Tax Professionals and Tax Evasion *

Marco Battaglini[†], Luigi Guiso[‡], Chiara Lacava[§], Eleonora Patacchini[¶]

January 30, 2020

Abstract

To study the role of tax professionals, we merge tax records of the entire population of sole proprietorship taxpayers in Italy for seven fiscal years with the respective audit files from the tax revenue agency. Exploiting quasi-random variation in audit policy, we first document that there is a robust correlation between a taxpayer's evasion and that of the other clients of the same tax practitioner. We then exploit the structure of our data set to study the mechanisms behind this phenomenon. We establish two results. First, taxpayers sort themselves into tax professionals with heterogeneous levels of tax evasion. Second, tax professionals generate informational externalities that affect their clients' tax compliance. The latter increases directly in response to personally experienced audits and indirectly following the audits of other clients of the accountant. While the direct effect of tax audits is decreasing over time, the indirect effect is increasing over time with a total cumulative marginal effect that amounts to 17% of that of the direct effect. The dynamic spillover generated by tax professionals is therefore an important channel of influence that ought to be considered in the evaluation and design of auditing schemes.

Keywords: tax enforcement, tax compliance, tax evasion, indirect treatment effects

JEL classification: K34, H26

*A previous version of this paper based on a subset of the data circulated with the title : Tax Professionals: Tax Evasion Facilitators or Information Hubs? We are very grateful to the Italian Revenue Agency for granting us access to the data. We are solely responsible for the ideas expressed in the paper. We thank Stephen Coate, Bernard Fortin, Roberto Galbiati, Philipp Kircher, Doug Miller, Matteo Paradisi, Alex Rees-Jones, Giulio Zanella, and participants to several seminars for valuable discussions.

†Cornell University and EIEF E-Mail: battaglini@cornell.edu

‡EIEF E-Mail: luigi.guiso@eief.it

§Goethe University Frankfurt E-Mail: lacava@econ.uni-frankfurt.de

¶Cornell University E-Mail: patacchini@cornell.edu

1 Introduction

The traditional literature on tax evasion focuses on the direct relationship between the tax authority and taxpayers: individuals are assumed as independent utility maximizers who trade off costs and benefits of violating the laws. It is, however, increasingly the case that the relationship is more complex, as it is mediated by tax professionals. This evolution in the relationship between taxpayers and the tax authorities is caused by the increasing complexity of the tax code and, at least for corporations and the upper tail of the income and wealth distribution, by globalization which stimulates the demand for professional advice by creating opportunities for avoidance and evasion.

In this context, understanding the role of tax professionals is key to minimizing the cost of compliance and making auditing more efficient. What do tax professionals do *beyond* helping their clients understand and apply the laws? Recent informal accounts suggest that tax professionals play a key role in the formation of their clients' expectations regarding enforcement probabilities and the risk of certain practices (see Braithwaite, 2005). They also help shape tax norms and ethical standards (see Smith and Kinsey, 1987; Raskolnikov, 2007). For example, Braithwaite (2005) finds from interviews of tax professionals that tax advisers played an important role in the diffusion of tax shelters with a "supply driven" contagion effect. A precise understanding of what tax professionals do is important for tax policy: information diffusion may be in the interest of both taxpayers and the tax authority; the diffusion of ambiguous ethical standards or the use of informational advantages to avoid controls certainly are not. Is it really true that tax professionals serve as informational and ethical hubs for their clients?

In this paper, we investigate these questions in the Italian context, where tax accountants provide a wide range of consulting services to their clients, including tax income reporting. Tax accountants are a regulated profession in Italy, requiring at least 3 years of postgraduate work experience under the guidance of a certified accountant and a certificate obtained after passing a nation-wide exam.

Our analysis is made possible by the use of an exclusive dataset. We merge individual level information from two separate administrative records from the Italian tax authorities: the return files for the universe of sole proprietorship taxpayers (approximately 4.7 million) from 2007 to 2013 and the audit files. The data provide detailed information about the taxpayer's reported income, demographic characteristics and audit history, including the outcome of any audit: i.e. the resulting assessed taxable income and the amount evaded, measured by the difference between assessed and reported income. Importantly, the data also provide information regarding whether a taxpayer employs a tax accountant: around

97% of taxpayers in our sample rely on the services of a tax accountant who is identified by a unique code by law. This allows us to match taxpayers with accountants and follow the history of the taxpayer-accountant relationship over time.

A natural concern when studying tax enforcement using administrative data is that audits are not randomly assigned across taxpayers. We are able to address this problem by exploiting a particular institutional feature of the Italian tax administration. In Italy tax audits can be initiated by two tax authorities, the *Italian Revenue Agency* (IRA), an administrative body in charge of tax collection and enforcement, and the *Guardia di Finanza* (GdF), a police force with a wide range of responsibilities including tax enforcement. While the latter naturally relies on "soft information" that we do not observe in the selection of audits, the former is in charge of a program of automated and desk assessments that are based on data in the *Anagrafe Tributaria* (the official Tax Registry), for which we have been granted exclusive access. We show that the audits originated by the IRA are random conditioning on observables and we therefore rely on them in our analysis.

We start our analysis documenting a positive and robust correlation between a taxpayer's evasion and that of own tax accountant, as measured by the average evasion of her/his clients over the entire period under analysis, excluding the taxpayer. The correlation disappears when relating own evasion with the evasion of similar taxpayers (in terms of location, business sector and other characteristics) served by different tax practitioners, suggesting a role of own tax practitioner in tax compliance. We study two mechanisms through which this correlation may arise, each one highlithing a potential specific role of the tax accountant. First, *self-selection* of taxpayers who sort themselves into accountants of heterogeneous tolerance for tax evasion. Seconds, *informational externalities* generated by the tax accountant activities. In the first case, the accountant serves as a tax evasion facilitator of tax evasion prone customers. In the second, as information hubs that allow sharing the audit experiences of other clients.

We use the panel structure of our dataset to test for the presence of the self-selection channel, and the related tax evasion facilitator role. We can indeed follow taxpayers as they, voluntarily or involuntarily, switch tax accountants. If taxpayers self-select into a tax accountant of their type, we expect the type of the accountant before the switch (as measured by the observed historical tendency of its clients to evade) to be correlated with the type of the accountant after the switch. We show this is indeed true both in the entire sample of taxpayers who switch accountants, and in the sub-sample of taxpayers who are forced to switch because their accountant exits the market (due to, e.g. retirement or closure).

To test for the *information externalities* mechanism, i.e. whether tax accountants also play an active role as informational hubs, we study whether the income reported by a taxpayer i at time t depends on whether other clients of his/her accountant j have been audited at $t-1$, and how this compares to the effect of a direct audit to i at $t-1$. We find that an audit of i at $t-1$ induces a 7.5% increase in the income reported at t; while an audit at $t-1$ to at least one other customer of accountant j induces an increase of 1.5% of i's reported income at t. This latter result is also robust to a placebo test: if we use audits to other taxpayers in comparable groups but with a different accountant, we see no relationship between a taxpayer's income declaration at t and the $t-1$ audits in the comparable group.

Sharing of information about audits of other taxpayers through the tax accountant also affects the taxpayer's decision to voluntarily switch accountants. We find that a given taxpayer is less likely to separate from their accountant when at least one other client of that accountant has been audited in the previous period, especially if the audited peers share common characteristics with the taxpayer. This is fully consistent with the informational hub role of the accountants. First, because the taxpayers may come to know that other clients have been audited because their accountant notifies them about the IRA activities. Second, because taxpayers that are notified are less likely to switch their accountant precisely because he is passing them valuable information. In sum, our evidence lends strong support to the idea that tax accountants play the dual role of informational hubs as well as of tax evasion facilitators for taxpayers prone to evasion.

We show that accountant-induced peer effects tend to magnify the effect of audits along two dimensions: first an audit on i has an effect on *all* peer-customers of i's accountant; second, the peer effects tend to persist over time. The effect on $i's$ currently reported income of a three-year old audit on other customers of the same accountant is not only positive but even larger than the one-year old effect. Even in the most conservative estimate, after three years, an audit on the other customers of the same accountant cumulatively increases reported income by 2%. In comparison, if taxpayer i is audited, the effect on reported income fades away with time: after three years it is 18% of the one year lagged effect and cumulatively is worth a 11.9% increase in reported income. Indeed, since the average number of clients is 31, the increase in reported income of non-audited clients of an audited accountant is more than three times the increase in reported income of the audited taxpayer. We find that this spillover is larger for clients of the same accountant in the same sector, age and size of business; it does not depend on whether there are other audited taxpayers geographically close, in the same business sector and of the same size but clients of another accountant. This suggests that the vehicle through which the spillover diffuses among taxpayers is the accountant, not geographic proximity, shared sector or business size; and that accountants selectively share with their customers the information that they gather from their activities. These results imply that accounting for information spillovers is key both when designing the audit policy as well as when assessing its effects on compliance.

Our work is at the intersection of two literatures that have remained mostly separated up until now. First, the relatively small literature on tax practitioners, and second, the literature on social spillovers in tax compliance. With respect to the first, the research on tax practitioners has traditionally focused on the role of tax accountants as providers of expert advice but has ignored the potential social spillovers between clients of the same practitioner, implicitly assuming that the tax accountant does not change the nature of the traditional direct relationship between the tax authority and an individual taxpayer.[1] The focus instead has fallen on the determinants of the choice of hiring a practitioner or not (Erard, 1993), the usefulness of tax practitioners (Slemrod, 1989), the effect on the level and type of compliance (Klepper et al., 1991, Erard, 1993), and the role played by practitioners in reducing uncertainty and costs of compliance (Scotchmer, 1989, Beck and Jung, 1989, Reinganum and Wilde, 1991). The role of tax accountants in collecting and distributing information has been suggested by anthropological and social studies, albeit informally (see, among others, Smith and Kinsey, 1987, Braithwaite, 2005, Raskolnikov, 2007).

The literature on social spillovers has focused on showing network externalities in compliance behavior relying mainly on lab or field data.[2] This literature, however, has not investigated the affiliation to a *common* tax practitioner as a source of network effects. In addition, we look at tax enforcement effects for small firms and professionals, which is an area where we know very little.[3]

A recent work that explicitly considers the role of tax professionals in tax compliance is Boning et al. (2018). They study the comparative effect of either the visit of an IRS Revenue

[1] See Andreoni et al. (1998) for a survey of this literature.

[2] Early work tested the hypothesis of social spillovers using laboratory experiments (see Fortin et al., 2007, Alm et al., 2009, and Alm et al., 2017). More recent research has used field data and experiments to investigate the role of spatial proximity. Studying compliance in TV license fees in Austria, Rincke and Traxler (2011) find that household compliance increases with enforcement in the vicinity. Galbiati and Zanella (2012) estimate social externalities of tax evasion in a model in which a social multiplier is induced by congestion of the auditing resources of local tax authorities. Del Carpio (2014) studies how compliance on property taxes in Peru depends on the perceived average compliance. Perez-Truglia and Troiano (2018) find evidence that tax delinquents respond to shaming penalties that increase the salience of their violations.

