Essays in Development Economics and Public Finance

by

Obeid Ur Rehman

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in the University of Michigan
2021

Doctoral Committee:

Professor Dean Yang, Chair
Associate Professor Achyuta Adhvaryu
Professor Tanya Rosenblat
Professor Joel Slemrod

Obeid Ur Rehman

obeidr@umich.edu

ORCID iD: 0000-0002-7357-9717

In the name of Allah, the most gracious, the most merciful.

To my parents, Zia Ur Rehman and Talat Zia,
my wife, Maham,

and to migrants around the world,
who leave their homes in search of a better life.

ACKNOWLEDGEMENTS

Although my name is on the front page, this work is the result of contributions from many individuals. Firstly, I am grateful to my parents for their continuous support. Their countless efforts allowed me the privilege to pursue this Ph.D.

I would like to thank my chair, Dean Yang, and my advisors, Tanya Rosenblat, Ach Adhvaryu, and Joel Slemrod, for their continued guidance. Joel and Mazhar Waseem were also my co-authors and working with them was a thorough learning experience. I also received valuable feedback over the years from Hoyt Bleakley, Charlie Brown, John Bound, and Martha Bailey. Asim Khwaja inspired me to pursue this Ph.D. and he has been a trusted mentor throughout.

Innovations for Poverty Action (IPA) Philippines was instrumental for my research on spousal information sharing. I would especially like to thank Yuna Liang and Peter Srouji from IPA Philippines for excellent research support. Laura Flak, Lauren Pulay, Julie Heintz, Cassie Kelly, and Hiba Baghdadi, from the Department of Economics, made every possible resource available for my success. I also gratefully acknowledge funding from the Graduate Research Fund of the Department of Economics for my projects.

The friends I made in Lorch Hall are my greatest asset from this experience. Merve Sariisik, Chiara Ferrero, Kimberly Conlon, and in particular my perpetual office-mates and confidants, Bhanu Gupta and Robert Venyige, made these years joyous (as much as a Ph.D. can be) and memorable.

Most of all, I would like to thank my wife, Maham, my constant champion and supporter. Her companionship through this journey and throughout life is the greatest blessing.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

ABSTRACT

This dissertation analyzes key challenges to maximizing the impact of development policies. It focuses on three sectors - migration, taxation, and education. Each of these sectors offers massive development potential but is constrained by some combination of information asymmetry, policy barriers, and limited resources. I study these challenges in a global context with projects in the UAE, Philippines, Pakistan, and the US. I use several context and question-specific research designs: conducting randomized controlled trials, analyzing natural experiments, and evaluating policy interventions. I combine survey and administrative data, with novel identification strategies and statistical techniques to estimate causal effects.

The first chapter studies spousal communication among transnational households –households where one spouse temporarily migrates for work. Despite regular communication between spouses, information asymmetry persists in these households. I analyze if this information asymmetry is caused by spouses strategically misreporting information to influence resource allocation in the household. Misreporting, by definition, involves purposefully falsifying information, making it challenging to identify. I address this challenge using a novel field experiment among Filipino migrants in the UAE and their spouses staying behind in the Philippines. I find that both migrants and their spouses staying behind have biased beliefs about each other's finances and these biases are the result of strategic misreporting. Spouses staying behind and some subgroups of migrants underreport their income to influence the remittance decision in their favor. The results show that addressing information asymmetry requires interventions that increase the ability of spouses to verify and monitor each other's reported information. However, the welfare impacts of such interventions are a priori ambiguous because better information sharing may reduce remittances.

The second chapter, co-authored with Joel Slemrod and Mazhar Waseem, evaluates two Pakistani programs to study the impact of public disclosure and social recognition of tax payments on tax compliance. Pakistan began revealing the 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 combine publicly disclosed and restricted administrative tax return data for the universe of tax filers to create an extended panel of tax records from 2006 to 2015.

Using empirical strategies based on name commonness for the public disclosure program and cutoffs in the social recognition program's eligibility criteria, we show that both programs induced strong compliance responses. Our results suggest that such programs can be important policy levers to mobilize resources, especially in weak-enforcement-capacity economies.

The third chapter studies the joint educational attainment and migration decisions of international students choosing to study in the US. Immigration critics argue that international students are primarily come to the US for employment and use their student status to bypass restrictions on employment-based migration. I use exchange rate variations to analyze these competing educational and employment incentives for potential international students. A depreciation of the home currency reduces educational incentives by increasing the relative cost of US education but increases employment incentives by making US income relatively more valuable. I find that the cost of education effect dominates the higher relative income effect. A depreciation of the home currency reduces the stock and flow of international students from that country.

CHAPTER I

Spousal Communication and Information Sharing: Evidence from Migrants and their Spouses

Abstract

Do spouses misreport information to each other to influence household decision-making? I analyze this question in the context of transnational households - where one spouse temporarily migrates for work and resource allocation decisions are made under significant information asymmetry. While the effects of this information asymmetry are well established, its persistence despite regular communication between spouses remains puzzling. Misreporting, by definition, involves purposefully falsifying information, making it difficult to identify. I address this challenge using a novel field experiment among Filipino migrants in the UAE and their spouses staying behind in the Philippines. I find that both migrants and their spouses staying behind have biased beliefs about each other's finances and these biases are the result of strategic misreporting. Spouses staying behind underreport their income to migrants by 31 percent. They hide income using the more difficult-to-catch strategy of underreporting known sources of income instead of hiding income sources altogether. Income is only hidden when migrants do not communicate about or demand control over the household's finances. The results are consistent with an income-sharing model where both spouses have private information and income hiding is constrained by the threat of punishment.

JEL Classification: D13, D82, J61, O15

Keywords: Asymmetric information, hidden income, migration, remittances

1.1 Introduction

Although most models of the household assume perfect information (Chiappori, 1988, 1992; Lundberg and Pollak, 1993), recent theoretical and empirical work has shown that spouses may have private information and may strategically use this information to influence resource allocation in the household (Ashraf, 2009; Castilla and Walker, 2013; Chen, 2006; De Laat, 2014). As a result, information sharing - when and how spouses share or conceal private information, plays a key role in determining household outcomes.

Although private information is a non-trivial concern for all households, it is especially relevant for transnational households - where one spouse temporarily migrates for work and sends a considerable portion of their income as remittances to the spouse staying behind. These remittances are economically significant intra-household transfers and the primary motivation for most temporary migration.[1] However, the remittance decision is made under considerable information asymmetry as migrants and their spouses staying behind have limited ability to observe or control each other's actions. This information asymmetry leads to migrants sending fewer remittances (Joseph, Nyarko and Wang, 2018; Ambler, 2015), accumulating lower savings (Ashraf et al., 2015), both migrants and their households spending resources on monitoring each other (De Laat, 2014), and biased beliefs about the returns to migration (Baseler, 2018; McKenzie, Gibson and Stillman, 2013).

While the effects of information asymmetry are well established, the persistence of information asymmetry, despite regular communication between migrants and their spouses staying behind, remains puzzling. While communication can reduce information frictions (Batista and Narciso, 2018), it can instead create or exacerbate information asymmetry if spouses purposefully misreport information to each other to influence the remittance decision in their favor.[2]

In this paper, I analyze if information asymmetry is caused by migrants and their spouses strategically misreporting information to each other. First, I document the extent of information asymmetry between migrants and their spouses staying behind across multiple margins. Next, I analyze if this asymmetry is the result of strategic misreporting. The defining characteristic of misreporting is that individuals are purposefully falsifying information making it inherently difficult to identify. I address this challenge using a novel experimental strategy to observe misreporting. Finally, I explore factors that exacerbate or mitigate misreporting.

[1] In 2016, remittances sent to developing countries amounted to USD $429 billion, roughly three times official development aid (World Bank, 2017).

[2] Ambler (2015) distinguishes between strategic information asymmetries that are created by strategic behavior and inadvertent information asymmetries that arise due to communication barriers. Based on her framework spousal communication can be used to create strategic information asymmetries or reduce inadvertent information asymmetries.

My research design is based on the simple idea that if migrants and their spouses strategically misreport information to each other, they will differentially report this information when it is observable to the other spouse compared to when it is not. I implement this design in the context of transnational households in the UAE-Philippines migration corridor. I invite married temporary Filipino migrants in the UAE and their spouses staying behind in the Philippines to separately participate in a survey. The survey is described to participants as a research activity to create awareness about the experience of Filipino migrants in the UAE and their households in the Philippines. In the survey, I collect data on income, expenses, and employment; with the migrant reporting their information in the UAE and the spouse staying back reporting theirs in the Philippines. I elicit the causal effect of spousal observability by experimentally varying whether the information reported by an individual in the survey is observable to their spouse. Participants in the control group are informed that their responses will be kept private and not shared with their spouses, whereas participants in the treatment group are informed that their responses will be observable to and shared with their spouses. Information is shared, based on the treatment status, after all surveys have been completed. Participants know their treatment status when the information is collected. If migrants and their spouses staying behind strategically misreport information to each other, I should observe differences in the information reported by the treatment and the control groups.

I document four findings. First, both migrants and their spouses staying behind have biased beliefs about each other's finances. Migrants underestimate their spouses' income and overestimate their spouses' expenses, whereas spouses staying behind underestimate migrants' less-observable non-wage benefits. Second, these biases are the result of strategic misreporting. Spouses staying behind underreport their monthly income by 31 percent (213 dirham or USD $58) when it is observable to the migrant, compared to their reported monthly income of 685 dirham (USD $186) when it is not observable to the migrant. This income hiding is even greater when the migrant also participates in the study. Third, misreporting is greater when information is more difficult to observe and less likely to be verified. Spouses staying behind hide income on the intensive margin by underreporting known sources of income instead of reporting zero income which would be easier for migrants to verify as misreporting. Income hiding only occurs when migrants do not demand control over or regularly communicate about the household's finances, making them less likely to verify reported information. Fourth, among both migrants and their spouses staying behind, women are more likely to hide income. These gender differences appear to be driven by men less frequently demanding control over and communicating about the household's finances.

The prior literature on asymmetric information in transnational households is comprised of non-experimental and experimental studies, focused on income hiding by the remittance sender. I expand on this work by first presenting a conceptual framework for and then robustly analyzing strategic misreporting on *both sides* of the remittance relationship, across *multiple* margins of the households *actual* finances. The key challenge in identifying strategic misreporting as the cause of information asymmetry in non-experimental settings is that spousal communication is not observed. These studies must infer strategic misreporting as the mechanism for other observed outcomes. Using this strategy, Seshan and Zubrickas (2017) show that wives in India underestimate their husbands' earnings in Qatar and the underestimation is associated with lower remittances. Joseph, Nyarko and Wang (2018) show that remittances from the UAE respond more to observable shocks in migrant's income than unobservable shocks. Baseler (2018) and McKenzie, Gibson and Stillman (2013) show that migrant-sending households in Kenya and potential migrants in Tonga, respectively, underestimate the returns to migration despite significant information flows between migrants and household members. In contrast to these studies, I directly observe spousal information-sharing. This allows me to cleanly identify strategic misreporting as the cause of information asymmetry in transnational households, ruling out alternate explanations such as unintentional information frictions caused by limited communication.

Lab and lab-in-field experiments also provide settings where spousal communication and decision-making has been directly observed.[3] These studies find that migrants respond to increased information sharing by sending more remittances. Salvadoran migrants in the US remit more when their choice of how much to remit is revealed to recipients (Ambler, 2015) and Filipino migrants in Italy remit more when they can label remittances with their intended purpose (De Arcangelis et al., 2015). However, as these studies accede, their findings may be limited to decision-making over one-time windfall gains. Households may treat income from unanticipated lottery winnings differently from their permanent income and the stakes involved in hiding or sharing these winnings may also be considerably lower. I move this research agenda forward by analyzing how results from these lab and lab-in-field settings translate when transnational households share information about their actual finances.

My results are also informative for the design and implementation of financial products and services for transnational households that leverage information sharing and control. There is a growing interest in these products that are motivated by the idea that greater control and information sharing will improve financial decision-making and outcomes but experimental evaluations have found mixed results.[4] Ashraf et al. (2015) show that Salvado-

[3] Ashraf (2009) and Castilla and Walker (2013) show that in co-residing households, spouses strategically use private information and lack of communication for personal gain.

[4] Field experiments have also been used to evaluate the impact of financial literacy and training programs targeted to transnational households to improve financial behaviors and decision-making. See Seshan and

ran migrants save more when they have access to bank accounts at home that offer greater control over savings. However, Ambler, Aycinena and Yang (2015) find no demand for a remittance product that channels funds directly to education, unless it is bundled with a subsidy. I find that the strategic behavior these products aim to address is limited to certain subgroups of transnational households, suggesting that these products would be most effective when targeted to these households. Importantly, my results also show that these subgroups can be identified from observable baseline characteristics.

The paper proceeds as follows: Section 2 presents a conceptual framework of the remittance and information sharing decisions. Section 3 provides details of the UAE-Philippines migration corridor. Section 4 and 5 describe the experimental design and data, respectively. Section 6 presents the empirical strategy and results. Section 7 discusses other motivations and strategies for income hiding and section 8 concludes.

1.2 Conceptual Framework: The Remittance & Information Sharing Decision

In this section, I present a conceptual framework of the remittance and information sharing decisions that builds on the frameworks developed by Joseph, Nyarko and Wang (2018) and Seshan and Zubrickas (2017) to incorporate opportunities of strategic misreporting on both sides of the remittance relationship. Couples face a trade-off between the benefits of strategically misreporting income, to influence the remittance decision in their favor, and the costs of punishment if they are caught lying. Remittances are the result of an income-sharing contract - increasing in the migrant's reported net income and decreasing in their spouse's reported net income. Although some portion of each spouse's income is common knowledge, migrants and their spouses have private information about their realized incomes which they report to each other. Each spouse can attempt to verify the other's report and punish the other spouse if they are caught lying. This framework generates predictions that are distinct from existing remittance models and can be empirically observed.

1.2.1 The Remittance Contract

Consider a transnational household where the migrant in the host country earns net income $y_M >0$, while their spouse in the home country earns net income y_S. Net income is income net of some specified subsistence expenditures and I refer to it as income from here on. Each spouse's income is comprised of an observable component y_i^o, which is common

Yang (2014), Gibson, McKenzie and Zia (2012), and Doi, McKenzie and Zia (2014).

knowledge across spouses, and a hidden component y_i^h, which is private information for each spouse, where $i = M$ or S. For migrants, while the terms of their contracts may be observable and common knowledge; their monthly working hours, bonuses, and consumption expenditure would be private information. Similarly for spouses staying behind, some portion of their income would be unobservable to the migrant. Each spouses realized income $y_i = y_i^o + y_i^h$, is therefore private information.

The income-sharing contract specifies that the migrant will share part of their income with their spouse as remittances, while in return, the spouse staying back will manage household and childcare responsibilities in the home country. This arrangement does not have to be an explicit contract and can instead be an implicit agreement or a social norm. Both the migrant and spouse send each other a report of their realized income $\tilde{y}_i$. The remittance amount r is a function of both of their reported incomes, $r(\tilde{y}_M, \tilde{y}_S)$. Remittances are increasing in the migrant's reported income, as high-income migrants are able and expected to remit more, and decreasing in the spouse's reported income, as high-income spouses have lower demand and need for remittances. Appendix 1.10.1 presents a model where this relationship is formally derived and shows that it exists for a range of income-sharing contracts with limited assumptions on the utility functions of migrants and spouses.

As remittances are based on reported information, migrants and spouses can attempt to verify each other's reports. Verification is imperfect and succeeds with probability $p_i(x_{i,-i}, c_i, \tilde{y}_i, y_i)$ that depends on migrant and spouse specific characteristics $x_{i,-i}$. Couples that monitor each other through regular communication about household finances, frequent visits, and support from relatives and peers have a greater probability of verifying each other's reports. Successful verification also depends on the amount spent on verification c_i, and the magnitude of the misreporting $(\tilde{y}_i - y_i)$.[5] If upon verification either spouse catches the other lying i.e. $\tilde{y}_i < y_i$, they can inflict a punishment, denoted by $P_i(\tilde{y}_i, y_i)$. The punishment may take the form of social or familial sanctions. In addition, the migrant may punish the spouse by sending fewer remittances in the future than specified by the contract, while the spouse may punish the migrant by refusing to carry out the migrant's specified tasks and responsibilities.

The migrant's utility is increasing in their income and decreasing in the remittances they have to send, the probability of being caught lying, and the punishment for lying. The migrant faces a trade-off between the benefit of underreporting their income and having to send fewer remittances and the cost of punishment if they are caught lying. The spouse's utility on the other hand is increasing in their income and the remittances they receive, while decreasing with the probability of being caught lying and the punishment for lying.

[5] An alternate but equivalent setup is that instead of increasing the probability of successful verification, these factors decrease the ability to keep the hidden portion of income private.

The spouse faces a similar trade-off between the benefit of underreporting their income and receiving more remittances and the cost of punishment if they are caught lying. Migrants and spouses decide how much income to report to each other and how much to spend on verifying each other's reported income.

1.2.2 Empirical Predictions from the Framework

This framework generates predictions that are distinct from existing altruism and exchange based remittance models (Lucas and Stark, 1985; Rapoport and Docquier, 2006; Yang, 2011).

First, spouses can directly influence remittances by strategically misreporting information to the migrant. Modeling remittances as a function of the spouse's reported income incorporates spousal demand for remittances in the framework (see Appendix 1.10.1 for details of such a model). This demand is an important feature of the remittance relationship and often a source of pressure on migrants. Migrant-sending households use remittances to insure against income shocks and therefore demand for remittances is directly impacted by changes in household income (Yang and Choi, 2007).

In existing exchange-based remittance models the effect of spousal income on remittances is ambiguous. These models limit the spouse's role to accepting or rejecting the terms of an agreement that specifies remittances as some function of only the migrant's income. In altruism-based remittances models, spouse's income negatively affects remittances. However, if remittances are purely altruistic there are no incentives on either side of the remittance relationship to hide income.[6]

Second, migrants can attempt to verify spousal reports at a cost. Migrants can spend significant resources to monitor their households (De Laat, 2014) and their limited ability to observe and control the household's decision-making is an important factor in the remittance decision (Ashraf et al., 2015). Existing remittance models however limit the verification decision to the spouse (and limit misreporting to the migrant).

Third, the likelihood of successful verification of spousal reports depends on individual characteristics and the resources spent on monitoring. In standard remittance models, income verification is perfect and incurs a fixed cost. The only choice for migrants is whether or not to incur the cost of verification.[7] However, couple's ability to monitor each other

[6] Altruism-based models define the migrant's utility as a function of the household's consumption, which in turn is a function of the household's income. However, to allow for spouses to strategically influence the remittance decision requires the additional assumption that spouses are not altruistic and they know that the migrant is altruistic.

[7] Joseph, Nyarko and Wang (2018) modify the standard model by allowing for two types of income with their respective verification costs, thereby also allowing households to choose which income to verify.

varies based on characteristics such as the size of their networks and the frequency of communication.

Finally, the relevant parameter in the remittance decision is net income, implying that both income and expenses can be misreported to impact remittances. Although this specification is not unique to my framework (see Seshan and Zubrickas (2017), Joseph, Nyarko and Wang (2018)), I highlight it here because the empirical research on misreporting has predominantly focused on income hiding. Overreporting expenses, which based on the conceptual framework has the same impact as income hiding, has not been analyzed before.

1.3 Context: Filipino Migrants in the UAE

In this section, I describe features of the UAE-Philippines temporary migration corridor that are important to studying communication among transnational households.

1.3.1 Immigrants in the UAE

The UAE provides the relevant host country institutional settings - a large remittance-sending migrant population, immigration policies that lead to the creation of transnational households, and a labor market that generates income fluctuations which migrants can strategically misreport. Some combination of these institutional settings is present in all countries that host temporary employment-based migrants.

The UAE is one of the largest temporary migration destinations and remittance sources in the world. 88 percent of the UAE's 9.6 million population are migrants. The migrant population has increased substantially over the last decades, from 3.3 million in 2005 to an estimated 8.6 million in 2019 (United Nations, 2019). The remittances these migrants sent amounted to 10.7 percent of the UAE's GDP in 2018 and made it the second-largest source of outward remittances (World Bank, 2018*a*).

Almost all migration to the UAE is temporary and employment-based. The immigration policy, known as *kefala* or sponsorship, is widely practiced in the Middle East. Visas are tied to employment status with a specific employer and do not offer any paths to legal permanent residence or citizenship. Employment contracts are only two years long but can be repeatedly renewed with the consent of both parties. When a contract ends or is terminated migrants must either obtain a new contract or return to their home country within 30 days.

The UAE's immigration policies and high cost of living lead to the creation of transnational households. Immigration requirements based on income and occupation prevent low-income migrants from inviting their spouses and children to the UAE - creating transnational households. Male migrants must have a monthly income of at least 4,000 dirham

(USD $1,089) to invite their spouses and children, while female migrants face stricter requirements.[8] Even when these income and employment conditions are not binding, the high cost of living in the UAE, relative to migrants' home countries, also dissuades them from inviting their spouses and children. Female migrants are again disproportionately impacted by these factors. Husbands cannot work in the UAE based on their wives' visa status and must acquire their own work visas. In contrast, wives of male migrants can work in the UAE based on their husbands' visa status.

Migrants working in the UAE experience fluctuations in their monthly income despite specific contracts. These fluctuations create additional opportunities to strategically misreport income that would not be present if incomes remained stable over the full contract cycle. Employment contracts are required to state the employee's remuneration, however, this is often only specified as the minimum required working hours and the corresponding total monthly wage. Joseph, Nyarko and Wang (2018) use administrative payroll data to show that migrants in the UAE experience substantial fluctuations in monthly wages caused by variations in working hours and overtime pay. Anecdotal evidence from focus groups suggests that migrants also experience income fluctuations due to delayed or missed paychecks.

1.3.2 Emigrants from the Philippines

The Philippines has a large, gender-balanced emigrant population. The remittances they send are a key component of the country's development policy and a significant proportion of Filipino households rely on these remittances to sustain themselves.

The Philippines has one of the largest emigrant populations in the world. In 2018, there were 2.3 million Filipino migrants (known locally as Overseas Filipino Workers or OFWs) worldwide. These migrants remitted USD $32.8 billion in 2017, making the Philippines the third-largest remittance-receiving country in the world. According to the Filipino government's nationally-representative Family Income and Expenditure Survey of 2009, 26 percent of households received remittances from abroad. The UAE was the second-largest destination and source of remittances for Filipino migrants, accounting for 15.7 percent of the total Filipino migrant population and 13 percent of the total remittances to the Philippines in 2017 (World Bank, 2018*b*). A key feature of Filipino migrants is their gender composition. In 2020, 56 percent of Filipino migrants were women. This is a much higher proportion of female migrants than most migrant-sending developing countries and allows the analysis of

[8]Female teachers, engineers, doctors, or other medical professionals have the same income requirements as men; however, women employed elsewhere are required to have a minimum monthly income of 10,000 dirham (USD $2,722) and even then each petition is decided on a case by case basis by the UAE immigration department. The income threshold for each category is reduced to 3,000 dirham (USD $817) or 8,000 dirham (USD $2,178) respectively if the migrant's accommodation is provided in-kind by their employer.

the interaction of gender with migration and remittance decisions.

1.4 Experimental Design

The experiment is designed to address the main challenges of studying information sharing in transnational households; observing communication about the household's actual finances among spouses spread across two countries. The experimental design is based on the simple idea that if couples strategically misreport information to each other, they will differentially report this information when it is observable to their spouse compared to when it is not observable. I implement this idea by separately surveying migrants and their spouses about their respective finances. In the survey, I experimentally vary if an individual's responses are observable to their spouse and use this variation to identify if spouses strategically misreport information to each other.

1.4.1 Sample

The study sample is comprised of migrants working in the UAE and their spouses living in the Philippines. This transnational sample allows me to analyze strategic misreporting from both sides of the remittance relationship.

The sample was drawn from a subject pool of participants of a separate study on the remittance behavior of migrants.[9] The subject pool consisted of migrant workers from the Philippines living and working in Dubai, UAE. Migrants were recruited between September and December 2018 via face-to-face intercepts in locations frequented by Filipino migrants in Dubai.[10] Migrants who expected to continue working in the UAE for the following 12 months and agreed to participate were enrolled in the subject pool. At enrollment, migrants were administered a baseline survey that collected information on demographics, remittance behavior, and contact details of their remittance recipients. I use this baseline data to identify and invite my study sample and to analyze selection into the study and heterogeneity in treatment effects.

