income covered by third-party reports can be close to one even if only a small percentage of tax returns are actually audited (Slemrod, 2007; Kleven *et al.*, 2011). Consistent with the model, the noncompliance rate of wage income is considerably lower than that of self-employment income, estimated in the United States to be 1% and 63%.

The deterrence model does not, though, explain all aspects of tax evasion, and does not take into account social and psychological factors.[14] These factors can be divided into three classes. First, there are factors that reduce utility in both states of the world. Guilt, for example, may cause psychological and emotional distress to a tax cheat even if the act of cheating remains undetected. Second are factors such as shame that reduce utility only if cheating gets detected (Erard & Feinstein, 1994). And, third, there are behavioral biases whereby the detection probability and penalty are systematically mis-estimated by taxpayers (Scholz & Pinney, 1995; Chetty, 2009).

The public disclosure program we examine potentially affects each of these factors. By facilitating whistle-blowing, it arguably raises both the real and perceived likelihood of detection. It may also intensify the guilt and shame felt by tax cheats, especially if reported income does not match consumption or wealth observed by peers. For these reasons, we expect the public disclosure to reduce evasion and increase tax payments. There is, however, some evidence, especially in the psychology literature, that the provision of information can sometimes backfire (see for example Schultz *et al.*, 2007). In our context, this suggests that some individuals may start paying less taxes after the public disclosure if they perceive others are paying even less. We investigate, and rule out, such a boomerang effect in our setting in section IV.A.

The TPHC program promotes compliance to the extent that social recognition of top taxpayers can induce pride and a sense of accomplishment. Individuals may also treat taxation as a position (Veblen) good, deriving utility from being seen as one of the richest in the country (Akerlof & Kranton, 2000).[15] The goodwill offered by the TPHC program can perhaps in some cases be monetized, as well. Individuals and firms may advertise their status as a top taxpayer to gain more consumers and sales. Due to these mechanisms, the costs of evasion jump up at the eligibility cutoff of the program. The resulting notch will induce taxpayers to locate on the eligible side of the cutoff, increasing the tax paid by agents close to the cutoff. Working in the opposite direction, some taxpayers may place negative value on the attention

[14] For example, in an influential survey of the tax compliance literature, Andreoni *et al.* (1998) write that "factors such as a moral obligation to be truthful, or the social consequences of being a known cheater, may add further enforcement incentives that are not accounted for in our models."

[15] It has been found that consuming goods associated with wealth provides utility to some individuals even if their consumption remains invisible to others (Bursztyn *et al.*, 2018).

the program provides.

III.B Empirical Strategy

We use difference-in-differences research designs to estimate the effects of the two programs on tax compliance. These designs are explained in greater detail below.

III.B.1 Public Disclosure Program

The public disclosure program was rolled out nationally, all at once. Therefore, the principal identification challenge in estimating its effects is to control for any trends or shocks that might affect tax reporting at the aggregate level and may coincide with the program. We achieve this by exploiting the variation in exposure to the program caused by the degree of uniqueness of a taxpayer's name. We define Name Frequency as the number of times a full name appears in the four years of the disclosed data. For example, the Name Frequency of the most frequent name in the data—Muhammad Aslam—is 15,598. Taking advantage of the observable differences in program intensity across taxpayers with different Name Frequency, we estimate regressions of the form

$$\log \text{TaxPaid}_{it} = \alpha_i + \beta \text{ treat}_i \times \text{after}_t + \lambda_t + u_{it}, \tag{3}$$

where α_i and λ_t are individual and year fixed effects, after_t is a dummy indicating 2012 or a later year, and treat_i is an indicator of the Name Frequency of individual i. We experiment with different Name Frequency cutoffs in our empirical specifications. The difference-in-differences (DD) coefficient of interest β captures the differential effect of the program, denoting the average additional tax paid in the post-program years by individuals with relatively low Name Frequency. In this and all subsequent specifications, we cluster standard errors at the individual level, the most aggregate level feasible in our setting (Abadie *et al.*, 2017; Bertrand *et al.*, 2004).

For β to have a causal interpretation, it must be shown that the interaction variable and the error terms are uncorrelated. Our treatment variable captures how unique a taxpayer's name is. But names are not randomly assigned. Instead, they are chosen by parents, perhaps with the help of close relatives and friends. Any measure of name uniqueness, therefore, could be correlated with parental traits such as income, education, and ethnicity. To control for such correlations, we always include individual fixed effects in our regressions. The parental traits, therefore, would influence our estimates only if their effect changes over time, in particular in 2012.

We offer three pieces of evidence to rule out this concern. First, exploiting the panel nature of data we show that there were no systematic differences between the compared groups in terms of their tax payments in the pre-program years. We show this through the following event-study regressions

$$\log \text{TaxPaid}_{it} = \alpha_i + \sum_{j=2007}^{2015} \gamma_j \ \text{treat}_i \times 1.(\text{year=}j)_t + \lambda_t + u_{it}. \tag{4}$$

The coefficients γ_js here capture the average difference in tax payment between the two groups in year j relative to the reference year 2006. For a variety of definitions of treatment, we show that the estimated γ_js remain trivial/insignificant in the pre-program years but become large and significant in the post-program years. While validating our empirical strategy, these results do not expressly rule out a contemporaneous macro event that affects the tax payments of more-uniquely-named individuals. Note that in most difference-in-differences setups this assumption remains untested and is presumed satisfied if the preexisting trends are parallel. But in our setting we can go one step further than the parallel-trends assumption to rule out this possibility more directly. As we note above, MPs in Pakistan are prominent in their communities and their constituencies are listed in the directory. The effectiveness of the disclosure is therefore plausibly independent of how conspicuous or obscure their name is. We show that β remains statistically indistinguishable from zero when equation (3) is estimated on the sample of MPs only. This result is consistent with our assertion that the estimated coefficient of interest is driven by the causal impact of disclosure, rather than by any residual correlation between the name and tax payment. In our final test, we estimate equation (3) on the pre-program periods only (2006-2011), pretending as if the program occurred in 2010 rather than the actual date of 2012. These placebo regressions always return trivial/statistically insignificant coefficients on the interaction term of interest.

Our primary population of interest are the self-employed individuals. The Pakistani tax code and our administrative data defines a taxpayer as self-employed if their salary income does not exceed 50% of their taxable income. Self-employment income, being self-reported and not subject to substantial cross-checking with third-party information reports, is the most amenable to manipulation. Tax compliance studies from around the globe show that the incidence and extent of noncompliance is the highest for the self-employed (see for example Slemrod, 2019 and Waseem, 2019). If the public disclosure program curtails tax evasion, the effect would be the strongest for this section of the population.

III.B.2 TPHC Program

The TPHC program recognizes and rewards the top 100 taxpaying corporations, partnerships, self-employed individuals, and wage-earners. If the incentives and recognition offered by the program are valued, taxpayers ranked just below 100 would attempt to get into the top 100 in the next year and taxpayers just above the cutoff would attempt to stay there. The discontinuous treatment would thus cause a spike in the growth of tax paid from year t to $t+1$ by taxpayers ranked around the eligibility cutoff of the program in year t. We test this hypothesis by estimating regressions of the following sort:

$$\Delta\text{log TaxPaid}_{it} = \alpha + \beta \ \text{treat}_i \times \text{after}_t + \lambda_t + u_{it}, \tag{5}$$

where λ_t are the year fixed effects and treat_i is a dummy indicating that taxpayer i was ranked in a window around the cutoff in year t. We begin with a narrow window around the cutoff and gradually widen it to determine whether, as expected, the effects of the program are concentrated close to the cutoff. The TPHC program was announced before the beginning of the tax year 2012. To respond to the program, however, the taxpayers needed to know their rank. We assume this was not possible before the publication of the first set of public disclosure data. For this reason, we consider 2013 as the first post-program year. We estimate equation (5) on a sample of the top 1000 taxpayers in each of the four categories. The principal identification concern in this setting is that income, and therefore tax liability, of top taxpayers may be trending differently than others for non-program reasons such as rising inequality. We rule out this concern through non-parametric event studies and placebo falsification exercises.

III.C Data

We use data from three different sources for our empirical analysis. First, we access the public disclosure data from the FBR's website. As we note above, this data set contains the name, numerical identifier, and tax paid by every taxpayer in Pakistan for the tax years 2012-2015. The data set for MPs includes the additional identifier of the constituency number. Second, we utilize administrative data from the FBR. The administrative data include income tax returns for the tax years 2006 to 2012 (the FBR stopped providing researchers access to tax returns after that) and a master register covering the whole sample period. The tax return data contains all the line items in the tax return form. The master register includes important taxpayer characteristics such as name, tax identifier, date of registration, and taxpayer type. The last variable lets us determine if a taxpayer is self-employed, a wage-earner, a corporation, or a partnership. Combining the administrative and disclosed data, we are able to construct a panel of all taxpayers in Pakistan from 2006 to 2015.

Pakistan runs an elaborate system of what is called tax withholding. A tax remittance responsibility is triggered by a number of transactions including wage payments. For some of such transactions (not including, e.g., employer withholding), the withheld tax is treated as the final discharge of liability. For example, income tax at the rate of 1% of the value is owed on all export transactions. The remittance is due at the time the payment is received and the withheld tax is deemed as the final discharge of liability: the taxpayer does not include income from the transaction in computing taxable income, nor is he or she allowed any refund or credit for the withheld tax. Tax payments reported in the disclosure data are the sum of the tax paid on taxable income and the tax paid at source (called "final tax paid" in the Pakistani tax code). We observe both these types of tax paid in the administrative data, and are thus able to construct a consistently-defined variable that captures tax payment of each taxpayer in all years included in the panel.

Table I presents summary statistics of our sample of self-employed individuals. Treatment group comprises individuals whose Name Frequency does not exceed 40. We first compare five moments of the distributions of taxable income, tax paid on taxable income, and tax paid at source for the two pre-program years across the treatment and control samples. In subsequent rows, we compare the mean of nine taxpayer traits across the two groups. Traits in rows 4-6 capture the intensity of the program. Since the program was rolled out electronically, taxpayers in cities with greater internet access were more exposed to it. On the other hand, taxpayers with multiple businesses or with a business in a city different from the city of residence were less exposed as linking the disclosed tax to the observed lifestyle is harder in such situations. Rows 7-9 of the table explore variation in risk aversion across the two groups. Early filers are expected to be more risk-averse, whereas men and younger individuals are expected to be less risk-averse than their counterparts (Borghans *et al.*, 2009; Albert & Duffy, 2012). And finally, rows 10-12 compare the knowledge of and responsiveness to taxation among the two groups.

Rows 1-3 of the table show that the two groups are fairly evenly distributed across the taxable income and the two tax-paid distributions. But, as expected, taxpayers with more unique names are different from the others along a few dimensions. For example, they are more likely to reside in a major city and less likely to be male or old. In our empirical strategy, these fixed traits are absorbed by the individual fixed effects. Table II explores if conditioning on these fixed effects removes the correlation between the treatment and the outcome of interest. We estimate a triple-difference version of model (3) on the pre-program years (2006-2011) only, pretending 2010-11 to be the post-program years. Clearly, the outcome is not correlated with the name-uniqueness once the individual fixed effects are included in the model. None of the triple-interaction coefficients in the nine specifications is significant at the conventional level in either the complete or the balanced panel sample. To further rule out the concern that our estimates are driven by differences in observables between the less- and more-common-named taxpayers, we also report results from specifications that include the full set of interactions between salient individual characteristics—region, gender, and age—with the year fixed effects.

IV Effects of the Public Disclosure Program

IV.A Intensive Margin

Event Study—Figure II shows the results from the estimation of equation (4). We restrict the sample to a balanced panel of self-employed individuals who file in every year from 2006 to 2015. The figure plots the estimated values of the γ_js from the equation along with 95% confidence intervals. Panels A-D feature four different definitions of treatment as indicated in the title of the panel. The first decile, first quartile, median, third quartile, and top decile of the Name Frequency distribution are 4, 6, 76, 1853, and 6091, respectively. Taxpayers in the first decile of the distribution, therefore, have literally unique names: their name appears 4 times in 4 years of data. To accentuate the comparison, we drop the middle part of the distribution in Panels C-D: second and third quartiles in Panel C and deciles 2-9 in Panel D. We also report the estimated coefficients γ_js and standard errors for all four specifications in this figure in a tabular form (see Table A.III). The results strongly support our empirical strategy. There are almost no pre-existing differences between the compared groups in terms of tax payments: for all the definitions of treatment, the γ_js are indistinguishable from zero for at least four of the five pre-program years. The tax payments of the two groups diverge exactly from the time the program takes effect. This divergence is sharp and persistent. It is also larger, the larger is the difference in exposure to the program. For example, the relative differences in Panel D (bottom vs. top decile) are almost double those in Panel B (below vs. above median).

All of the specifications show evidence of a dip in the treatment effect in 2013, the second year of the program. Although we cannot test it formally, we believe that the dip results from a mass media campaign launched by the Pakistani tax administration in 2014 to increase voluntary tax compliance in the country. The campaign began in mid-September and continued till October 31st, shortly before the deadline to file the 2013 tax return (Cyan *et al.* 2017).[16] During the campaign, the administration took out advertisements in television, radio, and newspapers and sent out mobile phone text messages telling prospective taxpayers how easy it was to file taxes and how important doing so was for national development. We feel that this campaign could conceivably have nudged even the control group taxpayers to increase their tax payments, reducing the gap between the two groups. No campaign of comparable intensity was launched in any other tax year.

Regression Results—Table III reports the regression results. We estimate equation (3) on the sample of self-employed individuals using four different definitions of treatment. To keep the control group fixed across all specifications, columns (1)-(6) drop taxpayers whose Name Frequency falls between the upper bound of the treatment and 40. All specifications include individual fixed effects and allow an

[16]The tax year 2013 in our paper refers to the year that runs from July 2013 to June 2014. Cyan *et al.* (2017) refer to it as the tax year 2014 in their paper.

unrestricted variance-covariance structure at the individual level (Bertrand *et al.*, 2004).

One concern in our setting is that the public disclosure may change the composition of the sample owing to the extensive margin response. Although the individual fixed effects mitigate this concern, we rule it out even further by estimating each specification on the balanced panel sample as well (even-numbered columns). Panel B provides a direct test of the validity of the research design, estimating each specification on the pre-program periods 2006-2011 only. We define the last two years in these placebo regressions as the post-program years.

The details of the regression results affirm the visual evidence presented above. The public disclosure induces individuals with relatively unique names to report on average around 9 log points more tax liability than others. This effect is statistically significant and remarkably stable across all specifications. As expected, it drops slightly as we widen the treatment window, allowing less distinctly named individuals to enter the treatment window, a finding we explore further in the next set of results. Panel B provides evidence that validates the empirical strategy, showing that the placebo coefficient capturing any pre-existing trends in tax payments across the compared groups is trivial/insignificant in all specifications. This indicates that leveraging the variation in exposure to the program based on name uniqueness indeed isolates the treatment effect of the program.

The evidence we have presented so far is consistent with our premise that the program intensity varies proportionally with the uniqueness of a person's name. Table IV explores this idea further. We now use a more continuous definition of treatment instead of a dichotomous one, exploring how the response varies across the Name Frequency distribution. The placebo specifications in columns (3)-(4) illustrate that no systematic relationship existed between the tax payment and the name of an individual before the program. However, a strong relationship appears after the program (columns 1-2), with self-employed taxpayers having more distinct names remitting significantly more tax. This effect is strongest at the left tail of the distribution, containing the most unique names. It declines monotonically as we move rightward and becomes indistinguishable from zero as the Name Frequency approaches 300. As we note above, we do not presume that taxpayers have a precise, objective idea of how common their name is. But life experiences of persons with a very common name such as Muhammad Aslam would have instilled subjective beliefs that their name affords virtual anonymity to them. The results in Table IV show that this threshold is apparently reached at about 300. Persons with such frequent names behave as if they are aware of the objective reality that linking the disclosed information to them through their name is virtually impossible.

In another check on our empirical strategy, we now show that no significant association exists between the name and tax payment for the sample of taxpayers who are (i) well-known and (ii) identified in the disclosed records through additional, publicly-known identifiers. Table A.IV presents the results. We replicate Table III, estimating equation (3) on the sample of MPs only. Because MPs fulfill conditions (i) and (ii), we do not expect the regressions to return significant DD coefficients. Reassuringly, the results are consistent with our expectations: the uniqueness of the name of an MP is not associated with a significantly higher or lower tax payment after the program in any of the eight specifications.