[3] An important related but distinct literature studies the role of third-party reported paper trails for enforcement. The idea is that taxpayer whose income is reported completely or partially by a third party have lower incentives to evade because their evasion can be easily detected. Using data from Denmark, Kleven et al. (2011) show that receiving a threat-of-audit letter from the tax authority has a significant effect on self reported income, but no effect on third party reported income. Pomeranz (2015) shows that companies that generate a VAT paper trail respond less to exogenously generated changes in their perceived audit probability. Changes in their perception, moreover, increase VAT payments to their suppliers. Pomeranz (2015) emphasizes production linkages across firms as a vehicle through which tax audits can spillover. We focus on spillovers generated by the common tax accountant even among firms that would otherwise be unrelated. Common to both is the importance of accounting for taxpayers interconnections when assessing tax enforcement.

Officer or of an informational letter from the IRS on firms suspected of noncompliance with the requirement to remit withheld income and payroll taxes, which are due every quarter. They look for network effects on other firms related to the targeted firm by geographical proximity, a parent-subsidiary link or, as we do, a shared tax preparer. They find neither network effects of letters nor network effects of visits on remittances more than a quarter from the visit irrespective of the network definition, but they find an effect, significant at the 10% level, on firm i's remittances of direct visits to other firms served by the same tax practitioner as i's in the previous quarter. This is an important finding consistent with ours, but it still leaves open the question of the existence of network effects in income tax returns. Tax returns are filed on a yearly basis, thus peer effects have an impact on returns only if they survive for more than one quarter. As mentioned above, we find that the effect of an audit on peers–arguably a more intrusive intervention than a mere visit–persists for years. One advantage of Boning et al. (2018)'s work with respect to ours is that it relies on a controlled experiment of over 12,000 firms suspected of violations. One advantage of our work is that we observe the universe of sole proprietorship taxpayers in Italy for seven years, thus allowing us to exploit the size of the dataset (over 20 million individual observations) and its panel structure to explore the *mechanisms* behind the observed network effects.

We proceed as follows. In Section 2 we provide an overview of our data and institutional background. In Section 3, we show evidence on the correlation between a taxpayer's income evasion and that of the other clients of her/his own tax preparer documenting its robustness to a vast set of model specification and estimation strategies. Guided by a simple model of the interaction between taxpayers and tax accountants (detailed in the Appendix), in Section 4 we study empirically the mechanisms thorough which the correlation can be generated. Section 5 concludes.

2 Institutional background and data

2.1 Institutional background

In Italy the administration of tax revenues is decentralized to two distinct agencies: the Italian Revenue Agency (*Agenzia delle Entrate*, IRA), an administrative body in charge of tax collection and tax enforcement; and the Financial Police *(Guardia di Finanza*, GdF in brief), a military police force responsible for dealing with financial crime, smuggling, performing customs and borders checks, and patrolling Italy's territorial waters which contributes to tax enforcement collaboration with the IRA.[4] Both agencies can initiate a tax audit but each follow different methodologies that leverage on their comparative advantage.

The IRA is almost exclusively in charge of designing the program of automated assessments that are routinely carried out by the agency based on information and data available on the Tax Registry *(Anagrafe Tributaria)*, elaborated via a number of different applications (see OECD, 2016, p. 48). The Tax Registry is a centralized database that identifies each taxpayer with a unique tax code (*Codice Fiscale*, the analog of the US Social Security Number) and associates it with a rich set of statistical information. Audits are chosen based on these records. The GdF, on the other hand, has a widespread presence in the territory, suitable for carrying out in-depth investigations aimed at uncovering a wide spectrum of illegal activities. This gives the GdF exclusive access to "soft" information which can be exploited to design its audit policy.

More specifically, since the institution of the *Studi di Settore* (Sector Studies) in 1993,[5] the Italian legislature has formalized the idea that IRA tax audits of medium-small businesses and practicioners should be primarily based on conformity of these types of taxpayers to expected ability to pay. The latter is defined at a narrow geographic and business sector level by a committee of experts on the basis of statistical studies. These studies gather information on homogeneous groups of taxpayers and provide estimates of (i) the minimum income they should produce and that are thus expected to file, and (ii) the range of several indicators of economic performance (D'Agosto et al., 2017). Based on these studies, for each tax filing the IRA elaborates two indexes of "conformity" to *Studi di Settore* that are used to target individual audits: an indicator variable called *Congruent*, and one called *Coherent*. A "congruent" tax filing is one whose reported income is above the estimated minimum income threshold identified in Study di Settore. A "coherent" tax filing is one with no indicator of economic performance outside the estimated ranges computed by *Studi di Settore* for a taxpayer of that type. Both indicators, as well as other information used by the IRA to select audits are in the data that the IRA has exclusively shared with us and that we describe in greater detail below. This gives us a direct insight into the IRA's auditing policy.

In sum, while IRA audit policy relies uniquely on "hard data" from its registries, the GdF audits are guided by insights from information gathered as part of its police activities. Below we exploit this institutional arrangement to identify a set of quasi-random audits, that are random conditional on the statistical information available to the IRA at the time when the audits are selected.

[4]A third agency involved with tax audits is the *Custom Agency*, but this agency is much smaller than the other two and not relevant to taxpayers in our sample.

[5]See law n. 427/1993 that introduced the *Studi di Settore* and law 146/1988 that regulates their use.

2.2 Data description

Our study relies on population-level data for all Italian sole proprietorship businesses. We merge information from two different administrative records from the IRA: returns files and audit files. In both cases, records are at the individual level and cover filings of incomes generated in seven fiscal years, from 2007 to 2013, reported between 2008 and 2014, and audited between 2009 and 2015. By law, tax filings older than five years cannot be audited, and so one needs at least five years of data to observe a full auditing cycle. Accordingly, our data contain three full sets of tax reports with completed audits. The data contains detailed information on all components of taxpayers reported income and his/her demographics (gender, age, marital status, detailed geographical location, detailed sector of activity). It also documents whether and when the taxpayer was audited and the tax filing year being audited as well as the agency that originated the audit, whether the IRA or the GdF . The data also contains the result of the audit with the assessed taxable income, and thus the amount of evasion found (if any) computed as the difference between reported and assessed income.[6] Importantly for our purposes, the data reports the identifier of the tax accountant/consultant that filed the taxpayer's tax statement.[7] Because we have population-level data we can trace all taxpayers that are served by the same tax consultant, observe taxpayers' mobility from one accountant to another, as well as accountant closures which force taxpayers to match with a new tax accountant. Because we know the location of the taxpayer and that of the accountant, we can portray the geography of the accountant clientele and identify accountants that are likely to be closer substitutes for the one currently serving a given taxpayer. As we will show, these properties are important for documenting peer effects and establishing their nature.

The taxpayers in our dataset are individuals who own a sole proprietorship, where no legal distinction is made between the enterprise and the sole owner. Table 1, panel A shows summary statistics. Overall, our sample contains almost 4.7 million taxpayers distributed in twenty regions. About 27% are women and the average age is 47 years. The average enterprise has been in operation for 13 years and, consistent with the small size of these businesses, employs 0.8 workers with relatively limited variability (standard deviation 3.2; 90th percentile: 2 employees). The average reported gross taxable annual income is 18,640 euros with relevant heterogeneity (standard deviation: 48,694 euros; 90th percentile 39,997 euros) partly reflecting differences across industries.

[6] The GdF has to communicate to the IRA all initiated audits together with their outcome. Thus all information on the audits, independently of who initiated it, is in the IRA archives.

[7] We drop observations with negative filed income (4.2% of the observations), taxpayers that do not use an accountant (2.5% of the observations) and filings of tax accountants (1.8% of the observations).

During the sample period, 289,434 taxpayers (6.2% of the total) were audited at least once and the vast majority of the audits (97%) were originated by the IRA and the rest by the GdF. The share of audited tax filings is around 1.9%. However, there is some variation over time and substantial heterogeneity in auditing probabilities across regions. Tax filings of larger firms, defined either in terms of the number of employees or the value of filed income, are more likely to be audited. Conditional on being audited, the fraction of filings with positive evasion is 66.45% (Table 1, panel B). The distribution of the share and amount of income not declared by evaders are shown in the two panels of Figure 3 and are both quite dispersed. Around 7.1%of audited taxpayers are found to have evaded all of the taxable income produced. The average amount evaded, conditional on evasion, is about 20,328 euros, 1.1 times the average income filed with the tax agency (Table 1, panel B). Conditional on auditing a taxpayer, the IRA can review any of the tax filings over the past five years. Hence, each year, the population of tax filings at risk of being audited comprises all tax filings up to five years old that have not been audited in the previous years (about 95.5 million filings in our sample). The share of audits over the population at risk is 0.41% (388,513 audits divided by 95.5 million cases).

Except for a small minority, all of the taxpayers rely on the services of a tax advisor. Accountants serve taxpayers geographically close to them: in our sample an accountant has 62% of the customers in the same municipality and almost all (96%) in the same region. Overall there are 107,069 tax accountants serving the 4.7 million taxpayers in our sample; hence, on average, a tax accountant serves 31 sole proprietorship taxpayers.[8] There is, however, considerable heterogeneity in the size of tax accountants, as shown in Figure 5, which plots the distribution of accountants in terms of the number of customers (panel A) and of the overall income filed by their customers (panel B). Over the sample period, we observe entries of new accountants and exits of existing ones. The average annual entry rate is 5.1% and the exit rate is 3.7%. Interestingly, while taxpayers tend to have long term relations with their tax accountant, some of them switch, sometimes as a consequence of closure of their accountant. Overall, we observe 18% of the customers switching accountants (Table 1, panel C) from one year to the next, with one-third of such switches following the closure of the accountant. In Section 5.2 we rely on switchers to test for sorting between taxpayers and accountants based on their propensity to evade taxes.

Figure 6 provides descriptive information on heterogeneity across accountants along two dimensions: the share of customers that are audited during the period (panel A) and the share of customers that are found to evade taxes, conditional on being audited (panel B).

[8]Needless to say, the customer base of a tax accountant is larger than this figure as they serve also incorporated firms as well as individual taxpayers.

In a given year, around 47% of the accountants have no customer audited. The rest of the distribution shows marked heterogeneity across accountants in the fraction of their customers audited. The empirical distribution of the share of evaders among the audited customers of each accountant also shows substantial heterogeneity and a long tail to the right–a few accountants have very large shares of evaders. The share of accountants with more than one-fourth of evaders among their audited clients is 81.9%, while the share of accountants with no evaders among their audited clients is 14.3%.

2.3 Quasi- random audit selection

We have been provided with *exclusive* information on the authority that originated the audit, IRA or GdF, and access to the *entire* hard data available at the time of deciding audits.[9] The data set includes hundreds of data points on taxpayers, including their demographic characteristics, their history (for instance of bankruptcies), information on their tax accountants, balance sheet indicators and more. Since, as previously discussed, the IRA relies on the same hard information that is available to us to define which tax filing to audit in a specific year, we obtain random variation in the IRA audit selection if we use a sufficiently large set of conditioning variables. In this section we show that once we condition on a selection of the available data, audit decisions are effectively random, thus ruling out that they are driven by additional "soft information" that we do not observe.