From this subject pool of Filipino migrants, I invited all married migrants, whose spouses were living in the Philippines and who had sent remittances to their spouse's household in the last year, to participate in this study. Separately I also invited their spouses in the Philippines to participate. These criteria produced an invited sample of 492 couples (984

[9] The subject pool was recruited as part of De Arcangelis and Yang (2019). Details of the subject pool recruitment are described in appendix 1.10.2

[10] The Filipino community in Dubai is highly concentrated in the Satwa and Rigga neighborhoods. Recruitment locations included metro stations, Filipino restaurants, retail stores, and remittance service provider branches in these neighborhoods.

individuals; half in the UAE, and half in the Philippines). Of these invited individuals, 159 migrants and 156 spouses participated in the study - a take-up rate of around 32% for both groups. This included 94 matched couples (both the migrant and their spouse participated), 65 cases where only the migrant participated, and 62 cases where only the spouse participated. Figure 1.1 shows the time-line of project activities along with the sample size at each stage.

1.4.2 Experimental Conditions

The experimental conditions were designed to identify strategic misreporting of the transnational household's actual finances.

Migrants and their spouses were invited to participate in a survey that was marketed as a research activity to improve information and awareness about the experience of Filipino migrant workers in the UAE and their migrant-sending households in the Philippines. In the survey, respondents were asked information about their finances and employment, and their beliefs about their spouse's finances; with the migrant reporting their information in the UAE and the spouse reporting theirs in the Philippines. Participants were informed that summary results of the data collected from the surveys would be shared with them when the study was completed. The surveys were conducted over the phone, separately for both migrants and their spouses between January and April 2019. Participants were aware that their spouses would also be separately invited to participate in the study. However, no details about their spouse's survey activity including; when they would be contacted, their participation status, or the questions they would be asked, were shared with participants.

To elicit the causal effect of spousal observability the experimental conditions varied whether an individual's responses to the relevant survey sections were observable to their spouse. During the survey, participants were first asked to report their beliefs about their spouse's finances. I use this information to document information asymmetry among spouses about each other's finances. After this section, each participant's treatment status was revealed to them. Participants in the treatment group were informed that the following survey section was designed as a joint activity with their spouse and that their responses in the following section would be shared with and observable to their spouse. In contrast, participants in the control group were informed that the following survey section was a separate activity for each spouse and that their responses would be kept private. Additional details of the experimental protocol including the treatment and control scripts read by surveyors to introduce the experimental survey section are described in Appendix 1.10.3.

In the experimental survey section, migrants and spouses reported their average monthly income and expenses. To ensure respondents did not report the transnational household's combined finances, migrants were specifically asked to report their income and expenses in the UAE while spouses were asked to report theirs in the Philippines. In addition, migrants were asked to exclude any remittances they sent from their reported expenses, and spouses were asked to exclude any remittances they received from their reported income. Respondents had the option to report information in either dirhams or philippine pesos. For the analysis, the responses have all been standardized to dirhams based on the exchange rate at the time of the survey to allow for comparisons.

Participants responded to the experimental survey section, which asked questions about their finances, knowing whether or not the information they were reporting would be shared with and observable to their spouse. Any difference in the information reported by the treatment and control groups is therefore the causal effect of spousal observability. The experimental design allows me to identify strategic misreporting using self-reported data, without observing participants' true finances (or the difference between their self-reported and true finances). Identification is driven by the *difference* in self-reported information when it is observable and not observable to the spouse. Any measurement issues related to self-reporting would equally affect the treatment and control groups and therefore not bias my results.

Random assignment was done at the couple level using the baseline survey data and was not stratified by any pretreatment characteristics.[11] Half of the invited participants were randomly assigned to the treatment group. Treatment status was assigned to all invited participants before they were contacted. Each participant was administered their treatment status specific survey. Participants were not informed that there were multiple treatment conditions, that treatment was assigned at the couple level, or the treatment status of their spouse.

1.5 Data

1.5.1 Descriptive Statistics

Table 1.1 shows summary statistics from the migrant baseline surveys for all invited migrants, participating migrants, and participating migrants by their treatment status. By design, all migrants in the sample are married. They are on average 37 years old and have two children. 69 percent of the invited migrant sample are men. Although less than a third of the sample is female, the proportion of women among Filipino migrants in the UAE is

[11] Treatment assignment was done at the couple level to avoid any household conflict from spouses being assigned different treatment status.

substantially higher than the proportion of women among migrants from other countries.[12]

Migrants whose spouses are in the Philippines generally have low incomes because of the income and employment requirements for family immigration described in section 1.3.1. A majority of the sample earned between 1,500 dirham (USD $408) and 4,500 dirham (USD $1,225) per month. (The income threshold for family immigration for male migrants and some female migrants is 4,000 dirham (USD $1,088) and 10,000 dirham (USD $2,722), respectively.) Migrants are primarily employed in the services, sales, and construction sectors. In terms of remittance behavior, all migrants have sent remittances to their spouses in the past year. 90 percent of migrants send remittances to their households every month and in almost all cases their spouse is their primary remittance recipient. The average monthly remittance is around 1,555 dirhams (USD $423) which corresponds to 40 percent of migrant's monthly income.[13] This matches findings in other studies that show that migrants with transnational households send a significant portion of their incomes as remittances. Migrants also report sending remittances to on average one other recipient over the last year. Other recipients include parents, siblings, in-laws, and other relatives.

Migrants are generally well settled in the UAE having lived there for an average of seven years. As employment contracts are two years long, the average stay of seven years implies that migrants stay for multiple contract cycles, either renewing with the same employer or switching employers. Contracts often include in-kind benefits such as food, housing, and annual flight tickets for migrants to visit their households. Most migrants visit their household in the Philippines once every year and the average duration since their last visit at the time of the survey was around two years.

To understand the level of communication and control over the household's finances at baseline, migrants were asked about their financial decision-making. Migrants report discussing household budgets with their spouses on average once per month. 43 percent of migrants say they would like more control over how their spouse spends remittances, while around half report that they instruct their spouses on how to spend remittances.

1.5.2 Selection and Balance

I test for selection into the study and selection into treatment to address concerns about external and internal validity of the experiment. I do not find evidence of either type of selection based on observables and discuss below how selection on unobservables may impact my treatment effect estimates.

[12]Other migrants primarily from South Asia are predominantly male. As a result, UAE had the highest gender imbalance in the world in 2015, with a male/female ratio of 2.2

[13]Based on the average income reported by the control group in the experimental survey. In the baseline survey migrants only reported their income range.

Individuals who hide or misreport information to their spouses may be less likely to participate in the study regardless of their treatment status. Participation involves reporting information to a research team that will also be in contact with their spouses. This may be enough of a deterrent from participating for anyone actively seeking to hide information from their spouse, causing me to underestimate misreporting. Alternately, individuals who are seeking information about their spouse's finances may participate in the study to access more information. Selection on this margin, however, is not associated with a participant's own misreporting and would not bias my estimated treatment effect. Although if these individuals are also more misinformed about their spouse's finances, I may overestimate information asymmetry at baseline.

I check for selection into the study using the migrant baseline data by comparing all migrants who were invited to participate in the study to those who participated. Columns (1) and (2) on table 1.1 show the means for the invited migrant sample and those who agreed to participate. Column (3) shows the p-value from the two-sided t-test of the equivalence of means of those who did and did not participate. I find no evidence of selection into the study based on observable remittance behavior or measures of communication and control over the household's finances. The only statistically significant difference is that participants were less likely to be male than non-participants and I control for this in my regression specifications.

Selection may also be based on treatment status. Individuals who want to hide information may be less likely to participate if they are assigned to the treatment group where survey responses would be shared with their spouse, again causing me to underestimate misreporting. As the treatment status was assigned before contacting migrants, I test for and find no evidence of selection into the study based on treatment assignment. Columns (4) and (5) of table 1.1 show group means by treatment status and column (6) shows the p-value from the two-sided t-test of the equivalence of these means.

Although I do not have baseline data for spouses, as a proxy, I use the migrant baseline to test for selection among migrants whose spouses participated in the study. The results are shown in table 1.11 which replicates table 1.1 for the sample of migrants whose spouses participated. I again do not find evidence of selection into the study or treatment.

1.6 Empirical Analysis and Results

Using a combination of descriptive and experimental results I show that there is significant information asymmetry between migrants and spouses and this asymmetry is driven by strategic misreporting. Information asymmetry and strategic misreporting is greater when information is more difficult to observe and less likely to be verified. Spouses and certain subgroups of migrants strategically underreport income to influence the remittance decision in their favor. Spouses hide income on the intensive margin by underreporting known sources of income instead of reporting zero income which would be easier for migrants to catch. Income is only hidden when migrants do not demand control over or regularly communicate about the household's finances, making them less likely to verify reported information.

1.6.1 Descriptive Analysis

1.6.1.1 Remittances and Net Income

First I show that remittances are increasing in the migrant's net income and decreasing in the spouse's net income, creating incentives for misreporting net income on both sides of the remittance relationship and validating a key feature of the conceptual framework presented in section 1.2.

Figure 1.2 shows scatter plots and the accompanying linear regression lines for monthly remittances plotted against the migrant's and spouse's reported net income. Remittances are reported in the migrant baseline survey, while the net income for each spouse is the difference between their reported monthly income and expenses in the experimental survey. The figures are drawn using data from only the control group, as income and expenses reported by the treatment group are affected by the treatment condition.

Panel A shows that remittances are positively correlated with migrant's net income. The linear regression line has a slope of 0.4, implying that migrants remit 40 percent of their reported net income and by underreporting net income they can decrease the amount of remittances they have to send. Panel B shows that remittances are negatively correlated with spouse's net income. The slope of the regression line in -0.5, implying that for each additional dirham of reported spousal net income, remittances decrease by 0.50 dirham. By underreporting net income spouses can increase the remittances they receive from migrants. Panel B also shows that a majority of spouses have negative net income. Their income, excluding any remittances they receive, is less than their expenses; highlighting that remittances are essential for these spouses to sustain their households in the Philippines.

1.6.1.2 Information Asymmetry

Next, I document the extent of information asymmetry between migrants and spouses. Prior work has primarily focused on information asymmetry of migrant's income among migrant-sending households (Baseler, 2018; Seshan and Zubrickas, 2017; Joseph, Nyarko and Wang, 2018). However, as the conceptual framework showed, both migrants and spouses have incentives to strategically misreport information to each other. In addition, the relevant parameter for the remittance decision is reported income net of expenses. Overstating expenses has the same impact as hiding income and is therefore also a plausible margin for strategic misreporting. By focusing only on income we may underestimate the true scope of information asymmetry in the transnational household. I expand on the literature by documenting information asymmetry; first, across multiple margins, and second, on both sides of the remittance relationship. I find that both migrants and spouses have biased beliefs about each other's finances. Migrants underestimate spouses' income and overestimate spouses' expenses, whereas spouses underestimate migrants' in-kind employment benefits.

In the experimental survey, in addition to reporting their own finances, migrants and spouses reported their beliefs about each other's finances. To measure information asymmetry, the comparison of each spouse's reported finances with the other's beliefs is visually shown in figures 1.3 and 1.4, and statistically analyzed in table 1.2. Again, the comparison is made using data from only the control group.[14]

Figure 1.3 shows spouse's beliefs about the migrant's finances. Spouses underestimate migrant's income by 16 percent (630 dirham or USD $172) and overestimate migrant's expenses by 15 percent (183 dirham or USD $50). Although these differences are large, because of significant variation in these measures they are not statistically significant. Migrants and spouses were also asked to report if migrants receive non-wage benefits. These benefits are a common and sizable component of migrant remuneration in the UAE, however, compared to wage income they are more difficult for spouses to observe. Panel II shows that spouses are not aware that migrants receive in-kind food, housing, transport and health care benefits, and these differences are all statistically significant (see table 1.2 for comparison on means).

Figure 1.4 shows information asymmetry among migrants about their spouse's finances. Despite the literature's focus on biased beliefs among migrant-sending households, I find strong evidence of biased beliefs among migrants. Migrants underestimate spouse's income and overestimate spouse's expenses by 29.5 and 22.3 percent respectively. Despite similar patterns to figure 1.3, these differences are larger and also statistically significant in both cases, highlighting that information asymmetry is greater among migrants. Figure 1.4 also shows that on average spouses' incomes, excluding any remittances they receive, are less than their expenses. Migrants are aware of and overestimate this gap in spouses' net income.

These results show the importance of analyzing information asymmetry across multiple margins from both sides of the remittance relationship. They also raise the question - why do migrants and spouses have biased beliefs about each other's finances? The directions of the bias (underestimating income and overestimating expenses) support the claim that

[14]Beliefs were elicited before the treatment assignment was revealed and are therefore not affected by treatment status.

these biases are caused by strategic behavior to influence the remittance decision. However, this evidence is only suggestive. Biased beliefs may exist for many reasons including a lack of communication or interest in financial issues. Money and finances are difficult topics to discuss for any household so biases may persist due to communication frictions without any strategic motivations. To causally identify whether these biased beliefs are the result of strategic misreporting I now analyze the results of the spousal observability experiment.

1.6.2 Experimental Analysis & Results

1.6.2.1 Specification

To identify strategic misreporting I estimate the following OLS regressions in the experimental results that follow:

$$Y_i = \alpha + \beta T_i + \gamma X_i + \epsilon_i \tag{1.1}$$

Y_i is the outcome of interest, either reported income or expenses. In the main specification, both outcomes are measured as average monthly amounts in dirhams. T_i is the treatment status indicator, X_i is a vector of controls from the baseline survey and ϵ_i is the error term adjusted for heteroskedasticity. The coefficient β is the average difference between the outcome when an individual's response is observable and not-observable to their spouse. I run regressions, separately for migrants and spouses to allow for heterogeneous responses to treatment, both without and with controls to improve the precision of my treatment estimates. I also report randomization inference p-values for the treatment estimates from 5,000 replications of the treatment assignment. All controls are from the migrant baseline survey and include demographic characteristics, baseline income-category dummies, and measures of monitoring and remittance behavior.

1.6.2.2 Main Results

Spouses hide income by strategically underreporting it when it is observable to the migrant. Income is hidden on the intensive margin by underreporting known sources of income and hiding is greater when the migrant also participates in the study, reinforcing the result that underreporting is driven by spousal observability. I do not find evidence of income misreporting by migrants. Neither migrants nor spouses misreport expenses.

Table 1.3 shows the treatment effect of migrant observability on spouse's reported monthly income and expenses. Columns (1) and (2) show the results for reported income without and with controls, respectively. Spouses in the treatment group underreport their income by 213 dirham (USD $58) which represents a 31 percent decrease from the control group's average income of 685 dirham (USD $186). The magnitude of the treatment effect almost exactly matches the magnitude of the migrants' bias in their belief about spouses' income (migrants underestimate spouses' income by 200 dirham or USD $54), supporting the claim that the information asymmetry is driven by strategic misreporting. In terms of the remittance relationship, this underreporting of income is associated with a 107 dirham (USD $29) or a 7 percent increase in monthly remittances, based on the relationship between spouse's reported net income and remittances shown in section 1.6.1.1.

Columns (3) and (4) show the results for reported expenses, without and with controls, respectively. I do not find evidence that spouses strategically misreport their expenses. This is not surprising because for both migrants and spouses, despite being a plausible margin for underreporting net income, overreporting expenses is likely to invite greater scrutiny and verification from the other spouse. Neither migrants nor spouses want to reward each other for greater spending. As a result of this moral hazard, reported expenses are more likely to be verified and any misreporting of expenses is more likely to be caught. Expenses are therefore not the preferred margin for misreporting net income.

Table 1.4 shows the results of spousal observability for migrants. Migrants in the control group reported an average monthly income of 3,809 dirham (USD $1,037) and expenses of 1,201 dirham (USD $327). Although for income the treatment coefficients are around 6 percent of the control mean, because of the large variation I do not find statistically significant evidence that migrants underreport income when it is observable to their spouse. These results persist after controlling for migrant baseline characteristics in columns (2). Similar to spouses, I do not find evidence that migrants misreport expenses.

The migrant results match the descriptive findings of limited information asymmetry among spouses about the migrant's finances. Spouse's beliefs were not statistically different from the migrant's reported income and expenses, providing suggestive evidence that migrants were either not misreporting information on these margins, or any misreporting was limited in magnitude. This descriptive evidence is corroborated by the experimental evidence that also does not show strategic misreporting by migrants.

To further analyze the relationship between misreporting and spousal observability I leverage the variation in migrant and spousal participation in the study. My sample includes migrants whose spouses did not participate, spouses whose migrants did not participate, and matched couples, where both the migrant and spouse participated. If an individual knows that their spouse is not participating in the study, treatment assignment will not affect spousal observability as reported information will remain private regardless of an individual's treatment status. Similarly, if an individual in the treatment group knows that their spouse is participating in the study, they may perceive a greater likelihood of their information being shared with and observable to their spouse.[15] While I do not inform individuals about their spouses' participation status, they may share this information directly. As a result, the intensity of the spousal observability treatment should be higher for matched couples and lower for cases where only the migrant or spouse participates.

To test for such a response, I re-estimate the treatment effects restricting the sample to matched couples - cases where both the migrant and their spouse participated in the study and therefore the effect of spousal observability should be greater. The results for spouses using the sample of matched couples are shown in table 1.5. Spouses hide more income when the migrant is also participating in the study. The treatment effect on reported income is larger in this subsample - 310 dirhams (USD $84) compared to 213 dirhams (USD $58) for the full sample. For expenses, I again do not find any evidence of misreporting by spouses. Table 1.6 shows matched couple results for migrants. The estimates are similar to the results for the full sample shown in table 1.4 and I again do not find evidence of strategic income or expense misreporting by migrants. The greater impact of participating as a matched couple on spouses may be driven by spouses being surveyed after migrants (see figure 1.1 for the project time-line). Although participants were not informed of the survey order, spouses are more likely to know the migrant's participation status because migrants were surveyed before them.

Spouses can hide income at the intensive margin, by underreporting but still reporting positive income from a source known by the migrant, or at the extensive margin, by reporting zero income and hiding income sources altogether. Based on the conceptual framework, income is less likely to be hidden when it is easier to verify. Verifying the existence of an income source is easier than verifying the amount of income earned from a known source. Income hiding at the extensive margin is therefore more likely to be caught because peers and other family members can also observe and verify the spouse's income sources for the migrant. In contrast, income hiding at the intensive margin is difficult to verify, even for other family members.

I test for and find that income hiding is driven by the intensive margin in figure 1.5 and table 1.7. Figure 1.5 shows the cumulative distribution of reported income, separately for migrants and spouses by treatment group. Panel A shows that in both the treatment and control groups about a third of spouses report zero income i.e. the remittances they receive are their only reported income source. Spouses do not hide income at the extensive margin

[15] Although participants in the treatment group were not told that the sharing of information with their spouse is conditional on their spouse's participation, they may still believe that information sharing is more likely if their spouse participates.

by differentially reporting zero income when it is observable to the migrant. The figure also shows that the distribution of spouses' reported incomes when it is observable to the migrant is always lower than the distribution when it is not observable to the migrant; i.e., the distribution when the response is observable stochastically dominates the distribution when the response is not observable. Table 1.7 presents an alternate specification, replicating the spouse's results from the main specification in table 1.3 with the outcome variable measured as the log of reported monthly income in dirhams. This specification drops spouses that report zero income, focusing exclusively on misreporting on the intensive margin. The results remain similar to table 1.3. Spouses underreport their income by 46 log points when it is observable to the migrant, showing that income hiding is driven primarily by the intensive margin of underreporting known sources of income which is harder for migrants to verify.[16]

1.6.3 Heterogeneity & Mechanisms: Who is Hiding and Why?

Based on the conceptual framework presented in section 1.2 strategic misreporting is motivated by the benefits of hiding income and constrained by the likelihood and punishment from being caught. It is therefore instructive to examine whether income hiding varies by characteristics associated with greater benefits from hiding and an increased ability of couples to verify each other's reported income. In this section, I analyze treatment effect heterogeneity by measures of communication and control over the household's finances, remittance behavior, and gender. I find that spouses only hide income when migrants do not demand control over or regularly communicate about the household's finances, making them less likely to verify reported information. High remittance sending migrants hide income to avoid sending even more remittances. Among both migrants and spouses, women are more likely to hide income. These gender differences also appear to be driven by men less frequently demanding control over and communicating about the household's finances.

1.6.3.1 Specification

I estimate the following modified regression:

$$Y_i = \alpha + \beta T_i + \lambda(T_i \times x_i) + \gamma X_i + \epsilon_i \tag{1.2}$$

$T_i \times x_i$ is the interaction between the treatment status and trait x_i. The coefficient β is now the average treatment effect for individuals that do not have trait x, λ is the difference

[16]Appendix table 1.12 shows the results for the log of migrant's reported monthly income and finds results similar to the main results in table 1.4. The results for the log of migrant's income and log of both migrant's and spouse's expenses are less informative because they are never zero.

between the average treatment effect of individuals that have and do not have trait x, and the sum of β and λ is the average treatment effect for individuals with trait x. I run regressions separately for migrants and spouses for each trait x. All regressions include the vector of controls X which always includes the main effect of trait x.

The results are shown in table 1.8 for spouses and table 1.9 for migrants. Each column is a separate regression, reporting the coefficient of the treatment indicator, the interaction of the treatment indicator with each trait, and the sum of the treatment and interaction coefficients. To allow comparisons, the first columns reproduces the main income results for spouses and migrants from column (2) in table 1.3 and 1.4 respectively. I again report heteroskedasticity-robust standard errors and randomization inference p-values.

1.6.3.2 Communication and control over finances

First I use three measures of the migrant's financial communication and control to analyze the impact of increased verification on income hiding.

A growing literature highlights that because of differences in the spending preferences of migrants and their spouses, migrants send fewer remittances when they cannot control how those remittances are spent (Ashraf et al., 2015; Yang, 2011; Chin, Karkoviata and Wilcox, 2015). I test whether migrants wanting more control over remittance spending impacts income hiding by migrants and spouses. The trait *control* is a dummy equal to one if at baseline the migrant reports wanting more control over how remittances are spent by their spouse.

Column (2) of table 1.8 shows that income hiding by spouses is entirely driven by spouses of migrants who do not demand more control over remittance spending. These spouses underreport their income by 573 dirham (USD $156) when it is observable to the migrant. On the other hand, spouses of migrants who demand control over remittance spending, do not underreport income when it is observable to the migrant. Migrants who want more control over remittance spending may communicate this demand to their spouses, alerting them to increased scrutiny from the migrant over the household's finances. This scrutiny would increase the likelihood of the migrant catching any misreporting, deterring spouses from hiding income. In contrast, spouses of migrants who do not report wanting more control and therefore do not face increased scrutiny, hide income. For migrants, column (2) of table 1.9 shows that wanting more control over remittance spending is not associated with greater income hiding.

I now move from analyzing cases where migrants *want* more financial control to cases where they *exercise* more financial control. The most basic form of financial control that migrants can exercise is communicating about the household's finances with their spouses -

asking and instructing spouses about where money is coming from and where it should be spent. In columns (3) and (4) of tables 1.8 and 1.9, I test whether increased communication about household finances impacts income misreporting by both spouses and migrants. *Instruct* and *budget* are dummies equal to one if at baseline the migrant reports instructing their spouse on how to use remittances and if the migrant discusses the household budget with their spouse more frequently than the median number of times (once every two months), respectively.

Table 1.8 shows that increased instruction and communication about finances from the migrant limits misreporting by spouses. The likelihood of the spouse's misreporting being caught is higher when the migrant regularly communicates about the household's finances. As a result, income hiding is entirely driven by spouses of migrants who do not exercise these traits. Spouses of migrants who do not instruct their spouses on remittance spending and less frequently discuss the household budget, underreport their income by 576 dirham (USD $157) and 597 dirham (USD $163) respectively. Underreporting income is only a beneficial strategy, as I find, when the migrant does not exercise control through communication. For migrants both measures are associated with lower misreporting, suggesting that increased communication also deters the migrant from hiding income. However, given the large variation in migrant's reported income, these effects are not statistically significant.

These three measures of communication and control are highly correlated with each other and proxy for the household's underlying relationship dynamics. Couples who have shared financial goals and actively communicate about and jointly make financial decisions are less likely to resort to income hiding to achieve their goals. These are important findings because they show that strategic income hiding is limited to certain subgroups of transnational households, that can be identified by observable baseline characteristics. This can improve the targeting and effectiveness of financial products and services for transnational households that leverage information sharing and control to impact financial decision-making.

1.6.3.3 Remittance Behavior

Based on the conceptual framework presented in section 1.2, income is primarily hidden to impact the remittance relationship, to either avoid or induce more remittances. I use two measures of remittance behavior to analyze its interaction with income hiding; *amount* is the average monthly remittances in dirham that migrants report sending to their spouses at baseline and *median* is an indicator equal to one if migrants report sending greater than the median monthly remittance amount of 1450 dirham (USD $395). The results for spouses and migrants are shown in column (5) and (6) of tables 1.8 and 1.9.

Both measures show that migrants that send higher levels of remittances underreport their income when it is observable by their spouse. For each additional dirham remitted, income is underreported by 0.77 dirham, and migrants who send higher than the median remittance underreport their income by 1,304 dirham (USD $355). This represents 34 percent of the income of the overall control group and 23 percent of the income of high remittance senders in the control group. For spouses, I find some evidence that those who receive low levels of remittances underreport their income to induce migrants to send more.