Another concern is that our definition of name commonness may conflate its true population measure with the return filing behavior. For example, our definition of Name Frequency assigns the same value to a full name appearing four times in a single year or once every post-reform year. While this concern is mitigated by the fact that the distribution of names in our sample is extremely stable across years (see Figure I-B), we address it more directly in Table A.V. We now define Name Frequency as 4 $\times$ the number of times a full name appears in a given year's data.[17] Unsurprisingly, we obtain very similar results. In a related robustness check, we use a local rather than the national measure of name commonness. We define Name Frequency as the number of times a full name appears in the four years of disclosed data in a district rather than nationally. District here denotes the district identified by the first five digits of the numeric tax identifier (CNIC), which was published along with names in the 2013-2015 tax directories. The additional information hidden in the numeric tax identifier could mean that even for people with the same names the degree of exposure varies depending upon the district they live in. Table A.VI reports the result of this exercise. The estimated response becomes stronger, although the difference from the baseline is not large. This result is not a surprise for at least two reasons. First, the significance of the first five digits of the CNIC namely that they identify the district the CNIC was registered in is not

[17]We multiply the number of occurrences of a full name in a given year's data by four to make this alternative definition more compatible with the one in our baseline specification.

commonly known. Second, the tax directories are in a PDF format and list taxpayers in the alphabetical order. Looking for taxpayers of a given district is therefore not straightforward, requiring search for the five digits throughout the document. For this reason, it remains true that the costs of linking an observation in the tax directory to a taxpayer are higher the more common nationally their name is.

Summary statistics presented in Table I show that our treatment and control samples are different along few dimensions. To show that our results are not driven by any difference in observables between the two groups, we estimate an augmented version of our baseline model (3). The augmented model includes the full set of interactions of three taxpayer characteristics—gender, age, and region—with the year fixed effects, allowing taxpayers with each characteristic their own time trend. These augmented models return qualitatively similar but somewhat smaller estimates than our baseline results; compare the results in Tables A.VII, A.VIII, and A.IX with Table III.

Table A.X shows the results of our final robustness check. We estimate equation (3) restricting the sample to self-employed taxpayers whose taxable income for the baseline year (2011) falls in the window indicated in the heading of the column. This check addresses the potential concern that taxpayers with common and uncommon names might be located in different areas of the income distribution and thus would be subject to different shocks. We have already shown in Table I that this is not the case, and that our treatment and control taxpayers are distributed fairly evenly across the taxable income distribution. The results in Table A.X confirm this. Even when taxpayers having baseline income within a window of PKR 100k are compared, the tax paid by unique-named taxpayers goes up significantly after the program relative to the others, although no such difference existed prior to the program (see the placebo exercise in Panel B of the table). Another important finding shown in the table is that the response declines as we move up the taxable income distribution, becoming insignificant as the income approaches PKR 400k. This finding is consistent with the recent theoretical literature that argues that large/high-income taxpayers have far less ability to engage in tax evasion (see Gordon & Li, 2009; Kopczuk & Slemrod, 2006; Kleven *et al.*, 2016).[18]

[18]Existing empirical results are also consistent with these theoretical models. Waseem (2019), for example, finds that the evasion rate for the self-employed in Pakistan is around 74% at the bottom of the taxable income distribution but reduces to 6% as the income approaches PKR 350k. Because the response to the public disclosure program captures a reduction in tax evasion, it is not surprising

Sign of the Effect of the Public Disclosure Response—Given the difference-in-differences research design, our estimates in Table III represent the relative difference in tax payments between less-common- and more-common-named self-employed that arises from pre- to post-program periods. Under the assumption that the program had a trivial or positive effect on the tax payments of more-common-named taxpayers, this approach delivers a lower bound on the true effect. We have taken this assumption for granted so far but test it formally now. This is worth checking because there is some evidence, especially in the psychology literature, that the provision of information can sometimes backfire (see for example Schultz *et al.*, 2007). In our context, backfiring means that some individuals may start paying less taxes after the public disclosure if they perceive others are paying even less.

To sign the average effect of the program for the universe of the self-employed, we compare their tax payments with those of wage-earners. The comparison is based on the assumption that the public disclosure is unlikely to affect the tax payments of wage-earners given that their income is third-party-reported.[19] We estimate both our event study and difference-in-differences models on the complete panel of taxpayers containing both self-employed and wage-earners, defining the former category as the treatment group. The event-study model (see Figure A.IV) shows that the preexisting trends of the two groups are not parallel: the double-difference coefficient is declining—almost linearly—in the pre-program years. This trend, however, reverses quite saliently in 2012, when the DD coefficient rises for the first time, illustrating that the tax payments of the self-employed go up relative to wage-earners in that year. This remains true if we drop less-common-named self-employed from the sample (see Panel B of the figure). After 2012, the DD coefficient starts declining again but at a significantly lower rate. The event study thus shows clear signs of a structural break in 2012.

Based on these results, we estimate a slightly modified version of our difference-in-differences model (3) where we control for the preexisting trends by allowing a separate linear time trend for each of the two groups. The result are in Table

that it becomes insignificant at the higher income levels.

[19]Third-party-reported income, as we argued above, is substantially less amenable to misreporting. In fact, Waseem (2019), uses Pakistani administrative data to show that the evasion of wage income in the country in the baseline years (2006-2011) was less than 1%. With such a near-perfect compliance at the baseline, the public disclosure is unlikely to affect the tax payments of wage-earners.

A.XI. The first two columns of the table report estimates from our baseline specification for both the complete and balanced panel samples. The rest of the columns are structured similarly to the first six columns of Table III. We include a triple-interaction term in these specifications that captures the additional effect of the program on less-common-named taxpayers. Three results in the table are noteworthy. First, the estimated double-difference coefficient is positive in all specifications. This captures the average effect of the program on all self-employed in the first two columns and the average effect of the program on the more-common-named self-employed in all others. Second, the estimated triple-difference coefficient is also positive in all specifications (it also has a fairly similar magnitude to what we estimate in Table III). This shows that the program has a stronger effect (around 12 log-points) on the less-common-named self-employed. Third, the estimated double-difference coefficient is negative and the estimated triple-difference coefficient is trivial in all placebo specifications. The latter finding is particularly important in our setting, showing that the tax payments of less-common- and more-common-named self-employed were evolving similarly in the pre-program years.

While the above analyses are based on stronger assumptions than those in our baseline specification, the combined evidence from both the event study and DD model is, we believe, sufficient to rule out any boomerang effect in our setting. The effect of the public disclosure is clearly positive even for the more-common-named self-employed. This implies that our estimates in Table III, as we argued above, have a lower-bound interpretation.

Heterogeneity—Table A.XII estimates a triple-difference version of model (3), exploring if the response varies across self-employed taxpayers with the nine traits listed in Table I. The first three of these traits, as we mention above, capture program intensity. The results are consistent with our expectations. Major-city residents with greater access to the internet and hence to the disclosed data respond more aggressively; multiple businesses owners, for whom there is greater ambiguity about their earnings, respond less aggressively. We do not observe either the residence or business city for roughly one-third of the population and very likely for this reason the triple-interaction coefficient in the second column, although of the expected sign, is insignificant. The next three columns of the table explore if the response varies with the likely correlates of the degree of risk aversion of a taxpayer.[20] The

[20]There is some evidence in literature that men and young are less risk-averse than their counter-

results of this exercise are inconclusive: all the triple-interaction coefficients are of the expected sign but insignificant. The last three columns of the table look for any variation in response across taxpayers with a varying degree of knowledge of or attention to the tax system or the ability to game the tax system. We find no differential response along these margins.

Revenue Effects—How much additional revenue did the public disclosure program generate? To answer this question credibly, it is important that we take into account response heterogeneity arising both from variations in taxpayer characteristics and from variations in treatment intensity. Our results in Table A.XII show that the most important trait along which the response varies is the location of the taxpayer. Based on this result, we divide taxpayers into 16 regions. These regions indicate the tax district taxpayers file their tax return in. We then estimate our model in Table IV separately for each region. We only retain the top six Name Frequency categories of taxpayers in the model as the response for other categories is not statistically different from zero. This approach effectively divides taxpayers into 96 (16×6) cells based on their location and treatment intensity. Combining the average estimate of the response in each cell with the tax paid by individuals in the cell, we estimate that an additional amount of PKR 29.2 billion was remitted in the post-program years as a result of the program. The self-employed in Pakistan paid a total amount of PKR 412.2 billion of income tax in these years. Thus, we conclude that the public disclosure caused a nearly 7% increase in aggregate revenue paid by the self-employed—the average treatment effect of the program. Note that the approach we follow assumes that the program had no effect on more-common-named taxpayers. But this is clearly not the case as shown by our results in Table A.XI. To this extent, our estimate has a lower bound interpretation.

IV.B Extensive Margin

Event Study—Public disclosure can also encourage tax filing by individuals with less common names. To probe this, we first present visual evidence. Figure III plots the log of the number of self-employed filers in the treatment and control groups from year 2006 to 2015. We normalize the outcome variable in both groups to 1 in

parts (Borghans *et al.*, 2009; Albert & Duffy, 2012). Similarly, individuals who habitually file their tax returns earlier than others are expected to be more risk averse.

2006 and track its evolution in the later years. As earlier, we consider four definitions of treatment indicated in the heading of each panel. To make the comparison more stark, we drop the middle portion of the distribution in Panels C-D as we did in Figure II. Plots show that the program did result in more filing by less-common-named taxpayers. This effect is qualitatively very similar to the intensive margin effect, although it is smaller in magnitude. The next section formalizes this result using the regression framework.

Regression Results—Table V reports the results from the following regressions

$$\log \mathrm{N}_{gt} = \alpha + \beta \ \text{treat}_g + \gamma \ \text{treat}_g \times \text{after}_t + \lambda_t + u_{gt}, \tag{6}$$

where N_{gt} is the log number of filers of group $g \in \{\text{treat}, \text{control}\}$ in year t. Columns (1)-(4) are constructed similarly to the corresponding columns of Table III, while columns (5)-(7) correspond to the three specifications in Figure IIIB-D. Panel B of the table conducts a placebo exercise, where we estimate the above equation on the pre-program periods only, treating 2010-11 as the two post-program years. Consistent with the visual evidence, none of these placebo coefficients is significant at the conventional level, illustrating that tax filing was evolving similarly in the compared groups. After the program, however, the tax filing of less-common-named taxpayers goes up relative to the more-common-named taxpayers. The DD coefficient is statistically different from zero in all specifications, showing that the program increased filing by around 1-2%.

In the above analyses, we measure the commonness of a name using the post-program data. One concern with this approach is that policy-induced increased filing by taxpayers of a given full name can mechanically make the name more common. If it occurs for less-common-named taxpayers, they would drop out of our treatment group defined on the basis of fixed Name Frequency thresholds. This would mechanically increase the number of control taxpayers and decrease the number of treated taxpayers in the post-program years, implying that the extensive margin response we report above is underestimated. To address this concern, we repeat our analysis using an alternative measure of name commonness. This alternative measure is based on the distribution of full names as it existed in the pre-program years. Figure A.V shows this distribution. Unsurprisingly, it is very similar to the post-program distribution. Figure A.VI and Table A.XIII replicate our baseline results using the alternative measure of name-commonness. As expected, the extensive margin response is now stronger. This result shows that our baseline results underestimate the extensive margin response and that the program could have increased filing by 4-5%.

V Effects of the TPHC Program

Figure IV provides non-parametric evidence on the effects of the TPHC program. The sample for this diagram includes corporations, partnerships, self-employed and wage-earners. We group taxpayers into 20-rank bins on the basis of their rank in year t. The upper bound of a bin is included in the bin so that, for example, the bin denoted by 40 in the horizontal axis includes the taxpayers ranked between 21 and 40 in each of the four categories. We then plot the average log change in tax paid from year t to $t+1$ in the bin. To increase the power of our analysis, we take the averages over three-year periods in Panel A and over the entire pre- and post-program periods in Panel B. Because we are plotting changes rather than levels, 2012 is the first post-program year in this analysis. If the program influences behavior, the post-program curves should be significantly higher than the pre-program ones around the cutoff of 100. The evidence in the diagram is consistent with this a priori reasoning: the post-reform earnings growth curve features a clear bump at the cutoff, suggesting that taxpayers located around the eligibility cutoff of the program do increase their tax payments in order to receive or continue to receive the benefits of the program.

Table VI formalizes this analysis. We estimate equation (5) on a sample of the top 1000 taxpayers in each of the four categories. We define taxpayers in a window around the eligibility cutoff of the program as treated, and look for any differential growth in tax liability reported by them relative to other taxpayers. In line with the visual evidence, the growth rate does spike up around the cutoff. For example, the DD coefficient in the first column shows that compared to the others, the yearly growth in tax liability reported by the 81-120 ranked taxpayers was on average 17 log points higher in the post-program years than it was in the pre-program years. This additional growth of 17 log points was sufficient to take a 120th ranked taxpayer into the top 100 of the distribution for any of the post-reform years, and thus corresponds intuitively to the notion that the response represents an effort by taxpayers around the eligibility cutoff of the program to become or remain eligible. The next columns of the table show that the response declines slightly as we widen the treatment window, suggesting that the effect is stronger closer to the cutoff.

To establish that our DD coefficient captures the causal effect of the program, we need to ensure that it is not driven by any differential trends resulting from, for example, rising inequality at the top. We take three steps to achieve this. First, we re-estimate each specification in the table by adding a treat$\times 1.$(year $\in \{2010, 2011\}$) interaction term. The coefficient on the term loosely captures any differences in the pre-existing trends across the compared groups. It is small and statistically insignificant in all the specifications. Second, we estimate our model on the pre-program period only (2006-2011), pretending that the program occurred in 2010. These placebo regressions, shown in Panel B, always return insignificant coefficients. Finally, we look for the effect of the program on very similar taxpayers unaffected by it. Table A.XIV conducts this exercise. The treatment window now contains taxpayers who are relatively far away from the eligibility cutoff of the program, on whose behavior we expect the program to have no influence. The results confirm this. None of the coefficients in the table is distinguishable from zero at the conventional level.

To increase the power of our analysis, we have so far combined all four categories of taxpayers in our estimation samples. Table A.XV decomposes the aggregate response. We now estimate our baseline specification (5) separately on the sample of top 1000 taxpayers of each of the four categories. The results show that the aggregate effect we report above is driven almost entirely by the behavior of corporations. Compared to the large and statistically significant effect on corporations, the program's effect on the other three categories of taxpayers is not different from zero.

These heterogeneous findings are perhaps not surprising. Of the four taxpayer types, corporations are perhaps in the best position to monetize the goodwill offered by the program. They can build their brands by advertising their status as one of the top taxpayers, translating the social recognition into higher sales and profits. Table A.XVI evaluates this explanation by exploring response heterogeneity across firms. Strikingly, firms that are likely to be more sensitive to their reputation—public-limited firms[21] and firms engaged in consumer sectors such as banking, food, and

[21] Public limited firms are corporations whose shares can be bought and sold by the general public

textile—respond aggressively to the program. In contrast, firms that are foreign-owned, face inelastic demand (pharma), or do not operate in the consumer sector (construction) seem unaffected. Although not all of the estimated interaction terms are statistically significant, the overall pattern is consistent with both our expectations and similar evidence from other contexts showing that big firms, in particular those in the consumer sector, are relatively more sensitive to their public image, especially in issues involving social responsibility and taxes (see for example Hanlon & Slemrod, 2009; Bénabou & Tirole, 2010; Graham *et al.*, 2013).[22]

Finally, we show that our results are not driven by any differences in observables across the treatment and control groups. Table A.XVII reports summary statistics of our TPHC sample containing the top 1000 corporations, comparing thirteen outcomes/characteristics across the treatment and control groups for the two baseline years. The comparison shows that the two groups are different along few dimensions. For example, treated corporations are more likely to be located in the three major cities of Pakistan than the control corporations. For every such characteristic where the difference between the means of the two groups is statistically significant in any of the two baseline years we run a robustness check, re-estimating our baseline model including the full set of interactions of the characteristic with the year fixed effects. The results are in Table A.XVIII. Reassuringly, the inclusion of these interaction terms, allowing firms with each characteristic their own time trend, does not alter our results. The placebo specifications always return a negative and insignificant coefficient, and the main regressions a positive, large, and statistically significant coefficient.