To this goal we propose two set of tests. First, we show that the selection of variables to which we condition our analysis is sufficient to describe the IRA's audit policy. Naturally, the IRA has not provided us with the exact algorithm they use to identify the specific tax filing to be audited but they shared with us the output of their main elaborations from the *Studi di Settore* and have guided us in the selection of the other relevant variables. We use a large number of natural variables including time, location fixed effects, the indexes of conformity from the *Studi di Settore*; the characteristics of the taxpayer and its firm (gender, marital status, years of activity, sector of activity, firm size etc); and the characteristics of the tax practitioner (including the number of clients and regional dispersion of the clients). To investigate whether our control function represents audit decisions well, we estimate the probability of being audited both using our selection of variables and using variables selected with machine-learning techniques and we show that the two approaches have comparable performances. We identify the relevant variables among the ones available in our data (more than 270) using the Least Absolute Shrinkage and Selection operator (LASSO); and we then

[9] Audits and investigations initiated by the GdF are transmitted to the IRA which collects in its archives all audits data, including whether the audit was initiated by the IRA or the GdF.

form a prediction using the selected variables.[10]

Table 2 reports goodness of fit statistics for LASSO probit predictions obtained using alternative shrinkage estimators. For each model, we report the number of non zero coefficients, the out-of-sample deviance and deviance ratio, and the pseudo R-squared. The table shows that the number of selected variables varies between 33 and 171 depending on the LASSO estimation used, with a pseudo R-squared in the range of 0.0516 and 0.0722. Our probit model includes 158 variables and produces for the testing sample a pseudo R-squared equal to 0.0556. The fact that this value lies within the range of the pseudo R-squared produced by the LASSO estimators shows that the predictive ability of our hand-curated set of variables is comparable to the ones chosen by the machine.

While the previous test shows we are using all the relevant information in the data set, a limitation is that we are conditioning on available data previous to the audit (that is if the audit is at time t, we condition on data up to time $t-1$). It may be that the audit at time t relies on new soft information generated at time t that is observed by the authorities but not by us. The problem with this is that we cannot condition on outcomes at time t or posterior to t at the individual level, since these outcomes are obviously affected by the audit to the extent that this data is generated after the audit. This motivates the following balance test. We identify a set of variables at the province/sector level that may be correlated with the potential soft information available to the authorities at the time of the audit. These include the average levels and growth rates of income from t to $t+1$, VAT taxable turnover, operating costs, the ratio between operating costs and the net value of production all at $t+1$, and also the share of evasion at $t+1$, based on future audits in the same province and business sector at $t+1$. The partition of the sample by geographic district and business sector is quite fine, since we have 110 distinct geographic districts and 21 business sectors. Each cell in this partition contains a median of about 400 tax returns.[11] We then test if the

[10] That is, we use post estimation LASSO to make sure we do not form predictions using penalized estimates of coefficients. See for technical details Hastie et al., 2015.

[11] The indicators listed above are informative of taxpayers tax compliance behavior and ability to pay back the debt in a specific location and business sector. For example, a higher growth rate in income from a year to the next may signal lower propensity to evade taxes, a higher level or growth rate in revenues, VAT taxable turnover and operational costs may signal firms in expansion, whereas a higher growth of the ratio between operational costs and the net value of production may signal distressed firms. A higher level of share of evasion may signal opportunities to hide income. There is ample casual evidence that the GdF uses "soft" information gained on the field to infer compliance habits of categories of taxpayers. These insights often translate in "campaigns" to audit specific professions (such as dentists and funeral homes, for instance) and in specific locations. For examples of campaigns on dentists and orthodontists, see *Repubblica* [1995] and *Il Giornale Trentino* [2014]. For examples on funeral homes, see *La Nuova Sardegna* [2011] and *Repubblica* [2019]. For examples of campaigns targeting specific regions, see *Corriere della Sera* [2012], reporting an auditing "blitz" in the popular ski resort of Cortina d'Ampezzo; and *Repubblica* [2012], reporting auditing campaign in popular seaside resorts near Rome.

expected values of these variables are systematically different in the population of audited and non audited taxpayers conditioning (and unconditioning) on the observable variables. We should see that unconditioning, the expected values are naturally different in audited vs. non-audited taxpayers. Conditioning on the observables, however, we expect the conditional expectations to be independent of the audit for the audits generated by the IRA and to be still dependent on the audit for the GdF (consistently with the former relying on the hard information in the Tax Registry and the latter relying on soft information).

Table 3 reports the OLS estimates of regression models where the levels at $t+1$ of the measures of economic performance are regressed on a dummy variable taking value 1 if the taxpayer is audited at t and zero otherwise.[12] Regressions in Panel A and B are shown for all audits, whether initiated by the IRA or the GdF, first unconditionally (panel A) and then conditioning on the information available in the Tax Registry archives when audits are decided (panel B). With no controls, the conditional expectation of the measures of future performance at the province/sector level (panel A) depends on the audit. For example, the audit dummy at t is significant for the level of income, total taxable revenues, VAT taxable turnover and operating costs (both in absolute and relative terms) at $t+1$. The estimated coefficient of the audit variable loses statistical significance in most but not all the regressions when conditioning on data available to the tax authorities at the time of deciding audits (panel B). Decomposing the audits between those chosen by the IRA and by the GdF explains why the audit remains significant after conditioning on observables. When we exclusively focus on audits chosen by the GdF, we observe that the audit dummy is significant for future levels of taxable revenues, VAT taxable turnover and operating costs at the 5% level (panel C) even after controlling for hard information.[13] This reflects the soft information gained on the ground by this police force. When instead we consider only audits chosen by the IRA and we control for the hard data available at t, conditional expectations of the economic variables at $t+1$ are independent from the audit variable (panel D). The control function absorbs all the information in the audit decision of the IRA, apart from a residual random component. Given the institutional design of tax revenue and enforcement agencies discussed in Section 2.1, this quasi-randomness of the audits decision by the IRA conditioning on the data in the Tax Registry should not be surprising. The rest of our study thus focuses on the 96% audits selected by the IRA and is based on regression models

[12]The OLS estimates of regression models using growth rates rather than levels as dependent variables show a similar evidence. They are available upon requests.

[13]This evidence suggests that the soft information available to the GdF allows them to direct audits towards firms with more opportunities to hide larger sums of income (as signaled by correlation with higher levels of evasion in own sector in the area and that with the levels of sales, VAT taxable turnover and operational costs) and that are relatively healty (as signalled by lower ratios of operational costs over the value of production), and thus more likely to pay back any due taxes and fines.

conditioning on the observed information.[14]

3 Exploratory evidence

We start by presenting evidence that tax accountants play a role in tax compliance. To this goal, we estimate the relationship between own tax evasion and the accountant tax evasion, as measured by the average tax evasion of the other clients of the same accountant over the entire period under analysis. We use several specifications of the linear regression model:

$$e_{ijk} = \alpha E_j + \beta controls_{ijk} + \varepsilon_{ijk} \tag{1}$$

where i denotes a tax filing, j denotes the tax accountant, k the fiscal year when the income is produced. The variable e_{ijk} is the share of income evaded by taxpayers i in fiscal year k, that is the difference between the income assessed during the audit and the declared income over the declared income. E_j is the average share of income evaded by the other customers of accountant j. The averages are computed over the entire period and exclude the evasion of taxpayer i for each year k. We include as controls the variables that guide the selection of tax filings during the audit process. We denote this vector $controls_{ijk}$. It contains characteristics of the audited tax filings together with the accountant's number of clients and number of provinces with at least one client.[15] If there is only one other filing different from i that is audited, the vector contains the characteristics of that tax filing. If more than one tax filing is audited, we include the average of their characteristics. The vector $controls_{ijk}$ also includes time fixed effects (capturing common trends in evasion), location fixed effects (picking up systematic differences in the propensity to evade across areas due e.g. to differences in audit costs), and business sector fixed effects (reflecting e.g. differences in ease of hiding income across sectors).[16] Finally, ε_{ijk} are i.i.d., mean 0 innovations.

Table 4 shows the OLS estimates of model 1 with different sets of controls in columns 1 to 3. The sample includes all tax accountants with at least another audited customer. Standard errors are clustered at the accountant level in all regressions. We start in column 1 by adding the gender of the taxpayer (a dummy equal to 1 if taxpayer is a woman), his/her

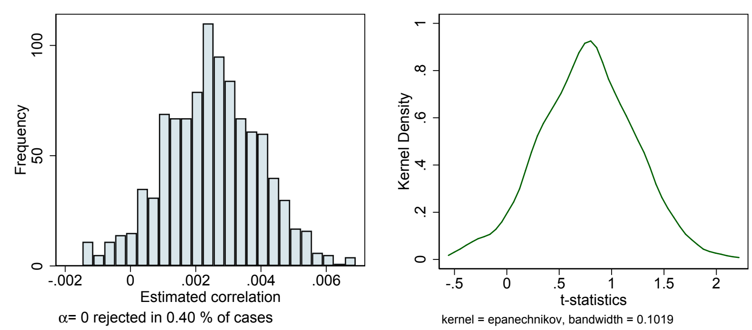
[14] Because the taxpayers audited by the GdF constitute less than 4% of the sample, the summary statistics reported in 1 remain roughly unchanged when excluding those audits from our sample. They are available upon request.

[15] These variables are used as proxies for accountant quality since they capture the accountant's ability to attract client and market coverage. Results remain unchanged if we also include the share of clients by age, firm sector, and firm size. Similarly, results are robust to controlling for the fraction of evaders on audited clients until the year t in which filing i is audited.

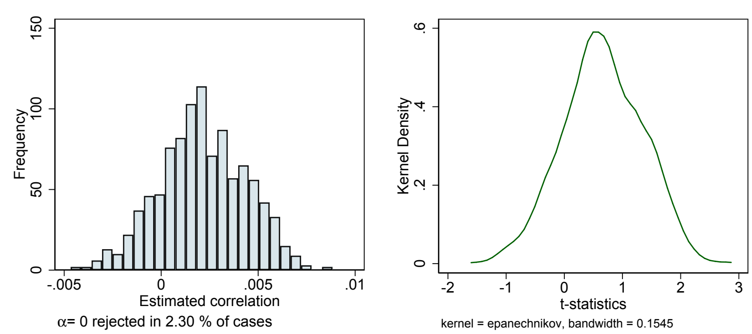
[16] Because an audit is a rare event (see table 1) there is not enough variation in accountant evasion over time to include in the model accountant fixed effects.

age and the age and size of the business he/she manages. In addition, we control for location using province dummies. Results reveal that women and younger people evade higher shares of income, while owners of larger and older firms evade smaller shares of income. The value of α is positive and highly statistically significant. In the second column, we add the indicators *Congruent* and *Coherent* developed by the IRA to gauge a taxpayer's evasion risk discussed above, and include among accountant's characteristics the number of clients and the number of provinces with at least one client (beside the averages of individual variables). As expected, being "congruent" and "coherent" correlates negatively and strongly with the share of evasion. The third column adds municipality fixed effects. Because Italy counts more than 8,000 municipalities, these very granular geographical controls ensure that the correlation between own evasion and the evasion of other clients is not a reflection of omitted local factors. The point estimate of the target effect α is slightly reduced across columns,

Panel A. Random accountant in same province and sector



Panel B. Random accountant in same province and size



Notes. The figures show the distribution of estimated coefficients and t-statistics for the OLS specification in Table 4, column 3 while randomly assigning accountants in the same province and with at least one client in the same sector as the taxpayer (panel A), and in the same province and decile of the taxpayer accountant's number of clients (panel B). The spillover estimate obtained in Table 4, column 3 is 0.132.