Income hiding by migrants that send higher remittances may be driven by an income effect. Higher remittances are associated with higher levels of income. At these income levels, migrants may be able to send a base level of remittances and still have significant money left over to hide, increasing the benefits from hiding it.[17] As my sample of migrants is primarily low-income this may also explain why I do not find evidence of income hiding by migrants on average. In addition, migrants that send higher levels of remittances may also face higher unmet demand for remittances from their spouses. This would encourage income hiding by increasing the likelihood that additional reported income would have to be shared.

1.6.3.4 Gender

Gender is an essential component of any analysis of household decision-making. However, the main challenge in analyzing the role of gender in transnational households has been the lack of gender balance among migrants (and spouses) in most settings. This makes it difficult to disentangle the impact of the role of each spouse in the transnational household as either the remittance sender or recipient, from their gender. In my sample, a third of migrants and two-thirds of spouses are women, allowing me to analyze the interaction of gender with strategic misreporting for both migrants and their spouses.

The trait *male* in column (7) of table 1.8 and 1.9 identifies male spouses and male migrants, respectively. I find that among both migrants and spouses, women underreport their income when it is observable to their spouse, whereas for male spouses and migrants the treatment effect of spousal observability is not statistically different from zero. Female spouses underreport their income by 323 dirham (USD $88) when it is observable to the migrant. This represents around 47 percent decrease from both the average reported income of the overall control group of 685 dirham (USD $186) and of women in the control group of 675 dirham (USD $183). Female migrants underreport their income by 578 dirham (USD $157) when it is observable to their spouse. This represents a 15 percent decrease from the average reported income of the overall control group of 3,809 dirham (USD $1,037) and a 17

[17] This matches Seshan and Zubrickas (2017) finding that information asymmetry of migrant's income among spouses, increases with the migrant's income.

percent decrease from the average reported income of women in the control group of 3,361 dirham (USD $915).

I use the migrant baseline data to further analyze gender differences in demographics, measures of communication and control over the household's finances, and remittance behavior. The results are shown in table 1.10. Female migrants earn lower incomes and send fewer remittances than male migrants. These differences stem from differences in employment - female migrants are less likely to be employed in the food-service and construction sectors and more likely to be employed in the personal service sector. However, these differences are unlikely to be the cause of greater income hiding by women. Lower incomes are more likely to be associated with less income hiding because low-income migrants have limited ability to send remittances and therefore limited incentives to hide their income (as discussed in section 1.6.3.3 *Remittance Behavior*). In addition, the findings of differential treatment effects for women among both migrants and spouses suggest that these differences are driven by factors that are not specific to the women's role in the transnational household as remittance sender or recipient and are instead broader.

Women's income hiding may instead be driven by gender norms about income and management of household finances. Filipino women are more likely to be the financial managers of the household, regardless of their income or occupation status, and are therefore more likely to want control of their husband's finances. Ashraf (2009) documents this norm for co-residing Filipino households and table 1.10 shows this norm persists for transnational Filipino households. Among migrants, women are more likely to instruct their spouse on remittance spending and more likely to want more control over the household's finances. As shown in section 1.6.3.2, these traits limit misreporting by the other spouse, therefore women may be better able to limit income hiding by men.

1.7 Discussion: Other Motivations, Strategies, and Implications for Welfare

In this section, I discuss potential motivations for income hiding other than remittances, strategies for income hiding other than purposeful misreporting, and the implications of income hiding on welfare.

The impact of income hiding on overall household welfare depends on spousal preferences and how hidden income is used and would have been used if it was not hidden. Without more information and assumptions about preferences and the counter-factual, the overall impact of income hiding on welfare cannot be determined. Even so, the results are still informative to policy discussions on household welfare. Policymakers have focused on interventions that facilitate greater remittances and reduce information asymmetries based on the positive impacts on a variety of measures of well being associated with each of these two outcomes (Yang, 2011). However, my results show that these two outcomes may be inconsistent with each other. The conceptual framework implies that if information asymmetry is reduced through a reduction in income hiding by spouses, remittances would also decrease, and vice versa. In such settings, the welfare impacts of information asymmetry reducing or remittance increasing interventions are a priori ambiguous.

Based on the conceptual framework and empirical evidence strategic misreporting is intended to influence remittance levels. However, misreporting may be intended to influence other aspects of the remittance relationship. Spouses may hide income if migrants view their participation in the labor force negatively. If based on the remittance contract, the spouse staying back takes on household and childcare responsibilities, higher reported incomes may signal to the migrant that spouses are not allocating enough time and effort to these tasks. Similarly, migrants may hide income but not reduce remittances to signal their altruism in sending a greater proportion of their income as remittances. Alternately, migrants may hide income to avoid sharing it with other household members. If spouses share the remittances they receive with other household or family members, migrants may hide income to avoid such sharing.

While the experimental design identifies purposeful misreporting, this may be an underestimate of strategic reporting behavior defined more broadly. Misreporting information is one of a range of actions that can be used to hide information from spouses. Instead of purposefully misreporting information individuals may avoid discussing financial matters or give incomplete information to their spouses. This passive misreporting, however, is no longer an option for individuals in the treatment group as their reported information will be shared with and observable to their spouses. They must either commit to hiding information and purposefully misreport it or report the truth. If spousal observability induces these passive misreporters to tell the truth, their prior misreporting will not be identified by the treatment effect. This passive misreporting may also explain why despite biased beliefs about expenses, I do not find evidence that migrants or spouses strategically misreport them.

1.8 Conclusion

This paper analyzes spousal communication in transnational households by eliciting the causal effect of spousal observability on reported information to identify if spouses strategically misreport information to each other. Research on information asymmetry between migrants and their households has primarily focused on income hiding by migrants and its impact on remittances and the perceived returns to migration. This is the first study that looks at strategic misreporting on both sides of the remittance relationship, across multiple margins of the household's actual finances.

I find that both migrants and spouses have biased beliefs about each other's finances. Spouses and certain subgroups of migrants strategically misreport information to each other. Misreporting is greater when information is more difficult to observe and less likely to be verified. The results are consistent with an income-sharing model where both spouses have private information and income hiding is constrained by the threat of punishment.

These results are important for policy-makers considering interventions to reduce information asymmetry among transnational households because they identify strategic misreporting as the cause of this asymmetry. Interventions that only increase communication between spouses would not be able to address this strategic behavior. In focus groups conducted before this study, almost all transnational couples reported communicating with each other daily through instant messages or phone calls. However, the results show that biased beliefs and the ability to strategically misreport information persist despite these significant improvements in communication technology. Addressing purposeful misreporting requires interventions that increase spouses' abilities to monitor and control each other's financial decision-making, including interventions that specifically increase communication about finances.

I also find that strategic misreporting is limited to certain subgroups of transnational households, implying that interventions to reduce strategic misreporting would be most effective when targeted to these households. Importantly, these subgroups can be identified from observable baseline characteristics of communication, monitoring, and control over financial decision-making.

1.9 Figures & Tables

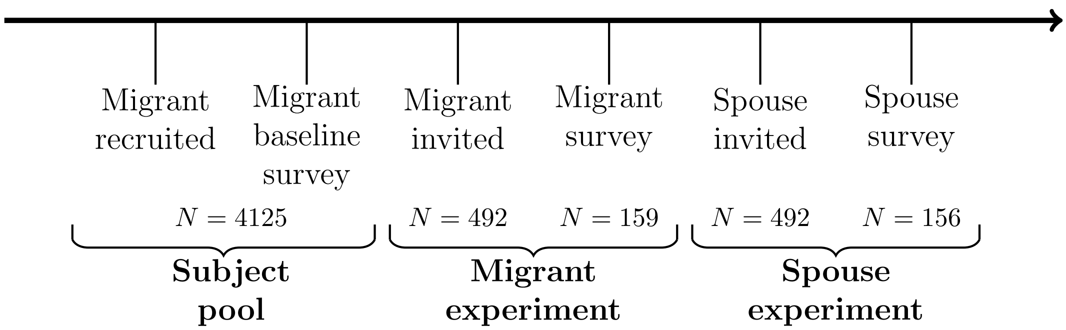
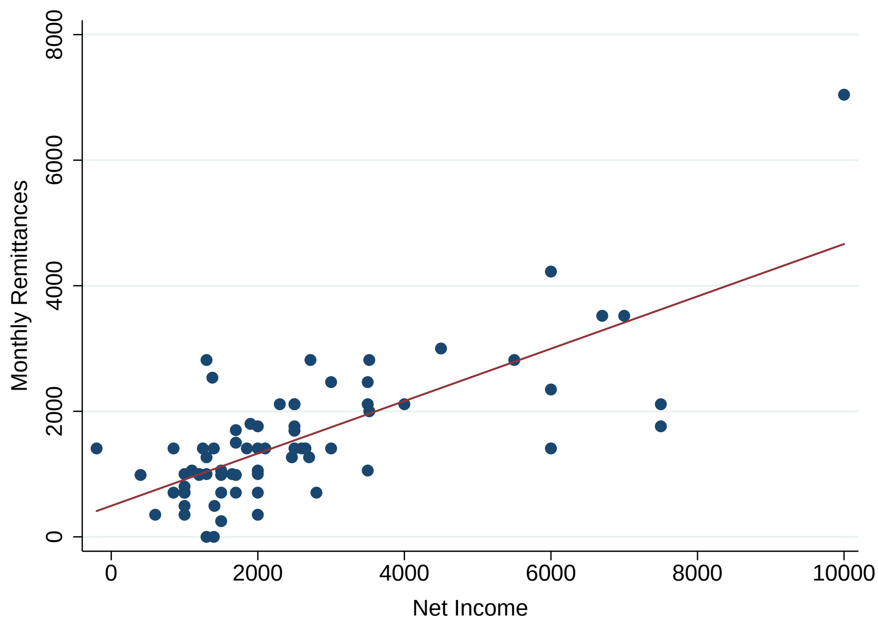


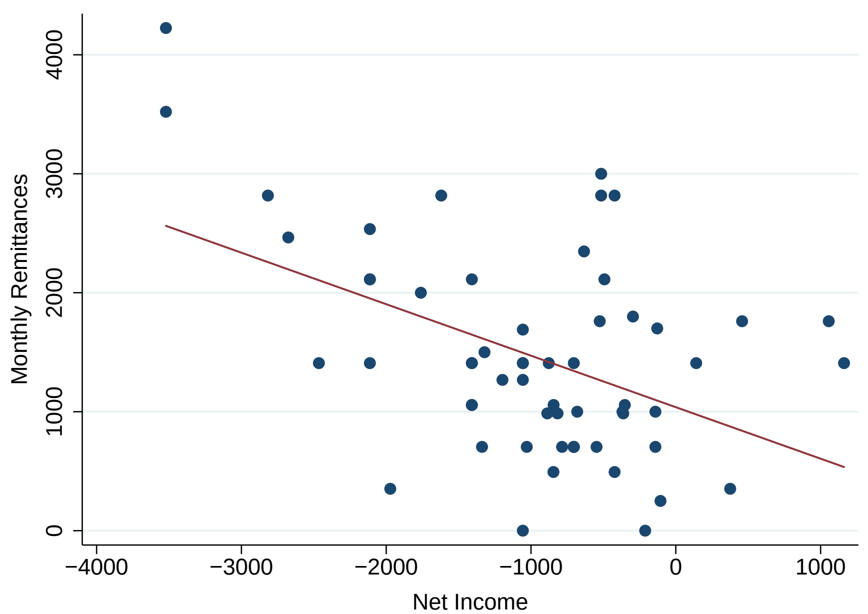
Figure 1.1: Project Timeline

Figure 1.2: Monthly Remittances Against Net Income

Panel A: Migrant's Net Income

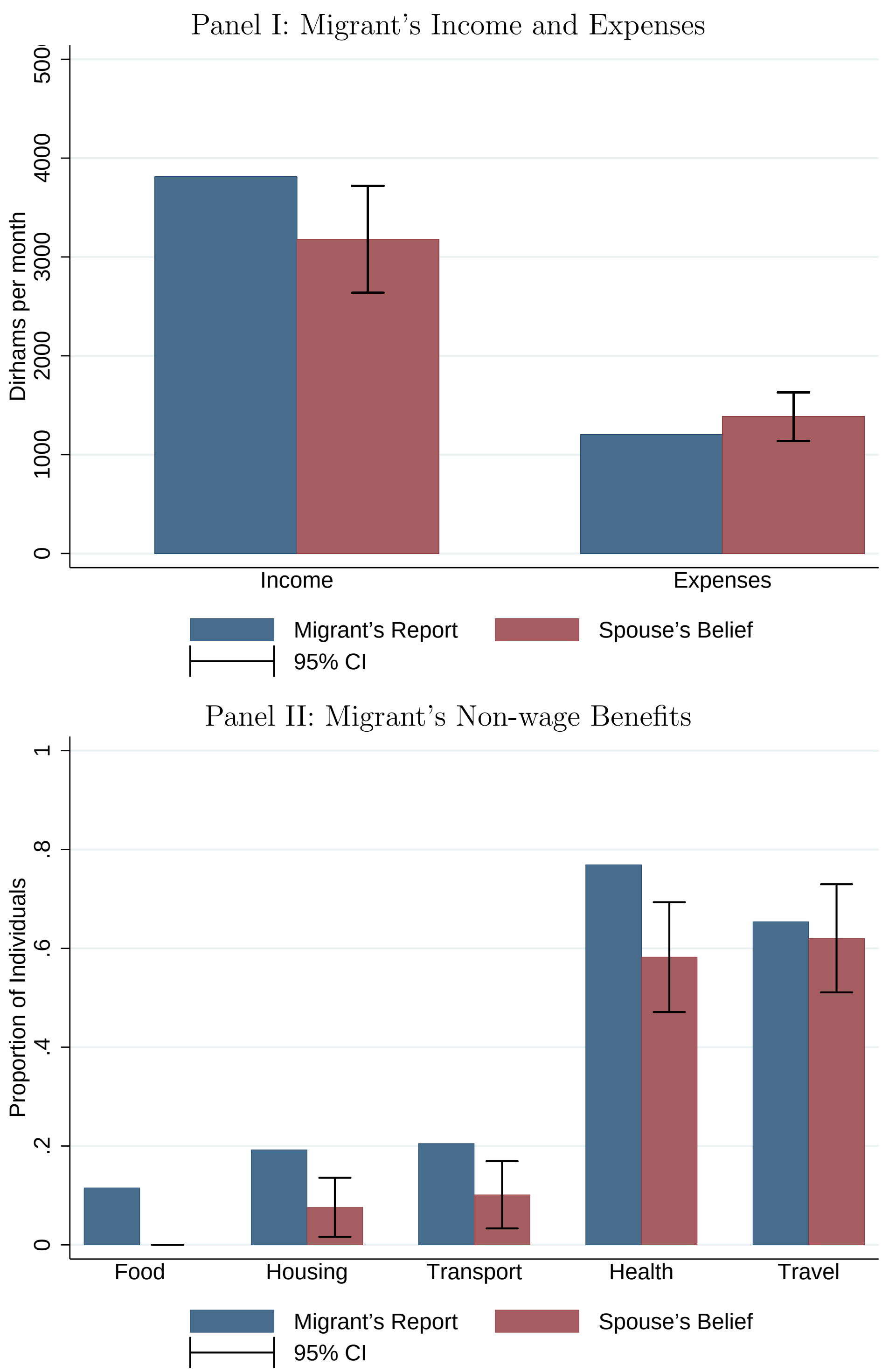


Panel B: Spouse's Net Income



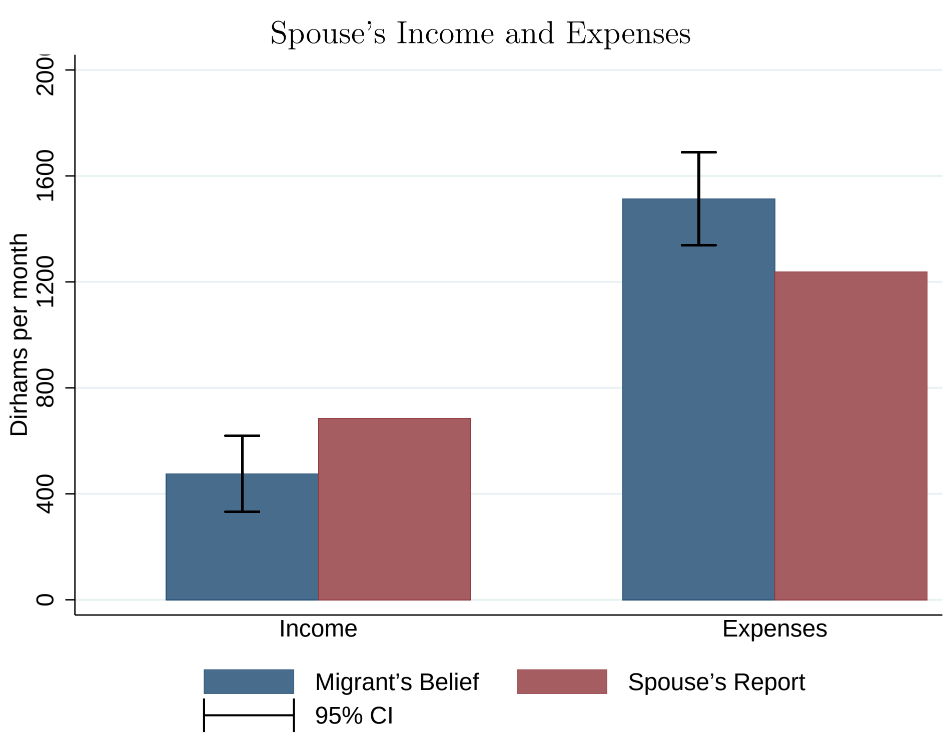
Notes: The figure shows scatter plots and linear regression lines for the control group. Remittances are measured in dirhams per month and were reported in the migrant baseline survey. Net income is reported monthly income net of reported monthly expenses, measured in dirhams. Panel A shows the relationship of remittances with migrant's reported net income and panel B shows it for spouse's reported net income.

Figure 1.3: Spouse's Information Asymmetry at Baseline



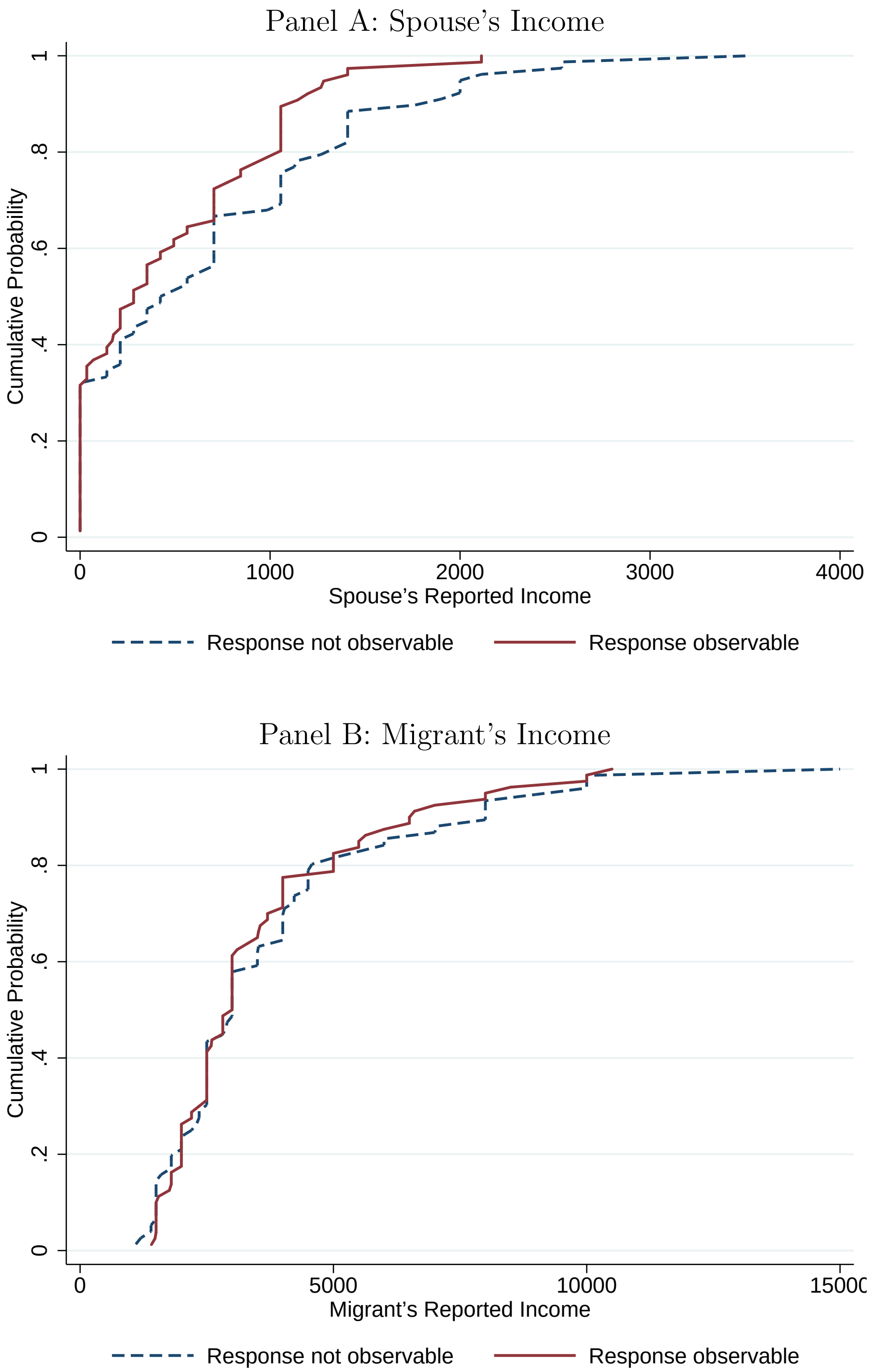
Notes: Vertical axis is dirhams per month in panel I and proportion of individuals reporting these migrant non-wage benefits in panel II.

Figure 1.4: Migrant's Information Asymmetry at Baseline



Notes: Vertical axis is dirhams per month.

Figure 1.5: Cumulative Distribution of Reported Income



Notes: Panels A and B plot the empirical cumulative distribution of the spouse's and migrant's reported incomes, respectively.

Table 1.1: Migrant Baseline Summary Statistics: Selection & Balance

	Selection			Balance		
	(1) Invited Sample	(2) Participating Sample	(3) Diff p-val	(4) Treat	(5) Control	(6) Diff p-val
Treatment	0.50	0.51	(0.82)	1.00	0.00	
Spouse participated		0.61		0.54	0.64	(0.21)
Demographics						
Male	0.69	0.62	(0.02)	0.63	0.60	(0.73)
Age	37.45	37.61	(0.74)	37.58	37.64	(0.96)
Children	1.95	2.09	(0.08)	2.02	2.17	(0.50)
Income range (AED/month)						
Less than 1,500	0.05	0.08	(0.09)	0.06	0.09	(0.51)
1,500 - 3,000	0.33	0.30	(0.27)	0.30	0.29	(0.98)
3,000 - 4,500	0.26	0.29	(0.39)	0.31	0.27	(0.59)
4,500 - 6,000	0.10	0.09	(0.63)	0.12	0.06	(0.20)
6,000 - 7,500	0.05	0.04	(0.63)	0.02	0.06	(0.23)
7,500 - 9,000	0.03	0.04	(0.46)	0.04	0.05	(0.66)
9,000 - 10,000	0.02	0.01	(0.48)	0.00	0.03	(0.16)
Greater than 10,000	0.06	0.06	(0.99)	0.09	0.04	(0.21)
Occupation						
Food & Personal Services	0.20	0.17	(0.22)	0.14	0.21	(0.25)
Sales	0.16	0.13	(0.15)	0.15	0.10	(0.39)
Construction & Maintenance	0.10	0.13	(0.15)	0.14	0.13	(0.89)
Administration	0.09	0.08	(0.38)	0.10	0.05	(0.26)
Communication and Control						
Years in UAE	7.05	6.79	(0.44)	6.93	6.65	(0.73)
Years since last visit		1.81		2.05	1.56	(0.12)
Visits per year		0.74		0.69	0.78	(0.29)
Relatives in UAE		2.69		2.43	2.95	(0.36)
Spouse HH members		3.29		3.32	3.26	(0.83)
Spouse lives with In-laws		0.30		0.34	0.26	(0.33)
Discuss budget (times per month)	1.10	1.06	(0.81)	1.37	0.78	(0.12)
Want more control of spending	0.43	0.42	(0.82)	0.46	0.39	(0.17)
Instruct spouse on spending	0.52	0.55	(0.24)	0.58	0.53	(0.49)
Remittance Behavior						
Spouse is main recipient	0.99	0.99	(0.50)	1.00	0.99	(0.32)
Other recipients	1.06	1.18	(0.18)	1.16	1.21	(0.84)
Remit monthly	0.90	0.91	(0.64)	0.93	0.89	(0.35)
Remittance (dirham/month)	1,555	1,449	(0.23)	1,330	1,517	(0.16)
N	492	159		81	78	

Notes: Columns (1) and (2) show means for all invited migrants and those who participated in the study, respectively. Column (3) shows the p-value from the two-sided t-test of equivalence of means between those who participated and those who were invited but did not participate in the study. Columns (4) and (5) show means within treatment and control groups, respectively. Column (6) shows the p-value from the two-sided t-test of equivalence of means between the treatment and control group.