How much additional revenue did the TPHC program generate? Combining our results in Table A.XVI with the tax paid by firms each year, we estimate that an additional amount of PKR 19.6 billion was remitted by firms ranked between 80 and 120 in the post-program years as a result of the program. This additional revenue is 1.5% of total income tax paid by the top 1000 corporations and 2.1% of total income tax paid by the top 100 corporations in these years. Taking into account

through the stock exchange. They are therefore more likely to care about their public image than private limited firms whose shares are not available to the public.

[22]One complementary mechanism driving the higher response by corporations could be the following. As we note above, the personal benefits of the program such as fast-track immigration are conferred on the CEO of the corporation. The burden of higher tax payments, on the other hand, falls on shareholders. If the oversight by the board of governors is weak, the agency problem can also result in a situation where the CEOs benefit at the cost of shareholders.

any response heterogeneity does not make a significant difference to these results. But the two estimates increase to 3% and 4.1% respectively if we consider a wider treatment window containing firms ranked between 50 and 150.

VI Conclusion

To mobilize resources, countries around the world are increasingly using programs that make tax information public, shame tax delinquents, and positively recognize top taxpayers. We analyze two such Pakistani programs to estimate their impacts on tax compliance and revenue. In the first of these programs, the government began revealing the tax liability reported by every taxpayer in the country. In the second program, the government began acknowledging and honoring top taxpayers in the country. These programs can encourage whistle-blowing, evoke shame and guilt, and inspire pride, promoting tax compliance. They could, conceivably, backfire, especially if they induce a perception that others are even less compliant.

We find that both programs elicited a substantial positive compliance response. The public disclosure caused on average a 9 log-points increase in the tax paid by individuals exposed to the program relative to the unexposed. The increase was larger the more intense was the exposure to the program. We do not find any evidence of the negative boomerang effect. The social recognition of top taxpayer also induced a substantial response. We find that the tax liability reported by treated taxpayers in the neighborhood of the program threshold went up by approximately 17 log-points. The average effect was largely driven by taxpayers for whom the reputational concerns from tax payments were first-order.

That these programs produce significant response has important implications. It shows that fear of detection and punishment as well as shame and pride may, in some settings, be meaningful determinants of behavior that economic models need to take into account. From a policy standpoint, the results show that public disclosure and social recognition of top taxpayers can be effective enforcement instruments. These programs cost little resources, and therefore can be a cost-effective complement to the other costly measures the governments undertake to deter non-compliance.

References

Abadie, Alberto, Athey, Susan, Imbens, Guido W, & Wooldridge, Jeffrey. 2017. *When should you adjust standard errors for clustering?* Tech. rept. National Bureau of Economic Research.

Akerlof, George A., & Kranton, Rachel E. 2000. Economics and Identity. *Quarterly Journal of Economics*, **115**(3), 715–753.

Albert, Steven M, & Duffy, John. 2012. Differences in risk aversion between young and older adults. *Neuroscience and neuroeconomics*, **2012**(1).

Allingham, Michael G., & Sandmo, Agnar. 1972. Income Tax Evasion: A Theoretical Analysis. *Journal of Public Economics*, **1**, 323–338.

Andreoni, James, Erard, Brian, & Feinstein, Jonathan. 1998. Tax Compliance. *Journal of Economic Literature*, **36(2)**, 818–860.

Benabou, Roland, & Tirole, Jean. 2003. Intrinsic and extrinsic motivation. *The review of economic studies*, **70**(3), 489–520.

Bénabou, Roland, & Tirole, Jean. 2010. Individual and corporate social responsibility. *Economica*, **77**(305), 1–19.

Bertrand, Marianne, Duflo, Esther, & Mullainathan, Sendhil. 2004. How much should we trust differences-in-differences estimates? *The Quarterly journal of economics*, **119**(1), 249–275.

Blank, Joshua D. 2014. Reconsidering Corporate Tax Privacy. *New York University Journal of Law & Business*, **11**, 31.

Bø, Erlend E., Slemrod, Joel, & Thoresen, Thor O. 2015. Taxes on the Internet: Deterrence Effects of Public Disclosure. *American Economic Journal: Economic Policy*, **7**(1), 36–62.

Borghans, Lex, Heckman, James J, Golsteyn, Bart HH, & Meijers, Huub. 2009. Gender differences in risk aversion and ambiguity aversion. *Journal of the European Economic Association*, **7**(2-3), 649–658.

Bursztyn, Leonardo, Ferman, Bruno, Fiorin, Stefano, Kanz, Martin, & Rao, Gautam. 2018. Status Goods: Experimental Evidence from Platinum Credit Cards. *The Quarterly Journal of Economics*, **133**(3), 1561–1595.

Castro, Lucio, & Scartascini, Carlos. 2015. Tax compliance and enforcement in the pampas evidence from a field experiment. *Journal of Economic Behavior & Organization*, **116**, 65–82.

Center for Investigative Reporting in Pakistan. 2012. *Representation Without Taxation*. Tech. rept.

Chetty, Raj. 2009. Is the Taxable Income Elasticity Sufficient to Calculate Deadweight Loss? The Implications of Evasion and Avoidance. *American Economic Journal: Economic Policy*, **1**(2), 31–52.

Cyan, Musharraf R., Koumpias, Antonios M., & Martinez-Vazquez, Jorge. 2017. The effects of mass media campaigns on individual attitudes towards tax compliance; quasi-experimental evidence from survey data in Pakistan. *Journal of Behavioral and Experimental Economics*, **70**, 10 – 22.

Del Carpio, Lucia. 2013. Are the Neighbors Cheating? Evidence from a Social Norm Experiment on Property Taxes in Peru Job Market Paper.

Dwenger, Nadja, & Treber, Lukas. 2018. Shaming for tax enforcement: Evidence from a new policy.

Dwenger, Nadja, Kleven, Henrik, Rasul, Imran, & Rincke, Johannes. 2016. Extrinsic and Intrinsic Motivations for Tax Compliance: Evidence from a Field Experiment in Germany. *American Economic Journal: Economic Policy*, **8**(3), 203–232.

Erard, Brian, & Feinstein, Jonathan. 1994. The Role of Moral Sentiments and Audit Perceptions in Tax Compliance. **49**(1), 70–89.

Fellner, Gerlinde, Sausgruber, Rupert, & Traxler, Christian. 2013. Testing Enforcement Strategies In The Field: Threat, Moral Appeal and Social Information. *Journal of the European Economic Association*, **11**(3), 634–660.

Glazer, Amihai, & Konrad, Kai. 1996. A Signaling Explanation for Charity. *American Economic Review*, **86**(4), 1019–1028.

Gordon, Roger, & Li, Wei. 2009. Tax structures in developing countries: Many puzzles and a possible explanation. *Journal of Public Economics*, **93**(7-8), 855–866.

Graham, John R, Hanlon, Michelle, Shevlin, Terry, & Shroff, Nemit. 2013. Incentives for tax planning and avoidance: Evidence from the field. *The Accounting Review*, **89**(3), 991–1023.

Hanlon, Michelle, & Slemrod, Joel. 2009. What does tax aggressiveness signal? Evidence from stock price reactions to news about tax shelter involvement. *Journal of Public Economics*, **93**(1-2), 126–141.

Hasegawa, Makoto, Hoopes, Jeffrey L, Ishida, Ryo, & Slemrod, Joel. 2012. The effect of public disclosure on reported taxable income: Evidence from individuals and corporations in Japan. *National Tax Journal*, **66(3)**, 571–608.

Hoopes, Jeffrey L, Robinson, Leslie, & Slemrod, Joel. 2018. Public tax-return disclosure. *Journal of Accounting and Economics*, **66**(1), 142–162.

Kettle, Stewart, Hernandez, Marco, Ruda, Simon, & Sanders, Michael. 2016. *Behavioral interventions in tax compliance: Evidence from Guatemala*. The World Bank.

Kleven, Henrik J., Knudsen, Martin B., Kreiner, Claus Thustrup, Pedersen, Søren, & Saez, Emmanuel. 2011. Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark. *Econometrica*, **79**(3), 651–692.

Kleven, Henrik Jacobsen, Kreiner, Claus Thustrup, & Saez, Emmanuel. 2016. Why can modern governments tax so much? An agency model of firms as fiscal intermediaries. *Economica*, **83**(330), 219–246.

Kopczuk, Wojciech, & Slemrod, Joel. 2006. Putting Firms into Optimal Tax Theory. *American Economic Review Papers and Proceedings*, **96**(2), 130–134.

Lenter, David, Slemrod, Joel, & Shackelford, Douglas. 2003. Public Disclosure of Corporate Tax Return Information: Accounting, Economics, and Legal Perspectives. *National Tax Journal*, **56**(4), 803–830.

Luttmer, Erzo F. P., & Singhal, Monica. 2014. Tax Morale. *Journal of Economic Perspectives*, **28**(4), 149–168.

Malik, Rabia. 2019. Transparency, Elections, and Pakistani Politicians' Tax Compliance. *Comparative Political Studies*.

Mascagni, Giulia. 2018. From the lab to the field: A review of tax experiments. *Journal of Economic Surveys*, **32**(2), 273–301.

Perez-Truglia, Ricardo. 2019. The effects of income transparency on well-being: Evidence from a natural experiment. NBER Working Paper No. w25622.

Perez-Truglia, Ricardo, & Troiano, Ugo. 2015. Shaming tax delinquents: Theory and evidence from a field experiment in the United States.

Scholz, John T., & Pinney, Neil. 1995. Duty, Fear, and Tax Compliance: The Heuristic Basis of Citizenship Behavior. *American Journal of Political Science*, **39**(2), 490–512.

Schultz, P Wesley, Nolan, Jessica M, Cialdini, Robert B, Goldstein, Noah J, & Griskevicius, Vladas. 2007. The constructive, destructive, and reconstructive power of social norms. *Psychological science*, **18**(5), 429–434.

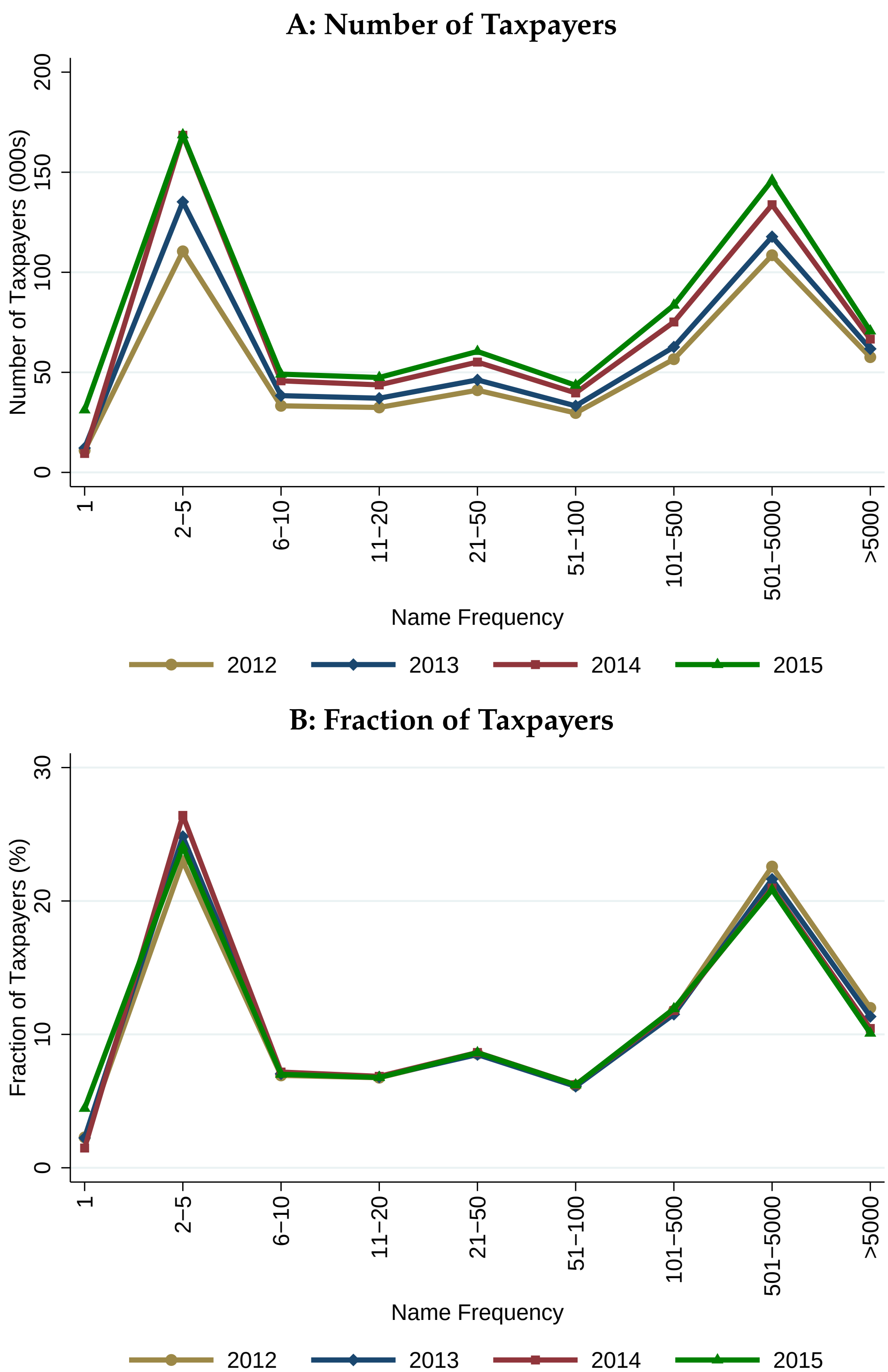
Slemrod, Joel. 2007. Cheating Ourselves: The Economics of Tax Evasion. *Journal of Economic Perspectives*, **21(1)**, 25–48.

Slemrod, Joel. 2019. Tax compliance and enforcement. *Journal of Economic Literature*, **57**(4), 904–54.

Slemrod, Joel, Blumenthal, Marsha, & Christian, Charles. 2001. Taxpayer Response to an Increased Probability of Audit: Evidence from a Controlled Experiment in Minnesota. *Journal of Public Economics*, **79**(3), 455–483.

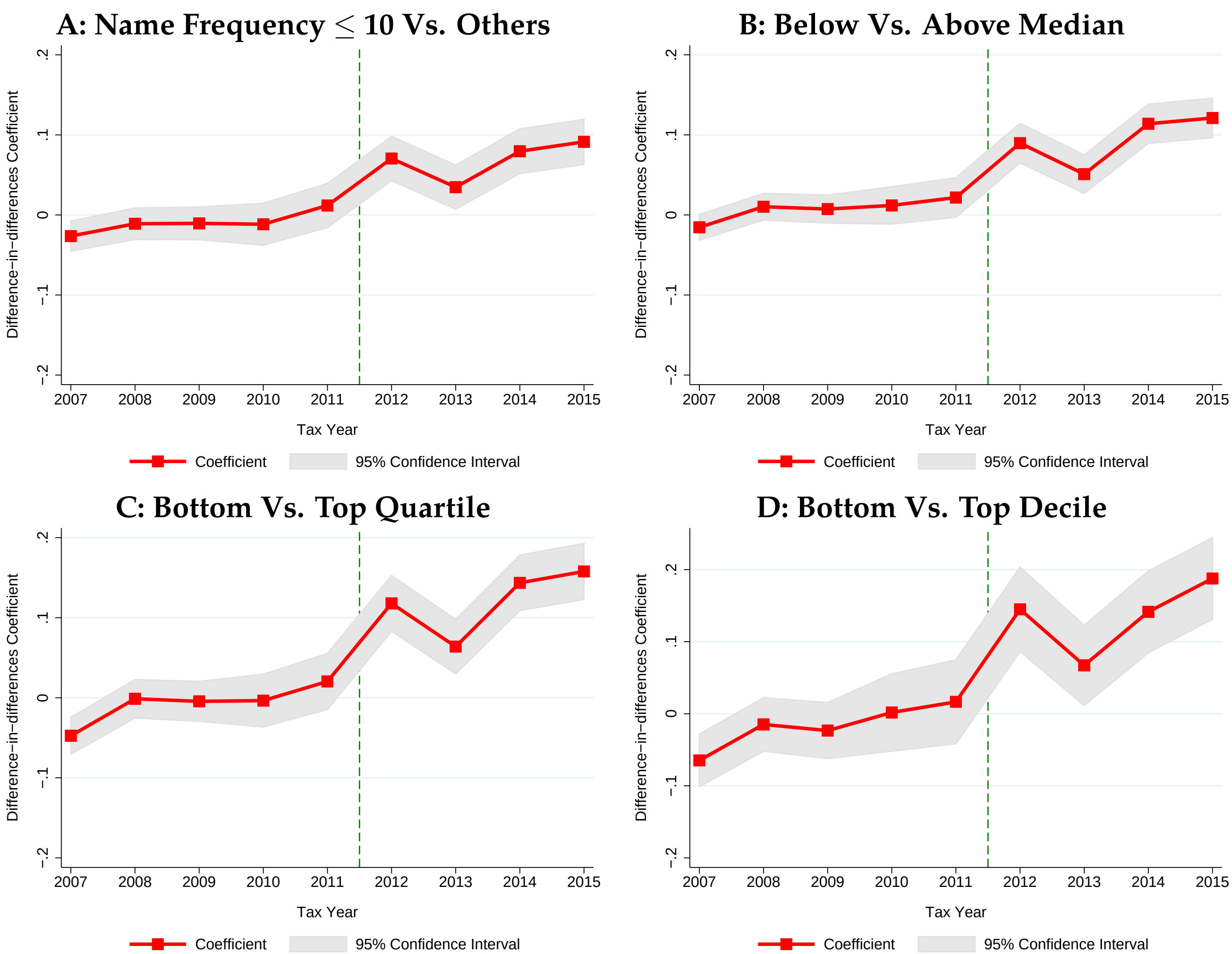
Waseem, Mazhar. 2019. Does Cutting the Tax Rate to Zero Induce Behavior Different from Other Tax Cuts? Evidence from Pakistan. Working Paper, University of Manchester.