Figure 1: Placebo regressions - spillover effect

but it remains strongly statistically significant (p-value 0.0047) and economically relevant. Looking at the specification with the more extensive sets of controls (column 3) one standard deviation increase in the average share of evasion of the the accountant is associated with an higher share of own evasion about 2.7% (about 8% of the mean). In columns 4 and 5 we estimate model 1 on a different sample and using a different estimation strategy. In column 5 we run the regression excluding tax accountant with fewer than 50 customers since average values may be misleading in very small groups. In column 4 we run our regression using a fractional probit specification (Papke and Wooldridge, 1996) since our dependent variable (the share of evasion) has a large number of boundary values equal to 0 or 1. Results show that the evidence remains unchanged in both columns 3 and 4.

We next run placebo regressions replacing E_j with the average share of evasion of the clients of a *different* but *similar* accountant located nearby the taxpayer who filed the tax return. We run 1000 regressions, each time randomly reassigning each taxpayer to a new accountant in the: i) same province and with at least one client in the same sector; or ii) same province and with similar number of clients (i.e. in the same decile of the accountant size distribution). Figure 1 shows the distribution of the estimated α parameter for these placebos, as well as the distribution of the t-statistic of the null $\alpha = 0$. The spillover parameter is significantly different from 0 only in 0.4% of the cases when the first assignment rule is used and in 2.3% of the cases using the second assignment rule. The conclusion is clear: the average share of evasion at accountants other than one's own bears no relation with own share of evasion except by chance. The correlation only arises when taxpayers share the same accountant. This evidence suggests that one's own accountant plays a specific role in tax compliance.

4 Two channels: self-selection and information externalities

The evidence presented in the previous section can be explained in two ways. First, becasue in Italy accountants are not liable for the evasion of their clients there could be sorting: taxpayers who are more willing to evade taxes look for accountants that facilitate these activities or at least are more tolerant in these respects. Taxpayers that are not interested in tax evasion, will not value these "qualities" in a tax accountant and will look instead for accountants who are religious about complying with tax laws. This story does not necessarily require that the tax accountant plays an active role in sharing information or ethical standards among its customers. A second story, is that tax accountants play an active role as information hubs: they collect information from their activities on the auditing strategy of the tax authorities, the cost/benefit trade-offs of tax evasion, and they share it with their customers thus directly affecting their decisions.

The policy implications of the two scenario are different. If only sorting is at work, than audits should be targeted to accountants with higher records of evasion among clients but all clients need to be audited to obtain tax compliance for each of them. If instead the activities of tax accountants generate the voluntary compliance of taxpayers who are not audited, then the number of audits can be reduced for any desired level of tax compliance. These indirect effects may not only help designing cost-effective auditing schemes but also inform the design of informational campaigns

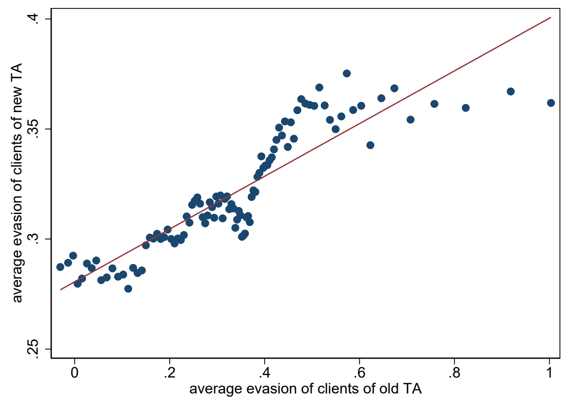
Note that these two stories are not mutually exclusive. In the appendix, we present a simple model of the interaction between taxpayers and tax accountants that features both channels. The model shows formally that both roles can independently or jointly explain the correlation in tax evasion between the customers of a tax accountant. The question whether we are in the presence of self-selection and sorting with a passive role of tax accountants, or of tax accountants as information hubs is an empirical one. We proceed to study this question in the next two sub-sections.

4.1 Self-selection into tax-evasion facilitators

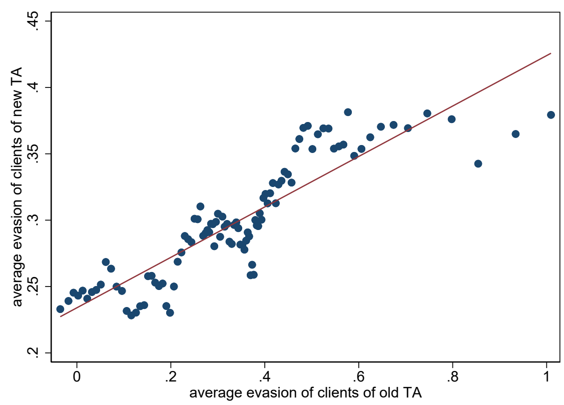
To examine whether self-selection can explain the correlation between own evasion and the evasion of other clients of the same accountant, we look at taxpayers who switch accountants.

First, we look at the correlation between the tax evasion of the accountant before the move and that of the accountant after the move. Sorting implies that, upon moving, a client should match with a new tax accountant with a similar tax-compliance propensity. Figure 2 shows a bin-scatter of the share of evasion of the old and new accountants, after we partial out year, sector and municipality fixed effects as well as a set of movers' characteristics (gender and age of taxpayer, size and duration of the business, year of move). These controls are primarily meant to mitigate possible issues related to the possibility that evasion rates have a local/sectoral component. The first panel shows the (non-parametric) relationship for the whole sample of movers (672,348 taxpayers for which we can compute the average evasion of the clients of the old and new accountant) and is unambiguously strongly positive. Because voluntary moving decisions may be triggered by a variety of reasons unrelated to tax evasion propensity, the second panel shows the same relationship but now computed on the sample of taxpayers who change accountants upon closure of their old one. In our sample, we indeed observe 350,839 forced switches due to the closure of the tax practitioner and for 206,064 of them we can calculate the average evasion of clients of both the old and the new accountant, that is they are cases for which both the old and new accountant have at least one client who has been audited in the observed period. The evidence on positive sorting remains unchnaged.

Panel A: All taxpayers switching accountant



Panel B: Taxpayers switching because of accountant's closure



Notes. The sample includes taxpayers changing accountant at least once in the observed period. The x-axes shows the average share of evasion of the clients of the accountant of origin binned in 100 quantiles. The y-axes reports the average share of evasion of the clients of the new accountant after partially out the characteristics of the taxpayer and his business, year and province fixed effects.

Figure 2: Sorting of taxpayers into accountants

Table 5 confirms the suggestive evidence of sorting presented in Figure 2 using OLS regressions of the evasion at the old accountant on evasion at the new accountant, controlling also for year of move fixed effects and the entire list of audit policy controls at the accountant level for both the old and the new accountants. In the first two columns we run the regressions on the whole sample with different location fixed effects, whereas in column 3 the regressions are run on the sample of movers after accountant closure. The last column presents the results when using a fractional probit estimator. The results in Table 5 show that irrespective of the controls used, sample and estimation strategy, the relationship between the share of evasion of old and that of the new tax accountant is positive and highly statistically significant.

The size of our sample allows us to refine this test further by running our regressions at the individual level and focusing on taxpayers that switch and were audited at least once *before* switching accountants. We can then measure tax evasion of the mover when he/she was served by the old accountant and tax evasion of the clients of the new tax practitioner before the move. Table 6 reports the results. Because we focus on audited switchers, the sample size shrinks to 32,385 taxpayers but remains large enough for reliable inference. The first two columns show regressions for the same specifications as in 5. The correlation between own evasion at the old accountant and average evasion of the clients at the new accountant (measured before the switch occurs) is positive, highly statistically significant and very similar in size to the slope values estimated in Table 5.

To deal with the possibility of endogenous switching we perform two exercises. First, in the third column of Table 6, we limit the sample to audited movers who filed with the previous tax accountant but were audited *after* switching. For example, we consider a taxpayer who was with accountant A until 2010, then switches to accountant B in that year and has his/her 2009 filing audited in 2012 when he is with the new accountant. In this way we deal with one particular source of endogenous switching: the one triggered by a taxpayer being audited (see next section for evidence). The estimated correlation is hardly affected. Second, in the third column of Table 6 we show the OLS estimates for the sample of movers after accountant closure. The sample size shrinks further to 6,533 observations and we lose some precision, but the estimate remains significant (at the 5% level) and of comparable size as in the other columns. In columns 5 and 6 we estimate model 1 when excluding tax accountant with less than 50 customers and using a fractional probit specification (Papke and Wooldridge, 1996), respectively. The evidence remain roughly unchanged across all columns.

Finally, we run two placebo tests to corroborate the evidence presented thus far. In the first, we match a switcher with a new, randomly selected tax accountant in the same province and sector.[17] In the second, we match the switcher with an accountant in the same province and decile of size of customer base as the old accountant. We repeat each test 1000 times, and each time we run the regression in the third column of Table 6 and record the coefficient on the average share of evasion of the previous accountant and its significance. Figure 7

[17] We present the results when defining an accountant in the same sector if the accountant has at least one client in the same sector as the accountant of the switcher. We obtain a similar picture if we use instead the same decile in the share of customers in the same sector.

shows the distribution of the estimated slope parameter and of the corresponding t-statistic. The graph shows that the estimates are small and centered around zero. The coefficients are statistically different from zero only in the 0.8% and 3.9% of the cases, respectively. Both values are much smaller than the actual estimate in Table 6. That is, randomly assigning switchers to another accountant never results in sorting as strong as the one implied by the actual new tax accountant. Overall, the evidence strongly supports the idea that tax accountants play an important role in facilitating tax evasion.

An alternative story would be that the observed correlation between one's evasion and that of others sharing the same tax accountant is determined by the quality of the tax consultant. In this case the results would be driven by errors instead of strategies. Bad quality consultants could make errors simultaneously for many of their clients and when they realize it – after an audit – they can correct it for all. This alternative scenario, however, is not consistent with the evidence in Figure 2, and Tables 5 and 6 that taxpayers moving to new tax consultants end up again with tax consultants whose clients are more likely to evade. In addition, as mentioned in Section 2 we always include in our regressions a set of observable accountant characteristics that are plausibly correlated with accountant quality, such as clientele size and market geographical coverage.

4.2 Tax accountants as information hubs

Once a tax accountant is selected, does she/he play any active role in diffusing auditing information among her/his clients? If the tax accountant acts as an information hub, reported income at t should change (probably increase) if other clients of the same accountant were audited at $t-1$. In addition, besides affecting filed income, tax audits may also affect a taxpayer's decision to switch accountants. We thus look at the effects of audits received by customer i and by other customers served by the same tax accountant on: 1. i's reported income in the years after audit; and on 2. i's decision to change tax accountant.

With respect to the first effect, the first column of Table 7 shows the estimation of a simple regression of log filed taxable income at time t where the only audit variable is an indicator equal to 1 if in the previous year at least one of the clients of the same accountant was audited (while excluding the taxpayer in question), labeled as *peer audit*. In the second column, we include an indicator for whether the taxpayer was audited at $t-1$, labeled as *own audit*. All estimates include taxpayer fixed effects. We are interested in the difference between a taxpayer's compliance behavior if the peers are treated and that of the same taxpayer if the peers happen to receive no treatment. We also control for a set of time-varying taxpayer and accountant observables (marital status, age, size of the business, years of activity, accountant's clientele size and geographical coverage) at the time the audit is received and time fixed effects. Of course, we always include also the audit control variables related to the own tax filing being audited and the average characteristics of those variables for the audited tax filings of peers if more than one peer has a tax filing audited (computed excluding the taxpayer).