Table 1.2: Information Asymmetry at Baseline

	Migrant's Report	Spouse's Report	Difference	
			Mean	p-val
	(1)	(2)	(3)	(4)
Panel A: Migrant's Information				
Income	3809.29	3178.31	630.98	(0.12)
Expenses	1201.39	1384.88	-183.49	(0.24)
Net Income	2634.66	2136.65	498.01	(0.17)
Employment Benefits				
Food	0.12	0.00	0.12	(0.00)***
Housing	0.19	0.08	0.12	(0.03)**
Transport	0.21	0.10	0.10	(0.07)*
Health	0.77	0.58	0.19	(0.01)**
Travel	0.65	0.62	0.03	(0.66)
Panel B: Spouse's Information				
Income	475.75	684.76	-209.01	(0.06)*
Expense	1513.50	1237.82	275.68	(0.02)**
Net Income	-975.50	-532.78	-442.71	(0.01)**
N	78	79		

Notes: Column (1) shows the means of migrant's reports of their own finances in panel A and the means of their beliefs about their spouse's finances in panel B. Column (2) shows the means of spouse's reports of their own finances in panel B and the means of their belief's about their migrant's finances in panel A. Column (3) shows the difference between the mean reports and beliefs. Column (4) shows the p-value from the two-sided t-test of equivalence of means between reports and beliefs. ***, **, * denote significance at the 1 percent, 5 percent and 10 percent levels

Table 1.3: Spouse's Reported Income & Expenses

	Income		Expenses	
	(1)	(2)	(3)	(4)
Response observable to migrant	-213.3**	-249.8**	-85.78	-95.00
	(104.5)	(114.0)	(101.4)	(93.35)
	[0.045]	[0.027]	[0.399]	[0.331]
Spouse is male		131.1		-324.6***
		(114.1)		(111.8)
Migrant's monthly remittances to spouse		-0.0711		0.0443
		(0.0652)		(0.0843)
Migrant's remittance: Above median		-151.0		74.92
		(170.8)		(171.5)
Migrant wants more control over remittance spending		4.704		-3.092
		(133.5)		(120.9)
Migrant instructs spouse about remittance spending		-210.8		-4.754
		(135.4)		(127.3)
Migrant discusses budget with spouse: Frequency above median		-5.847		-53.18
		(130.3)		
Mean when response not observable to migrant	684.8***		1,238***	
	(85.89)		(77.93)	
Observations	154	154	152	152
R-squared	0.026	0.121	0.005	0.266

Notes: Robust standard errors in parentheses. ***, **, * denote significance at the 1 percent, 5 percent and 10 percent levels, respectively. Randomization inference p-values for 5000 replications of the treatment assignment are shown in square brackets. Outcome variable is reported monthly income in dirham in columns (1)-(2) and reported monthly expenses in dirham in columns (5)-(8). Columns (2) and (4) include migrant baseline income category dummies. Monthly remittance are measured in dirham.

Table 1.4: Migrant's Reported Income & Expenses

	Income		Expenses	
	(1)	(2)	(3)	(4)
Response observable to spouse	-247.9	-190.1	72.89	41.38
	(379.3)	(212.8)	(132.7)	(98.27)
	[0.521]	[0.418]	[0.584]	[0.674]
Migrant is male		428.3**		-49.78
		(215.7)		(93.83)
Migrant's monthly remittances to spouse		0.641*		0.137
		(0.352)		(0.122)
Migrant's remittance: Above median		-418.1		-179.5
		(509.1)		(196.9)
Migrant wants more control over remittance spending		-292.5		-133.3
		(247.3)		(106.3)
Migrant instructs spouse about remittance spending		490.6*		122.6
		(270.6)		(104.0)
Migrant discusses budget with spouse: Frequency above median		440.8*		-12.41
		(255.4)		(119.3)
Years since migrant last visited spouse		-63.06		19.15
		(60.47)		(26.12)
Mean when response not observable to spouse	3,809***		1,201***	
	(296.5)		(97.07)	
Observations	155	155	156	156
R-squared	0.003	0.709	0.002	0.581

Notes: Robust standard errors in parentheses. ***, **, * denote significance at the 1 percent, 5 percent and 10 percent levels, respectively. Randomization inference p-values for 5000 replications of the treatment assignment are shown in square brackets. Outcome variable is reported monthly income in dirhams in columns (1)-(2) and reported monthly expenses in dirhams in columns (3)-(4). Columns (2) and (4) include migrant baseline income category dummies and dummies for participation and treatment status in accompanying experiment in the appendix. Monthly remittances are measured in dirham.

Table 1.5: Spouse's Reported Income & Expenses: Matched Couples

	Income		Expenses	
	(1)	(2)	(3)	(4)
Response observable to migrant	-309.6**	-248.6**	-19.73	-26.20
	(120.3)	(119.5)	(123.4)	(111.9)
	[0.013]	[0.089]	[0.878]	[0.833]
Spouse is male		101.8		-213.7
		(127.3)		(143.2)
Migrant's monthly remittances to spouse		-0.0799		0.307**
		(0.152)		(0.139)
Migrant's remittance: Above median		90.84		-182.1
		(296.8)		(256.5)
Migrant wants more control over remittance spending		-139.3		-61.51
		(174.6)		(180.4)
Migrant instructs spouse about remittance spending		15.43		-20.05
		(164.0)		(181.4)
Migrant discusses budget with spouse: Frequency above median		-157.7		86.78
		(179.3)		(157.5)
Mean when response not observable to migrant	670.3***		1,124***	
	(102.6)		(96.16)	
Observations	93	93	91	91
R-squared	0.065	0.188	0.000	0.266

Notes: Robust standard errors in parentheses. ***, **, * denote significance at the 1 percent, 5 percent and 10 percent levels, respectively. Randomization inference p-values for 5000 replications of the treatment assignment are shown in square brackets. Outcome variable is reported monthly income in dirham in columns (1)-(2) and reported monthly expenses in dirham in columns (5)-(8). Columns (2) and (4) include migrant baseline income category dummies. Monthly remittances are measured in dirham.

Table 1.6: Migrant's Reported Income & Expenses: Matched Couples

	Income		Expenses	
	(1)	(2)	(3)	(4)
Response observable to spouse	-254.6	-360.0	-9.991	-104.9
	(438.9)	(297.4)	(161.2)	(150.2)
	[0.562]	[0.174]	[0.953]	[0.411]
Migrant is male		422.5		0.338
		(267.9)		(116.5)
Migrant's monthly remittances to spouse		0.166		0.0874
		(0.299)		(0.231)
Migrant's remittance: Above median		325.6		49.92
		(485.4)		(272.5)
Migrant wants more control over remittance spending		-137.8		-191.9
		(345.9)		(149.8)
Migrant instructs spouse about remittance spending		102.9		125.8
		(310.7)		(139.1)
Migrant discusses budget with spouse: Frequency above median		782.0**		149.1
		(347.1)		(162.4)
Years since migrant last visited spouse		17.19		48.49*
		(51.81)		(24.82)
Mean when response not observable to spouse	3,697***		1,207***	
	(311.8)		(110.5)	
Observations	94	94	94	94
R-squared	0.004	0.763	0.000	0.574

Notes: Robust standard errors in parentheses. ***, **, * denote significance at the 1 percent, 5 percent and 10 percent levels, respectively. Randomization inference p-values for 5000 replications of the treatment assignment are shown in square brackets. Outcome variable is reported monthly income in dirham in columns (1)-(2) and reported monthly expenses in dirham in columns (5)-(8). Columns (2) and (4) include migrant baseline income category dummies. Monthly remittances are measured in dirham.

Table 1.7: Log of Spouse's Reported Income & Expenses

	Log Income		Log Expenses	
	(1)	(2)	(3)	(4)
Response observable to migrant	-0.462**	-0.544***	-0.0604	-0.0795
	(0.177)	(0.181)	(0.106)	(0.104)
	[0.009]	[0.003]	[0.568]	[0.446]
Spouse is male		0.143		-0.315**
		(0.199)		(0.125)
Migrant's monthly remittances to spouse		0.000409**		0.0000241
		(0.000163)		(0.0000647)
Migrant's remittance: Above median		-0.897**		0.102
		(0.387)		(0.145)
Migrant wants more control over remittance spending		0.367		-0.0601
		(0.233)		(0.127)
Migrant instructs spouse about remittance spending		-0.484**		-0.0224
		(0.206)		(0.115)
Migrant discusses budget with spouse: Frequency above median		-0.0657		-0.0381
		(0.222)		(0.110)
Mean when response not observable to migrant	6.639***		6.945***	
	(0.110)		(0.0733)	
Observations	105	105	152	152
R-squared	0.062	0.246	0.002	0.215

Notes: Robust standard errors in parentheses. ***, **, * denote significance at the 1 percent, 5 percent and 10 percent levels, respectively. Randomization inference p-values for 5000 replications of the treatment assignment are shown in square brackets. Outcome variable is log of reported monthly income in dirham in columns (1)-(2) and log of reported monthly expenses in dirham in columns (5)-(8). Columns (2) and (4) include migrant baseline income category dummies. Monthly remittances are measured in dirham.

Table 1.8: Spouse's Reported Income: Heterogeneous Treatment Effect

		Communication & Control			Remittance Behaviour		Gender
		Control	Instruct	Budget	Amount	Median	Male
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treatment	-249.8**	-572.7***	-576.1***	-597.7***	-368.9**	-195.8	-323.3**
	(114.0)	(148.0)	(187.8)	(196.1)	(181.6)	(138.3)	(148.7)
	[0.027]	[0.000]	[0.003]	[0.003]	[0.049]	[0.152]	[0.032]
Treat x Trait		721.5***	579.7**	525.9**	0.0846	-162.0	221.0
		(203.8)	(226.9)	(227.0)	(0.0936)	(227.9)	(212.0)
		[0.001]	[0.014]	[0.031]	[0.384]	[0.521]	[0.320]
Treatment + (Treat x Trait)		148.7	3.549	-71.75	-368.8	-357.8*	-102.3
		(151.3)	(129)	(128.1)	(181.5)	(188.3)	(157.3)
		[0.332]	[0.979]	[0.585]	[0.049]	[0.063]	[0.518]
Controls	Y	Y	Y	Y	Y	Y	Y
Observations	154	154	154	154	154	154	154
R-squared	0.121	0.187	0.163	0.154	0.125	0.124	0.127

Notes: Robust standard errors in parentheses. ***, **, * denote significance at the 1 percent, 5 percent and 10 percent levels, respectively. Randomization inference p-values for 5000 replications of the treatment assignment are shown in square brackets. The outcome variable is reported monthly income in dirhams. The Trait variable is defined in the title of each column. Control, Instruct, and Budget, are dummies equal to one if the migrant reports wanting more control over how remittances are spent, instructing their spouse on how to spend remittances and discussing the household budget with their spouse more than the median frequency. Male is a dummy equal to one if the spouse is male. Remittance amount is monthly remittances in dirhams and Remittance Median is a dummy equal to one if the migrant reports sending more remittances than the median amount. All regressions include the controls used in the main results.

Table 1.9: Migrant's Reported Income: Heterogeneous Treatment Effect

		Communication & Control			Remittance Behaviour		Gender
		Control	Instruct	Budget	Amount	Median	Male
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treatment	-190.1	-52.08	-595.0*	-516.5	905.9**	231.5	-575.8**
	(212.8)	(366.0)	(353.5)	(481.7)	(401.9)	(217.9)	(287.0)
	[0.418]	[0.888]	[0.120]	[0.264]	[0.093]	[0.328]	[0.053]
Treat x Trait		-300.4	711.4*	499.8	-0.770**	-1,304**	656.1
		(475.3)	(415.3)	(547.8)	(0.299)	(577.6)	(423.7)
		[0.495]	[0.141]	[0.344]	[0.046]	[0.021]	[0.141]
Treatment + (Treat x Trait)		-352.4	116.4	-16.70	905.1	-1072**	80.31
		(250.5)	(246.8)	(226.1)	(401.7)	(508.4)	(302.8)
		[0.164]	[0.687]	[0.945]	[0.093]	[0.038]	[0.807]
Controls	Y	Y	Y	Y	Y	Y	Y
Observations	155	155	155	155	155	155	155
R-squared	0.709	0.709	0.714	0.711	0.734	0.724	0.713

Notes: Robust standard errors in parentheses. ***, **, * denote significance at the 1 percent, 5 percent and 10 percent levels, respectively. Randomization inference p-values for 5000 replications of the treatment assignment are shown in square brackets. The outcome variable is reported monthly income in dirhams. The Trait variable is defined in the title of each column. Control, Instruct, and Budget, are dummies equal to one if the migrant reports wanting more control over how remittances are spent, instructing their spouse on how to spend remittances and discussing the household budget with their spouse more than the median frequency. Male is a dummy equal to one if the migrant is male. Remittance amount is monthly remittances in dirhams and Remittance Median is a dummy equal to one if the migrant reports sending more remittances than the median amount. All regressions include the controls used in the main results.

Table 1.10: Differences by Gender: Migrant Baseline

	(1) Men	(2) Women	(3) Difference Mean	p-val
Demographics				
Age	38.49	36.20	-2.29	(0.05)*
Children	2.12	2.05	-0.07	(0.73)
Income range (AED/month)				
Less than 1,500	0.05	0.11	0.06	(0.18)
1,500 - 3,000	0.23	0.39	0.16	(0.04)*
3,000 - 4,500	0.31	0.26	-0.04	(0.55)
4,500 - 6,000	0.13	0.03	-0.10	(0.02)*
6,000 - 7,500	0.03	0.07	0.03	(0.34)
7,500 - 9,000	0.04	0.05	0.01	(0.81)
9,000 - 10,000	0.01	0.02	0.01	(0.75)
Greater than 10,000	0.09	0.02	-0.08	(0.03)*
Occupation				
Services	0.16	0.18	0.02	(0.78)
Food Services	0.14	0.02	-0.13	(0.00)**
Personal Services	0.02	0.16	0.14	(0.01)**
Sales	0.11	0.15	0.04	(0.53)
Construction & Maintenance	0.18	0.05	-0.13	(0.01)**
Administration	0.05	0.11	0.06	(0.18)
Monitoring and Control				
Years in UAE	7.32	5.95	-1.37	(0.08)
Years since last visit	2.02	1.48	-0.54	(0.04)*
Visits per year	0.73	0.74	0.00	(0.96)
Relatives in UAE	2.40	3.13	0.73	(0.24)
Spouse HH members	3.49	2.97	-0.52	(0.07)
Spouse lives with In-laws	0.27	0.34	0.07	(0.37)
Discuss budget per monthy	0.53	0.61	0.08	(0.31)
Want more control of spending	0.36	0.58	0.22	(0.01)**
Instruct spouse on spending	0.50	0.70	0.20	(0.01)*
Remittance Behavior				
Spouse is main recipient	0.99	1.00	0.01	(0.32)
Other recipients	1.14	1.25	0.10	(0.64)
Remit monthly	0.91	0.91	-0.01	(0.91)
Remittance (AED/month)	1700.63	1173.48	-527.15	(0.00)**
N	98	61		

Notes: Columns (1) and (2) show means for male and female migrants who participated in the study, respectively. Column (3) shows the difference of means between male and female migrant participants. Column (4) shows the p-value from the two-sided t-test of equivalence of means of male and female migrant participants. ***, **, * denote significance at the 1 percent, 5 percent and 10 percent levels

1.10 Appendix

1.10.1 Remittances and Income: Comparative Statics

I present an income-sharing model that adapts the exchange-based model from Rapoport and Docquier (2006) and shows that remittances are increasing in the migrant's income and decreasing in the spouse's income.

Consider two agents, the migrant (M) and their spouse (S). Each agent has their respective pre-transfer income (y) and consumption (c). The migrant sends a remittance (r) to their spouse while in return the spouse provides household or childcare services (h). Each agent derives utility (U_i for i=M, S) from their own consumption with diminishing marginal utility ($u_i'(c_i) > 0$ and $u_i'(c_i) < 0$). Each agent's consumption increases in their income ($c_i'(y_i) > 0$ for i=M, S). The migrant's consumptions decreases, while the spouse's consumption increases with remittances ($c_M'(r) < 0$ and $c_S'(r) > 0$). The migrant also derives increasing utility from the service provided by the spouse ($u_M'(h) > 0$), whereas the spouse experiences dis-utility of effort from providing the service ($u_S'(h) < 0$).

Both the migrant and spouse must accept the terms of the income-sharing contract. Suppose that to accept a contract the spouse requires a minimum compensating utility of ($\bar{U}_S$). This utility level is based on the spouse's expectations of the level consumption they will enjoy and the level of household and childcare services they will have to be provide as a result of the temporary migration of the other spouse ($\bar{U}_S(\bar{c}_S, \bar{h})$). If the migrant remits the minimum amount such that the spouse will accept, the remittance amount must be such that:

$$U_S(c_S(y_S, r), h) \geq \bar{U}_S(\bar{c}_S, \bar{h})$$

Solving this acceptance constraint with equality, r may be expressed as $r = r(y_S)$. The implicit function theorem therefore implies:

$$\frac{\partial r}{\partial y_S} = -\frac{\frac{\partial U_S(c_S, h)}{\partial c_S} . \frac{\partial c_S}{\partial y_S}}{\frac{\partial U_S(c_S, h)}{\partial c_S} . \frac{\partial c_S}{\partial r}} = \frac{\frac{\partial c_S}{\partial y_S}}{\frac{\partial c_S}{\partial r}} < 0$$

The minimum amount of remittances that the spouse is willing to accept decreases with the spouse's income. If the spouse's propensity to consume from income and remittances is the same, i.e. remittances are completely fungible, this becomes a one-to-one relationship. A similar acceptance constraint can de derived for the migrant for the maximum amount they are willing to remit to keep a minimum compensating utility ($\bar{U}_M(\bar{c}_M, \bar{h})$).

$$U_M(c_M(y_M, r), h) \geq \bar{U}_M(\bar{c}_M, \bar{h})$$

Again solving for the comparative static:

$$\frac{\partial r}{\partial y_M} = -\frac{\frac{\partial U_M(c_M, h)}{\partial c_M} . \frac{\partial c_M}{\partial y_M}}{\frac{\partial U_M(c_M, h)}{\partial c_M} . \frac{\partial c_M}{\partial r}} = \frac{\frac{\partial c_M}{\partial y_S}}{\frac{\partial c_M}{\partial r}} > 0$$

As the migrant's consumption decreases with remittances, the maximum amount of remittances that the migrant is willing to send is increasing with the migrant's income.

1.10.2 Subject pool recruitment protocol

The subject pool of migrants was recruited as part of De Arcangelis and Yang (2019). The subject pool is comprised of migrant workers from the Philippines living and working in Dubai, United Arab Emirates (UAE). Migrants were recruited via face-to-face intercepts in locations frequented by Filipino workers in Dubai. Participants had to answer yes to the following screening questions to enroll in the subject pool:

1. Do you expect to continue working in Dubai for the next twelve months?
2. To participate, you will need to download a mobile application we developed called "Padalapp" that allows OFWs (Overseas Filipino Workers) to record and keep track of their remittances. Are you willing to download the smartphone app Padalapp using our pocket wifi?
3. Are you willing to commit to participating for the whole 12-month study period starting from today?
4. Do you agree to complete the weekly one-question surveys for the next 12 months?
5. Do you agree to receive phone calls, SMS, and FB messages from the research team for the next 12 months? We will only contact you for the purpose of facilitating this study.
6. Do you agree for us to invite your household in the Philippines (we will identify the household respondent later in this interview) to also participate in this study?

Individuals answering yes to all the above questions were then asked to sign a consent form to join the subject pool. Participants were administered a short face-to-face baseline survey to collect baseline characteristics of participants and their households in the Philippines.

To identify the relevant remittance-receiving household in the Philippines, participants were asked to name (and provide contact information for) an individual in the Philippines who would be the recipient of a US$500 lottery prize (implemented by the study among subject pool participants). The participants choice identifies an individual (referred to as the target beneficiary) and household (referred to as the target household) in the Philippines whose well-being is important to the participant. Subject pool participants who identified their spouses as either their target beneficiaries or as a member of their target household were invited to participate in the spousal communication experiment.

An overlapping subset of the migrants in the subject pool also participated in the randomized labeled remittances intervention in De Arcangelis and Yang (2019). In my analysis, I control for the migrants' participation and treatment status, conditional on participation, in this intervention.

1.10.3 Experiment Protocol: Scripts

Introduction SMS - Prior to being called for the survey, participants were sent the following text message from a number identified as IPA (Innovations for Poverty Action):

Hello PARTICIPANT_NAME, I am SURVEYOR_NAME, a surveyor from Innovations for Poverty Action. You have been participating in our study about OFWs remittance behaviour in UAE. We would like to invite you to participate in a new survey about the migration experience of OFWs in UAE and their households in the Philippines. The survey will take about 30 min of your time. By participating you will help inform fellow Filipinos about the migration experience and also learn from their experience. Would you be available at DAY and TIME? If so I would call you then and tell you more details about it.

Introduction Call - Surveyors introduced the study using the following script as part of the consent process:

I would like to invite you to participate in a research study on the migration experience of OFWs in UAE and their households in the Philippines. The purpose of this study is to learn about the experience of OFWs and how we can better inform OFWs and their households about the costs and benefits of living and working in the UAE. By taking part in this study you will learn about these important issues and will also be helping inform fellow Filipinos about them.

If you choose to participate, you will be asked to complete a survey that covers your demographic and financial information. This survey will take approximately 15 minutes of your time. We will also call your spouse and invite them to participate in this study. At the end of the study, we will share our results with you and your spouse, which will include information about the average income, expenditures of OFWs in the UAE and their spouses in the Philippines.

Treatment Status - The treatment status was revealed during the survey using the following scripts when the surveyor reached the experimental survey section:

Control Group: Now I would like to ask some questions about your experience in the UAE. As I mentioned, we will be sharing with you and your spouse the summary results from this section. Keep in mind that your individual responses will NOT be shared with your spouse or anyone else. This is a separate activity with each spouse and because of the rules of this activity, we will not share your individual responses to the following questions with your spouse. Your individual responses will be kept private.

Treatment Group: Now I would like to ask some questions about your experience in the UAE. As I mentioned, we will be sharing with you and your spouse the summary results from this section. Keep in mind that your individual responses WILL also be shared with your spouse. This is a joint activity with your spouse and because of the rules of this activity, we will share your individual responses to the following questions with your spouse. Your individual responses will not be private.

Table 1.11: Migrant Baseline Summary Statistics of Participating Spouses: Selection & Balance

	Selection			Balance		
	(1) Invited Sample	(2) Participating Sample	(3) Diff p-val	(4) Treat	(5) Control	(6) Diff p-val
Treatment	0.50	0.49	(0.73)	1.00	0.00	
Spouse participated		0.61		0.58	0.63	(0.50)
Demographics						
Male	0.69	0.65	(0.23)	0.67	0.63	(0.62)
Age	37.45	37.79	(0.49)	37.61	37.96	(0.76)
Children	1.95	1.95	(0.99)	1.88	2.01	(0.47)
Income range (AED/month)						
Less than 1,500	0.05	0.06	(0.54)	0.08	0.04	(0.28)
1,500 - 3,000	0.33	0.31	(0.53)	0.28	0.34	(0.38)
3,000 - 4,500	0.26	0.31	(0.13)	0.32	0.30	(0.87)
4,500 - 6,000	0.10	0.09	(0.50)	0.14	0.04	(0.02)*
6,000 - 7,500	0.05	0.05	(0.69)	0.01	0.08	(0.06)
7,500 - 9,000	0.03	0.05	(0.42)	0.03	0.06	(0.27)
9,000 - 10,000	0.02	0.02	(0.91)	0.01	0.03	(0.58)
Greater than 10,000	0.06	0.06	(0.93)	0.07	0.06	(0.95)
Occupation						
Food & Personal Services	0.20	0.17	(0.19)	0.20	0.14	(0.34)
Sales	0.16	0.14	(0.49)	0.09	0.19	(0.08)
Construction & Maintenance	0.10	0.13	(0.20)	0.12	0.14	(0.70)
Administration	0.09	0.10	(0.55)	0.14	0.06	(0.10)
Communication and Control						
Years in UAE	7.05	6.66	(0.24)	6.37	6.94	(0.45)
Years since last visit		1.70		1.98	1.46	(0.19)
Visits per year		0.68		0.59	0.76	(0.10)
Relatives in UAE		2.53		2.14	2.88	(0.28)
Spouse HH members		3.28		3.32	3.24	(0.82)
Spouse lives with In-laws		0.32		0.33	0.31	(0.83)
Discuss budget per month	1.10	1.05	(0.78)	1.26	0.85	(0.32)
Want more control of spending	0.43	0.43	(0.93)	0.45	0.41	(0.60)
Instruct spouse on spending	0.52	0.55	(0.32)	0.55	0.54	(0.92)
Remittance Behavior						
Spouse is main recipient	0.99	0.99	(0.71)	1.00	0.97	(0.16)
Other recipients	1.06	1.12	(0.51)	0.89	1.34	(0.04)*
Remit monthly	0.90	0.89	(0.55)	0.93	0.86	(0.15)
Remittance (dirham/month)	1,555	1,513	(0.65)	1,421	1,601	(0.36)
N	492	155		76	79	

Notes: Columns (1) and (2) show means for the migrants of all invited spouses and those spouses who participated in the study, respectively. Column (3) shows the p-value from the two-sided t-test of equivalence of means between those who participated and those who were invited but did not participate in the study. Columns (4) and (5) show means within treatment and control groups, respectively. Column (6) shows the p-value from the two-sided t-test of equivalence of means between the treatment and control group.