FIGURE I: DISTRIBUTION OF NAMES



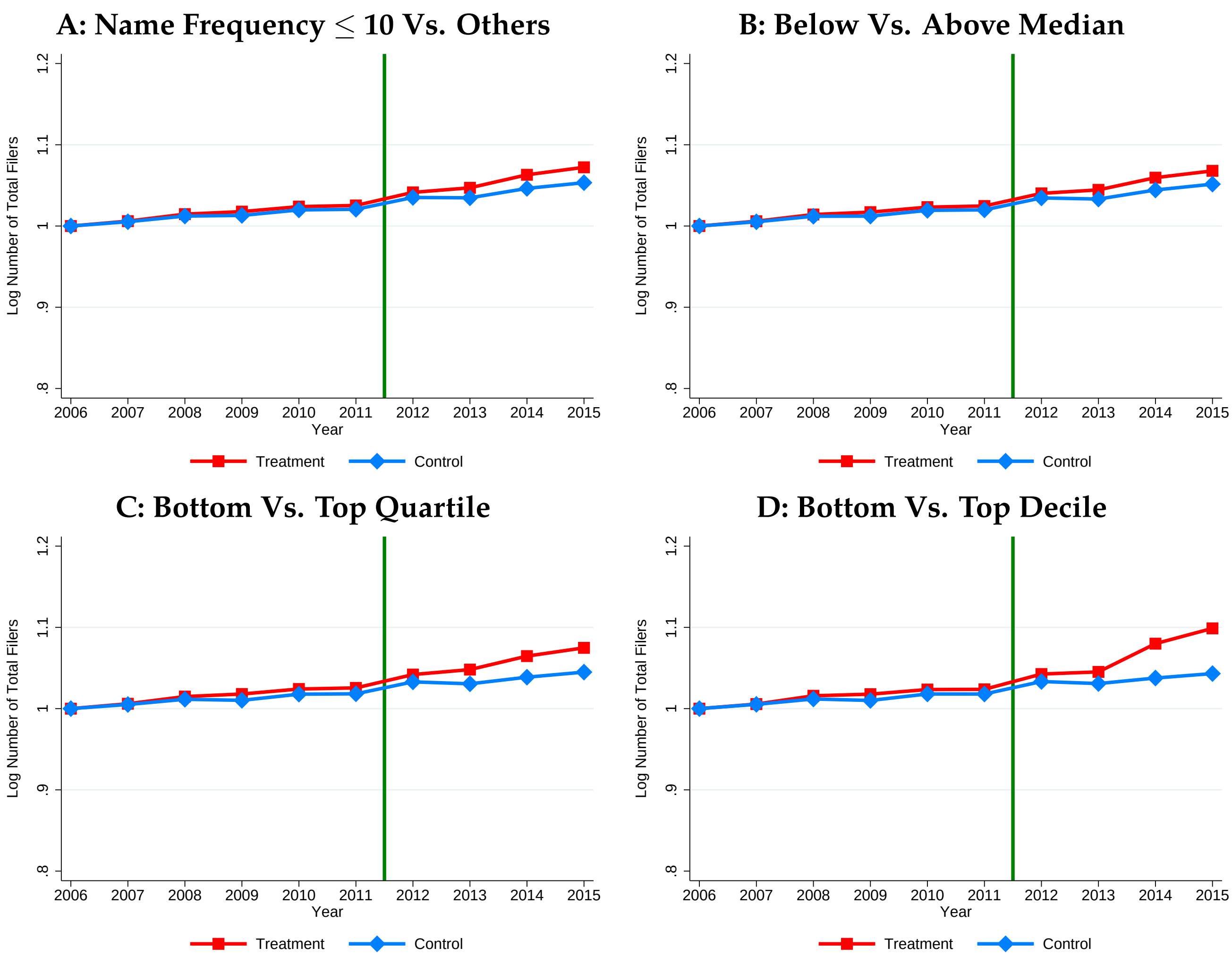
Notes: The figure illustrates the distribution of full names in Pakistan. We define Name Frequency as the number of times a full name appears in the disclosure data for the years 2012-2015. The Name Frequency of 4, for example, means that the full name appears four times in four years of data. The two panels plot the distribution of the variable. Each marker in panel A denotes the number of individuals in year t whose Name Frequency falls in the interval indicated in the horizontal axis. Panel B plots the fraction of taxpayers in place of the number. We treat all English variants of an Urdu name as one. For example Muhammad spelled as Mohammad, Mohammed, or Muhammed is treated as one name. The algorithm we use to clean such spelling variations is described in Appendix A.1.

Figure II: Intensive Margin Response to the Public Disclosure Program



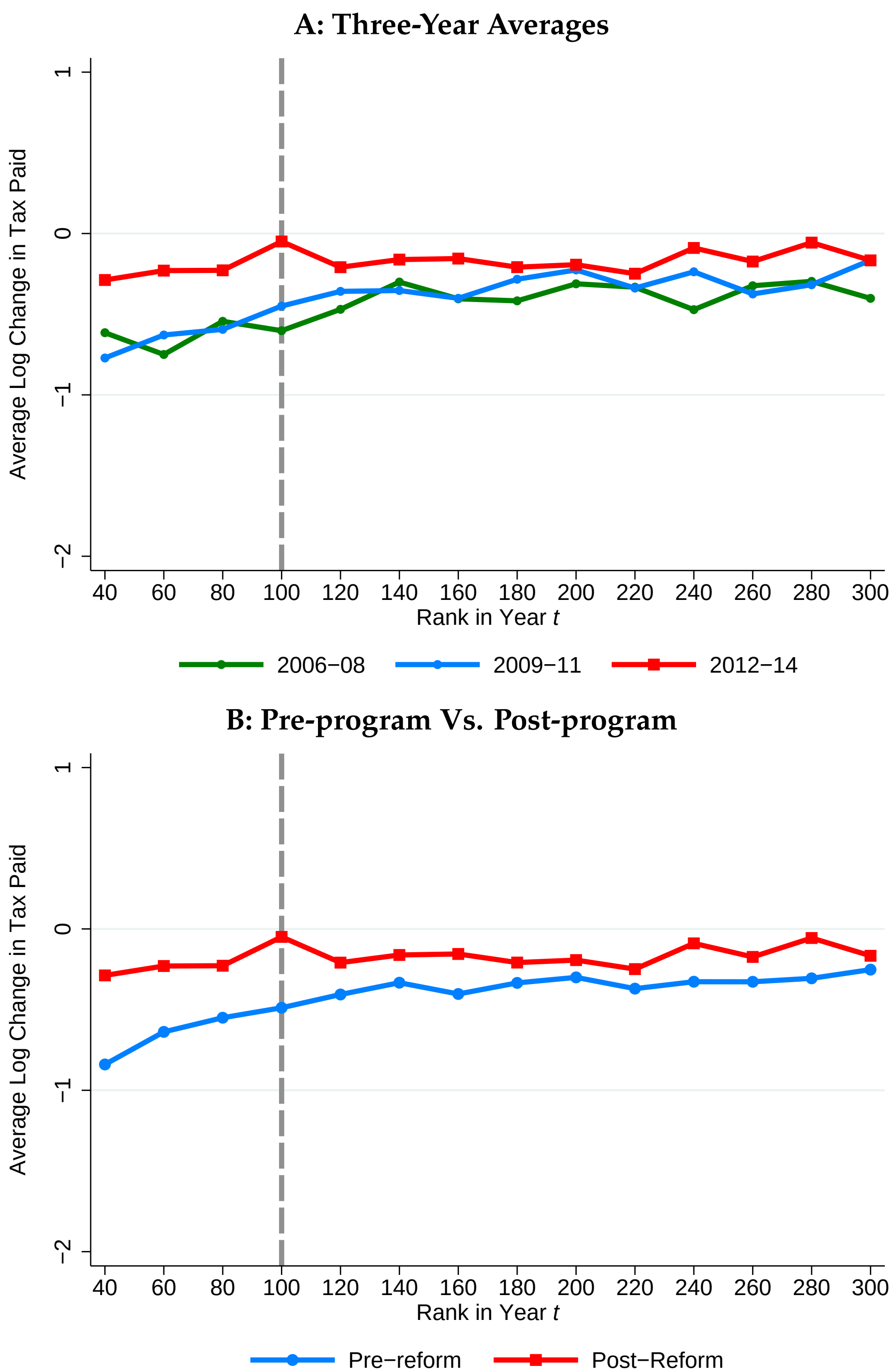
Notes: The figure plots the coefficients γ_js and 95% confidence interval around them from the event study equation (4). We estimate the equation on a balanced panel sample of self-employed taxpayers, who file in all years from 2006 to 2015. The definitions of the treatment and control groups are provided in the title of each panel. For example, for Panel A all observations where full name of the taxpayer appears at the most ten times in the four years' disclosure data are considered as treated; the rest of the taxpayers serve as the control group. The first decile, first quartile, median, third quartile, and top decile of the Name Frequency distribution are 4, 6, 76, 1853, and 6091, respectively. For Panels C-D, we drop observations in the middle of the distribution: the middle two quartiles in Panel C and the middle eight deciles in Panel D. The standard errors have been clustered at the individual level. Vertical lines demarcate the time from which the public disclosure begins to have an effect on the tax paid by individuals.

Figure III: Extensive Margin Response to the Public Disclosure Program



Notes: The figure plots the log of the number of treatment and control self-employed tax filers from 2006 to 2015. We normalize the log of the number of filers in each group to one in 2006 and track its evolution in the next nine years. The definitions of the treatment and control groups are provided in the title of each panel. For example, for Panel A all observations where full name of the taxpayer appears at the most ten times in the four years' disclosure data are considered as treated; the rest of the taxpayers are considered as the control group. For Panels C-D, we drop observations in the middle of the distribution: the middle two quartiles in Panel C and the middle eight deciles in Panel D. Vertical lines demarcate the time from which the public disclosure begins to have an effect on the tax paid by individuals.

Figure IV: **Response to the TPHC Program**



Notes: The figure explores the response to the TPHC program. We rank taxpayers in each of the four categories—self-employed, wage-earners, partnerships, and corporations—on the basis of tax paid by them in period t, group them into 20 rank bins, and plot the average log change in tax paid from period t to $t+1$ in the bin as a function of the rank in period t. Panel A takes the average over three-year periods; Panel B over the entire pre- and post-program periods. The upper bound of the bin is always included in the bin. For example, the bin indicated by 40 includes 21-40 ranked taxpayers of each category. The vertical line demarcates the eligibility cutoff of the program.

Table I: **Summary Statistics**

	2011		2010	
	Treatment	Control	Treatment	Control
	(1)	(2)	(3)	(4)
1. Taxable Income:				
25th percentile	12.281	12.255	12.044	12.017
Median	12.560	12.516	12.304	12.255
Mean	12.505	12.459	12.306	12.248
75th percentile	12.723	12.680	12.554	12.497
90th percentile	12.899	12.766	12.766	12.612
2. Tax on taxable income:				
25th percentile	10.271	10.244	10.091	10.070
Median	10.521	10.494	10.337	10.264
Mean	11.064	11.015	10.737	10.567
75th percentile	11.845	11.884	11.081	10.531
90th percentile	12.848	12.613	12.520	12.155
3. Tax at source:				
25th percentile	9.502	9.517	9.287	9.259
Median	10.917	10.943	10.625	10.540
Mean	10.915	10.984	10.678	10.687
75th percentile	12.411	12.475	12.132	12.162
90th percentile	13.699	13.804	13.450	13.526
4. Major city	0.462	0.336	0.458	0.334
	(0.001)	(0.001)	(0.001)	(0.001)
5. Business in other city	0.123	0.123	0.123	0.123
	(0.001)	(0.001)	(0.001)	(0.001)
6. Multiple businesses	0.158	0.131	0.157	0.129
	(0.001)	(0.001)	(0.001)	(0.001)
7. Male	0.919	0.986	0.924	0.986
	(0.001)	(0.000)	(0.001)	(0.000)
8. Early filer	0.615	0.642	0.554	0.543
	(0.001)	(0.001)	(0.001)	(0.001)
9. Young	0.545	0.507	0.521	0.485
	(0.002)	(0.002)	(0.002)	(0.002)
10. Buncher	0.049	0.054	0.044	0.046
	(0.000)	(0.000)	(0.000)	(0.000)
11. Strictly dominated choice	0.018	0.016	0.022	0.019
	(0.000)	(0.000)	(0.000)	(0.000)
12. Revised return	0.002	0.002	0.003	0.003
	(0.000)	(0.000)	(0.000)	(0.000)

Notes*:* The table presents summary statistics for the treatment and control groups of self-employed taxpayers. Treatment group comprises individuals whose Name Frequency does not exceed 40. We first compare five moments of the log of taxable income, tax paid on taxable income, and tax paid at source distributions for the two pre-program years across the two groups. Rest of the rows present the mean and standard error of nine taxpayer traits, all defined as dummy variables. The definitions of these dummy variables are provided in Appendix A.2 of the paper.

TABLE II: **BALANCE OF TREATMENT CONTROL SAMPLES**

	Major City	Business in Other City	Multiple Businesses	Male	Early Filer	Young	Buncher	Dominated	Revised Return
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A: Complete Panel (2006-2011)									
treat × after	0.002	0.001	0.010	0.012	-0.000	-0.016	0.002	0.014	0.014
	(0.008)	(0.009)	(0.007)	(0.006)	(0.009)	(0.014)	(0.008)	(0.006)	(0.006)
treat × trait × after	0.003	-0.011	-0.012	-0.001	0.021	-0.017	0.025	-0.001	0.070
	(0.013)	(0.026)	(0.019)	(0.044)	(0.013)	(0.021)	(0.013)	(0.030)	(0.058)
Observations	1,484,133	917,213	1,484,174	1,482,108	1,430,873	574,137	1,496,374	1,496,374	1,496,374
B: Balanced Panel (2006-2011)									
treat × after	-0.007	-0.004	0.007	0.007	-0.001	-0.010	0.004	0.009	0.009
	(0.010)	(0.011)	(0.008)	(0.008)	(0.011)	(0.017)	(0.011)	(0.008)	(0.008)
treat × trait × after	0.023	-0.020	-0.016	-0.028	0.016	-0.038	0.010	0.027	0.060
	(0.016)	(0.034)	(0.024)	(0.058)	(0.016)	(0.026)	(0.015)	(0.034)	(0.064)
Observations	837,536	486,993	837,550	837,147	807,171	288,788	840,469	840,469	840,469
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table illustrates that conditional on the individual fixed effects the evolution of our outcome variable is independent of taxpayer traits shown in the column headings, listed in Table I, and defined in A.2. We estimate a triple-difference version of model (3) on the pre-program years 2006-2011, defining the last two years as the $after$ years. The sample is all self-employed taxpayers. Treatment here is defined as an individual whose Name Frequency does not exceed 40. To avoid making strong functional form assumptions all traits are introduced into the equation nonparametrically, as dummy variables. The model includes a full set of double-interaction terms. Panel B reports the results for a balanced panel sample, where we include only the taxpayers who file in all years included in the sample.