The results in column 1 show that the effect of audits of peers is positive and highly statistically significant. Having at least one other client of the same accountant audited triggers a 2.1% increase in filed income in the following year.[18] When an indicator for whether a taxpayer was audited at $t-1$ is added in the regression (column 2), the effect of other customers' audits is somewhat smaller but retains its statistical significance. If a taxpayer was audited at $t-1$, then in the subsequent year he/she reports an higher income to the tax authority, roughly 7.5% higher. The last column of Table 7 shows that audits of different clients of the same accountant are not serially correlated. This suggests that previous year audits on other clients do not mechanically affect the chance of being audited. Still, they provide useful information and affect future declared income because, in the presence of uncertainty on the audit probability, they are used by the tax accountants to update their beliefs on the audit probabilities of their respective clients.

The fact that the effect of own audit is much larger than that of peer audit is not surprising. One's own history of tax audits is obviously the first source of information one draws on to learn about the IRA audit policy and this has a much stronger effect on filed income than other customers' audits. The magnitude of the effects of peers may instead reflect heterogeneity in types among the clients of the accountant. Peer audits may be more informative about own risk of being audited the more similar is the taxpayer to the audited peer. This is because tax accountants may inform their customers that the tax authority is auditing taxpayers with a specific set of individual characteristics. Taxpayers with those specific characteristics may react more strongly, while taxpayers with different characteristics may feel un-targeted and, as a result, not alter their reported income. The average response will then depend on the distribution of these characteristics across clients. To test whether tax accountants share audit information especially with other customers that are similar to the audited taxpayers along some dimensions, we add to the specification interactions between *peer audit* and dummy variables with values 1 if at least one of the audited peers is of the same sector, age, or business type of the taxpayer. Business type is defined according to the EU Commission Recommendation 2003/36, and adopted by the IRA as reference for

[18]A similar positive and significant correlation is obtained if we use the precise number or share of clients audited at $t-1$, including the taxpayer in question, which is part of the set of signals observed by the tax accountant.

fiscal purposes, to define micro-enterprises all businesses with less than 10 employees. Table 8 shows the estimation results. It appears that indeed the reported income is higher for clients of the same accountant with similar traits to those of the audited taxpayers.

An alternative story is that the taxpayer comes to know about these audits directly from the audited peers and not through the tax accountant. To investigate this possibility, we replace the peer audit and similarity dummies with analogous dummies that indicate if other individuals that are *not* clients of the taxpayer's accountant, live in the same province and have similar characteristics (same sector, same age or same business type) have been audited in the previous year. We find that these variables have no explanatory power, suggesting again that it is the information disseminated by one's own accountant that really matters.

The next question that the richness of our data allows us to investigate is whether accountants share information about the audits of own customers with other accountants. To test whether accountants share auditing information, we estimate a regression model like the one in the second column of Table 7 but replace *peer audit* with a dummy equal to 1 if in the previous year at least one of the clients of a *different* but *similar* accountant nearby was audited (labeled as *non-peer audit*). Because we do not know who is in communication with whom, we randomly assign each taxpayer to another accountant in the following pairs: *i*) the same province and sector, *ii*) the same province and decile of size of customer base, *or iii*) the same province and decile of share of audited clients that have evaded. We run 1000 regressions with a new randomly assigned accountant each time. If accountants are informationally connected, we should see a significant effect of audits of other accountants' clients in a large fraction of these regressions. Figure 8 plots the distribution of the estimated parameter and of the corresponding t-statistic. In the vast majority of the cases, 98.9% in case i), 97.6% in case ii) and 99% in case iii), we find no effect on reported income at t of the share of audited customers of a different accountant at $t-1$. Overall we thus find little support for the idea that accountants form an information-sharing network.

In Table 9 we investigate further the information mechanism by examining the persistence of the information effect. In the first column, we include the three lagged values of own audit at $t-1$, at $t-2$ and $t-3$ while controlling for other clients' audits at $t-1$. Interestingly, the effect of own audits is significant at all lags but the size decays over time, albeit slowly: the effect of a three-year old audit on current reported income is still 39% of the effect of a one-year old audit. The cumulative effect of an audit after three years is to increase reported income by 16.6%–twice as much as the one-year lagged effect. In this specification, the effect of an audit of other clients in the last year is significant and of the same size as in Table 7. In the second column, we also allow the audits of peers to affect reported income with lags of up to three years. The three lags are all positive and highly statistically significant. Importantly, once they enter together their size increases considerably. Perhaps most interestingly, the effect of the other audits observed by the accountants on taxpayer reported income is larger for older audits. One potential explanation for this result is that information disseminates with lags. A perhaps more plausible explanation is that details about the IRA policy are revealed as the audits unfold after they have already been notified. The variable for one-year lagged audits of others only captures the information about the IRA's notification of an audit to the taxpayers (and to the accountant), while the two- and three-year old audits also reveal what the IRA investigates. This additional information allows the tax accountants to infer more about the IRA auditing policy. Both because estimated coefficients are larger and because several lags matter, the cumulative effect of the information spillover increases reported income by 10 percent, which is about 60% of the direct effect. When we include audit policy controls for both own and peer audits and for all the different lags the magnitude of the estimated coefficients slightly decreases. The effects, however, remain statistically significant and follow the same patterns over time. The indirect effect is about 17% of the direct effect. This is a non-negligible effect, since the indirect effect is at work for the entire population.

To get a better sense of the quantitative importance of the information channel on reported income, consider increasing the number of audits by one unit for each tax accountant. Our estimates imply that the total cumulative direct effect, over three years, on the reported income of these taxpayers amounts to EUR 1,315 millions, and the information spillover effect amounts to EUR 731 millions – approximately 56% the direct effect.[19]

We now turn to study the effects of tax audits on accountant switches. In Table 10 we study whether tax audits may affect taxpayers' decision to switch accountants in addition to affecting filed income. We use a probit regression model with the same specification of the model used in Table 7 (column 2) but where the dependent variable is a dummy variable equal to 1 if the taxpayer has switched accountants in year t. The results in the first column reveal that the effect of other clients' audits is negative: a taxpayer is less likely to switch accountants if other customers of his/her accountant have been audited. This effect is also quite sizable. If at least one other client of the accountant has been audited, the probability of the taxpayer switching accountants falls by about 6 percentage points.

[19] These effects are estimated as follows. Let α_i be the marginal direct effect of own audit at lag $i = 1, 2, 3$; and let β_i be the marginal effect of the *others* audits at lag $i = 1, 2, 3$. The direct effect is estimated as *Number of audits*$\times \sum \alpha_i \times$*Average income of audited* $=$ 377,113$\times$0.119$\times$29,300 $=$ 1,315 millions euros, where the average income figure is that of the audited taxpayers from Table 1, panel B. The cumulative spillover effect of an extra control is equal to *Number of affected clients of a tax account*ant$\times(\sum \beta_i)\times$*Average income* $=$30.11$\times$0.02)$\times$18,628 $=$ 11,218 euros, where the average income is that of the total sample (Table 1, panel A). The total spillover is obtained by multiplying this number by the number of tax accountants affected by audits (65,133). The total indirect effect is 731 millions euros.

While this negative effect may, prima facie, sound implausible, it is fully consistent with the information dissemination role of the tax accountant. Indeed, taxpayers can come to know that other clients have been audited because their accountant notifies them about the IRA's activities. Taxpayers that become aware of this are less likely to switch accountants for two possible reasons. The first possible reason is the well-known behavioral phenomenon known as the "gambler's fallacy," that is the mistaken belief that, if something happens more frequently than normal during a given period, it will happen less frequently in the future. The second and perhaps more important reason is that the taxpayer who has not been targeted by the audit may appreciate the fact that the tax accountant shares valuable information about the audits with other customers. The effect of own audit is instead positive and significant: a taxpayer that has been audited this year is more likely to switch accountants next year. The increase in the probability of switching is about 0.5 percentage points. One plausible interpretation of this positive effect is that the tax audit signals some incompetence of the accountant to the taxpayer. [20] The other columns of Table 10 enrich the specification by adding interactions between the indicator for audits of other customers and indicators for similarity between the others and the taxpayer to test whether the decision to switch is sensitive to information that is more relevant to the taxpayer's characteristics. Indeed we find that audits on peers have a stronger negative effect on the probability of switching when there is at least one audited customer who is similar to the taxpayer either in terms of sector of activity or business size or age.

Finally, Table 11 studies dynamic effects of audits of others and own audit on the switching decision by adding lags of these variables. The specification is the same as in Table 9 but with a different dependent variable. Results reveal that the effect of own audit is always positive at all lags but fades away with time, while the effect of audits of others is always negative at all lags and its absolute size is either constant or increasing over time. This pattern of effects is very similar to the response in reported income documented in Table 9. The last column reports the results when repeating the estimation on the sample of those whose tax accountant is still active to make sure that the switching decision is not triggered by accountant closure. The results are qualitatively unchanged.

[20] This phenomenon is not necessarily inconsistent with the gambler's fallacy hypothesis described above. However, the disappointment associated with the fact of being audited may overwhelm the effect of the gambler's fallacy in this case.

5 Conclusions and policy implications

Tax codes in advanced countries have become increasingly complex, creating scope for experts' advice. We argue that, depending on the role played, tax intermediaries can have profound effects on the nature of the relationship between tax authorities and taxpayers. Tax accountants may help taxpayers take advantage of the complexity of tax rules and game tax authorities by offering taxpayer-specific counseling on how to minimize income reporting within or outside the boundaries of the tax code. The implication is the emergence of a market for tax advisors where (some) accountants specialize in offering evasion advice to evasion-prone taxpayers. We find strong evidence that evasion-prone taxpayers match with evasion prone-tax accountants, implying that indeed some accountants specialize as tax-evasion facilitators. A smart tax authority should then invest resources to learn the accountants' types, diverting attention from the taxpayers to their intermediaries and auditing with higher probability clients of more evasion-prone accountants. This breaks the direct link between the tax authority and the taxpayers assumed in the traditional literature on tax evasion and compliance (e.g. Allingham and Sandmo, 1972; Graetz et al., 1986). In these models, absent tax intermediaries, taxpayers comply only because they can be audited with some probability and punished if found non-compliant. With tax intermediaries, taxpayers can also be disciplined by the audits of other clients of their own accountant. Accountants may act as information hubs: taxpayers can learn about the tax authority's policy because accountants can pool the audit experiences of many customers over many years and share this information with each of their clients. From the point of view of the taxpayer, this speeds up learning about the tax authority policy function, providing an additional incentive to rely on tax accountants. From the point of view of the tax authority, auditing one taxpayer can, through the information disseminated by the accountant, affect the compliance of the other clients. We find evidence that this is indeed the case. Reported income not only responds positively to a directly experienced audit but also to the audits of the other customers of one's own tax accountant. The size and pattern of responses to the two types of audits is telling: taxpayers' response to own audits is strong on impact but its effect is short-lived and decreases rapidly with time. The response to other clients' audits is milder on impact but persists unchanged over time. One interpretation is that own audits have much greater salience than others' audits, but salience vanishes as distance from the audit increases and a new audit that would maintain high salience is rare. On the other hand, at each point in time, accountants are much more likely than single taxpayers to observe an audit. Passing on this information to their clients increases audit salience. Accountants have the ability to keep track of all previous audits of their clients: information accumulates and becomes more precise as time lapses. Understanding the dynamic response to direct and indirect exposure to audits is both intriguing and of practical relevance to evaluate the effects of audit policies and improve their design. Our analysis moves a first step in this direction.