Table 1.12: Log of Migrant's Reported Income & Expenses

	Log Income		Log Expenses	
	(1)	(2)	(3)	(4)
Response observable to spouse	-0.0294	-0.0285	0.119	0.102
	(0.0876)	(0.0512)	(0.124)	(0.110)
	[0.743]	[0.599]	[0.339]	[0.318]
Migrant is male		0.122**		-0.0235
		(0.0572)		(0.110)
Migrant's monthly remittances to spouse		0.00008*		0.000012
		(0.00005)		(0.00069)
Migrant's remittance: Above median		0.0376		0.0481
		(0.0896)		(0.151)
Migrant wants more control over remittance spending		-0.0466		-0.188*
		(0.0579)		(0.113)
Migrant instructs spouse about remittance spending		0.0385		0.178*
		(0.0615)		(0.101)
Migrant discusses budget with spouse: Frequency above median		0.111*		-0.0412
		(0.0631)		(0.121)
Years since migrant last visited spouse		-0.0168		0.0237
		(0.0186)		(0.0237)
Mean when response not observable to spouse	8.068***		6.814***	
	(0.0664)		(0.0946)	
Observations	155	155	156	156
R-squared	0.001	0.696	0.006	0.489

Notes: Robust standard errors in parentheses. ***, **, * denote significance at the 1 percent, 5 percent and 10 percent levels, respectively. Randomization inference p-values for 5000 replications of the treatment assignment are shown in square brackets. Outcome variable is log of reported monthly income in dirhams in columns (1)-(2) and log of reported monthly expenses in dirhams in columns (3)-(4). Columns (2) and (4) include migrant baseline income category dummies. Monthly remittances are measured in dirham.

CHAPTER II

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

From a work with Joel Slemrod and Mazhar Waseem

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 all taxpayers 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 the public disclosure caused a 9 log-points and the social recognition program a 17 log-points increase in the tax payments of agents exposed to the program. Our results suggest that such programs can be important policy levers to mobilize additional resources.

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

2.1 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 and 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 and 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 the 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 tax compliance through a number of channels. 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 and Kranton, 2000;

[1]Dwenger and 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 and Troiano, 2018).

[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.

Glazer and 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 and 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 parents 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 effects change 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 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, Slemrod and Thoresen (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, Robinson and Slemrod, 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 and 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 and Troiano (2018) 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 and Scartascini (2015) run a similar experiment in Argentina and Kettle et al. (2016) in Guatemala (please see Slemrod, Blumenthal and Christian, 2001, Fellner, Sausgruber and Traxler, 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

[3]Please see Lenter, Slemrod and Shackelford (2003); Blank (2014); Perez-Truglia (2020) for the non-tax effects of public disclosure.

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.

2.2 Context

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

2.2.1 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]

[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 $t+1$.

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 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

[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".

[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.

taxpayer, and therefore there could be no meaningful disclosure without it.[10]

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, 2020). 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.

2.2.2 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 program 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);

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

(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.

2.2.3 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 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 2.1 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 En-

[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."

glish 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.

2.3 Conceptual Framework

2.3.1 Social and Psychological Motivations in Tax Compliance

Economists have traditionally modeled tax evasion as if it were a choice under uncertainty (Allingham and 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

$$\max_{e} \quad (1-p)\,.u\left[(1-\tau)z + \tau e\right] + p.u\left[(1-\tau)z - \theta e\right]. \tag{2.1}$$

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.

$$\frac{u'(c_A)}{u'(c_{NA})} = \frac{(1-p)\,\tau}{p\theta}, \tag{2.2}$$

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. For example, in the United States the noncompliance rate of wage income is estimated to be 1%, whereas that of self-employment income is around 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 and Feinstein, 1994). And, third, there are behavioral biases whereby the detection probability and penalty are systematically mis-estimated by taxpayers (Scholz and 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 that others are paying even less. We investigate, and rule out, such a boomerang effect in our setting in section 2.4.1.

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

[14] For example, in an influential survey of the tax compliance literature, Andreoni, Erard and Feinstein (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."

taxation as a position (Veblen) good, deriving utility from being seen as one of the richest in the country (Akerlof and 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 the program provides.

2.3.2 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.

2.3.2.1 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

$$\log \text{TaxPaid}_{it} = \alpha_i + \beta \, \text{treat}_i \times \text{after}_t + \lambda_t + u_{it}, \tag{2.3}$$

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, Duflo and Mullainathan, 2004).

[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).

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

$$\log \text{TaxPaid}_{it} = \alpha_i + \sum_{j=2007}^{2015} \gamma_j \ \text{treat}_i \times 1.(\text{year}{=}j)_t + \lambda_t + u_{it}. \tag{2.4}$$

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 (2.3) 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 (2.3) 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, 2020). If the public disclosure program curtails tax evasion, the effect would be the strongest for this section of the population.

2.3.2.2 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:

$$\Delta\log \text{TaxPaid}_{it} = \alpha + \beta \ \text{treat}_i \times \text{after}_t + \lambda_t + u_{it}, \tag{2.5}$$

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 (2.5) 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.

2.3.3 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 A.III 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 and 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 A.IV 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 (2.3) 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.

2.4 Effects of the Public Disclosure Program

2.4.1 Intensive Margin

Event Study —Figure 2.2 shows the results from the estimation of equation (2.4). 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.V). 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, Koumpias and Martinez-Vazquez 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 2.1 reports the regression results. We estimate equation (2.3) 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 unrestricted variance-covariance structure at the individual level (Bertrand, Duflo and Mullainathan, 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 A.VI explores

[16] The tax year 2013 in our paper refers to the year that runs from July 2013 to June 2014. Cyan, Koumpias and Martinez-Vazquez (2017) refer to it as the tax year 2014 in their paper.

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 A.VI 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.VII presents the results. We replicate Table 2.1, estimating equation (2.3) 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 2.1-B), we address it more directly in Table A.VIII. 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

[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.

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.IX 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 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 A.III 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 (2.3). 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.X, A.XI, and A.XII with Table 2.1.

Table A.XIII shows the results of our final robustness check. We estimate equation (2.3) 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 A.III 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.XIII 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 and Li, 2009; Kopczuk and Slemrod, 2006; Kleven, Kreiner and Saez, 2016).[18]

[18]Existing empirical results are also consistent with these theoretical models. Waseem (2020), for example,

Sign of the Effect of the Public Disclosure Response—Given the difference-in-differences research design, our estimates in Table 2.1 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 that 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 (2.3) 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 A.XIV. 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 2.1. We include a triple-interaction term in these specifications that captures the additional effect of the program on less-common-named taxpayers. Three results

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 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 (2020), 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.

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 2.1). 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 2.1, as we argued above, have a lower-bound interpretation.

Heterogeneity—Table A.XV estimates a triple-difference version of model (2.3), exploring if the response varies across self-employed taxpayers with the nine traits listed in Table A.III. 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 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 from variation in both taxpayer characteristics and treatment intensity. Our results in Table A.XV show that the most important trait along which the response varies

[20] There is some evidence in literature that men and young are less risk-averse than their counterparts (Borghans et al., 2009; Albert and Duffy, 2012). Similarly, individuals who habitually file their tax returns earlier than others are expected to be more risk averse.

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 A.VI 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.XIV. To this extent, our estimate has a lower bound interpretation.

2.4.2 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 2.3 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 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 2.2. 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 2.2 reports the results from the following regressions

$$\log \mathrm{N}_{gt} = \alpha + \beta \ \mathrm{treat}_g + \gamma \ \mathrm{treat}_g \times \mathrm{after}_t + \lambda_t + u_{gt}, \tag{2.6}$$

where N_{gt} is the log number of filers of group $g \in \{\mathrm{treat}, \mathrm{control}\}$ in year t. Columns (1)-(4) are constructed similarly to the corresponding columns of Table 2.1, while columns (5)-(7) correspond to the three specifications in Figure 2.3B-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 the 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.XVI 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%.

2.5 Effects of the TPHC Program

Figure 2.4 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 2.3 formalizes this analysis. We estimate equation (2.5) 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.XVII 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.XVIII decomposes the aggregate response. We now estimate our baseline specification (2.5) 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.XIX 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 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 and Slemrod, 2009; Bénabou and 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.XX 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.XXI. 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.XIX 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 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

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

consider a wider treatment window containing firms ranked between 50 and 150.

2.6 Conclusion

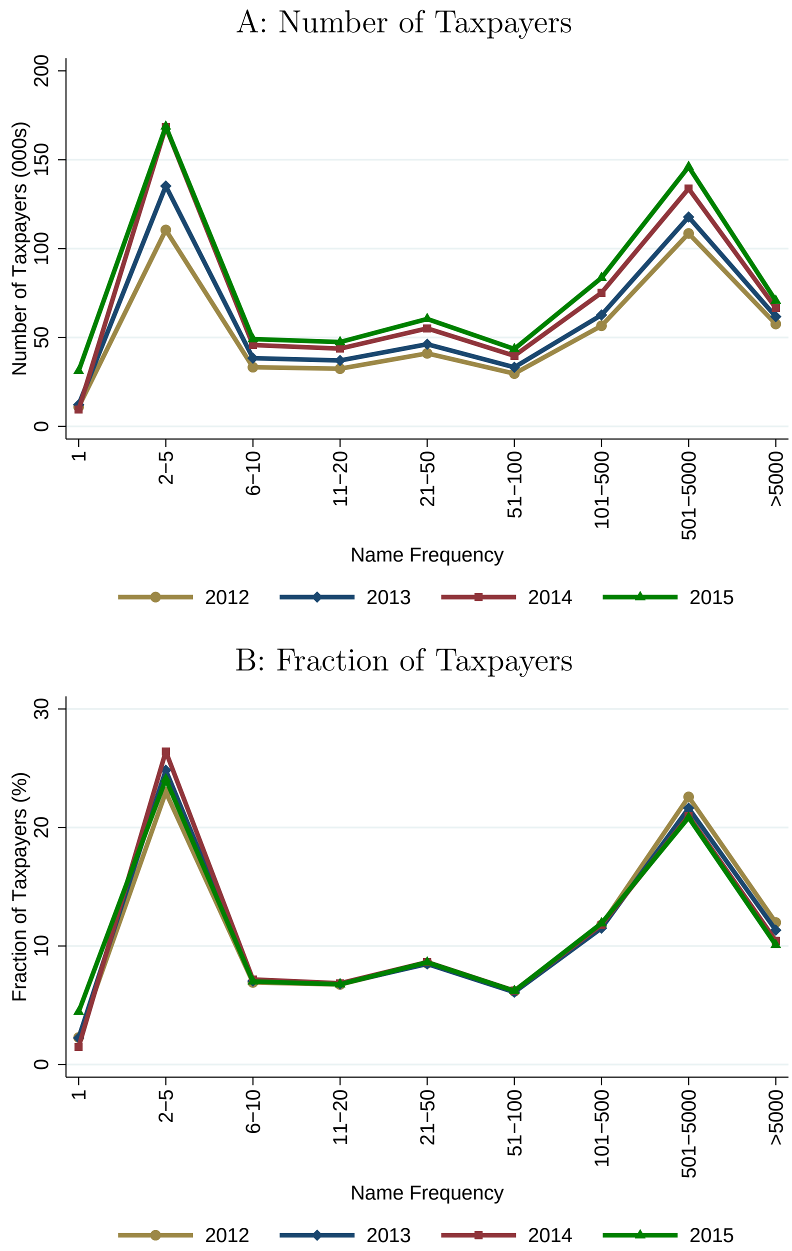
To mobilize resources, countries around the world are increasingly using programs that make tax information public, shame tax delinquents, and positively recognize top taxpayers. We analyze two such Pakistani programs to estimate their impacts on tax compliance and revenue. In the first of these programs, 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 noncompliance.

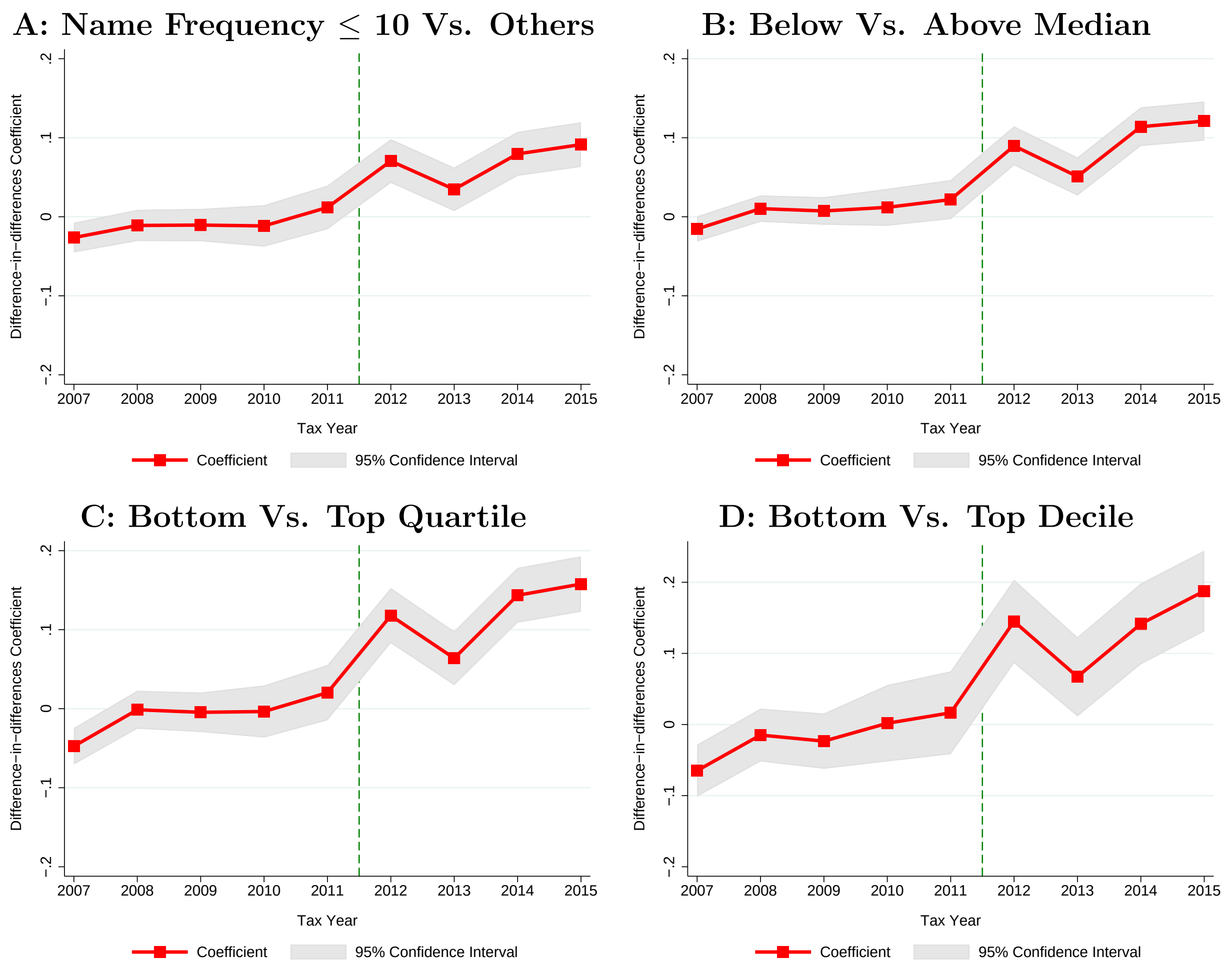
2.7 Figures & Tables

Figure 2.1: Distribution of Names



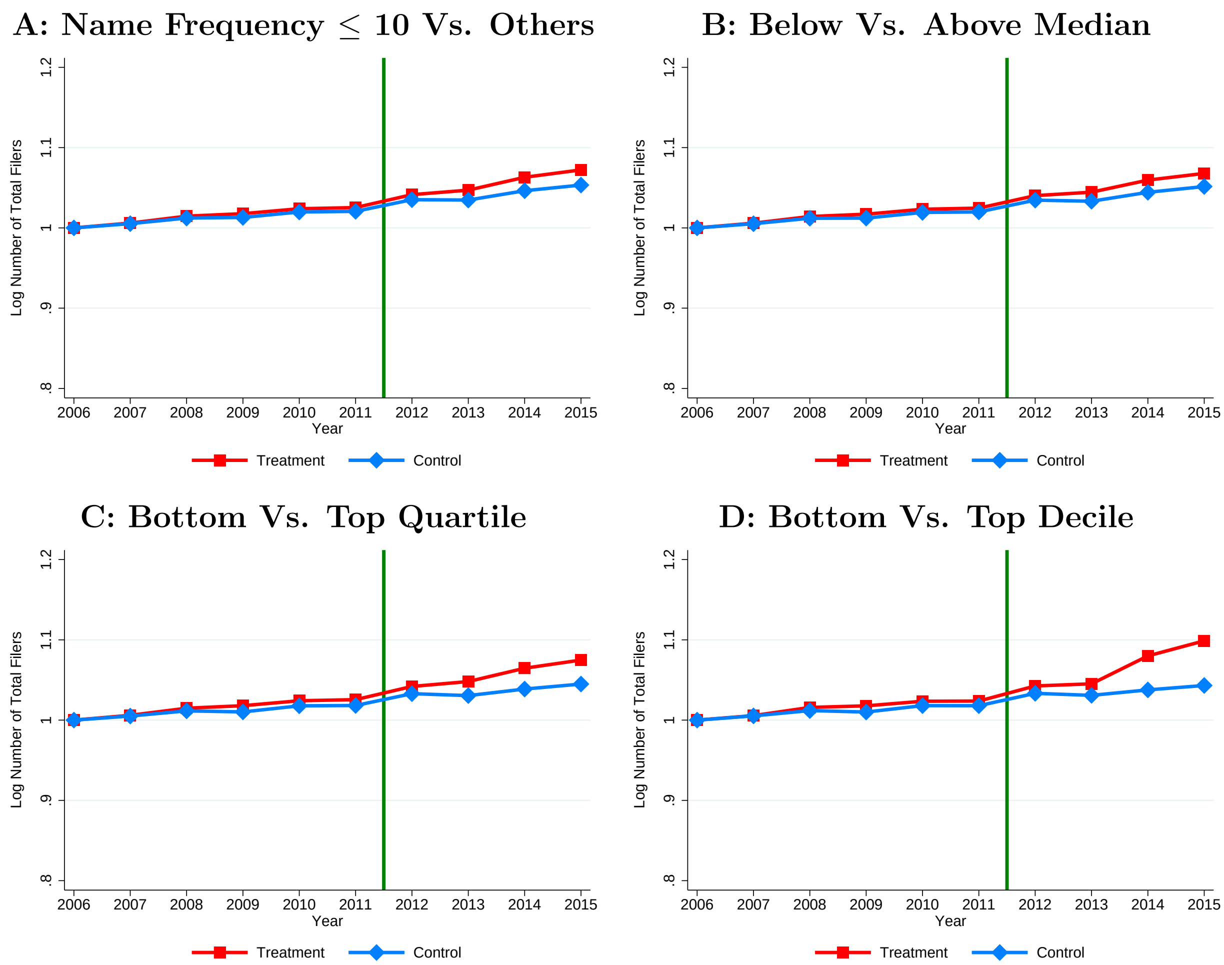
Notes: The figure illustrates the distribution of full names in Pakistan. We define Name Frequency as the number of times a full name appears in the disclosure data for the years 2012-2015. The two panels plot the distribution of the variable. Each marker in panel A denotes the number of individuals 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.

Figure 2.2: Intensive Margin Response to the Public Disclosure Program



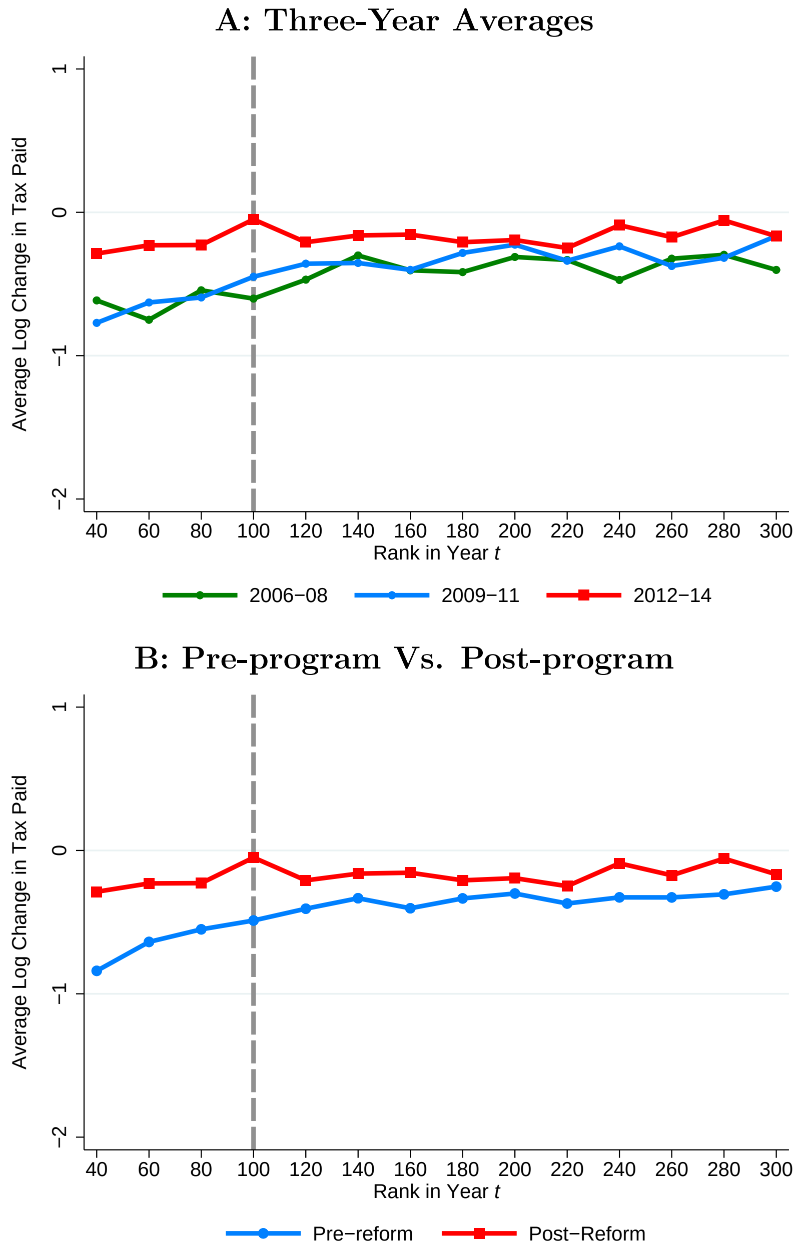
Notes: The figure plots the coefficients γ_js and 95% confidence interval around them from the event study equation (2.4). We estimate the equation on a balanced panel sample of self-employed taxpayers, who file in all years from 2006 to 2015. The definitions of the treatment and control groups are provided 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 2.3: 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 2.4: 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 2.1: 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 (2.3). 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 2.2: 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 (2.6). 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 2.3: 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 (2.5). 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.

2.8 Appendix

2.8.1 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.

2.8.2 Definition of Variables

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

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

3. **Multiple businesses.** The taxpayer owns more than one businesses.

4. **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.