Table III: **Intensive Margin Response to the Public Disclosure Program**

	Treat: Name Frequency							
	≤ 10		≤ 20		≤ 30		≤ 40	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Main Regression (2006-2015)								
treat $\times$ after	0.094	0.093	0.090	0.089	0.089	0.086	0.088	0.086
	(0.006)	(0.009)	(0.005)	(0.008)	(0.005)	(0.008)	(0.005)	(0.008)
Observations	2,430,002	773,038	2,614,754	833,675	2,720,267	868,250	2,792,270	891,420
B: Placebo Regression (2006-2011)								
treat $\times$ after	0.009	0.005	0.013	0.009	0.013	0.010	0.014	0.010
	(0.007)	(0.008)	(0.006)	(0.008)	(0.006)	(0.008)	(0.006)	(0.008)
Observations	1,307,541	734,269	1,403,240	787,845	1,458,457	818,942	1,496,374	840,469
Sample:								
Balanced Panel	No	Yes	No	Yes	No	Yes	No	Yes
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table reports the estimates from equation (3). For Panel A, we estimate the equation on a sample containing all self-employed individuals for the period 2006-2015. The definition of the treatment variable is provided in the title of each column. The dummy variable takes the value 1 if the Name Frequency of an individual does not exceed the cutoff indicated in the title. To maintain a fixed control group across all columns, we drop taxpayers with Name Frequency between 10 and 40 in Columns (1) to (6). Even-numbered columns restrict the sample to a balanced panel of taxpayers, who file in all years included in the sample. Treatment cutoffs of 10 and 40 correspond to the 33rd and 46th percentiles of the Name Frequency distribution for our baseline specification and 30th and 44th percentiles for our balanced-panel specification. Panel B reports the results from parallel placebo regressions, where the sample is restricted to tax years 2006 to 2011, with the last two years defined as the post-program years. Standard errors are in parenthesis, which have been clustered at the individual level.

TABLE IV: **PUBLIC DISCLOSURE RESPONSE ACROSS THE NAME DISTRIBUTION**

	Baseline Specification (2006-2015)		Placebo Specification (2006-2011)	
	(1)	(2)	(3)	(4)
Name Freq $\in (0,\ 50] \times$ after	0.107	0.105	0.020	0.013
	(0.005)	(0.008)	(0.007)	(0.008)
Name Freq $\in (50,\ 100] \times$ after	0.067	0.069	0.014	0.003
	(0.011)	(0.016)	(0.014)	(0.016)
Name Freq $\in (100,\ 150] \times$ after	0.061	0.080	0.027	0.036
	(0.015)	(0.023)	(0.019)	(0.023)
Name Freq $\in (150,\ 200] \times$ after	0.050	0.046	0.029	0.034
	(0.019)	(0.029)	(0.025)	(0.030)
Name Freq $\in (200,\ 250] \times$ after	0.043	0.011	0.014	-0.005
	(0.021)	(0.031)	(0.026)	(0.032)
Name Freq $\in (250,\ 300] \times$ after	0.045	0.022	-0.014	-0.027
	(0.022)	(0.033)	(0.028)	(0.036)
Name Freq $\in (300,\ 350] \times$ after	0.047	0.086	0.032	0.042
	(0.025)	(0.038)	(0.032)	(0.039)
Name Freq $\in (350,\ 400] \times$ after	0.037	0.039	0.028	0.021
	(0.027)	(0.041)	(0.037)	(0.043)
Name Freq $\in (400,\ 450] \times$ after	0.035	0.017	0.017	0.029
	(0.026)	(0.039)	(0.033)	(0.041)
Observations	2,792,270	891,420	1,496,374	840,469
Sample:				
Balanced Panel	No	Yes	No	Yes
Individual Fixed Effects	Yes	Yes	Yes	Yes

Notes*:* The table explores how the intensive margin response to the public disclosure program varies across the name distribution. We estimate an augmented version of equation (3), including the nine interaction terms shown above. The equation is estimated on a sample of all self-employed individuals. The control group in these regression are the self-employed whose Name Frequency exceeds 450. The coefficient on each interaction terms accordingly captures the average additional tax paid (in log points) by the self-employed with Name Frequency falling in the interval as a result of the program. Columns (1) and (2) report the results for the baseline specification containing periods 2006-2015, both for the complete and balanced panels. Columns (3) and (4) estimate the specifications on the pre-program years only, defining the years 2010 and 2011 as the post-program period. Standard errors are in parenthesis, which have been clustered at the individual level.

Table V: **Extensive Margin Response to the Public Disclosure Program**

	Treat: Name Frequency						
	≤ 10	≤ 20	≤ 30	≤ 40	$\leq$ Median	$\leq$ 1st Quartile	$\leq$ 1st Decile
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A: Main Regression (2006-2015)							
treat $\times$ after	0.0117	0.0106	0.0101	0.0097	0.0094	0.0163	0.0265
	(0.0027)	(0.0024)	(0.0023)	(0.0022)	(0.0022)	(0.0041)	(0.0089)
B: Placebo Regression (2006-2011)							
treat $\times$ after	0.0027	0.0027	0.0026	0.0025	0.0024	0.0038	0.0026
	(0.0018)	(0.0017)	(0.0017)	(0.0016)	(0.0016)	(0.0026)	(0.0027)

Notes: The table reports the estimates from equation (6). The equation is estimated on a sample of all self-employed individuals. The outcome variable here is the log number of filers in group g in year t. Panel A estimates the equation on the period 2006-2015. The definition of the treatment variable is provided in the title of each column. The dummy variable takes the value 1 if the Name Frequency of an individual does not exceed the cutoff indicated in the title. To maintain a fixed control group across columns (1)-(4), we drop taxpayers with the Name Frequency between 10 and 40 in columns (1) to (3). In columns (6) and (7) we drop the middle part of the distribution: the middle two quartiles in column (6) and the deciles 2-9 in column (7). Panel B reports the results from parallel placebo regressions, where the sample is restricted to tax years 2006 to 2011, with the last two years defined as the post-program years. Standard errors are in parenthesis.

Table VI: Response to the TPHC Program

	Treat: Rank							
	$\in (80,\ 120]$		$\in (70,\ 130]$		$\in (60,\ 140]$		$\in (50,\ 150]$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Main Regression (2006-2014)								
treat $\times$ after	0.166	0.138	0.171	0.161	0.136	0.126	0.140	0.128
	(0.075)	(0.077)	(0.062)	(0.064)	(0.054)	(0.055)	(0.048)	(0.049)
treat $\times$ 1.(year $\in$ {2010,2011})		-0.163		-0.060		-0.058		-0.070
		(0.151)		(0.126)		(0.115)		(0.105)
Observations	32,047	32,047	32,047	32,047	32,047	32,047	32,047	32,047
B: Placebo Regression (2006-2010)								
treat $\times$ after	0.019		0.010		-0.086		-0.090	
	(0.120)		(0.102)		(0.091)		(0.081)	
Observations	17,208		17,208		17,208		17,208	

Notes: The table reports the results from the equation (5). We estimate the equation on a sample containing top 1000 taxpayers of each of the four categories of taxpayers, corporations, partnerships, self-employed, and wage-earners. The treatment variable here denotes taxpayers ranked in period t in a window around the eligibility cutoff of the program. The exact length of the treatment window is indicated in the title of each column. Given that we measure the outcome variable here in changes rather than levels, the first post-program year is 2012. Panel A estimates the equation on years 2006-2014. Panel B runs parallel placebo regressions on years 2006-2010, with the last two years defined as the post-program years. Columns (2), (4), (6) and (8) test the parallel trend assumption by including a $treat \times 1.(year \in \{2010, 2011\}$ interaction into the regression. Standard errors are in parenthesis, which have been clustered at the individual level.

A Online Appendix

A.1 Name Cleaning Algorithm

Identifying Potential Spelling Variations in Pakistani Names

Most Pakistani names are derived from Arabic, Persian or Turkish. Like Urdu, these languages are (or were) written in variants of the Arabic script. As a result the spelling variations in Pakistani names arise mainly because of standard issues in transliterating Arabic script into English.

The most common issue is the spelling of transliterated vowel sounds. As there are no standardized rules for transliteration each vowel sound can be spelled in many different ways. In Urdu, shorter vowel sounds are not indicated through separate letters. So, for example, the name Muhammad in Urdu is spelled with only four letters - MHMD. In transliterating the name to English there is considerable discretion as to what English vowels will be used for the sound in each syllable. The first syllable can be spelled as M, MA, MO, MU, MUA, MOU, MU; the second syllable as HAM, HUM, HOM, and the third syllable as MED, MAD, MD. The various combinations of these syllables generates multiple spellings for the same name.

In Urdu, some longer vowel sounds are indicated through specific letters. However the spelling issue still persists in these cases because of a lack of transliteration rules. For example the name Mehmood in Urdu is spelled with five letters - MHMUD. The added vowel represents the "oo" sound as in "rude" but it can be spelled in English as either U OO OU or UO.

Secondly, in Urdu elongated sounds or sounds that are repeated across syllables are not indicated through double letters (as is often the case in English) but are also expressed through accent marks. Again taking the case of the name Muhammad, the middle "m" sound is repeated but spelt with a single letter in Urdu. In English the repeated sound can be spelled as M or MM depending on whether the spelling is based on the Urdu spelling or the phonetic sound.

So for a given Urdu name, the vowel and repeated sounds imply potential spelling variations which we use to identify variants of the same name.

Standardizing Full Names

The tax directory published by the Federal Board of Revenue (FBR) lists each taxpayer's full name. We combine the tax directories for all "Individual" taxpayers for 2012-2015 to get an exhaustive list of all full names that have ever appeared in the disclosure data. We then split the full names, based on spaces or hyphens, into the different (given or family) single names they constitute. This gives us a master list of all distinct single names in the data.

Given the possible spelling variations we manually work through this master list to identify the English variants of the same Urdu names. By convention, certain spellings of names have become more common and widely used. Each name variant is standardized to the most common spelling used for that name in the data. After the spellings of the single names are standardized we combine them back again to create standardized full names. The name frequency measures we use in the analysis are based on these standardized full names.

A.2 Definition of Variables

(*i*) **Major city.** The taxpayer reports an address in one of the three major cities—Karachi, Lahore, and Islamabad—of Pakistan.

(*ii*) **Business in other city.** The taxpayer conducts business in a city different from where he or she resides.

(*iii*) **Multiple businesses.** The taxpayer owns more than one businesses.

(*iv*) **Early filer.** The taxpayer files their return relatively early. The dummy variable takes the value 1 if the taxpayer filed their return for year t before the median filing date for the year.

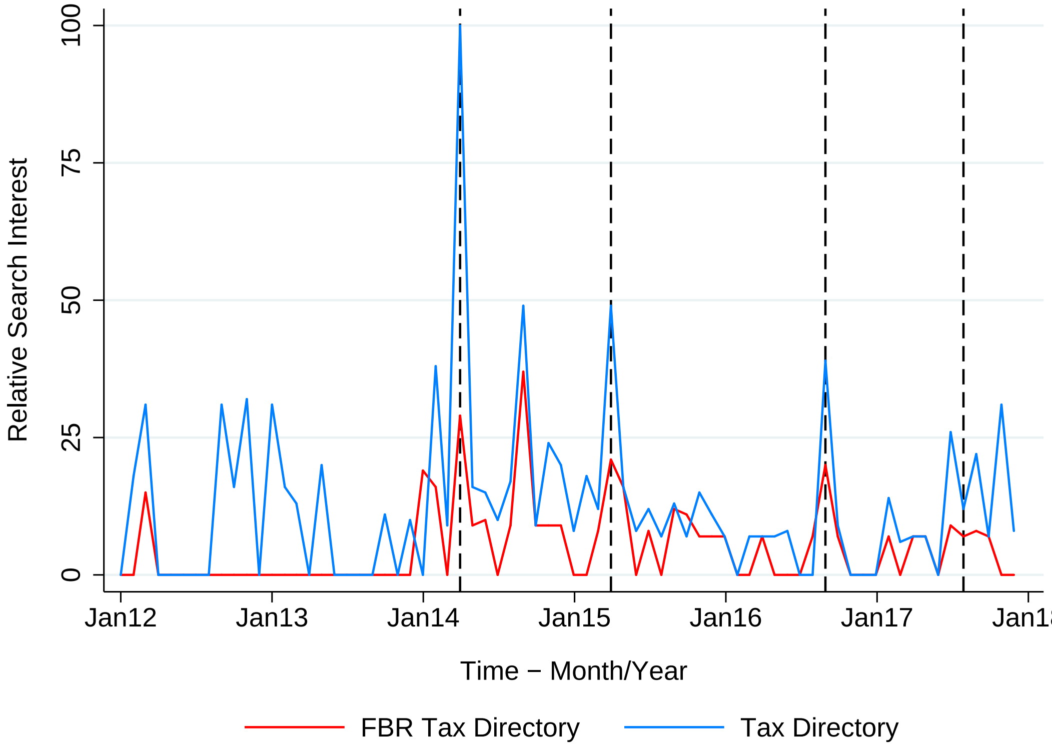
(*v*) **Young.** If the taxpayer is younger than the median income tax filer for the year t.

(*vi*) **Buncher.** If the taxpayer reported income at or within a window of ten thousand PKR below any notch in the corresponding tax schedule.

(*vii*) **Strictly dominated choice.** If the taxpayer reported income within the strictly dominated region above any notch in the corresponding tax schedule.

(*viii*) **Revised return.** If the taxpayer filed a revised return for the given tax year t.

Figure A.I: **Google Search Interest**

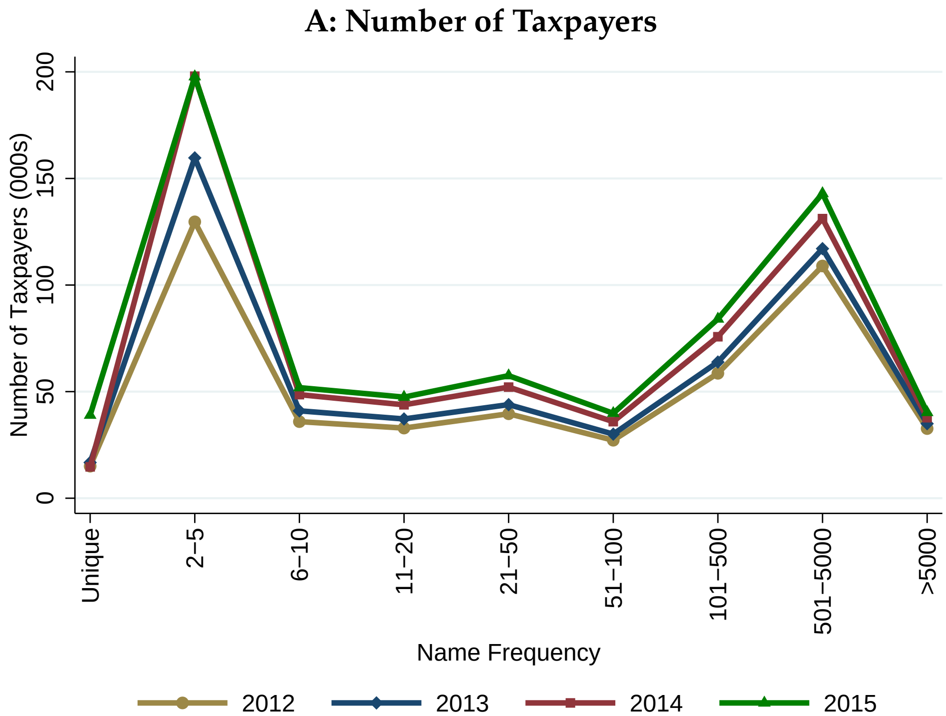


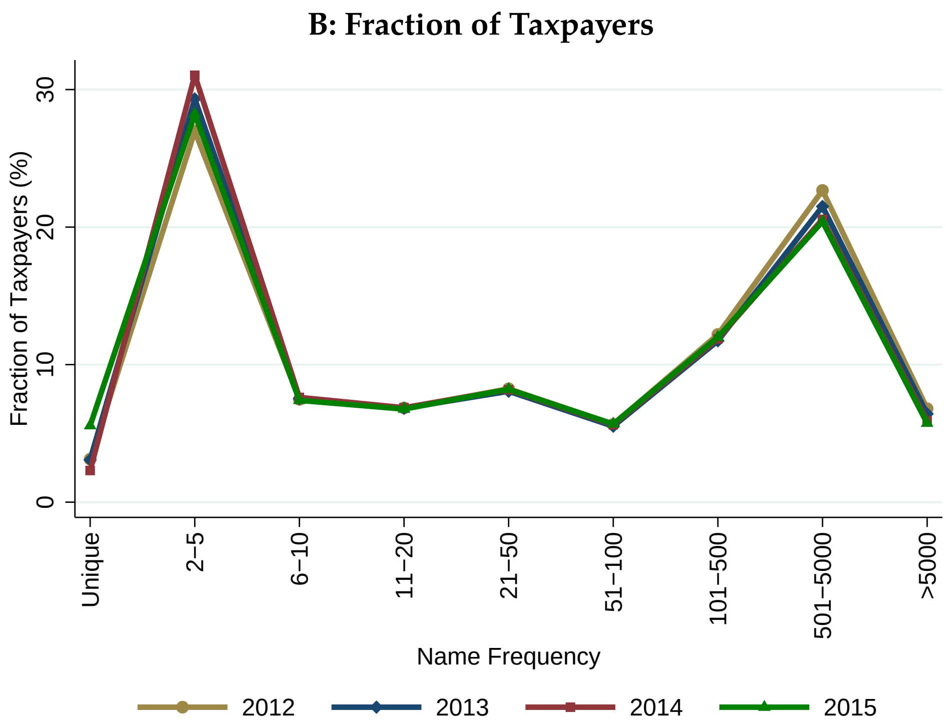
Notes: The figure plots Google Trends data for the monthly search interest in Pakistan for the terms "FBR Tax Directory" and "Tax Directory" from January 2012 to January 2018. The data is normalized by time and location and scaled on a range of 0 - 100 to compare relative popularity. The data point with the highest search queries within the specified time and location is given a score of 100 and other points are scored relative to it. Vertical lines demarcate the months in which the tax directories were released. Directories for tax years 2012, 2013, 2014 and 2015 were released in April 2014, April 2015, September 2016 and August 2017 respectively.