References

Allingham, M. G. and A. Sandmo (1972). Income tax evasion: a theoretical analysis. *Journal of Public Economics 1*(3-4), 323–338.

Alm, J., K. M. Bloomquist, and M. McKee (2017). When you know your neighbour pays taxes: information, peer effects and tax compliance. *Fiscal Studies 38*(4), 587–613.

Alm, J., B. R. Jackson, and M. McKee (2009). Getting the word out: enforcement information dissemination and compliance behavior. *Journal of Public Economics 93*(3-4), 392–402.

Andreoni, J., B. Erard, and J. Feinstein (1998). Tax compliance. *Journal of Economic Literature 36*(2), 818–860.

Beck, P. J. and W.-O. Jung (1989). The role of tax preparers in tax compliance. *Policy Sciences 22*(2), 167–194.

Boning, W. C., J. Guyton, R. H. Hodge, J. Slemrod, and U. Troiano (2018). Heard it through the grapevine: direct and network effects of a tax enforcement field experiment. *NBER WP 24305*.

Braithwaite, J. (2005). Globalisation, redistribution and tax avoidance. *Public Policy Research 12*(2), 85–92.

D'Agosto, E., M. Manzo, S. Pisani, and F. M. D'Arcangelo (2017). The effect of audit activity on tax declaration: Evidence on small businesses in Italy. *Public Finance Review 46*(1), 29–57.

DeGroot, M. H. (1970). *Optimal Statistical Decisions*. John Wiley and Sons.

Del Carpio, L. (2014). Are the neighbors cheating? Evidence from a social norm experiment on property taxes in Peru. *mimeo, Princeton University*.

Erard, B. (1993). Taxation with representation: an analysis of the role of tax practitioners in tax compliance. *Journal of Public Economics 52*(2), 163–197.

Fortin, B., G. Lacroix, and M.-C. Villeval (2007). Tax evasion and social interactions. *Journal of Public Economics 91*, 2089–2112.

Galbiati, R. and G. Zanella (2012). The tax evasion social multiplier: evidence from Italy. *Journal of Public Economics 96*(5), 485 – 494.

Graetz, M. J., J. F. Reinganum, and L. Wilde (1986). The tax compliance game: toward an interactive theory of law enforcement. *Journal of Law Economics and Organization 2*(1), 1–32.

Hastie, T., R. J. Tibshirani, and M. Wainwright (2015). *Statistical Learning with Sparsity: The Lasso and Generalizations*. Chapman and Hall/CRC.

Klepper, S., M. Mazur, and D. Nagin (1991). Expert intermediaries and legal compliance: the case of tax preparers. *Journal of Law and Economics 34*(1), 205–29.

Kleven, H. J., M. B. Knudsen, C. T. Kreiner, S. Pedersen, and E. Saez (2011). Unwilling or unable to cheat? Evidence from a tax audit experiment in Denmark. *Econometrica 79*(3), 651–92.

OECD (2016). Italy's tax administration: A review of institutional and governance aspects. Technical report, OECD Publishing.

Papke, L. E. and J. M. Wooldridge (1996). Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *Journal of Applied Econometrics 11*, 619–632.

Perez-Truglia, R. and U. Troiano (2018). Shaming tax delinquents. *Journal of Public Economics 167*, 120–137.

Pomeranz, D. (2015). No taxation without information: Deterrence and self-enforcement in the value added tax. *American Economic Review 105*(8), 2539–69.

Raskolnikov, A. (2007). The cost of norms: tax effects of tacit understandings. *University of Chicago Law Review 74*(2).

Reinganum, J. F. and L. L. Wilde (1991). Equilibrium enforcement and compliance in the presence of tax practitioners. *Journal of Law, Economics and Organization 7*(1), 163–181.

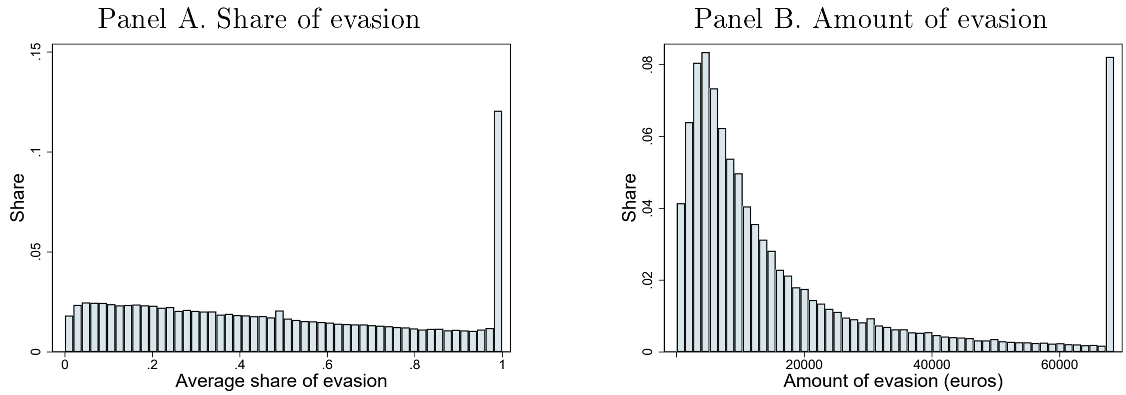
Rincke, J. and C. Traxler (2011). Enforcement spillovers. *The Review of Economics and Statistics 93*(4), 1224–1234.

Scotchmer, S. (1989). Who profits from taxpayer confusion? *Economics Letters 29*(1), 49–55.

Slemrod, J. (1989). The return to tax simplification: an econometric analysis. *Public Finance Quarterly 17*(1), 3–27.

Smith, K. W. and K. A. Kinsey (1987). Understanding taxpaying behavior: a conceptual framework with implications for research. *Law and Society Review 21*(4), 639–663.

A Figures



Notes: Taxpayers with positive evasion. In panel B, values of evasion above the 95-*th* percentile are winsorized.

Figure 3: Distribution of evaders by share and amount of tax evasion

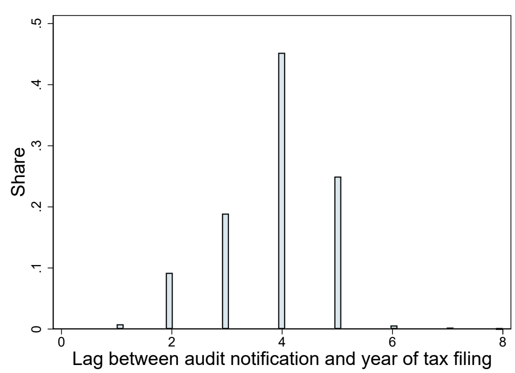
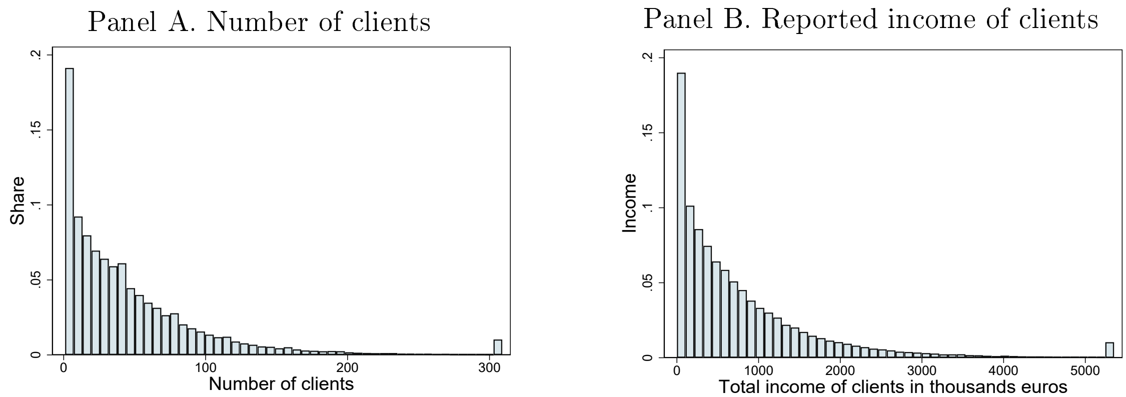
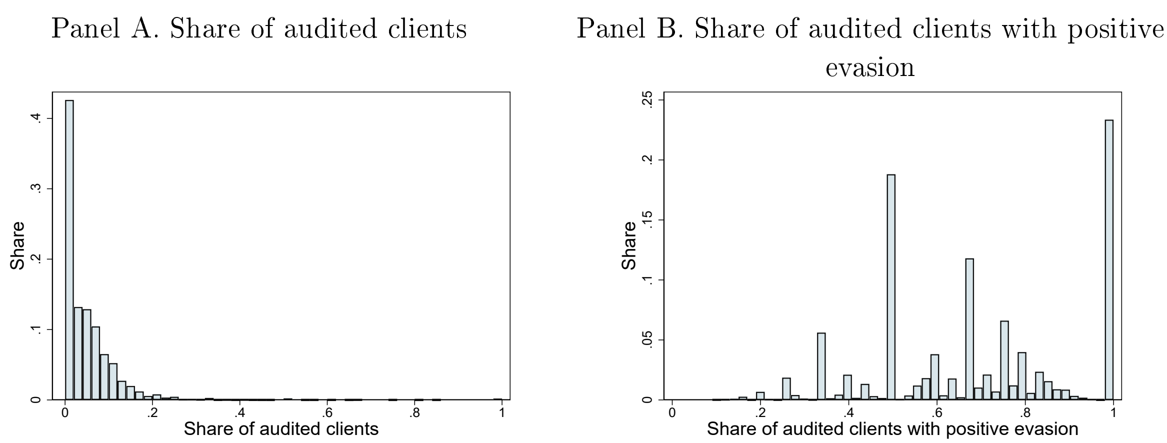


Figure 4: Distribution of audited tax filings by age



Notes: In panel A, the x-axis is the number of unique clients over the entire period; in panel B it is the sum of the average income of the clients over the entire period. For legibility, the mass above the 99-*th* percentile is collapsed in a single class.

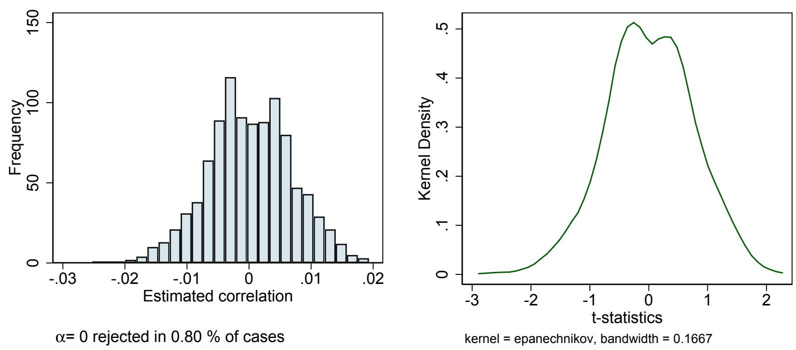
Figure 5: Distribution of tax accountants by size



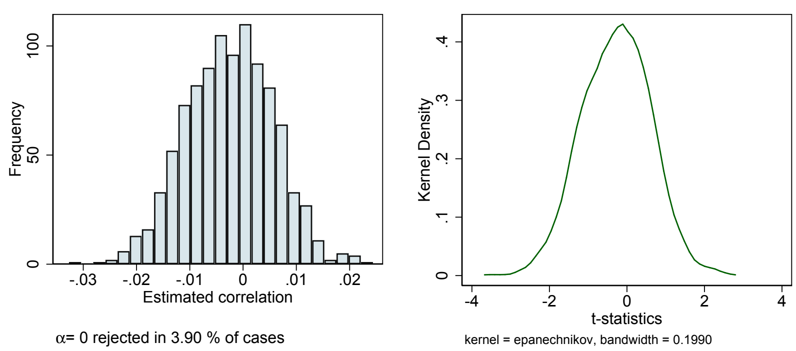
Notes: Panel A includes all tax accountants. In panel B, the sample includes all accountants with at least two clients audited and at least one client with positive evasion.