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

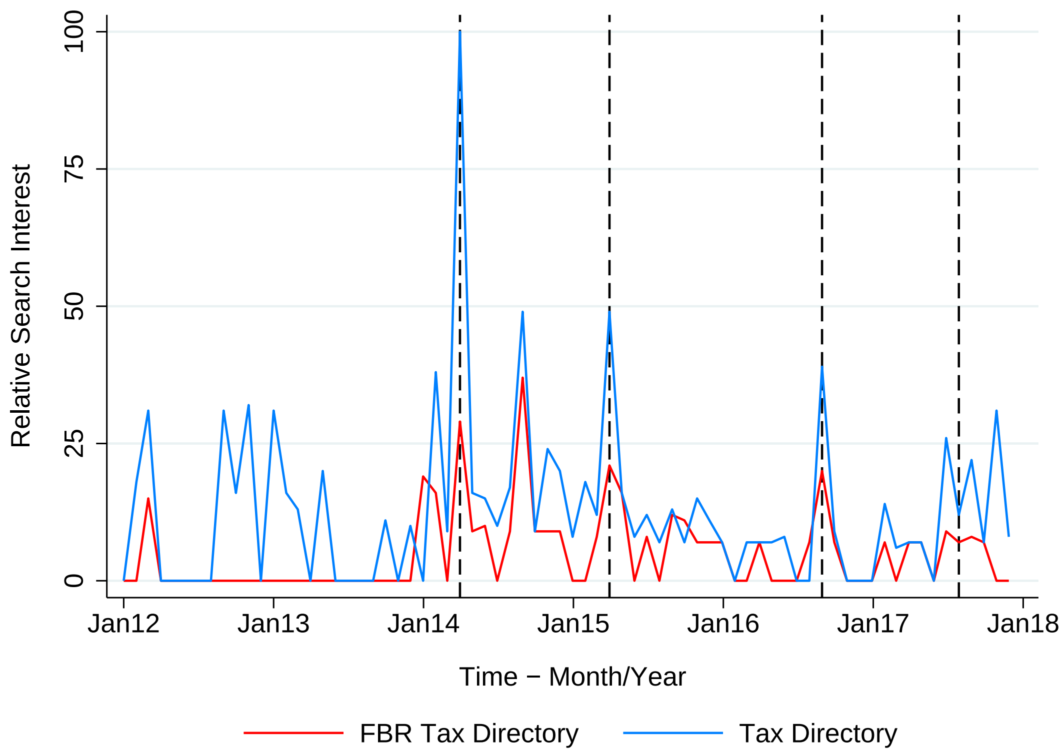
6. **Buncher.** If the taxpayer reported income at or within a window of ten thousand PKR below any notch in the corresponding tax schedule.

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

8. **Revised return.** If the taxpayer filed a revised return for the given tax year t.

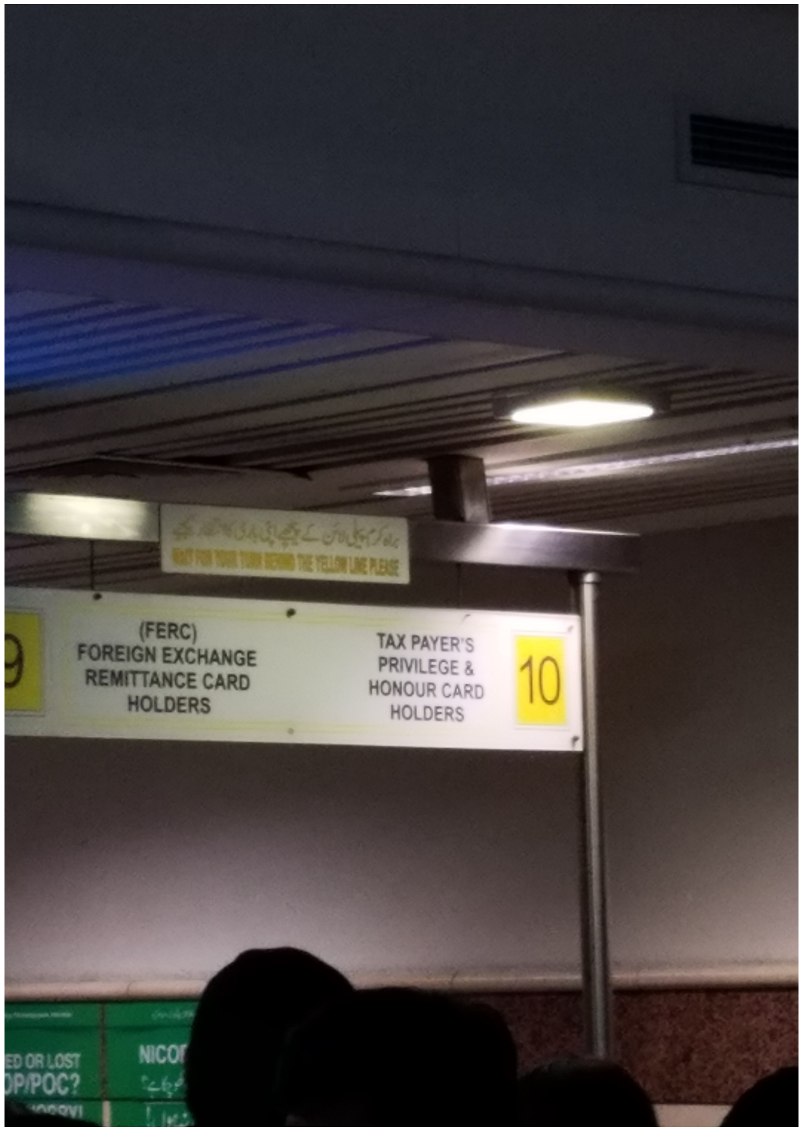
2.8.3 Appendix: Figures & Tables

Figure 2.5: Google Search Interest



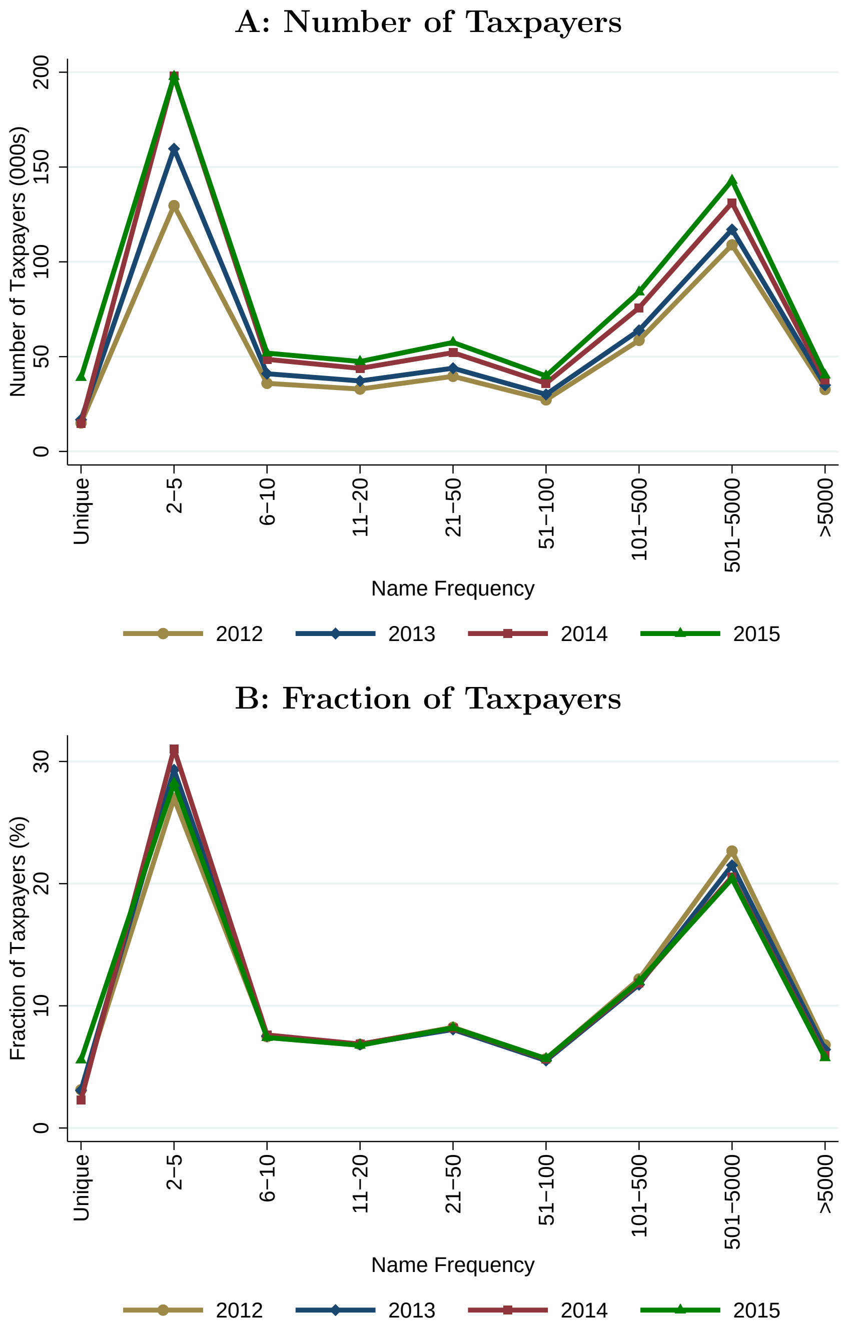
Notes: The figure plots Google Trends data for the monthly search interest in Pakistan for the terms "FBR Tax Directory" and "Tax Directory" from January 2012 to January 2018. The data is normalized by time and location and scaled on a range of 0 - 100 to compare relative popularity. The data point with 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 2.6: Special Immigration Counter for TPHC Holders



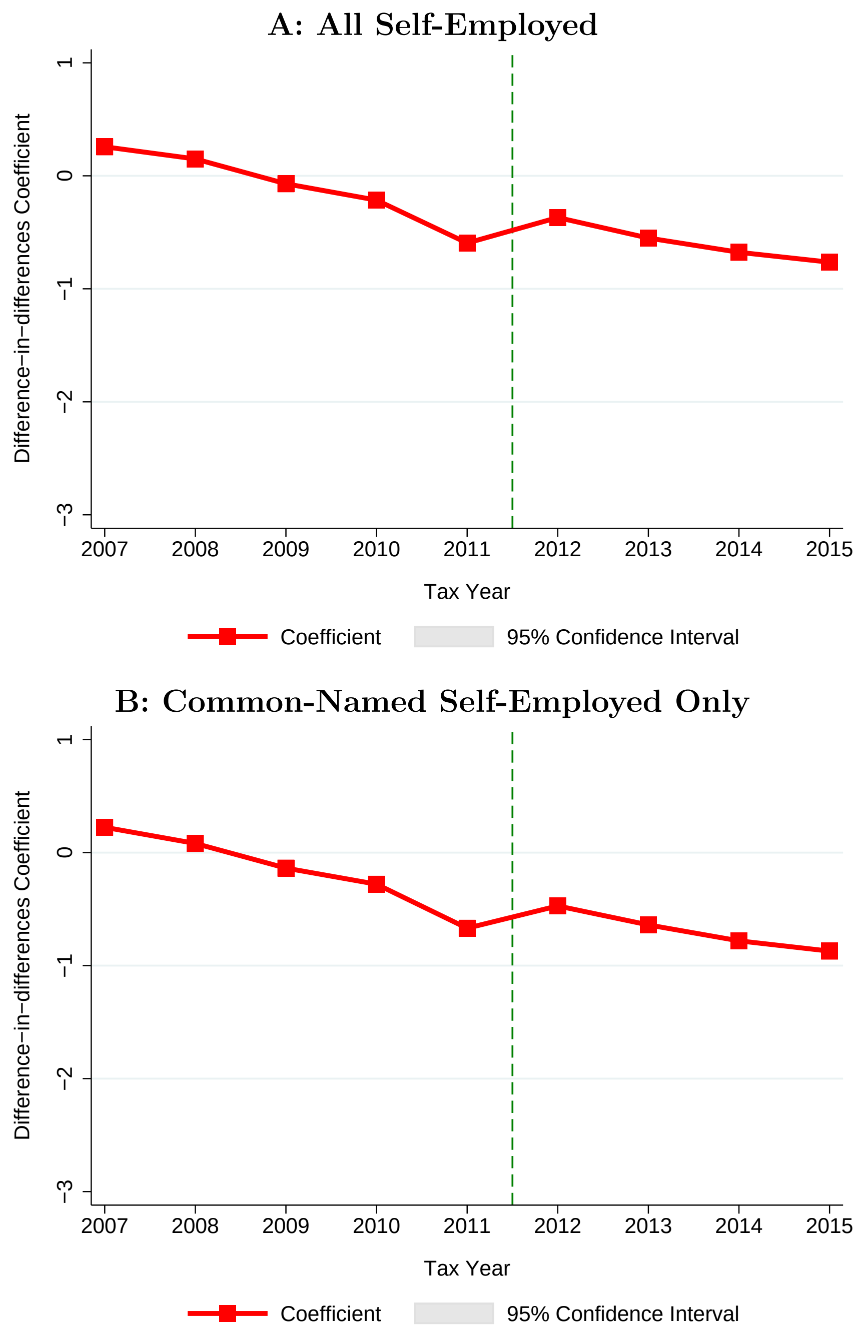
Notes: The figure shows the picture of special immigration counter at the Allama Iqbal International Airport, Lahore. The picture was taken in the summer of 2018.

Figure 2.7: 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.

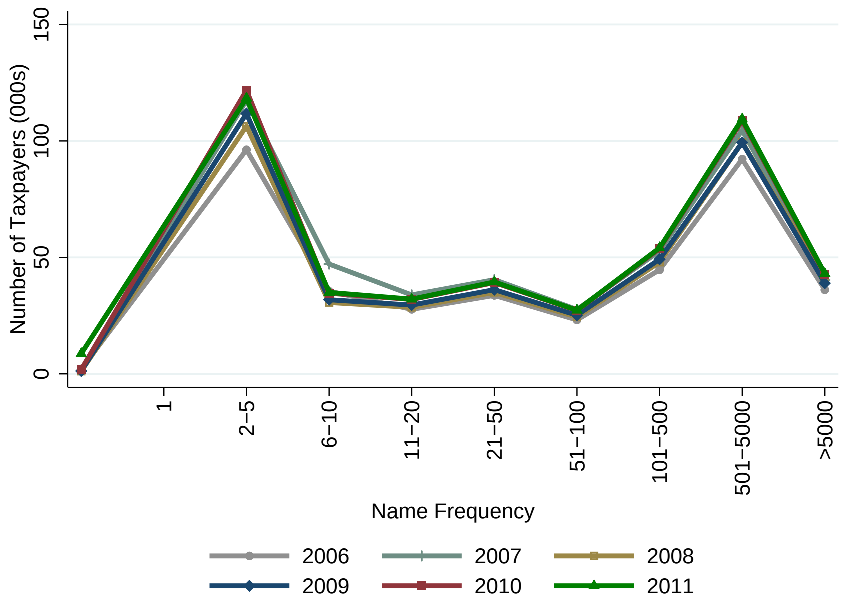
Figure 2.8: Evolution of Tax Payments – Self-Employed Vs. Wage-Earners



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 2.9: Distribution of Names – Pre-Program Years (2006-2011)

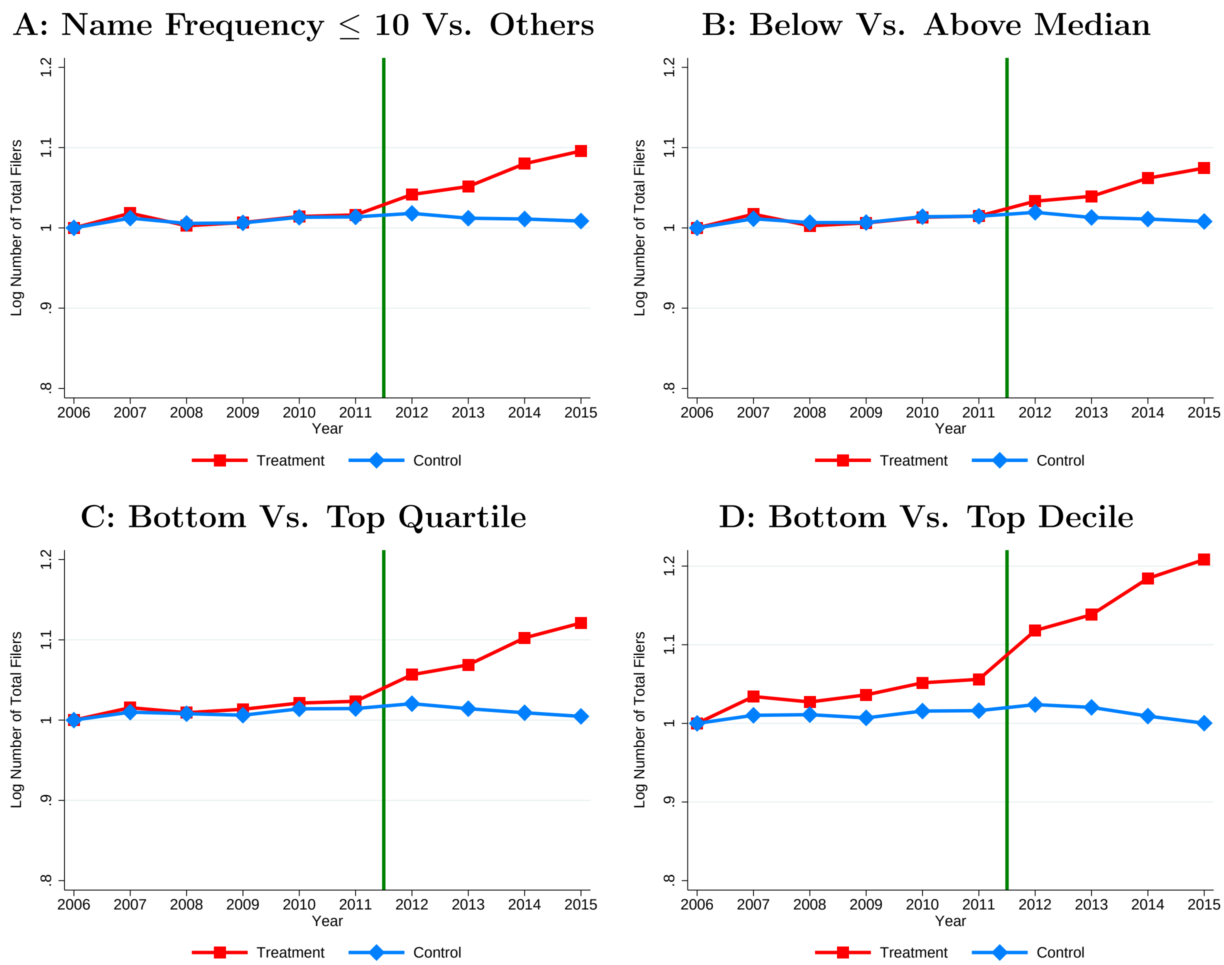
A: Number of Taxpayers



B: Fraction of Taxpayers

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 2.10: 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 2.1: 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 2.2: 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 2.3: 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 2.8.2 of the paper.

Table 2.4: 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 2.3, and defined in 2.8.2. 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 2.5: 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 × 2007	-0.026	-0.015	-0.047	-0.065
	(0.010)	(0.008)	(0.012)	(0.019)
treat × 2008	-0.011	0.010	-0.001	-0.015
	(0.010)	(0.009)	(0.012)	(0.019)
treat × 2009	-0.011	0.007	-0.005	-0.023
	(0.010)	(0.009)	(0.013)	(0.020)
treat × 2010	-0.012	0.012	-0.004	0.002
	(0.013)	(0.012)	(0.017)	(0.027)
treat × 2011	0.012	0.022	0.020	0.017
	(0.014)	(0.013)	(0.018)	(0.030)
treat × 2012	0.071	0.090	0.118	0.145
	(0.014)	(0.013)	(0.018)	(0.030)
treat × 2013	0.035	0.051	0.064	0.067
	(0.014)	(0.012)	(0.017)	(0.028)
treat × 2014	0.080	0.114	0.144	0.141
	(0.014)	(0.013)	(0.018)	(0.029)
treat × 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 2.6: 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 2.7: 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 I 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 2.8: 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 × 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 × 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 I 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 × 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 I. 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 2.9: 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 I 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 I.

Table 2.10: 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 I 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 2.11: Intensive Margin Response to the Public Disclosure Program – With Age × 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	0.079	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 × 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 × 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 I 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 2.12: 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 I 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 2.13: 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 I 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 2.14: 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 × 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 × after × 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 × 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 × after × 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 I, 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 2.15: 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 × 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 × trait × 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 I. 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 × 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 2.16: Extensive Margin Response to the Public Disclosure Program – Baseline Frequency

	Treat: Name Frequency						
	$\leq$ 10	$\leq$ 20	$\leq$ 30	$\leq$ 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 II 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 II 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 2.17: 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 $\times$ 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 $\times$ 1.(year $\in$ {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 $\times$ 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 III, 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 2.18: 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 $\in$ {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 2.19: 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 2.18. 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 2.20: 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 2.8.2 and Table 2.19.

Table 2.21: Response to the TPHC Program – Robustness

	(1)	(2)	(3)	(4)	(5)	(6)
A: Main Regression (2006-2014)						
treat × 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 × 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 × 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 2.8.2 and Table 2.19. Columns (1) reports results from the baseline specification. It is the same as column (7) of Table 2.18. The other columns add interaction terms to this baseline specification. Standard errors are in parenthesis, which have been clustered at the firm level.

CHAPTER III

International Student Demand for Higher Education in the US

Abstract

International students in the US are a mobile, high-skilled, and politically contentious migrant group. Their incentives significantly differ from other migrants, yet there is little analysis on the determinants of international student demand for higher education in the US. I use exchange rate variations to analyze the effect of competing educational attainment and employment incentives on the stock and flow of international students to the US. A depreciation of the home currency reduces educational incentives by making US education relatively more expensive but increases employment incentives by making US income relatively more valuable. I find that the cost of education effect dominates. A one percent depreciation of the home currency reduces the stock and flow of international students from that country by 0.11 percent and 0.25 percent, respectively. This response is driven entirely by international undergraduate students, whereas international graduate student enrollment does not respond to exchange rate variations.

JEL Classification: F22, I21

Keywords Intentional students, exchange rates

3.1 Introduction

Studies on migration often rely on the comparison of binary groups: migrants versus non-migrants; foreign-born versus local-born; citizens versus non-citizens. However, these broad categories may include diverse subgroups that have different decision and incentive structures. Migrants broadly defined can include foreign-born naturalized citizens, permanent residents, temporary workers, students, and refugees. It is important to parse through these subgroups and study them individually, to better understand their incentives and better inform immigration policy. This paper studies an important subgroup of temporary migrants—international students.

International students are a large, mobile, high-skilled, and politically contentious migrant group. In 2015 there were around one million international students in the United States, the largest this group has ever been. This was also a 10 percent increase from the previous year, making it the highest growth rate in the international student population in the US in the past 35 years (Institute of International Education, 2015). Compared to other migrant categories international students are highly skilled; they either have a college degree (graduate students) or are in the process of acquiring one (undergraduate students). The decisions of these students have important implications for debates in the migration and labor literature: the impact of shocks in the US or the home country on migrant populations, the home country brain gain/brain drain effect of high skilled emigration, the revenue and student composition effects on US universities and the positive productivity or negative labor market effects of high skilled immigration to the US. International students are also politically contentious because immigration critics argue that these students primarily migrate to the US for employment and use their student status to bypass restrictions on employment-based migration. Understanding the factors that affect the decisions of international students is therefore central to understanding these debates.

This paper studies the international student demand for higher education in the US. I analyze the effect of changes in relative price and home country income levels on the stock and flow of international students coming to the US and how this effect differs by academic level. I use exchange rate variations to analyze the competing educational attainment and employment incentives of international students. A depreciation of the home currency reduces educational incentives by making US education relatively more expensive but increases employment incentives by making US income relatively more valuable. In addition, I use variations in GDP per capita to study the effect of changes in home country income levels.

I use student visa data from the US State Department and survey data from the Institute of International Education (IIE) to create a panel of international student stocks and flows from 1996-2014. I find that a 1 percent increase in the exchange rate (a depreciation of home currency) leads to a 0.25 percent decrease in the flow of international students and a 0.11 percent decrease in the stock of international students from that country. The higher relative cost of tuition and living expenses as a result of the depreciation decreases the number of international students coming to the US. This response is driven entirely by international undergraduate students, whereas international graduate student enrollment does not respond to exchange rate variations. A 1 percent increase in the GDP per capita leads to a 0.32 percent increase in the flow of international students from that country. This suggests that US education for international students is a normal good with demand increasing as income levels increase. I analyze the interaction of GDP and exchange rate shocks and find that student flows from lower-income countries are more sensitive to exchange rate shocks. This suggests that binding budget and credit constraints are an important factor of international student demand for US education.