Figure A.II: **Special Immigration Counter for TPHC Holders**

Notes: The figure shows the picture of special immigration counter at the Allama Iqbal International Airport, Lahore. The picture was taken in the summer of 2018.

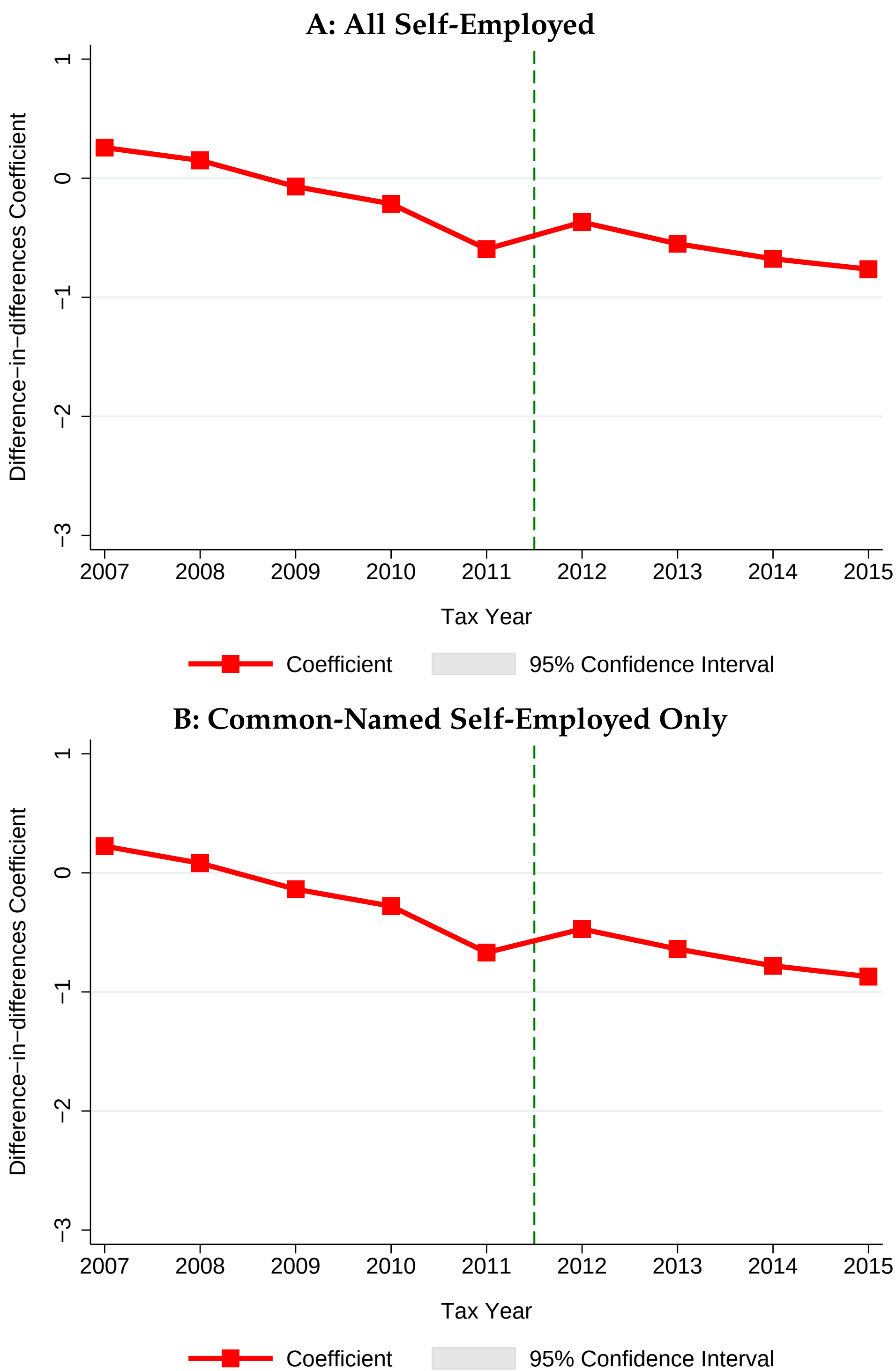
Figure A.III: **Distribution of Names – Original Spelling**





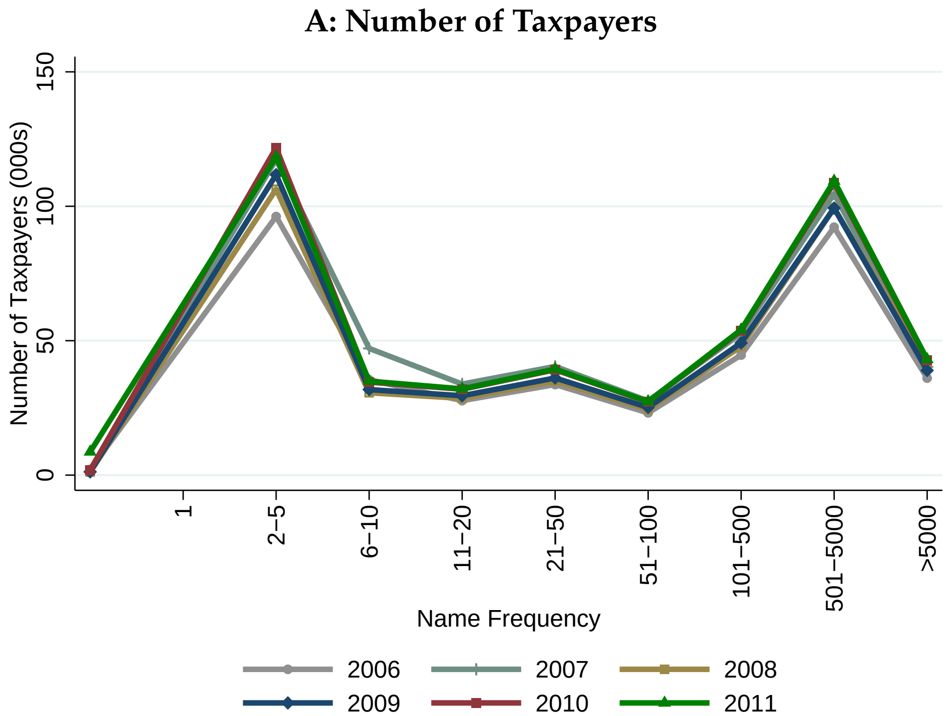
Notes: The figure illustrates the distribution of full names in Pakistan. We define Name Frequency as the number of times a full name appears in the disclosure data for the years 2012-2015. The Name Frequency of 4, for example, means that the full name appears four times in four years of data. The two panels plot the distribution of the variable. Each marker in panel A denotes the number of individuals in year t whose Name Frequency falls in the interval indicated in the horizontal axis. Panel B plots the fraction in place of the number. Here, we treat all English variants of an Urdu name as distinct names. For example Muhammad, Mohammad, Mohammed, and Muhammed are treated as distinct names.

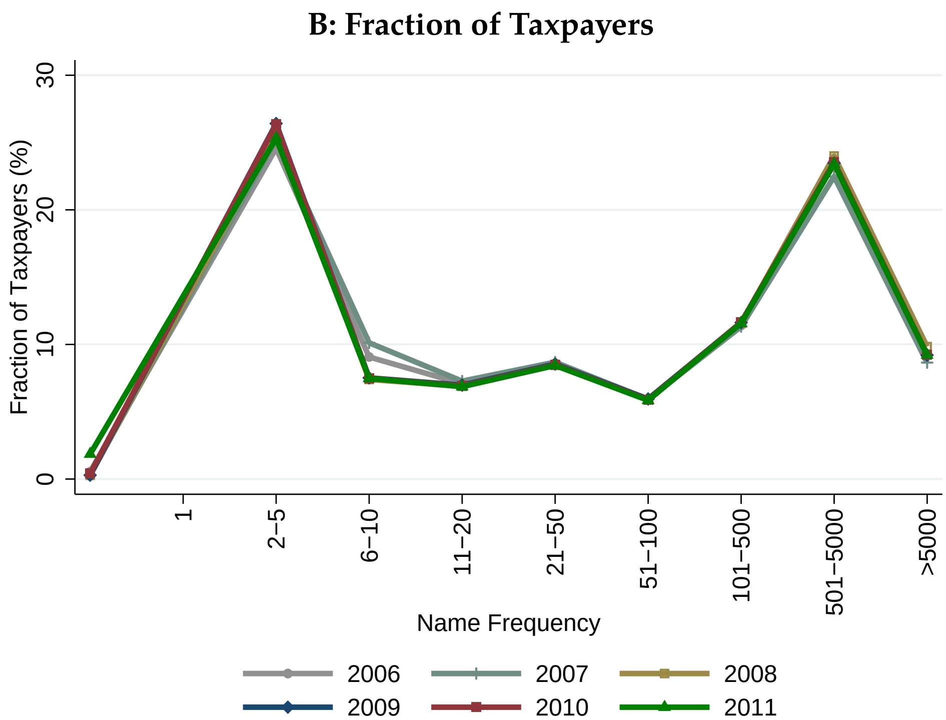
Figure A.IV: **Evolution of Tax Payments – Self-Employed Vs. Wage-Earners**



Notes: The figure compares the evolution of tax payments across self-employed and wage-earners. We plot the coefficients γ_js and 95% confidence intervals around them from the event study equation (4). The equation is estimated on the complete panel of taxpayers containing both self-employed and wage-earners. We define self-employed as the treated group. Panel A includes all self-employed, whereas Panel B drops the self-employed with Name Frequency less than or equal to 10. Note that the 95% confidence interval around the DD coefficient is so tight that it is barely visible. Vertical lines demarcate the time from which the public disclosure begins to have an effect on the tax paid by the self-employed.

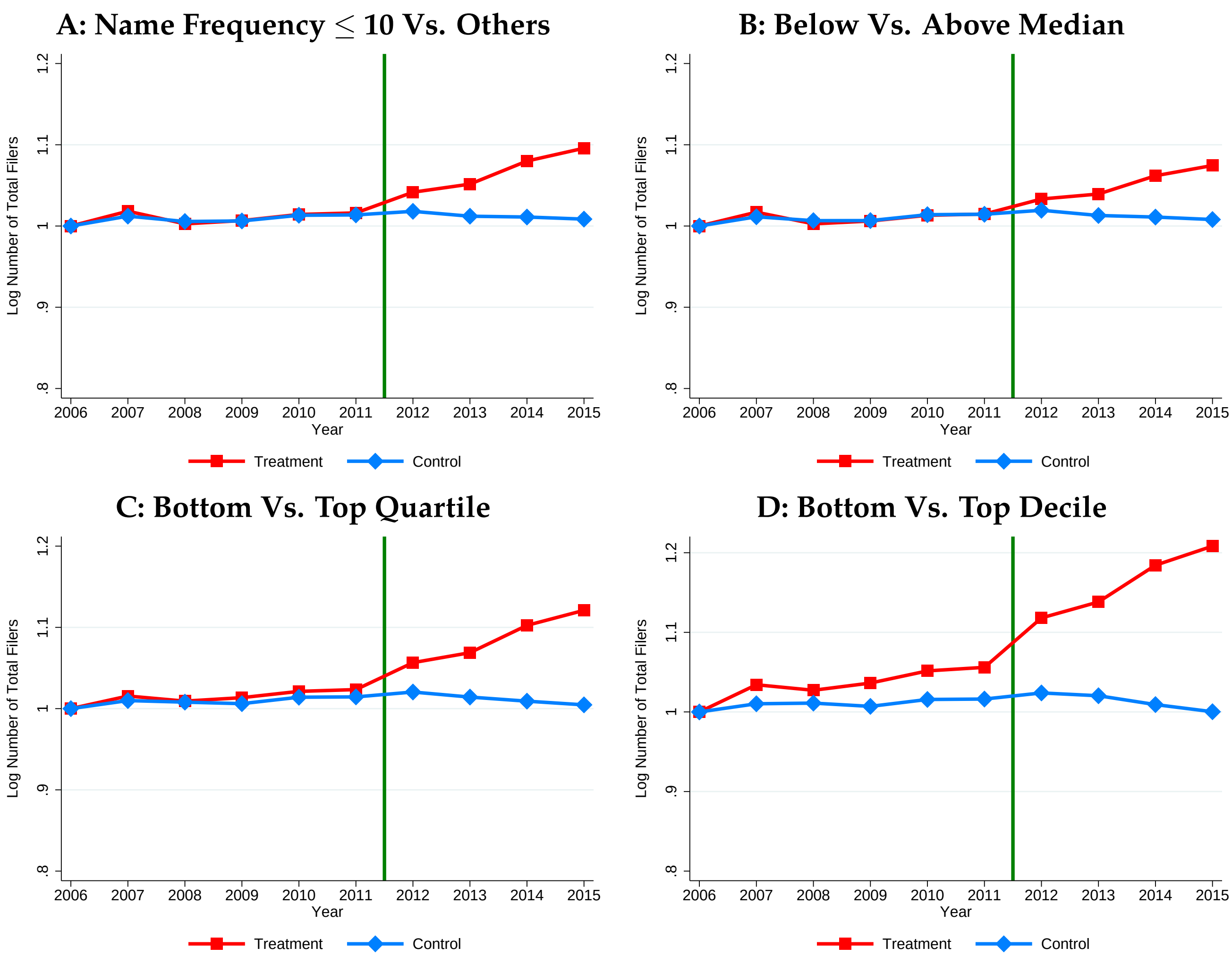
Figure A.V: Distribution of Names – Pre-Program Years (2006-2011)





Notes: The figure illustrates the distribution of full names in Pakistan, as it existed at the baseline. We define Name Frequency as the number of times a full name appears among tax filers in the six baseline years 2006-2011. We normalize this measure of Name Frequency by a factor of 4/6 to make it compatible with the measure we use for all other results, where we measure Name Frequency as the number of times a full name appears in the four years of disclosed data 2012-2016. The Name Frequency of 4, for example, in this figure means that the full name appears six times in the six years of data. The figure replicates the two panels of Figure I using this alternative definition of Name Frequency.