Figure 6: Distribution of tax accountants by share of audited clients

Panel A. Random accountant in same province and sector



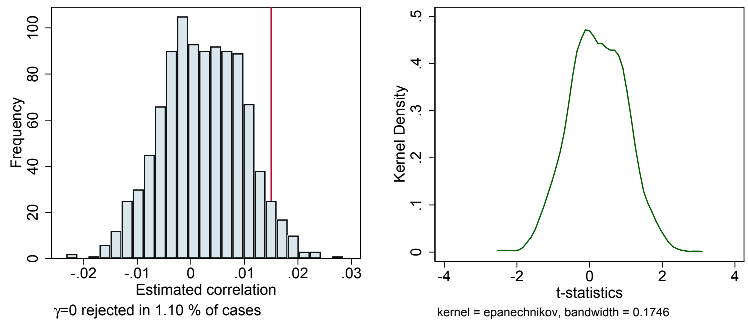
Panel B. Random accountant in same province and same size



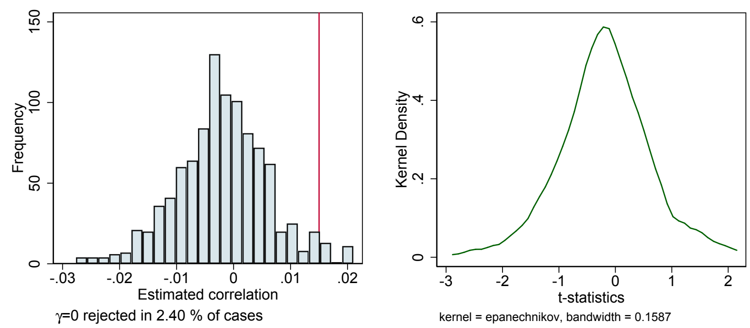
Notes: The figures show the distribution of estimated coefficients and t-statistics from the OLS specification in Table 6, column 3, while randomly assigning accountants in the same province and with at least one client in the same sector as the taxpayer (panel A), and in the same province and decile of the accountant's number of clients (panel B). The estimate of the sorting effect obtained in Table 6, column 3 is 0.055.

Figure 7: Placebo regressions - sorting effect

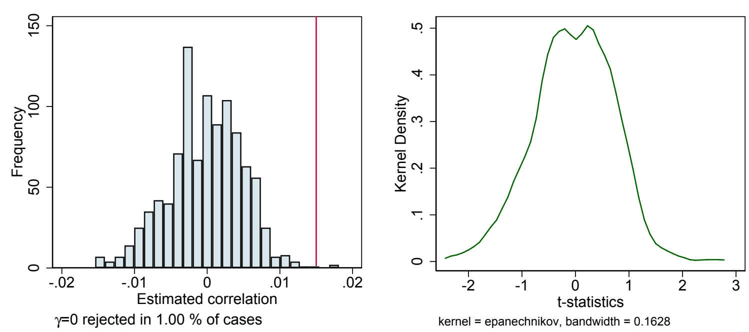
Panel A. Random accountant in same province and sector



Panel B. Random accountant in same province and with same size



Panel C. Random accountant in same province and with same share of evaders



Notes: Figures show the distribution of estimated coefficients γ and t-statistics from the OLS specification in Table 7, column 2, while randomly assigning accountants in the same province and with at least one client in the same sector as the taxpayer (panel A), and in the same province and decile of the taxpayer accountant's number of clients (panel B), and in the same province and decile of number of evaders over the number of audited clients (panel C). The red line represents the actual estimate obtained in Table 7, column 2, which is 0.015.

Figure 8: Placebo regressions - pathway of the spillover effect

B Tables

Table 1: Descriptive statistics

A. Taxpayers

Taxpayers: 4,697,751 Filings: 20,324,271	mean	median	sd	10th pct	90th pct
Woman	0.27	0	0.44	0	1
Married	0.65	1	0.48	0	1
Age	46.77	46	12.45	32	63
Years of activity	13.50	12	10.43	1	29
N. employees	0.83	0	3.19	0	2
Mover	0.07	0	0.25	0	0
Filed income (euros)	18,640.10	10,515	48,693.71	0	39,997
By industry:					
Agriculture	4,286.36	845	28,116.77	0	9,775
Retail	15,175.54	10,002	26,548.88	0	32,342
Manufacturing	16,174.32	13,164	23,7854.85	246	29,856
Private services	23,914.77	12,270	69,011.16	0	49,607
Public services	47,356.02	34,044	54,771.84	2,819	99,356

B. Audits and Evasion

Audited filings: 388,513 ; audited taxpayers: 289,434 Audits by *Guardia di Finanza*: 11,400 (2.93%) Taxpayers audited at least once: 6.16% Taxpayers with positive evasion: 66.45%	mean	median	sd	10th pct	90th pct
Yearly % audited tax filings	1.89	2.42	1.28	0.19	3.35
Yearly % tax filings with positive evasion	52.56	62.90	18.06	20.39	69.65
Yearly % not congruent tax filings	35.06	33.22	5.75	29.47	45.17
Yearly % not coherent tax filings	51.79	51.73	3.85	46.01	56.39
Yearly % taxpayers audited more than once	7.71	7.92	0.82	6.43	8.51
Age of audited tax filings	3.87	4	0.96	2	5
Filed income among audits (euros)	29,602.44	13,560	100,753.95	0	60,327
Evaded income (euros)	20,328.02	4,053	143,316.44	0	36,256
Evaded income among evaders (euros)	32,688.89	10,139	180,624.23	2,524	56,673
Share of evasion on total income	32.59	19.64	35.22	0	94.24

C. Accountants

Accountants: 107,069	mean	median	sd	10th pct	90th pct
N. taxpayers per accountant	31.13	17.57	106.51	2	64
% clients in the same municipality	61.65	61.29	23.14	30.50	96.28
% clients in the same province	89.83	93.77	11.97	73.13	100
% clients in the same region	95.88	98.60	7.18	88.81	100
% clients audited	4.81	3.21	7.28	0	11.54
% evaders on clients	3.11	1.35	5.65	0	8.11
% evaders on audited	64.31	66.67	34.31	0	100
% of new accountants in a year	5.14	5.19	0.72	4.03	6.21
% of closing accountants year	3.67	3.57	0.34	3.35	4.22

Table 2: LASSO model selection

Model	Sample	Non-zero coefficients	Deviance	Deviance Ratio	Pseudo-Rsq
Minimum BIC	Training	33	.0576	.0532	.0532
	Testing		.0561	.0481	.0516
Cross Validation	Training	171	.0565	.0717	.0692
	Testing		.0554	.0597	.0722
Adaptive LASSO	Training	126	.0565	.0707	.0706
	Testing		.0555	.0588	.0717

Notes: Goodness of fit measures of alternative probit LASSO models of a binary variable with value one if the filing is audited on all available information on filings (273 variables). The sample includes all filings at risk of audit in any possible year of audit. Because of the computationally intensive LASSO procedure and the very large size of our sample, the model selection exercise is performed on a 1% random extraction of the sample. Models are trained of a sample of 503,721 observations (50% of the random sample) and then tested out-of-sample on the remaining 503,720 observations. Postselection coefficients are considered. The table reports the number of non-zero coefficients selected by each model and different measures of fit (Hastie et al., 2015). The last column displays the pseudo R-squared of probit models using the covariates selected by the LASSO. The range of reported pseudo R-squared includes that of a probit model using as covariates the audit policy controls in our baseline estimates (158 variables, Pseudo R-squared in the testing sample equal to 0.0556).

Table 3: Balance tests

Panel A: All audits						
Dep. Var. at *t+1*:	Income	Share of evasion	Tot. taxable revenues	VAT taxable turnover	Operating costs	Operating costs/NPV
Audited at *t*	-705.504**	-0.005*	2,450.539***	1,881.596***	2,141.219***	1.531***
	(299.423)	(0.003)	(768.985)	(712.426)	(708.736)	(0.331)
Audit Policy Controls	no	no	no	no	no	no

Panel B: All audits						
Dep. Var. at *t+1*:	Income	Share of evasion	Tot. taxable revenues	VAT taxable turnover	Operating costs	Operating costs/NPV
Audited at *t*	-4.328	0.002*	226.286	-12.746	159.856	-0.013
	(31.300)	(0.001)	(262.234)	(147.688)	(109.292)	(0.113)
Audit Policy Controls	yes	yes	yes	yes	yes	yes

Panel C: Audits by GdF						
Dep. Var. at *t+1*:	Income	Share of evasion	Tot. taxable revenues	VAT taxable turnover	Operating costs	Operating costs/NPV
Audited at *t*	-39.033	0.006**	1,169.509**	898.333*	684.581*	-0.742**
	(85.342)	(0.002)	(579.2)	(541.044)	(371.77)	(0.355)
Audit Policy Controls	yes	yes	yes	yes	yes	yes

Panel D: Audits by IRA						
Dep. Var. at *t+1*:	Income	Share of evasion	Tot. taxable revenues	VAT taxable turnover	Operating costs	Operating costs/NPV
Audited at *t*	-1.053	0.002	141.499	112.603	-94.774	0.053
	(31.884)	(0.001)	(262.49)	(109.159)	(145.493)	(0.129)
Audit Policy Controls	yes	yes	yes	yes	yes	yes

Notes. OLS estimates with standard errors clustered at the province and sector level (in parentheses). The estimates displayed in each row are for separate regressions in which the dependent variable is the variable named in the heading column by province and sector and the independent variable is displayed in the row. NPV stands for net present value. Audit policy controls, include the characteristics of the tax filing, taxpayer and tax accountants at the time of the filing, sector and province fixed effects, year of filing fixed effects, and age of filing fixed effects. *, **,*** denote statistical significance at the 10, 5 and 1 percent level.