This paper builds on the literature that studies the impact of exchange rate and income shocks on migrant outcomes. McKenzie, Theoharides and Yang (2014) study the impact of destination country GDP shocks on migrant flows and wages. Yang (2006), Yang (2008), Nekoei (2013), and Abarcar (2017) use exchange rate variations to identify effects on return migration, remittances and labor migrant outcomes. Yang (2006), Yang (2008) and Abarcar (2017) use exchange rate variations generated by the Asian financial crises, whereas Nekoei (2013), similar to this paper, uses exchange rate variations over a longer time horizon. Although this literature analyzes a range of migrant outcomes, the results in each case are primarily driven by the relative income benefits that migrants receive from GDP shocks or home country currency depreciations (relative to the destination/host country currency). Yang (2006) and Abarcar (2017) find reduced returned migration, Yang (2008) finds higher remittances sent home, and Nekoei (2013) finds a decrease in labor supply as a result of the depreciation of the home currency. In contrast to this literature, I find that for international students the cost effect of a depreciation of the home currency dominates. A depreciation in the home currency reduces the number of international students coming to the US due to higher tuition costs and tighter budget constraints. This result highlights how international students as a subgroup differ significantly from other migrant groups and therefore warrant specific attention.

There is a growing literature on the incentives of international students to study in the US. Rozensweig (2006) proposes two competing models to explain international student demand for education; a constrained domestic schooling model that is based on the lack of schooling options in the home country and a migration model that is based on the increased probability of international students finding employment in the US. In the paper closest to this work, Bird and Turner (2014) look at the effects of costs, home country opportunities, and incomes on international student enrollment and show that Rozensweig's two models are not mutually exclusive. I build on this work by developing a model that incorporates both domestic schooling and migration considerations. I use the predictions from the model to explain my empirical findings.

Other work on international students focuses on the *effects* of international student enrollment. Bound et al. (2020) study the relation between university funding and international student enrollment. They find that a decline in state appropriations to public universities is associated with an increase in international enrollment showing that US universities may rely on international enrollment to supplement their funding. Several papers study at the effect of international student enrollment on domestic enrollment (Shih (2017) Shen (2016)) A common theme that emerges from the literature is that international students differ significantly by academic level and program type and results are often not generalizable to all international students. I also find a large difference in the response of undergraduate and graduate students. The effect of exchange rates is driven entirely by undergraduate students whereas graduate enrollment does not respond to exchange rate variation. I find suggestive evidence that this difference is driven by lower out-of-pocket costs for graduate students.

The remainder of the paper is structured as follows: Section 2 describes the conceptual framework to understand the decision of international students to study in the US. Section 3 describes the data and empirical methodology. Results are presented in Section 4, followed by robustness checks in Section 5. Section 6 concludes.

3.2 Conceptual Framework

In this section, I present a conceptual framework of the decision of a potential international student to study in the US. The framework analyzes how international students' demand to study in the US is affected by incomes and relative price. The student chooses to study in the US if it generates a positive net return and if they can afford the cost of the degree. These two conditions can be expressed as follows:

$$\Big(pY_{US} + (1-p)Y_H\Big) - \Big(\tau_{US} + c_{US,H} + \phi_H\Big) > 0 \tag{3.1}$$

$$\alpha_H \geq \tau_{US} \tag{3.2}$$

Keeping track of currencies is important in this framework as changes in relative price, are induced by changes in the exchange rate. The subscripts denote the currency in which each component is measured; US and H represent US dollars and the student's home currency, respectively.

Equation 3.1 states that the net return from higher education in the US should be positive. The first group of terms in equation 3.1 are the benefits of studying in the US; future incomes Y, earned either in the US or in the home country. They are weighed by p, the probability of finding employment in the US. This probability can be modeled as a function of various parameters, for example, a student's preferences or ability. However for simplicity, as it does not affect the analysis here, I do not expand on the functional form or parameters of p.[1]

The second group of terms in equation 3.1 are the costs of studying in the US divided into three categories; tuition costs, moving and living costs, and opportunity costs. τ_{US} is a student's direct out-of-pocket tuition cost of attending school in the US. c represents all other costs associated with studying in the US, such as moving and living expenses. Finally, ϕ is the opportunity cost of studying in the US and depends on the student's outside option. If the outside option is employment in the home country, the opportunity cost is forgone income, whereas if the outside option is higher education in the home country, the opportunity cost is foregone benefits of this education.[23] Considering both the costs and benefits, students will choose to study in the US if the net return from this decision is positive.

The second condition is that students can afford a US education. Equation 3.2 is the student's budget constraint. α represents their (or their family's) total wealth/savings. The constraint implies that a student must be able to pay all associated out-of-pocket tuition costs upfront. Said differently, students face a borrowing constraint where they are not allowed to borrow for out-of-pocket tuition expenses. If the conditions in both equations 3.1 and 3.2 are met students will choose to study in the US.

This conceptual framework generates testable predictions about international student enrollment. First, for any given level of future returns, students that face higher out-of-pocket tuition costs will be less likely to study in the US. Higher tuition will decrease the net return from education and more importantly, higher tuition costs will result in a tighter budget constraint making it more likely that US higher education is not feasible. Second, students with higher family wealth/assets/savings are more likely to study in the

[1]By modeling p as a function of education in the US, Rosenzweig's migration model, where studying in the US increases the probability of finding future employment in the US, can be incorporated into this framework.

[2]The constrained domestic schooling model from Rozensweig (2006) can be incorporated in the opportunity cost. Students from countries with a limited number of high-quality higher education institutes will have a lower opportunity cost of attending school in the US as opposed to those from countries that have many high-quality universities.

[3]Other outside options, such as international education in other countries may also be incorporated in the framework. This would only change the relevant exchange rate to consider for the analysis

US. Again this effect operates through a looser budget constraint. Third, students with higher opportunity costs will also be less likely to study in the US. A high opportunity cost may be in the form of better employment or educational opportunities in the home country.

Finally, we can also analyze the effect of a change in the exchange rate. A change in the exchange rate e, changes the value of all the components in our framework that are measured in dollars (denoted by the subscript US) as follows:

$$p\frac{\partial Y_{US}}{\partial e} - \Big(\frac{\partial \tau_{US}}{\partial e} + \frac{\partial C_{US}}{\partial e}\Big)$$

An increase in the exchange rate (a depreciation of the home currency against the dollar) has three effects. First, the cost of education in the US increases, as the relative price of tuition and living expenses increases due to the stronger dollar. Secondly, the budget constraint faced by students becomes tighter, again the relative price of tuition has increased whereas the value of a student's assets/savings has not changed. Lastly, the value of potential future income in the US increases. This relative income effect is important because it is the driving force in all the results related to the impact of exchange rates on migrants. For international students, however, this relative income effect will be relatively small. Firstly because it is discounted by p the probability of finding employment in the US. Secondly, the potential income will also only be realized many years in the future, adding to the uncertainty and further reducing its value. The overall result is that for international students, unlike other migrants, a depreciation of the home currency will on net be a negative shock.

3.3 Data & Empirical Methodology

To test the empirical predictions from the conceptual framework presented above, I analyze both the stock and flow of international students in the US. I use administrative student (F1) visa data from the US State Department to measure the flow of international students to the US. The State Department reports the total number of student (F1) visas issued to citizens of each country, in each fiscal year. All international students are required to have a valid student visa to enter the US. To obtain a visa, the student must first accept an admission offer at an accredited educational institution. The institution confirms the student's enrollment to the State Department by issuing the student an I-20 form. This form must then be presented by the student to a US consulate or embassy when they apply for a student visa.[4]. To obtain the visa students must also show that they have the required

[4]Canadian citizens do not require a stamped visa in their passport. However, their status as international students is established and recorded by the State Department when they present their I-20 form at a US port of entry

resources and funding to pay for the program in which they have been admitted. Insufficient (proof of) funds is grounds for visa denial. Each issued student visa, therefore, represents an international student enrolled at a US university in a given year.[5]

Using administrative visa data to measure international student flows has two main advantages over survey-based measures - first, visa statistics represent the universe of international students coming to the US, and second, the student status of these individuals is directly and accurately observed. In contrast, survey-based measures of international student flows often rely on demographic information to infer immigration status. International students, for example, maybe identified using their place of birth.This strategy however may produce biased estimates as foreign-born individuals may be naturalized US citizens or permanent residents.

A drawback of the administrative visa data is that it does not identify students by their degree type or academic level. Based on the conceptual framework and the literature, the response of international students to relative price variations will vary significantly by their academic level. To study the difference in response of undergraduate and graduate students I complement the visa data, with data from the Open Doors survey conducted by the Institute of International Education (IIE). IIE has been conducting this census of international students in the United States since 1948. In the fall of each academic year, IIE surveys campus officials in over 2500 regionally accredited institutions of higher education in the US. Institutions report the number of international students currently enrolled by their country of origin and academic level (Institute of International Education, 2015).

The IIE data, which I refer to as my total enrollment measure, represents the total stock of international students in the US. Although this data is not the universe of international students studying in the US, it provides the most comprehensive available statistics on international students by country of origin and academic level. Enrollment is divided into three categories; undergraduate, graduate, and others. Undergraduates include students pursuing a bachelor's degree. Graduate students include all masters, (MBA, LLM, etc.) doctoral, and professional degree (JD, MD) students, while the "other" category includes non-degree and English language students. The "other" category accounts for a very small percentage of total enrollment and is excluded from the analysis. These academic level categories, although insightful, may still include diverse subgroups. For example doctorate, master's, and MBA students are grouped in the graduate student category but they have

[5]In certain cases, an F1 visa may not represent a *new* student. A continuing student may be issued a new visa if, the student is still enrolled in an academic program, their visa has expired, they have traveled outside the US and need to return. Student visa durations are determined by reciprocity agreements with foreign governments. Students from most countries are granted 5-year student visas, which means that most student visas remain valid throughout a student's degree program.

very different incentive structures. Doctoral students have longer degrees but better funding opportunities which translate to lower out-of-pocket expenses. Masters and MBA students in contrast have shorter degrees but limited funding options. In addition, the opportunity costs of students looking to pursue each of these degrees will also vary significantly.

To analyze international student demand I set up a panel of students from 90 countries from 1996 to 2014. These countries account for at least 95% of international students in the US in any given year.[6] Table 3.1 shows descriptive statistics for the panel. The average total enrollment by country and year is around 6,700 students and the average undergraduate and graduate enrollment are both around 2,900.[7] The average number of visas issued is around 3,600. There is huge variation in enrollment with each measure having a standard deviation much larger than its mean. The maximum for each measure also shows that the distribution has a very extended right tail. Chinese students in 2014 account for the maximum value of each measure; the number of student visas, total enrollment, undergraduate and graduate enrollment.

The Open Doors survey also records the primary source of funding for all students. For students that have multiple sources, it records the source that has the largest share. Although the information is not split by country of origin, it still provides some insight into how students vary by academic level. Table 3.2 shows the primary source of funding for international students by academic level in our data. 80 percent of undergraduate students are financed by personal and family funds. In contrast, less than 50 percent of graduate students are primarily funded by personal funds, although it is still the largest funding source for them. The difference between these two groups is mainly driven by US college and university funding for graduate students. This matches the observation that universities generally allocate a higher proportion of their funds toward graduate programs. Combined, these two sources contribute around 85 percent of the funding for both graduate and undergraduate students. The funding distribution suggests that undergraduates who are more often self-funded would face higher out-of-pocket tuition costs.

The overall trend of international student flows and stocks is shown in Figure 3.1. Both the total enrollment and the number of visas issued each year have been steadily increasing over the past decade. These trends are not unique to this period. International student enrollment has been steadily growing in the US over the last half-century. In 2015 there were around 1 million international students in the US (Institute of International Education, 2015).

[6] Some countries had to be dropped due to missing data on exchange rates and other controls

[7] Undergraduate and graduate categories do not add up to total enrollment because of non-degree and English language students

Although the two measures trend similarly, the effect on total enrollment lags the number of visas issued by a year. This is because of the timing of students obtaining visas and starting their degrees. Visas are counted by fiscal year (October of the year to the following September), and enrollment is measured at the start of the fall semester. This implies that in the data, students starting school in a particular academic year are receiving visas in the previous fiscal year. 2001 for example saw a large decline in the number of student visas issued after 9/11. These represent visas issued from 1st October 2001 to 30th September 2002. The effect of this on total enrollment shows up in 2002, in the form of a leveling off of total enrollment in the fall of 2002.

The total enrollment line can be broken down into undergraduate and graduate enrollment. Those trends are shown in Figure 3.2. Undergraduate and graduate enrollment have been similar throughout the decade, both growing steadily. Undergraduate enrollment saw larger declines in the 2000s but it has also grown faster since then. The net result is that the relative size of undergraduate and graduate enrollment is similar at both the start and end of our panel.

These similar and stable overall trends hide significant heterogeneity at the country level. To illustrate this Figure 3.3 shows the trends in undergraduate and graduate enrollment of four of the largest international student origin countries. Among these enrollment of Chinese and Indian students has seen very large increases in the past decade. The growth in Chinese students has been at both the undergraduate and graduate level, in contrast, the growth of Indian students has been almost entirely driven by graduate students. The bottom panels include Japan and Germany, countries that have seen a decline in student enrollment. Again the decline in German students has been at both the graduate and undergraduate level whereas the decline in Japanese students has been mostly driven by undergraduate enrollment. These examples show that at the country level, there has been a lot of variation in demand for US education, which I analyze in the following sections.

3.3.1 Specification

I run the following empirical specification:

$$ln(STUDENTS)_{it} = \alpha + \beta ln(ER)_{it} + \lambda ln(GDP)_{it} + \psi X_{it} + \delta_i + \gamma_t + \epsilon_{it}$$

The outcome $ln(STUDENTS)$ is the natural log of either the flow (student visas) or the stock (total enrollment) of international students from country i in year t. The two main explanatory variables are the exchange rate and GDP per capita. Both directly come out of the conceptual framework developed earlier. ER is the nominal exchange rate in terms of the currency of country i per US dollar. An increase in ER, therefore, represents a depreciation of the home currency and an appreciation of the dollar. The exchange rate is measured as the annual average exchange rate of each fiscal year, to match the visa and student enrollment data time-lines. GDP is GDP per capita measured in the home currency. X is a vector of time and country varying controls. Exchange rate, GDP, and other macroeconomic controls are obtained from the World Bank, World Development Indicators database. Country and time fixed effects are included to account for country and time-invariant characteristics respectively. All variables are included as natural logs, allowing the coefficients to be interpreted as elasticities. Correct identification of the coefficients in this model requires the error to be uncorrelated with the regressors. Said differently, this requires that any country and time-varying factors that affect international student enrollment and are correlated with the exchange rate and/or GDP per capita should be controlled for.

Exchange rates are notoriously difficult to predict in the short run. They deviate largely from expectations based on market fundamentals and are often best modeled as random walks. For this reason, the effect of exchange rate variations is often well-identified in such settings, which is also why they have been extensively used in the literature. Despite this, there may be concerns that exchange rates and GDP per capita are merely proxies for other underlying macroeconomic changes that could affect students' decisions to study in the US. To check for this I control for other macroeconomic indicators such as inflation and unemployment and analyze if their addition has any effect on the coefficients of exchange rate and GDP per capita.

International politics and US foreign policy are also important determinants of who gets student visas to study in the US. Visa restrictions may be an important additional constraint for potential students. Variation in the flow of international students may be a function of varying visa policies. If these are correlated with other determinants of international student demand our estimates will be biased. To address this concern I add a time and country-specific "Visa Regime" control. This is the log of all non-student visas issued to a country in a given year. It is a measure of the total flow of people between the two countries and thus captures other determinants of getting a visa, including political and social conditions.

3.4 Results

Table 3.3 shows the results for the flow of international students. The dependent variable is the natural log of the number of student visas issued. Regressing student flows on only the exchange rate generates a small insignificant coefficient as shown in column 1. The coefficient on the exchange rate increases in size and becomes significant when GDP per capita is added to the regression in column 2. This shows that the exchange rate and GDP per capita, both of which affect international student flows, are positively correlated. The correct specification should therefore include both these explanatory variables. The signs of these coefficients are as expected from our conceptual framework. A 1 percent increase in the exchange rate (a depreciation of home currency) leads to a 0.25 percent decrease in the flow of international students from that country. The higher relative price of tuition and living expenses decreases the number of international students coming to the US. Whereas a 1 percent increase in the GDP per capita leads to a 0.32 percent increase in the flow of international students. This suggests two mechanisms; US education for international students is a normal good with demand increasing as income levels increase, or this may be evidence of budget and borrowing constrained potential students. Higher income levels loosen the budget constraint increasing the number of international students for whom studying in the US is now feasible.

The Visa Regime control is added in column three. The coefficient of Visa Regime is positive, significant, and larger than the coefficients for both GDP per capita and exchange rates. This is evidence that visa policy does play a role in determining where international students come from. Adding Visa Regime as a control does not affect the coefficient on the exchange rate, the GDP coefficient however drops to about half its size. This means that our Visa Regime variable is correlated with GDP per capita. Visa Regime is thus capturing a mix of visa policy and other factors promoting the flow of people between countries such as income levels. Individuals from richer countries can afford more international travel; confounding Visa Regime and income levels. Despite this, it is still informative to add Visa Regime as a control to be convinced that our results are not being driven by visa policy.

In the fourth column, I add other macroeconomic controls including inflation and unemployment. These are included to ensure that the effects we are attributing to the exchange rate and GDP per capita are not being driven by other macroeconomic movements in the economy. Adding these controls does not significantly change any of our coefficients of interest, which supports our identification.

This specification can also be used to analyze the results for total international student enrollment as well as undergraduate and graduate enrollment separately. These results are shown in Table 3.4. The first column is the same as the fourth column in table 3.3 and is presented here for comparison. In the second column I run the same specification on log total enrollment. I find that all estimates have the same sign but are smaller in magnitude. This implies that the stock measure is less responsive than the flow measure to both exchange rates and GDP per capita. Change in the flow measure represents an increase or decrease in new international students enrolling in US universities. In contrast change in the stock measure of total enrollment may either be caused by changes in the flow of new students, or some response from international students already enrolled in US universities. The only margins that currently enrolled international students have to respond are either choosing to drop out or somehow changing the duration of their degree. The smaller coefficients in column two suggest that there is limited scope for adjustment on these margins and the main response comes from the flow of new students.

The main advantage of the enrollment measure is that I can analyze undergraduate and graduate enrollment separately. These results are shown in columns 3 and 4. The striking difference between undergraduate and graduate enrollment is the coefficient on exchange rates. I find that the exchange rate effects are driven entirely by undergraduate enrollment. For graduate students, the effect is small and not statistically different from zero.

Does the exchange rate effect on international enrollment vary by income levels? Based on the conceptual framework the answer is yes. At high income/asset levels, the student's budget constraint is less likely to bind. As a result, exchange rate variation will only cause a price effect, changing the net return from international education. In contrast for low-income students closer to the constraint, exchange rate variation, in addition to causing a price effect on the net return, may push students beyond their budget constraint. That is, university tuition may become prohibitively expensive for these students. Overall this would lead to exchange rate variation having a greater effect on international student enrollment at lower income levels. To explore this question I add the interaction of GDP per capita and Exchange Rate to my specifications. The results are shown in Table 3.5. Columns 1, 3, 5 and, 7 are the same as Table 3.4 and are repeated here for comparison. Columns 2, 4, 6, and 8 add the interaction term to each specification. To help analyze the overall effect of exchange rates (combining the interacted and non-interacted terms), the bottom of the table reports the magnitude and standard error of these effects at different levels of GDP per capita[8]

For both the flow and stock of international students, the interaction of GDP per capita and the exchange rate is positive and significant. As predicted by the framework, students from lower-income countries are more sensitive to exchange rate variation. Overall in column 2, the exchange rate effect on the flow of international students remains negative and significant at GDP per capita levels one standard deviation above and below the mean. For total student enrollment, the exchange rate effect is not distinguishable from zero when GDP per capita is a standard deviation above the mean. Columns 6 and 8 show the result separately for undergraduate and graduate enrollment. The results for undergraduate enrollment follow

[8]The interpretation of the coefficients on the non-interacted variables; GDP per capita and Exchange Rate, are not intuitive as all variables are in logs

the same patterns as total enrollment. Graduate enrollment however does not respond to exchange rate variations even at GDP per capita levels above or below the mean. Although in column 8 the coefficient on the exchange rate is negative and significant and the interaction term is positive and significant, the overall exchange rate effects computed at the bottom of the table are all indistinguishable fore zero. The magnitude of the coefficients suggests that the exchange rate may have an effect on graduate enrollment at the tails of the income distribution but does not have an effect at and around the mean.

How can we interpret this overall lack of response of graduate enrollment? There are a couple of plausible explanations. If the budget constraint for graduate students is not binding, either due to low out-of-pocket tuition costs or high levels of assets/savings, the price effect of exchange rate variation may be too small to induce a change in their decision. In addition, if graduate students have higher potential future incomes, the exchange rate effects may again be too small to cause a change in their decisions. As Table 3.2 shows, there is evidence that graduate students are less likely to be self-funded, which suggests that this result may be driven by lower out-of-pocket tuition costs. However, as I do not observe tuition costs I cannot definitively test whether this is the mechanism generating the result. It is clear however that as a group graduate students differ significantly from undergraduates.

3.5 Robustness

3.5.1 Alternate Specifications

To test whether the results are dependent on my empirical specification I run two alternate specifications. In the first, I include quadratic terms for Exchange Rate and GDP per capita. This is to test if a linear specification (as in the baseline results) is a good fit or if a more flexible functional form with quadratic terms is needed. The results are presented in table 3.6. Again to analyze the overall effect of Exchange Rates and GDP per capita, the total effect at the average levels of each variable is given at the bottom of the table. Comparing the total effects at the mean exchange rate and GDP per capita to the baseline results in table 3.3 I find that the Exchange Rate effect is slightly larger in the quadratic specification (0.3 compared to 0.26), whereas the GDP effect is almost the same. This suggests that although the quadratic terms are statistically significant for both these variables, the total effect as estimated by a linear specification is similar to the estimate of the quadratic specification and thus a reasonable functional form.

In addition, I also run a first difference specification. First difference specifications are also used to control for unobserved heterogeneity at a unit level, in this case, the country level. Although the relative efficiency on fixed effect and first difference estimators depends on assumptions about the serial correlation of errors, both fixed effects and first difference estimators are unbiased if the strict exogeneity of the error term holds. Therefore a check of this condition is to run both models and see if they generate similar estimates. Finding different results would imply that either one or both specifications are biased, showing that the exogeneity of the error term has been violated.

The specification is as follows:

$$\Delta ln(STUDENTS)_{it} = \alpha + \beta\Delta ln(ER)_{it} + \lambda\Delta ln(GDP)_{it} + \psi\Delta X_{it} + \gamma_t + \epsilon_{it}$$

All variables are differenced and country fixed effects are dropped. Table 3.7 shows the results of the first difference specification. The signs and significance of the results are similar to the original specification results in Table 3.3 suggesting that our model is correctly specified.

3.5.2 Lagged Effects

Another issue to consider is the relevant timing of the exchange rate variation. The main specification uses contemporaneous changes to the exchange rate. However, it is plausible that it may take individuals some time to respond to exchange rate variation. This would imply that lagged exchange rate is the relevant explanatory variable. Secondly, the effect of exchange rate variation may also be persistent. So an exchange rate shock may affect enrollment in the current year as well as future years.

To analyze this timing issue I run the main specification with combinations of lagged exchange rates included. The results are presented in Table 3.8. The exchange rate and its lags are highly correlated, which increases the standard errors of our estimates. The results show that the contemporaneous exchange rate has the largest effect and it is the only effect that is significant in multiple specifications. I interpret this as evidence that the contemporaneous exchange rate is the relevant explanatory variable for international student demand.

3.6 Conclusion

In this paper, I examine international students' demand for higher education in the US. I propose a simple conceptual framework for the decision of international students to study in the US. The framework combines education and employment incentives proposed by the literature to explain international student demand and generates testable predictions about the response of international student enrollment. I then compare these predictions with my empirical results.

I find that a 1 percent increase in the exchange rate (a depreciation of home currency) leads to a 0.25 percent decrease in the flow of international students from that country. The higher cost and tighter budget constraint reduce the number of students studying in the US. The results are entirely driven by undergraduate students. Graduates students, perhaps due to lower out-of-pocket costs, do not respond to changes in the exchange rate. I also find that higher income levels in the home country lead to a higher flow of international students. A 1 percent increase in the GDP per capita leads to a 0.32 percent increase in the flow of international students. This income effect suggests that students have budget and borrowing constraints, which at low-income levels prevents them from being able to study in the US.

The response of international students to exchange rate shocks differ from other migrant categories. While the literature finds that working migrants benefit from the increase in their relative incomes, international students are negatively affected due to higher tuition costs and tighter budget constraints. The result highlights that international students are a unique migrant category with very different incentive structures. They should not be combined with other migrant groups when developing or analyzing immigration policy as they are likely to respond very differently.