Figure A.VI: Extensive Margin Response to the Public Disclosure – Baseline Frequency



Notes: The figure conducts a robustness check on our extensive margin result. We replicate Figure III using an alternative definition of Name Frequency, measuring it as the number of times a full name appears among the tax filers in the six baseline years 2006-2011. We multiply this measure of Name Frequency with a factor of 4/6 to make it compatible with the definition used in Figure III and our other results. We plot the log of the number of treatment and control self-employed tax filers from 2006 to 2015. We normalize the log of the number of filers in each group to one in 2006 and track its evolution in the next nine years. The definitions of the treatment and control groups are provided in the title of each panel. For example, for Panel A all observations where the Name Frequency using our alternative definition does not exceed 10 are considered as treated; the rest of the taxpayers are considered as the control group. For Panels C-D, we drop observations in the middle of the distribution: the middle two quartiles in Panel C and the middle eight deciles in Panel D. Vertical lines demarcate the time from which the public disclosure begins to have an effect on the tax paid by individuals.

Table A.I: **Timeline of the Public Disclosure Program**

Date (1)	Event (2)
Sep-Dec, 2012	Investigative reports alleging tax noncompliance by MPs begin appearing in the press
December, 2012	First CIRP report published. It publishes the data that formed the basis of earlier investigative reports, cataloging tax noncompliance of MPs elected in the 2008-2013 election cycle of Pakistan
December, 2012	The Federal Tax Ombudsman orders the FBR to begin disclosing the tax paid by every public office holder in the country
January, 2013	The leading opposition party and eventual election winner, PML-N, issue election manifesto, pledging the public disclosure of tax paid by all taxpayers in the country
May 11, 2013	General elections
June 30, 2013	Tax year 2012 ends
December 15, 2013	Final date for filing of 2012 tax return
December, 2013	Second CIPR report published. It documents the tax payments of MPs who won during the 2013 elections
February 28, 2014	MPs' directory for tax year 2012 published
April 15, 2014	All taxpayers' directory for tax year 2012 published
June 30, 2014	Tax year 2013 ends
April 10, 2015	MPs' and all taxpayers' directories for tax year 2013 published
June 30, 2015	Tax year 2014 ends
June 30, 2016	Tax year 2015 ends
September 9, 2016	MPs' and all taxpayers' directories for tax year 2014 published
July 27, 2017	MPs' directory for tax year 2015 published
August 11, 2017	All taxpayers' directory for tax year 2015 published

Notes*:* The table report the timeline of important events in the public disclosure program. The date each event listed in column (2) occurred is given in column (1). Pakistani tax year runs from July to June. Tax year indicated by t in this paper runs from July t to June $t+1$. The first CIRP report indicated in the second row is available here; the second report indicated in the eighth event is available here. Tax directories of all years can be downloaded from here.

Table A.II: **Structure of Pakistani Legislature**

House	Total Seats	Directly Elected	Reserved		
			Women	Minorities	Technocrats
(1)	(2)	(3)	(4)	(5)	(6)
National Assembly	342	272	60	10	-
Senate	104	66	17	4	17
Punjab Assembly	371	297	66	8	-
Sind Assembly	168	130	29	9	-
KP Assembly	124	99	22	3	-
Balochistan Assembly	65	51	11	3	-
Total	1174	915	205	37	17

Notes*:* The table shows the composition of the Pakistani legislature. National Assembly and Senate are the two houses at the Federal level. Pakistan has four provinces: Punjab, Sind, Khyber Pakhtoonkhwah (KP), and Balochistan. Each province has its own legislature. The legislative powers are divided between the federation and provinces by the constitution. Seats are reserved for women and religious minorities (non-Muslims) in every house and for technocrats in Senate. Reserved seats are filled through a proportional representation system.

TABLE A.III: **INTENSIVE MARGIN RESPONSE TO THE PUBLIC DISCLOSURE PROGRAM – DYNAMICS**

	Treat: Name Frequency			
	$\leq$ 10	$\leq$ Median	$\leq$ First Quartile	$\leq$ First Decile
	(1)	(2)	(3)	(4)
treat $\times$ 2007	-0.026	-0.015	-0.047	-0.065
	(0.010)	(0.008)	(0.012)	(0.019)
treat $\times$ 2008	-0.011	0.010	-0.001	-0.015
	(0.010)	(0.009)	(0.012)	(0.019)
treat $\times$ 2009	-0.011	0.007	-0.005	-0.023
	(0.010)	(0.009)	(0.013)	(0.020)
treat $\times$ 2010	-0.012	0.012	-0.004	0.002
	(0.013)	(0.012)	(0.017)	(0.027)
treat $\times$ 2011	0.012	0.022	0.020	0.017
	(0.014)	(0.013)	(0.018)	(0.030)
treat $\times$ 2012	0.071	0.090	0.118	0.145
	(0.014)	(0.013)	(0.018)	(0.030)
treat $\times$ 2013	0.035	0.051	0.064	0.067
	(0.014)	(0.012)	(0.017)	(0.028)
treat $\times$ 2014	0.080	0.114	0.144	0.141
	(0.014)	(0.013)	(0.018)	(0.029)
treat $\times$ 2015	0.091	0.121	0.158	0.188
	(0.014)	(0.013)	(0.018)	(0.029)
Observations	891,420	891,420	451,158	242,944
Sample:				
Balanced Panel	Yes	Yes	Yes	Yes
Individual Fixed Effects	Yes	Yes	Yes	Yes

Notes*:* The table reports the coefficients γ_js along with standard errors from our event study equation (4). These coefficients and the 95% confidence intervals around them are plotted in Figure II. We estimate the equation on a balanced panel sample of self-employed taxpayers, who file in all years from 2006 to 2015. The definitions of the treatment and control groups are provided in the title of each column. For example, for column (1) all observations where full name of the taxpayer appears at the most ten times in the four years' disclosure data are considered as treated; the rest of the taxpayers serve as the control group. The first decile, first quartile, median, third quartile, and top decile of the Name Frequency distribution are 4, 6, 76, 1853, and 6091, respectively. For columns (3) & (4), we drop observations in the middle of the distribution: the middle two quartiles for column (3) and the middle eight deciles for column (4). The standard errors have been clustered at the individual level.

Table A.IV: **Intensive Margin Response to the Public Disclosure Program – Placebo**

	Treat: Name Frequency							
	≤ 10		≤ 20		≤ 30		≤ 40	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: 2006-2015								
treat × after	-0.255	-0.302	-0.221	-0.227	-0.230	-0.254	-0.226	-0.235
	(0.181)	(0.233)	(0.179)	(0.229)	(0.179)	(0.228)	(0.178)	(0.227)
Observations	4,818	1,345	5,147	1,469	5,334	1,507	5,452	1,544
B: 2006-2011								
treat × after	-0.178	-0.183	-0.131	-0.093	-0.148	-0.119	-0.148	-0.121
	(0.183)	(0.245)	(0.182)	(0.245)	(0.180)	(0.243)	(0.179)	(0.242)
Observations	1,521	770	1,621	838	1,680	862	1,713	883
Sample:								
Balanced Panel	No	Yes	No	Yes	No	Yes	No	Yes
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes*:* The table illustrates that the name of a taxpayer does not influence their tax payment as long as the effectiveness of the disclosure is independent of the name. We replicate Table III on a sample of MPs only. As MPs are (i) well-known and (ii) identified in the disclosed data directly through their constituency numbers, their exposure to the program does not depend upon how common their name is. As earlier, the definition of the treatment variable is provided in the title of each column. The dummy variable takes the value 1 if the Name Frequency of the MP does not exceed the cutoff indicated in the title. To maintain a fixed control group across all columns, we drop MPs with Name Frequency between 10 and 40 in Columns (1) to (6). Panel B reports the results from a parallel placebo regression, where the sample is restricted to tax years 2006 to 2011, with the last two years defined as the post-program years. Even-numbered columns restrict the sample to a balanced panel of MPs, who file in all years included in the sample. Standard errors are in parenthesis, which have been clustered at the individual level.

Table A.V: **Intensive Margin Response to the Public Disclosure Program – Alternative Definition of Name Frequency**

	Treat: Name Frequency							
	≤ 10		≤ 20		≤ 30		≤ 40	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Main Regression (2006-2015)								
treat $\times$ after	0.098	0.094	0.093	0.092	0.092	0.091	0.091	0.088
	(0.006)	(0.009)	(0.005)	(0.008)	(0.005)	(0.008)	(0.005)	(0.008)
Observations	2,394,847	764,796	2,621,675	837,306	2,704,406	863,405	2,792,270	891,420
B: Placebo Regression (2006-2011)								
treat $\times$ after	0.014	0.010	0.018	0.014	0.018	0.014	0.017	0.013
	(0.007)	(0.008)	(0.006)	(0.008)	(0.006)	(0.008)	(0.006)	(0.008)
Observations	1,288,038	723,868	1,406,460	789,856	1,449,905	814,280	1,496,374	840,469
Sample:								
Balanced Panel	No	Yes	No	Yes	No	Yes	No	Yes
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table reports the estimates from equation (3). We replicate Table III using an alternative definition of the variable Name Frequency. Instead of defining Name Frequency as the number of times a full name appears in the four years of disclosed data (2012-2015), we define it as 4 $\times$ the number of times a full name appears in the 2012 disclosed data. We multiply the number of occurrences of a name in 2012 by four to make this alternative definition of Name Frequency more compatible with the one in our baseline specification. Other than this change of definition, the table is constructed exactly similar to Table III. We obtain similar results if we use any other post-disclosure year 2013-2015 in place of 2012 used here to define Name Frequency.

Table A.VI: **Intensive Margin Response to the Public Disclosure Program – District Level Frequency**

	Treat: Name Frequency							
	≤ 10		≤ 20		≤ 30		≤ 40	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Main Regression (2006-2015)								
treat $\times$ after	0.114	0.108	0.108	0.101	0.104	0.097	0.101	0.094
	(0.006)	(0.009)	(0.006)	(0.009)	(0.006)	(0.009)	(0.006)	(0.009)
Observations	2,351,532	742,873	2,582,045	821,373	2,708,553	863,146	2,792,270	891,420
B: Placebo Regression (2006-2011)								
treat $\times$ after	-0.013	-0.020	-0.014	-0.021	-0.016	-0.022	-0.016	-0.021
	(0.008)	(0.009)	(0.008)	(0.009)	(0.007)	(0.009)	(0.007)	(0.009)
Observations	1,251,402	701,393	1,378,358	773,611	1,449,162	813,524	1,496,374	840,469
Sample:								
Balanced Panel	No	Yes	No	Yes	No	Yes	No	Yes
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes*:* The table reports the estimates from equation (3). We replicate Table III using a district level measure of the variable Name Frequency. We now define Name Frequency as the number of times a full name appears in the four years of disclosed data at the district rather than the national level. The district here denotes the district indicated by the first five digit of the Computerized National Identity Card (CNIC) of the taxpayer. This CNIC was reported along with the full name in the disclosed data for the years 2013-2015. Other than this change of definition, the table is constructed exactly similar to Table III.

Table A.VII: **Intensive Margin Response to the Public Disclosure Program – With Gender × Year Fixed Effects**

	Treat: Name Frequency							
	≤ 10		≤ 20		≤ 30		≤ 40	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Main Regression (2006-2015)								
treat × after	0.094	0.093	0.090	0.089	0.088	0.086	0.088	0.087
	(0.006)	(0.009)	(0.005)	(0.008)	(0.005)	(0.008)	(0.005)	(0.008)
Observations	2,430,002	773,038	2,614,754	833,675	2,720,267	868,250	2,792,270	891,420
B: Placebo Regression (2006-2011)								
treat × after	0.005	0.001	0.009	0.006	0.010	0.006	0.011	0.007
	(0.007)	(0.009)	(0.007)	(0.008)	(0.006)	(0.008)	(0.006)	(0.008)
Observations	1,307,541	734,269	1,403,240	787,845	1,458,457	818,942	1,496,374	840,469
Sample:								
Balanced Panel	No	Yes	No	Yes	No	Yes	No	Yes
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Gender × Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes*:* The table reports results from an augmented version of equation (3). We now include a full set of interactions of a dummy indicating gender of the taxpayer with the year fixed effects. The table replicates Table III using this augmented model. The definition of the treatment variable is provided in the title of each column. The dummy variable takes the value 1 if the Name Frequency of an individual does not exceed the cutoff indicated in the title. To maintain a fixed control group across all columns, we drop taxpayers with Name Frequency between 10 and 40 in Columns (1) to (6). Even-numbered columns restrict the sample to a balanced panel of taxpayers, who file in all years included in the sample. Panel B reports the results from parallel placebo regressions, where the sample is restricted to tax years 2006 to 2011, with the last two years defined as the post-program years. Standard errors are in parenthesis, which have been clustered at the individual level.

Table A.VIII: **Intensive Margin Response to the Public Disclosure Program – With Age $\times$ Year Fixed Effects**

	Treat: Name Frequency							
	≤ 10		≤ 20		≤ 30		≤ 40	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Main Regression (2006-2015)								
treat $\times$ after	0.079	0.077	0.074	0.073	0.072	0.070	0.072	0.070
	(0.006)	(0.009)	(0.005)	(0.008)	(0.005)	(0.008)	(0.005)	(0.008)
Observations	2,430,002	773,038	2,614,754	833,675	2,720,267	868,250	2,792,270	891,420
B: Placebo Regression (2006-2011)								
treat $\times$ after	-0.004	-0.007	-0.001	-0.003	-0.001	-0.004	0.000	-0.003
	(0.007)	(0.008)	(0.006)	(0.008)	(0.006)	(0.008)	(0.006)	(0.008)
Observations	1,307,541	734,269	1,403,240	787,845	1,458,457	818,942	1,496,374	840,469
Sample:								
Balanced Panel	No	Yes	No	Yes	No	Yes	No	Yes
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Young $\times$ Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes*:* The table reports results from an augmented version of equation (3). We now include a full set of interactions of the dummy variable *young* with the year fixed effects. The dummy variable indicates that the age of the taxpayer is less than the median age. The table replicates Table III using this augmented model. The definition of the treatment variable is provided in the title of each column. The dummy variable takes the value 1 if the Name Frequency of an individual does not exceed the cutoff indicated in the title. To maintain a fixed control group across all columns, we drop taxpayers with Name Frequency between 10 and 40 in Columns (1) to (6). Even-numbered columns restrict the sample to a balanced panel of taxpayers, who file in all years included in the sample. Panel B reports the results from parallel placebo regressions, where the sample is restricted to tax years 2006 to 2011, with the last two years defined as the post-program years. Standard errors are in parenthesis, which have been clustered at the individual level.

Table A.IX: **Intensive Margin Response to the Public Disclosure Program – With Region × Year Fixed Effects**

	Treat: Name Frequency							
	≤ 10		≤ 20		≤ 30		≤ 40	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Main Regression (2006-2015)								
treat × after	0.072	0.067	0.069	0.065	0.069	0.064	0.070	0.065
	(0.006)	(0.009)	(0.005)	(0.008)	(0.005)	(0.008)	(0.005)	(0.008)
Observations	2,384,729	769,876	2,566,965	830,292	2,670,952	864,750	2,741,975	887,857
B: Placebo Regression (2006-2011)								
treat × after	-0.002	-0.004	0.003	0.002	0.004	0.002	0.005	0.004
	(0.007)	(0.009)	(0.007)	(0.008)	(0.006)	(0.008)	(0.006)	(0.008)
Observations	1,273,370	725,804	1,367,056	778,950	1,421,027	809,708	1,458,172	831,038
Sample:								
Balanced Panel	No	Yes	No	Yes	No	Yes	No	Yes
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tax Office Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tax Office × Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

The table reports results from an augmented version of equation (3). We now include a full set of interactions of the region dummies with the year fixed effects. The region dummy indicates the district the taxpayer's registered office is located in. There are 25 such regions in our data. The table replicates Table III using this augmented model. The definition of the treatment variable is provided in the title of each column. The dummy variable takes the value 1 if the Name Frequency of an individual does not exceed the cutoff indicated in the title. To maintain a fixed control group across all columns, we drop taxpayers with Name Frequency between 10 and 40 in Columns (1) to (6). Even-numbered columns restrict the sample to a balanced panel of taxpayers, who file in all years included in the sample. Panel B reports the results from parallel placebo regressions, where the sample is restricted to tax years 2006 to 2011, with the last two years defined as the post-program years. Standard errors are in parenthesis, which have been clustered at the individual level.

Table A.X: **Intensive Margin Response to the Public Disclosure Program – By Baseline Taxable Income**

	Baseline Taxable Income:					
	$\in (0,\ 100k]$	$\in (100k,\ 200k]$	$\in (200k,\ 300k]$	$\in (300k,\ 400k]$	$\in (400k,\ 500k]$	$\in (500k,\ 600k]$
	(1)	(2)	(3)	(4)	(5)	(6)
A: Main Regression (2006-2015)						
treat $\times$ after	0.075	0.083	0.061	0.058	0.014	-0.026
	(0.059)	(0.018)	(0.009)	(0.010)	(0.028)	(0.056)
Observations	26,071	197,583	575,312	447,856	60,784	14,442
B: Placebo Regression (2006-2011)						
treat $\times$ after	0.058	0.019	0.005	-0.029	-0.072	-0.069
	(0.046)	(0.010)	(0.021)	(0.024)	(0.036)	(0.078)
Observations	44,234	760,496	104,403	38,149	21,214	5,214
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes

Notes*:* The table explores how the intensive margin response to the public disclosure program varies across the taxable income distribution. We replicate the specification in Column (7) of Table III restraining the sample to taxpayers whose taxable income in the baseline year (2011) was within the interval indicated in the heading of each column. The treatment variable takes the value 1 if the Name Frequency of an individual does not exceed 40. Panel B reports the results from parallel placebo regressions, where the sample is restricted to tax years 2006 to 2011, with the last two years defined as the post-program years. The baseline year for these regression is 2009. Standard errors are in parenthesis, which have been clustered at the individual level.

TABLE A.XI: **EVOLUTION OF TAX PAYMENTS – SELF-EMPLOYED VS. WAGE-EARNERS**

	Baseline Specification		Unique: Name Frequency					
			≤ 10		≤ 20		≤ 30	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Main Regression (2006-2015)								
SE $\times$ after	0.254	0.151	0.203	0.098	0.199	0.094	0.195	0.093
	(0.004)	(0.007)	(0.004)	(0.008)	(0.004)	(0.008)	(0.004)	(0.007)
SE $\times$ after $\times$ unique			0.135	0.132	0.128	0.126	0.126	0.123
			(0.003)	(0.005)	(0.003)	(0.004)	(0.003)	(0.004)
Observations	5,314,786	1,471,400	4,599,189	1,268,359	4,967,881	1,373,193	5,175,705	1,432,297
B: Placebo Regression (2006-2011)								
SE $\times$ after	-0.075	-0.048	-0.082	-0.049	-0.083	-0.051	-0.083	-0.051
	(0.005)	(0.007)	(0.005)	(0.007)	(0.005)	(0.007)	(0.005)	(0.007)
SE $\times$ after $\times$ unique			0.009	-0.001	0.013	0.003	0.014	0.005
			(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)
Observations	2,812,445	1,345,896	2,439,465	1,168,511	2,630,688	1,258,902	2,739,410	1,310,975
Sample:								
Balanced Panel	No	Yes	No	Yes	No	Yes	No	Yes
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes*:* The table compares the evolution of tax payments across self-employed and wage-earners. We report results from estimating equation (3) on samples containing both wage-earners and self-employed, defining the latter category of taxpayers as the treatment group. Panels (1) & (2) report results from the baseline specification. Columns (3)-(8) add an additional term into the model. The additional term interacts the double-difference term with a dummy indicating that the self-employed has a relatively unique name. The dummy variable takes the value 1 if the Name Frequency of the self-employed does not exceed the cutoff indicated in the title. To make the analyses in this table compatible with that in Table III, we drop taxpayers with Name Frequency between 10 and 40 in Columns (3) to (8). Even-numbered columns restrict the sample to a balanced panel of taxpayers, who file in all years included in the sample. Panel B reports the results from parallel placebo regressions, where the sample is restricted to tax years 2006 to 2011, with the last two years defined as the post-program years. Standard errors are in parenthesis, which have been clustered at the individual level.

TABLE A.XII: **HETEROGENEITY IN INTENSIVE MARGIN RESPONSE TO THE PUBLIC DISCLOSURE PROGRAM**

	Major City	Business in Other City	Multiple Businesses	Male	Early Filer	Young	Buncher	Dominated	Revised Return
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
treat $\times$ after	0.066	0.068	0.090	0.137	0.075	0.050	0.083	0.088	0.089
	(0.006)	(0.007)	(0.005)	(0.038)	(0.009)	(0.011)	(0.007)	(0.005)	(0.005)
treat $\times$ trait $\times$ after	0.032	-0.007	-0.068	-0.052	0.017	-0.018	0.004	0.003	-0.019
	(0.010)	(0.021)	(0.016)	(0.038)	(0.014)	(0.017)	(0.010)	(0.025)	(0.051)
Baseline Coefficient	0.088	0.068	0.088	0.088	0.081	0.049	0.088	0.088	0.088
	(0.005)	(0.007)	(0.005)	(0.005)	(0.007)	(0.008)	(0.005)	(0.005)	(0.005)
Observations	2,767,938	1,780,777	2,767,995	2,763,734	1,628,762	1,329,391	2,792,270	2,792,270	2,792,270
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes*:* The table explores heterogeneity in the intensive margin response to the public disclosure program. We estimate a triple-difference version of equation (5) to see how the response varies across taxpayers of different traits. Treatment here is defined as an individual whose Name Frequency does not exceed 40, so the estimates correspond to the specification in Column (7) of Table III. To avoid making strong functional form assumptions all traits are introduced into the equation nonparametrically, as dummy variables. The dummy variable in the first column indicates if the taxpayer belongs to Karachi, Lahore, or Islamabad; in the second column if the taxpayer has business in a city different from the one he resides in; in the third column if the taxpayer has more than one businesses; in the fourth column if the taxpayer is a male, in the fifth column if the taxpayer routinely files her return before the median filing date; in the sixth column if the taxpayer is younger than the median tax filers; in the seventh column if the taxpayer bunched at any of the notches in the 2006-09 tax system of Pakistan; in the eighth column if the taxpayer was in a dominated region above any of the notches; and in the final column if the taxpayer filed a revised return in any of the pre-program periods. We do not observe some of the traits for the whole sample. The Baseline Coefficient reports the treat $\times$ after coefficient in equation (5) for the restricted sample for which we observe the trait. Standard errors are in parenthesis, which have been clustered at the individual level.

Table A.XIII: **Extensive Margin Response to the Public Disclosure Program – Baseline Frequency**

	Treat: Name Frequency						
	≤ 10	≤ 20	≤ 30	≤ 40	$\leq$ Median	$\leq$ 1st Quartile	$\leq$ 1st Decile
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A: Main Regression (2006-2015)							
treat $\times$ after	0.054	0.046	0.043	0.041	0.039	0.070	0.125
	(0.012)	(0.011)	(0.010)	(0.010)	(0.010)	(0.014)	(0.022)
B: Placebo Regression (2006-2011)							
treat $\times$ after	0.000	-0.000	-0.000	-0.000	-0.001	0.005	0.021
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.010)

Notes*:* The table conducts a robustness check on our extensive margin results. We replicate Table V using an alternative definition of Name Frequency, measuring it as the number of times a full name appears among the tax filers in the six baseline years 2006-2011. We multiply this measure of Name Frequency with a factor of 4/6 to make it compatible with the definition used in Table V and our other results. The table reports the estimates from equation (6). The equation is estimated on a sample of all self-employed individuals. The outcome variable here is the log number of filers in group g in year t. Panel A estimates the equation on the period 2006-2015. The definition of the treatment variable is provided in the title of each column. The dummy variable takes the value 1 if the normalized value of Name Frequency of an individual does not exceed the cutoff indicated in the title. To maintain a fixed control group across columns (1)-(4), we drop taxpayers with the Name Frequency between 10 and 40 in columns (1) to (3). In columns (6) and (7) we drop the middle part of the distribution: the middle two quartiles in column (6) and the deciles 2-9 in column (7). Panel B reports the results from parallel placebo regressions, where the sample is restricted to tax years 2006 to 2011, with the last two years defined as the post-program years. Standard errors are in parenthesis.

Table A.XIV: **Response to the TPHC Program – Placebo**

	Treat: Rank							
	$\in (150,\ 200]$		$\in (200,\ 250]$		$\in (250,\ 300]$		$\in (300,\ 350]$	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Main Regression (2006-2014)								
treat × after	-0.029	-0.001	0.027	0.054	-0.004	0.019	-0.021	-0.003
	(0.068)	(0.076)	(0.065)	(0.072)	(0.058)	(0.065)	(0.066)	(0.071)
treat × 1.(year ∈ {2010,2011})		0.079		0.083		0.065		0.054
		(0.098)		(0.085)		(0.081)		(0.093)
Observations	32,047	32,047	32,047	32,047	32,047	32,047	32,047	32,047
B: Placebo Regression (2006-2010)								
treat × after	0.084		0.025		-0.040		0.058	
	(0.100)		(0.092)		(0.094)		(0.094)	
Observations	17,208		17,208		17,208		17,208	

Notes*:* The table tests the validity of the research design used to estimate the TPHC response. We estimate equation (5) on a sample containing top 1000 taxpayers of each of the four categories of taxpayers, corporations, partnerships, self-employed, and wage-earners. But in distinction to Table VI, the treatment variable here denotes taxpayers who are not affected by the program, being too far away from its eligibility cutoff. The exact length of the treatment window used here is indicated in the title of each column. Given that we measure the outcome variable here in changes rather than levels, the first post-program year is 2012. Panel A estimates the equation on years 2006-2014. Panel B runs parallel regressions on years 2006-2010, with the last two years defined as the post-program years. Columns (2), (4), (6) and (8) test the parallel trend assumption by including a $treat \times 1.(year \in \{2010, 2011\}$ interaction into the regression. Standard errors are in parenthesis, which have been clustered at the individual level.

Table A.XV: **Response to the TPHC Program – By Taxpayer Category**

	Treat: Rank $\in (80,\ 120]$							
	Self-Employed		Wage-Earners		Partnerships		Corporations	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Main Regression (2006-2014)								
treat × after	-0.033	0.013	0.215	0.276	0.036	0.089	0.412	0.267
	(0.205)	(0.241)	(0.143)	(0.172)	(0.105)	(0.114)	(0.115)	(0.129)
treat × 1.(year ∈ {2010,2011})		0.130		0.176		0.144		-0.444
		(0.221)		(0.254)		(0.102)		(0.206)
Observations	7,619	7,619	7,914	7,914	8,185	8,185	8,329	8,329
B: Placebo Regression (2006-2010)								
treat × after	0.231		0.173		0.120		-0.387	
	(0.278)		(0.258)		(0.116)		(0.225)	
Observations	3,993		4,241		4,420		4,554	

Notes: The table breaks down the TPHC response by taxpayer category. We estimate equation (5) separately for each category of taxpayers. These categories are indicated in the title of each column. The sample for each regression includes top 1000 taxpayers of the corresponding category in each year included in the sample. The treatment variable here denotes taxpayers of the category ranked 81-120 in the given year. Given that we measure the outcome variable here in changes rather than levels, the first post-program year is 2012. Panel A estimates the equation on years 2006-2014. Panel B runs parallel placebo regressions on years 2006-2010, with the last two years defined as the post-program years. Columns (2), (4), (6) and (8) test the parallel trend assumption by including a $treat \times 1.(year \in \{2010, 2011\}$ interaction into the regression. Standard errors are in parenthesis, which have been clustered at the individual level.

Table A.XVI: **Heterogeneity in Response to the TPHC Program**

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
treat × after	0.412	0.356	0.501	0.369	0.399	0.369	0.462	0.427
	(0.115)	(0.214)	(0.124)	(0.115)	(0.116)	(0.119)	(0.124)	(0.119)
treat × after × public		0.091						
		(0.255)						
treat × after × foreign owned			-0.793					
			(0.295)					
treat × after × banking				1.241				
				(0.718)				
treat × after × food					0.389			
					(0.583)			
treat × after × textile						0.114		
						(0.272)		
treat × after × pharma							-0.573	
							(0.233)	
treat × after × construction								-0.342
								(0.394)
Observations	8,329	8,329	8,329	8,329	8,329	8,329	8,329	8,329

Notes*:* The table explores heterogeneity in corporate firms' response to the TPHC program. We estimate the triple-difference version of model (5), adding the interaction terms shown above. Columns (1) reproduces column (7) of Table A.XV. The other columns add interaction terms to this baseline specification. The dummy variable *public* denotes a public-limited corporation; *foreign owned* a completely-owned subsidiary of a foreign firm; and *food*, *textile*, *pharma*, and *construction* the industry the firm operates in. Standard errors are in parenthesis, which have been clustered at the firm level.

TABLE A.XVII: SUMMARY STATISTICS – TPHC SAMPLE

	2011		2010	
	Treatment	Control	Treatment	Control
	(1)	(2)	(3)	(4)
1. Taxable Income	18.389	17.005	20.505	18.506
	(0.546)	(0.140)	(0.059)	(0.092)
	[0.014]	[0.014]	[0.000]	[0.000]
2. Tax Paid on Taxable Income	17.165	15.847	19.434	17.403
	(0.635)	(0.149)	(0.091)	(0.097)
	[0.044]	[0.044]	[0.000]	[0.000]
3. Final Tax Paid	12.730	13.070	13.908	13.132
	(1.023)	(0.211)	(0.605)	(0.208)
	[0.745]	[0.745]	[0.226]	[0.226]
4. Major city	0.925	0.834	0.950	0.882
	(0.042)	(0.012)	(0.035)	(0.010)
	[0.039]	[0.039]	[0.063]	[0.063]
5. Early filer	0.700	0.707	0.625	0.528
	(0.073)	(0.015)	(0.078)	(0.016)
	[0.926]	[0.926]	[0.220]	[0.220]
6. Young Firm	0.375	0.548	0.525	0.518
	(0.078)	(0.016)	(0.080)	(0.016)
	[0.029]	[0.029]	[0.929]	[0.929]
7. Public Limited	0.450	0.335	0.825	0.555
	(0.080)	(0.015)	(0.061)	(0.016)
	[0.158]	[0.158]	[0.000]	[0.000]
8. Foreign Owned	0.075	0.050	0.100	0.067
	(0.042)	(0.007)	(0.048)	(0.008)
	[0.559]	[0.559]	[0.494]	[0.494]
9. Bank	0.000	0.004	0.050	0.029
	(0.000)	(0.002)	(0.035)	(0.005)
	[0.045]	[0.045]	[0.555]	[0.555]
10. Food	0.050	0.043	0.050	0.071
	(0.035)	(0.007)	(0.035)	(0.008)
	[0.837]	[0.837]	[0.561]	[0.561]
11. Textile	0.125	0.146	0.025	0.136
	(0.053)	(0.011)	(0.025)	(0.011)
	[0.701]	[0.701]	[0.000]	[0.000]
12. Pharma	0.075	0.015	0.100	0.023
	(0.042)	(0.004)	(0.048)	(0.005)
	[0.154]	[0.154]	[0.111]	[0.111]
13. Construction	0.125	0.111	0.025	0.053
	(0.053)	(0.010)	(0.025)	(0.007)
	[0.802]	[0.802]	[0.280]	[0.280]

Notes: The table presents summary statistics of our TPHC sample containing top 100 tax paying corporations in each year. The treatment variable here denotes corporations ranked between 80 and 120 in period t . Each row compares the mean value of the variable across the two groups for the two pre-program years. We report standard error of the mean in parenthesis and the p-value of the test of equality of two means in square brackets. The definitions of the variables are provided in Appendix A.2 and Table A.XVI.

Table A.XVIII: **Response to the TPHC Program – Robustness**

	(1)	(2)	(3)	(4)	(5)	(6)
A: Main Regression (2006-2014)						
treat $\times$ after	0.412	0.402	0.416	0.400	0.383	0.363
	(0.115)	(0.114)	(0.115)	(0.114)	(0.114)	(0.115)
Observations	8,329	8,329	8,329	8,329	8,329	8,329
B: Placebo Regression (2006-2010)						
treat $\times$ after	-0.387	-0.366	-0.392	-0.254	-0.360	-0.332
	(0.225)	(0.221)	(0.224)	(0.224)	(0.219)	(0.224)
Observations	4,554	4,554	4,554	4,554	4,554	4,554
Trait:						
	-	Major City	Young Firm	Public Limited	Bank	Textile
Trait $\times$ Year Fixed Effects	No	Yes	Yes	Yes	Yes	Yes

Notes: The table conducts a robustness check on our TPHC program results. We report results from an augmented version of equation (5). The augmented model includes the full set of interactions of the dummy variable *Trait* with the year fixed effects. We report results for five different traits indicated in the second-last row of the table. The definition of these traits are provided in Appendix A.2 and Table A.XVI. Columns (1) reports results from the baseline specification. It is the same as column (7) of Table A.XV. The other columns add interaction terms to this baseline specification. Standard errors are in parenthesis, which have been clustered at the firm level.