3.7 Figures & Tables

Figure 3.1: Trends in the Stock and Flow of International Students in the US

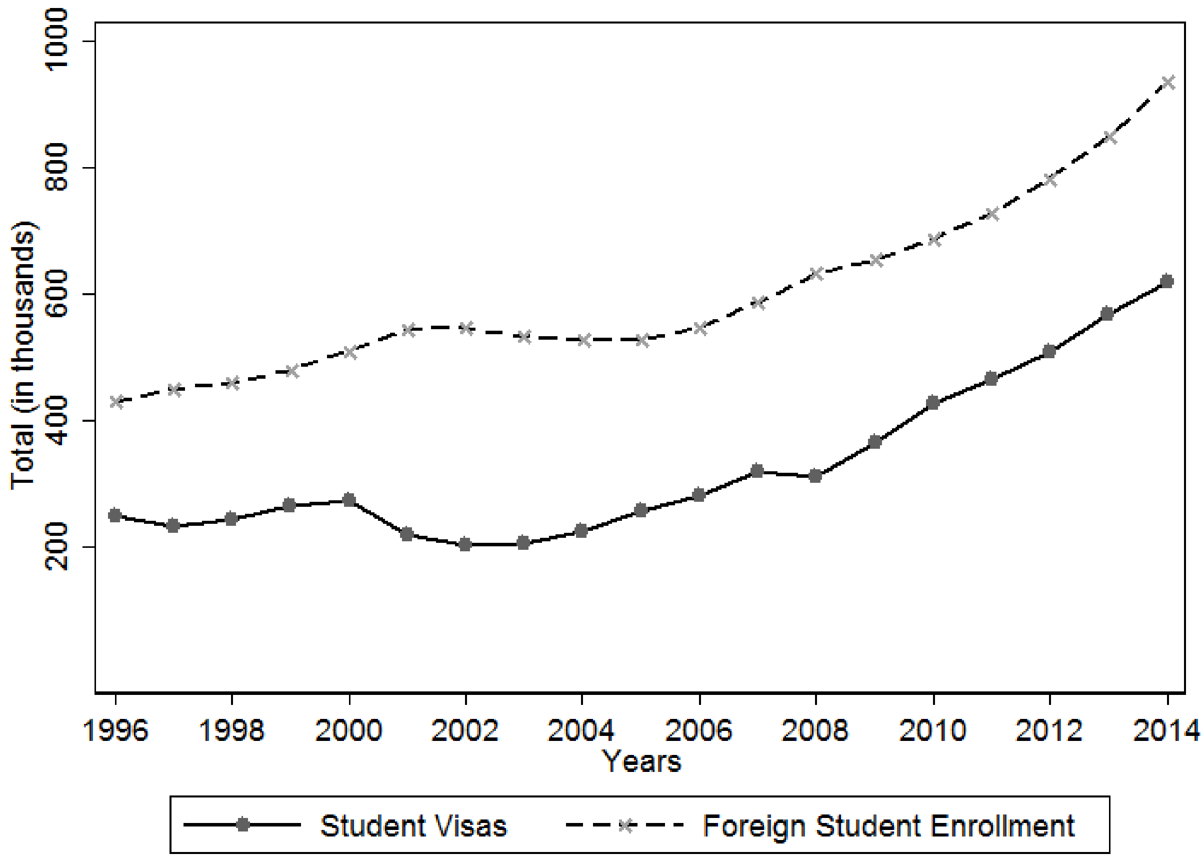


Figure 3.2: Total International Student Enrollment by Degree Type

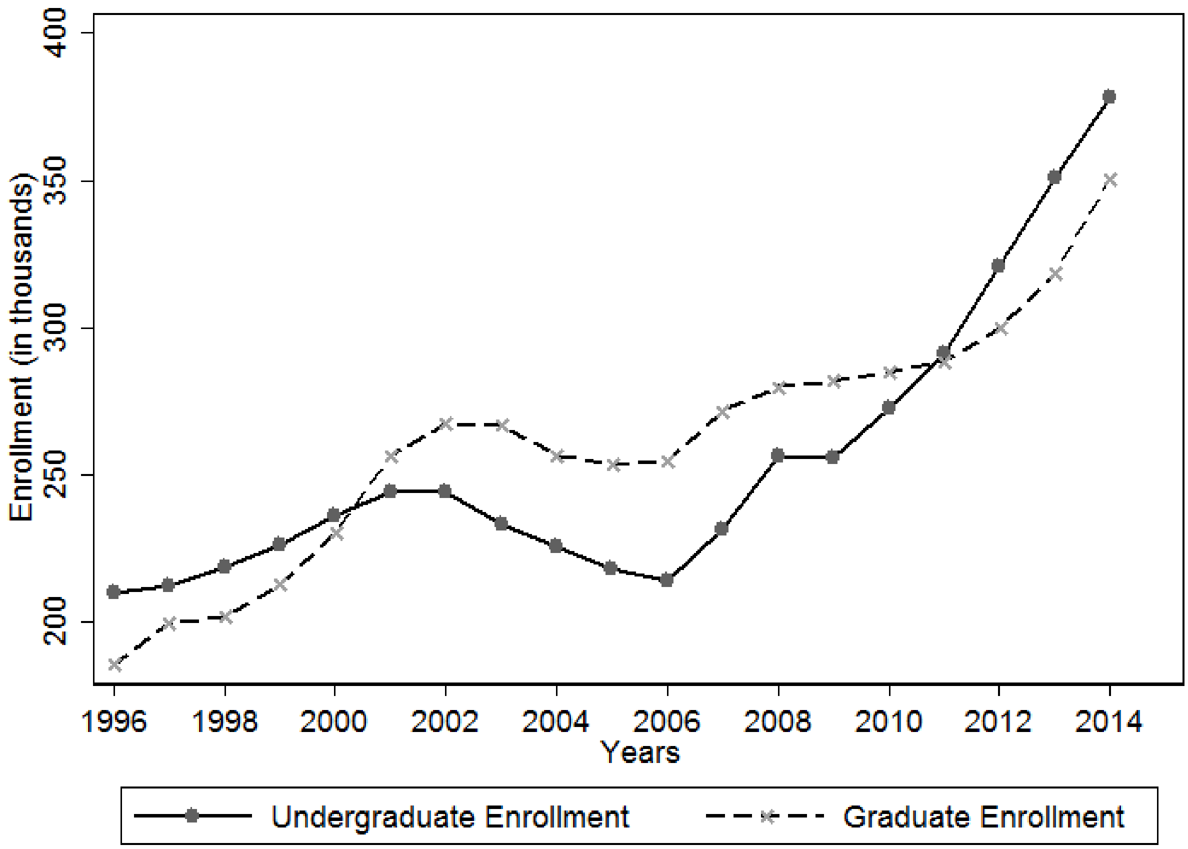


Figure 3.3: Country Level Heterogeneity in International Student Enrollment by Degree Type

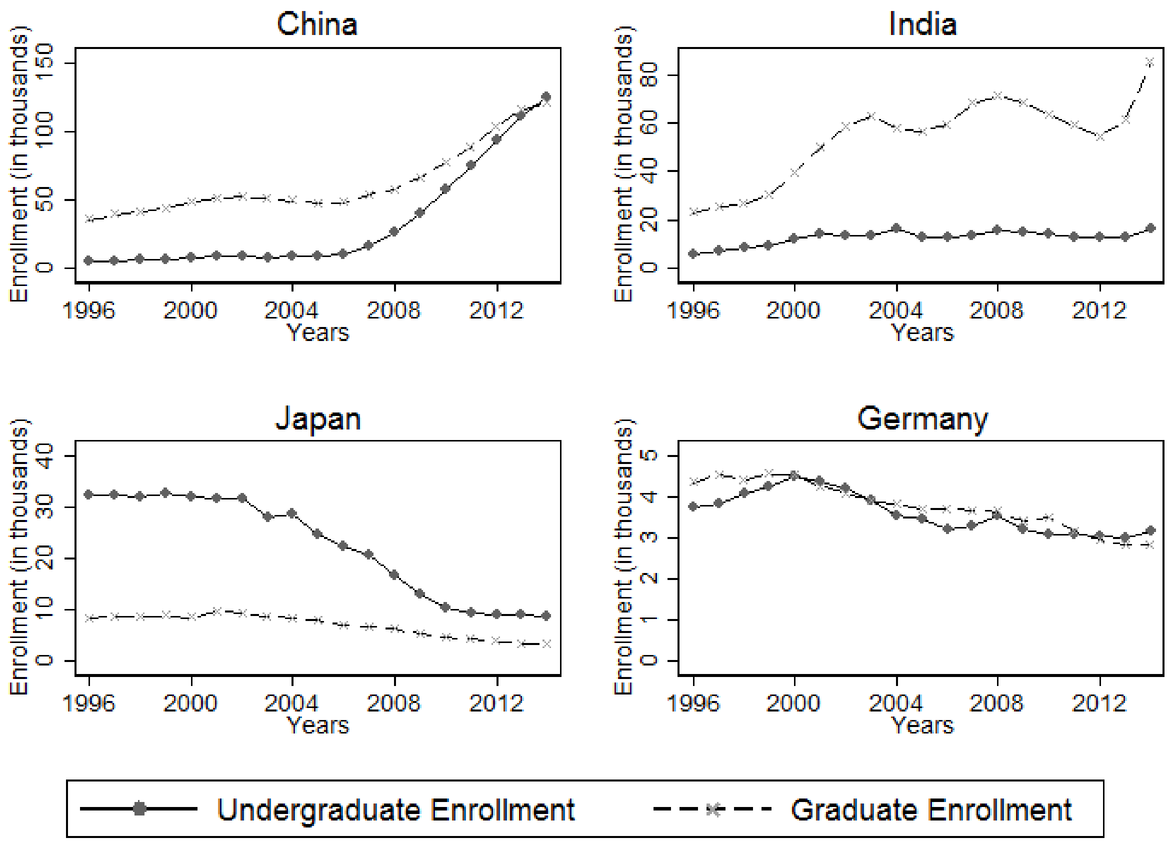


Table 3.1: Descriptive Statistics

	Mean	Std Dev	Min	Max
Student Visas	3,653	13,728	58	274,460
Total Foreign Student Enrollment	6,678	18,309	175	304,040
Undergraduate Enrollment	2,841	6,895	96	124,552
Graduate Enrollment	2,912	9,498	46	120,331

Table 3.2: Funding Sources for International Students as a Percentage of Total Funding (1995 to 2014)

	All	Undergraduate	Graduate	Other
Personal & Family	67.1	80.8	47.2	65.2
U.S College or University	18.9	8.2	37.7	5.8
Home Government/University	4.5	3.9	5.4	3.9
Foreign Private Sponsor	2.7	2.5	2.8	3.1
U.S Private Sponsor	2.7	3.2	2.4	0.9
Current Employment	2.3	0.3	1.7	17.5
Other Sources	0.8	0.3	1.3	2.0
U.S Government	0.6	0.6	0.8	0.3
International Organizations	0.5	0.3	0.7	1.3

Table 3.3: Effect of Exchange Rate and GDP Variation on the Flow of Foreign Students

	(1)	(2)	(3)	(4)
	Student Visas			
ExchangeRate	-0.0527	-0.253***	-0.262***	-0.265***
	(0.0416)	(0.0652)	(0.0688)	(0.0677)
GDPpc		0.317***	0.175***	0.221***
		(0.0515)	(0.0500)	(0.0529)
VisaRegime			0.384***	0.366***
			(0.0415)	(0.0428)
Observations	1,679	1,679	1,679	1,643
Controls	N	N	N	Y
Country & Time FE	Y	Y	Y	Y

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Controls include Inflation, Unemployment

The dependent variable is natural log of the number of student visas issued

All regressors are also in natural logs

Table 3.4: Result by Student Academic Level

	(1)	(2)	(3)	(4)
	F Visas	Total Enrol	Undegrad	Grad
ExchangeRate	-0.265***	-0.110**	-0.196***	0.0218
	(0.0677)	(0.0489)	(0.0533)	(0.0502)
GDPpc	0.221***	0.185***	0.199***	0.205***
	(0.0529)	(0.0366)	(0.0452)	(0.0355)
VisaRegime	0.366***	0.230***	0.294***	0.213***
	(0.0428)	(0.0282)	(0.0343)	(0.0264)
Observations	1,643	1,643	1,637	1,637
Controls	Y	Y	Y	Y
Country & Time FE	Y	Y	Y	Y

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Controls include Inflation, Unemployment

The outcome variable is natural log of each measure

All regressors are also in natural logs

Table 3.5:
Interaction Effect of GDP and Exchange Rates

	(1) F Visas	(2) F Visas	(3) Total Enrol	(4) Total Enrol	(5) Undergrad	(6) Undergrad	(7) Grad	(8) Grad
ExhcangeRate	-0.265***	-0.420***	-0.110**	-0.355***	-0.196***	-0.296***	0.0218	-0.219***
	(0.0677)	(0.105)	(0.0489)	(0.0776)	(0.0533)	(0.0735)	(0.0502)	(0.0721)
GDPpc	0.221***	0.138**	0.185***	0.0541	0.199***	0.150***	0.205***	0.0858**
	(0.0529)	(0.0620)	(0.0366)	(0.0442)	(0.0452)	(0.0507)	(0.0355)	(0.0414)
ExchangeRate x GDPpc		0.0137*		0.0217***		0.00847*		0.0206***
		(0.00708)		(0.00523)		(0.00438)		(0.00514)
Observations	1,643	1,643	1,643	1,643	1,637	1,637	1,637	1,637
Country Controls	Y	Y	Y	Y	Y	Y	Y	Y
Country & Time FE	Y	Y	Y	Y	Y	Y	Y	Y
Total Exchange Rate Effect At								
1 St Dev Above Mean GDPpc		-0.237***		-0.0657		-0.183***‘		0.0548
		(0.0674)		(0.0488)		(0.0543)		(0.0469)
Mean GDPpc		-0.269***		-0.115**		-0.202***		0.00801
		(0.0662)		(0.0464)		(0.0550)		(0.0489)
1 St Dev Below Mean GDPpc		-0.300***		-0.164***		-0.221***		-0.0387
		(0.0690)		(0.0461)		(0.0554)		(0.0479)

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. Controls include Inflation, Unemployment, VisaRegime
The outcome variables and regressors are all expressed in logs. Columns 1, 3, 5 and 7 are the same as Table 4

Table 3.6: Checking Quadratic Form of Exchange Rate and GDPpc

	(1)	(2)	(3)	(4)
		Student Visas		
ExchangeRate	-0.0996	-0.340***	-0.447***	-0.435***
	(0.0643)	(0.0753)	(0.104)	(0.107)
ExchangeRate^2	0.0153***	0.0102*	0.0203***	0.0188**
	(0.00586)	(0.00520)	(0.00709)	(0.00738)
GDPpc		-0.0243	-0.119	-0.0587
		(0.0951)	(0.0922)	(0.0939)
GDPpc^2		0.0164***	0.0137***	0.0124***
		(0.00430)	(0.00403)	(0.00404)
Observations	1,706	1,679	1,679	1,643
Country Controls	N	N	N	Y
Country & Time FE	Y	Y	Y	Y
Total Exchange Rate Effect At				
Mean Exchange Rate	0.00482	-0.270***	-0.308***	-0.307***
	(0.0416)	(0.0609)	(0.0696)	(0.0706)
Total GDPpc Effect At				
Mean GDPpc		0.340***	0.186***	0.216***
		(0.0511)	(0.0492)	(0.0505)

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses.
Controls include Inflation, Unemployment, VisaRegime
The outcome variables and regressors are all expressed in logs

Table 3.7:
Alternate Specification: First Difference

	(1)	(2)	(3)	(4)
		Δ FVisas		
Δ ExchangeRate	-0.196***	-0.262***	-0.188***	-0.163***
	(0.0539)	(0.0580)	(0.0567)	(0.0539)
Δ GDPpc		0.251***	0.144**	0.0601
		(0.0619)	(0.0608)	(0.0657)
Δ VisaRegime			0.327***	0.337***
			(0.0497)	(0.0456)
Observations	1,616	1,590	1,590	1,504
Controls	N	N	N	Y
Time FE	Y	Y	Y	Y

Notes: Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Controls include Inflation, Unemployment

Table 3.8: Checking Exchange Rate Lags

	(1)	(2)	(3)	(4)
		F Visas		
ExchangeRate_t	-0.265***	-0.176	-0.201*	-0.191*
	(0.0677)	(0.146)	(0.108)	(0.100)
$\text{ExchangeRate}_{t-1}$		-0.102		-0.0226
		(0.135)		(0.146)
$\text{ExchangeRate}_{t-2}$			-0.0803	-0.0678
			(0.0972)	(0.108)
Observations	1,643	1,643	1,624	1,624
Controls	Y	Y	Y	Y
Country & Time FE	Y	Y	Y	Y

Notes: Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1. Controls include GDPpc, VisaRegime, Inflation, Unemployment

BIBLIOGRAPHY

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey Wooldridge. 2017. "When Should You Adjust Standard Errors For Clustering?" National Bureau of Economic Research.

Abarcar, Paolo. 2017. "The Return Motivations of Legal Permanent Migrants: Evidence from Exchange Rate shocks and Immigrants in Australia." *Journal of Economic Behavior & Organization*, 144: 62–77.

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

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

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

Ambler, Kate. 2015. "Don't tell on me: Experimental evidence of asymmetric information in transnational households." *Journal of Development Economics*, 113: 52–69.

Ambler, Kate, Diego Aycinena, and Dean Yang. 2015. "Channeling Remittances to Education: A Field Experiment among Migrants from El Salvador." *American Economic Journal: Applied Economics*, 7(2): 207–232.

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

Ashraf, Nava. 2009. "Spousal control and intra-household decision making: An experimental study in the Philippines." *American Economic Review*, 99(4): 1245–77.

Ashraf, Nava, Diego Aycinena, Claudia Martnez A., and Dean Yang. 2015. "Savings in Transnational Households: A Field Experiment among Migrants from El Salvador." *The Review of Economics and Statistics*, 97(2): 332–351.

Baseler, Travis. 2018. "Hidden Income and the Perceived Returns to Migration: Experimental Evidence from Kenya." *Working Paper*.

Batista, Catia, and Gaia Narciso. 2018. "Migrant remittances and information flows: evidence from a field experiment." *The World Bank Economic Review*, 32(1): 203–219.

Benabou, Roland, and Jean Tirole. 2003. "Intrinsic and extrinsic motivation." *The Review of Economic Studies*, 70(3): 489–520.

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

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics*, 119(1): 249–275.

Bird, Kelli, and Sarah Turner. 2014. "College in the States: Foreign Student Demand and Higher Education Supply in the U.S." *EdPolicyWorks Working Paper Series*, 23.

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

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

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

Bound, John, Breno Braga, Gaurav Khanna, and Sarah Turner. 2020. "A Passage to America: University Funding and International Students." *American Economic Journal: Economic Policy*, 12(1): 97–126.

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

Castilla, Carolina, and Thomas Walker. 2013. "Is ignorance bliss? The effect of asymmetric information between spouses on intra-household allocations." *American Economic Review: Papers & Proceedings*, 103(3): 263–68.

Castro, Lucio, and Carlos Scartascini. 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."

Chen, Joyce J. 2006. "Migration and imperfect monitoring: implications for intrahousehold allocation." *American Economic Review*, 96(2): 227–231.

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.

Chiappori, Pierre-Andr. 1988. "Rational Household Labor Supply." *Econometrica*, 56(1): 63–90.

Chiappori, Pierre-Andr. 1992. "Collective Labor Supply and Welfare." *Journal of Political Economy*, 100(3): 437–467.

Chin, Aimee, Léonie Karkoviata, and Nathaniel Wilcox. 2015. "Impact of bank accounts on migrant savings and remittances: evidence from a field experiment." *Working Paper*.

Cyan, Musharraf R., Antonios M. Koumpias, and Jorge Martinez-Vazquez. 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.

De Arcangelis, Giuseppe, and Dean Yang. 2019. "Labeled Remittances." *AEA RCT Registry*, https://doi.org/10.1257/rct.2554–3.1.

De Arcangelis, Giuseppe, Majlinda Joxhe, David McKenzie, Erwin Tiongson, and Dean Yang. 2015. "Directing remittances to education with soft and hard commitments: Evidence from a lab-in-the-field experiment and new product take-up among Filipino migrants in Rome." *Journal of Economic Behavior & Organization*, 111: 197–208.

De Laat, Joost. 2014. "Household allocations and endogenous information: The case of split migrants in Kenya." *Journal of Development Economics*, 106: 108–117.

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

Doi, Yoko, David McKenzie, and Bilal Zia. 2014. "Who you train matters: Identifying combined effects of financial education on migrant households." *Journal of Development Economics*, 109: 39 – 55.

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

Dwenger, Nadja, Henrik Kleven, Imran Rasul, and Johannes Rincke. 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, and Jonathan Feinstein. 1994. *The Role of Moral Sentiments and Audit Perceptions in Tax Compliance.* Vol. 49.

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

Gibson, John, David McKenzie, and Bilal Zia. 2012. "The Impact of Financial Literacy Training for Migrants." *The World Bank Economic Review*, 28(1): 130–161.

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

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

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

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

Institute of International Education. 2015. "International Students at All Institutions 2001/02 -2014-15." *Open Doors Report on International Educational Exchange.*

Joseph, Thomas, Yaw Nyarko, and Shing-Yi Wang. 2018. "Asymmetric information and remittances: evidence from matched administrative data." *American Economic Journal: Applied Economics*, 10(2): 58–100.

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

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

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

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

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

Lucas, Robert E. B., and Oded Stark. 1985. "Motivations to Remit: Evidence from Botswana." *Journal of Political Economy*, 93(5): 901–918.

Lundberg, Shelly, and Robert A. Pollak. 1993. "Separate Spheres Bargaining and the Marriage Market." *Journal of Political Economy*, 101(6): 988–1010.

Luttmer, Erzo F. P., and Monica Singhal. 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.

McKenzie, David, Caroline Theoharides, and Dean Yang. 2014. "Distortions to the International Migrant Labor Market: Evidence from Filipino Migrationa and Wage Responses to Destination Country Economic Shocks." *American Economic Journal: Applied Economics*, 6: 49–75.

McKenzie, David, John Gibson, and Steven Stillman. 2013. "A land of milk and honey with streets paved with gold: Do emigrants have over-optimistic expectations about incomes abroad?" *Journal of Development Economics*, 102: 116 – 127.

Nekoei, Arash. 2013. "Immigrants' Labor Supply and Exchange Rate Volatility." *American Economic Journal: Applied Economics*, 5: 144–164.

Perez-Truglia, Ricardo. 2020. "The effects of income transparency on well-being: Evidence from a natural experiment." *American Economic Review*, 110(4): 1019–54.

Perez-Truglia, Ricardo, and Ugo Troiano. 2018. "Shaming tax delinquents." *Journal of Public Economics*, 167: 120–137.

Rapoport, Hillel, and Frdric Docquier. 2006. "The Economics of Migrants' Remittances." In . Vol. 1. 1 ed., , ed. S. Kolm and Jean Mercier Ythier, Chapter 17, 1135–1198. Elsevier.

Rozensweig, Mark. 2006. "Global Wage Differentials and International Student Flows." *Brookings Trade Forum*, 6: 57–87.

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

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

Seshan, Ganesh, and Dean Yang. 2014. "Motivating migrants: A field experiment on financial decision-making in transnational households." *Journal of Development Economics*, 108(C): 119–127.

Seshan, Ganesh, and Robertas Zubrickas. 2017. "Asymmetric Information about Migrant Earnings and Remittance Flows." *World Bank Economic Review*, 31(1): 24–43.

Shen, Ying. 2016. "The impacts of the influx of new foreign undergraduate students on US higher education." *Journal of Economic Literature*.

Shih, Kevin. 2017. "Do International Students Crowd-out or Cross-subsidize Americans in Higher Education?" *Journal of Public Economics*, 156: 170–184.

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

United Nations. 2019. "International Migrant Stock 2019."

Waseem, Mazhar. 2020. "Does Cutting the Tax Rate to Zero Induce Behavior Different from Other Tax Cuts? Evidence from Pakistan." *Review of Economics and Statistics*, 102(3): 426–441.

World Bank. 2017. "Migration and Development Brief 27: Migration and Remittances."

World Bank. 2018*a*. "Annual Remittances Data 2018."

World Bank. 2018*b*. "Bilateral Remittance Matrices 2017."

Yang, Dean. 2006. "Why do Migrants Return to Poor Countries? Evidence from Philippine Migrants Responses to Exchange Rate Shocks." *Review of Economics and Statistcs*, 88: 715–735.

Yang, Dean. 2008. "International Migration, Remittances and Household Investment Evidence from Philippine Migrants Exchange Rate Shocks." *Economic Journal*, 118: 591–630.

Yang, Dean. 2011. "Migrant Remittances." *Journal of Economic Perspectives*, 25(3): 129–52.

Yang, Dean, and HwaJung Choi. 2007. "Are Remittances Insurance? Evidence from Rainfall Shocks in the Philippines." *The World Bank Economic Review*, 21(2): 219–248.