Table 4: Tax evasion spillovers

Dep. Var.: Own Evasion	(1)	(2)	(3)	Robustness checks (4)	(5)
Accountant Evasion	0.163***	0.155***	0.132***	0.182***	0.153***
	(0.006)	(0.006)	(0.006)	(0.010)	(0.005)
Audit policy controls own					
Woman entrepreneur	0.015***	0.008***	0.008***	0.008***	0.008***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Married entrepreneur	-0.030***	-0.025***	-0.025***	-0.027***	-0.025***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Entrepreneur 31-50 y.o.	-0.029***	-0.020***	-0.019***	-0.015***	-0.015***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)
Entrepreneur >50 y.o.	-0.022***	-0.019***	-0.019***	-0.012***	-0.017***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)
Years of activity	-0.002***	-0.002***	-0.002***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Firm size: 1-5	-0.024***	-0.015***	-0.015***	-0.014***	-0.015***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Firm size: 6-10	-0.023***	-0.005	-0.005	-0.006	-0.007**
	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)
Firm size: 11-15	-0.027***	-0.009	-0.008	-0.005	-0.010**
	(0.005)	(0.005)	(0.005)	(0.007)	(0.005)
Firm size: 16-20	-0.029***	-0.009	-0.008	0.003	-0.009
	(0.008)	(0.008)	(0.008)	(0.010)	(0.008)
Firm size: >20	-0.034***	-0.016**	-0.014*	-0.007	-0.020***
	(0.008)	(0.008)	(0.008)	(0.010)	(0.007)
Congruent		-0.109***	-0.108***	-0.104***	-0.104***
		(0.002)	(0.002)	(0.002)	(0.002)
Coherent		-0.096***	-0.095***	-0.094***	-0.101***
		(0.002)	(0.002)	(0.002)	(0.002)
Accountant: n. of clients/1000		0.002***	0.002***	0.002***	-0.001
		(0.000)	(0.000)	(0.000)	(0.001)
Accountant: n. of provinces		0.001***	0.001***	0.001**	0.001**
		(0.000)	(0.000)	(0.000)	(0.000)
Year of filing FE	yes	yes	yes		yes
Age of filing FE	yes	yes	yes		yes
Sector FE	yes	yes	yes		yes
Province FE	yes	yes			yes
Municipality FE			yes	yes	
Audit policy controls accountant	yes	yes	yes	yes	yes
Sample: Accountants with >50 clients				yes	
R-squared	0.097	0.143	0.171	0.178	
Pseudo R-squared					.067
N.Obs.	322,043	322,043	321,576	195,433	333,207

Notes: Columns 1 - 4 report OLS estimates, column 5 reports marginal effects of fractional probit model estimates. Standard errors are clustered at the tax accountant level (in parentheses). The sample includes all audited tax filings of clients of accountants with at least another audited client. Audit policy controls at the level of accountant are computed as averages of the variables listed in Table 3 for the audited clients. Baseline categories: age of entrepreneur: 30 years or younger, firm size: no employee. *, **,*** denote statistical significance at the 10, 5 and 1 percent level.

Table 5: Sorting channel - Aggregate evidence

Dep. Var.: Old accountant evasion	(1)	(2)	(3)	(4)
New accountant evasion	0.055***	0.042***	0.103***	0.106***
	(0.006)	(0.005)	(0.015)	(0.016)
Audit policy controls old accountant	yes	yes	yes	yes
Audit policy controls new accountant	yes	yes	yes	yes
Year of move FE	yes	yes	yes	yes
Province FE	yes			yes
Municipality FE		yes	yes	
Sample: Old accountant closed			yes	yes
R-squared	0.181	0.218	0.340	
Pseudo R-squared				.033
N.Obs.	672,332	672,144	205,353	206,064

Notes: Columns 1 - 3 report OLS estimates, column 4 reports marginal effects of fractional probit model estimates. Standard errors are clustered at the tax accountant level (in parentheses). The sample includes taxpayers who changed accountant at least once and whose old and new accountant have at least one client audited during the sample period. Audit policy controls at the level of accountant are computed as averages of the variables listed in Table 3 for the audited filings of clients compiled before the move. *, **,*** denote statistical significance at the 10, 5 and 1 percent level.

Table 6: Sorting channel - Disaggregated evidence

					Robustness checks	
Dep. Var.: Evasion of mover before move	(1)	(2)	(3)	(4)	(5)	(6)
New accountant evasion before move	0.056***	0.048***	0.055***	0.050**	0.074***	0.067***
	(0.009)	(0.010)	(0.012)	(0.021)	(0.017)	(0.017)
Audit policy controls mover	yes	yes	yes	yes	yes	yes
Audit policy controls new accountant	yes	yes	yes	yes	yes	yes
Year of move FE	yes	yes	yes	yes	yes	yes
Province FE	yes					yes
Municipality FE		yes	yes	yes	yes	
Sample: No audit before move			yes			
Sample: Old accountant closed				yes		yes
Sample: Accountant with >50 clients					yes	
R-squared	0.127	0.237	0.256	0.299	0.270	
Pseudo R-squared						.065
N.Obs.	32,385	30,788	20,069	6,533	16,769	7,882

Notes: Columns 1 - 4 and 6 report OLS estimates, column 5 reports marginal effects of fractional probit model estimates. Standard errors are clustered at the tax accountant level (in parentheses). The sample includes taxpayers who changed accountant at least once and were audited at least once before the switch. Audit policy controls at the level of accountant are computed as averages of the variables listed in Table 3 for the audited filings of clients compiled before the move. Audit policy controls mover are the averages of characteristics of the audited filings compiled by the mover before the move. *, **,*** denote statistical significance at the 10, 5 and 1 percent level.

Table 7: Information sharing channel

Dep. Var.:	(1) Income at t (logs)	(2) Income at t (logs)	(3) Audit at t
Peer audit at t-1	0.021***	0.015***	-0.000
	(0.003)	(0.003)	(0.000)
Own audit at t-1		0.075***	-0.156***
		(0.004)	(0.001)
Controls at t			
Married entrepreneur	0.354***	0.302***	0.003***
	(0.007)	(0.007)	(0.000)
Age of entrepreneur	-0.115***	-0.115***	-0.006***
	(0.002)	(0.001)	(0.000)
Years of activity	0.102***	0.072***	0.003***
	(0.002)	(0.001)	(0.000)
Size of the firm	0.025***	0.016***	0.001***
	(0.002)	(0.001)	(0.000)
Accountant: n. clients/1000	-0.021***	-0.021***	-0.000
	(0.007)	(0.005)	(0.000)
Accountant: n. of provinces	0.001	0.001	-0.000
	(0.001)	(0.001)	(0.000)
Year of filing FE	yes	yes	yes
Taxpayer FE	yes	yes	yes
Audit policy controls peer	yes	yes	yes
Audit policy controls own		yes	yes
R-squared	0.679	0.686	0.340
N.Obs.	15,921,793	13,928,480	13,928,480

Notes: OLS estimates with standard errors clustered at the tax accountant level (in parentheses). The sample includes all tax filings. Time-varying characteristics in the year of filing are added. Audit policy controls for peer and own audit include the average characteristics listed in Table 3 of the tax filings audited in the previous year. *, **,*** denote statistical significance at the 10, 5 and 1 percent level.

Table 8: Information sharing channel: peers vs accountant

Dep. Var.: Income at t (logs)	(1)	(2)	(3)	(4)	(5)	(6)
Peer audit at t-1	0.013***	0.012***	0.015***	0.015***	0.015***	0.015***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Own audit at t-1	0.071***	0.073***	0.074***	0.075***	0.075***	0.075***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Peer audit at t-1 same sector	0.022***					
	(0.004)					
Peer audit at t-1 same type		0.012***				
		(0.003)				
Peer audit at t-1 same age			0.022***			
			(0.006)			
Non-peer audit at t-1 nearby same sector				0.048		
				(0.068)		
Non-peer audit at t-1 nearby same type					-0.051	
					(0.265)	
Non-peer audit at t-1 nearby same age						0.053
						(0.050)
Controls at t						
Married entrepreneur	0.302***	0.302***	0.302***	0.302***	0.302***	0.302***
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Age of entrepreneur	-0.114***	-0.114***	-0.115***	-0.115***	-0.115***	-0.115***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Years of activity	0.072***	0.072***	0.072***	0.072***	0.072***	0.072***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Size of the firm	0.016***	0.017***	0.016***	0.016***	0.016***	0.016***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Accountant: n. of clients	-0.021***	-0.021***	-0.021***	-0.021***	-0.021***	-0.021***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Accountant: n. of provinces	0.001	0.001	0.001	0.001	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Year of filing FE	yes	yes	yes	yes	yes	yes
Taxpayer FE	yes	yes	yes	yes	yes	yes
Audit policy controls peer audit	yes	yes	yes	yes	yes	yes
Audit policy controls own audit	yes	yes	yes	yes	yes	yes
R-squared	0.686	0.686	0.686	0.686	0.686	0.686
N.Obs.	13,928,480	13,928,480	13,928,480	13,928,480	13,928,480	13,928,480

Notes: OLS estimates with standard errors clustered at the tax accountant level (in parentheses). The sample includes all tax filings of taxpayers filing in two consecutive years. “Peer” denotes other clients of the *same* accountant, “Non-peer” denotes other clients of a *different* accountant with similar characteristics. Time-varying characteristics in the year of filing are added. Audit policy controls for peer and own audit include the average characteristics listed in Table 3 of the tax filings audited in the previous year. *, **,*** denote statistical significance at the 10, 5 and 1 percent level.

Table 9: Memory of information

Dep. Var.: Income at t (logs)	(1)	(2)	(3)
Peer audit at t-1	0.012***	0.021***	0.006*
	(0.003)	(0.004)	(0.003)
Peer audit at t-2		0.041***	0.006*
		(0.004)	(0.004)
Peer audit at t-3		0.038***	0.008**
		(0.004)	(0.004)
Own audit at t-1	0.083***	0.083***	0.078***
	(0.005)	(0.006)	(0.006)
Own audit at t-2	0.049***	0.051***	0.027***
	(0.005)	(0.006)	(0.006)
Own audit at t-3	0.033***	0.032***	0.014**
	(0.005)	(0.005)	(0.005)
Controls at t			
Married entrepreneur	0.147***	0.145***	0.162***
	(0.009)	(0.009)	(0.009)
Size of the firm	0.013***	0.013***	0.010***
	(0.001)	(0.001)	(0.001)
Accountant: n. of clients/1000	-0.042***	-0.051**	-0.015*
	(0.016)	(0.021)	(0.008)
Accountant: n. of provinces	0.006***	0.006**	0.000
	(0.002)	(0.002)	(0.001)
Year of filing FE			yes
Taxpayer FE	yes	yes	yes
Audit policy controls peer at t-1, t-2, t-3			yes
Audit policy controls own at t-1, t-2, t-3			yes
R-squared	0.739	0.743	0.746
N.Obs.	7,526,420	7,044,423	7,044,423
F tests - P-values			
$H_0 : \beta^{own}_{t-1} = \beta^{own}_{t-2} = \beta^{own}_{t-3}$	.000	.000	.000
$H_0 : \beta^{peer}_{t-1} = \beta^{peer}_{t-2} = \beta^{peer}_{t-3}$	-	.000	.914
$H_0 : \beta^{peer}_{t-2} = \beta^{peer}_{t-3}$	-	.567	.710

Notes: OLS estimates with standard errors clustered at the tax accountant level (in parentheses). In column 1 the sample includes taxpayers filing in two consecutive years; in column 2 taxpayers filing in four consecutive years. Time-varying characteristics in the year of filing are added. Audit policy controls for peer and own audit include the average characteristics listed in Table 3of the tax filings audited in the previous years. *, **,*** denote statistical significance at the 10, 5 and 1 percent level.

Table 10: Audit and accountant switches

Dep. Var.: Moved to a new accountant at t	(1)	(2)	(3)	(4)
Peer audit at t-1	-0.062***	-0.054***	-0.044***	-0.060***
	(0.003)	(0.002)	(0.002)	(0.002)
Own audit at t-1	0.005***	0.010***	0.010***	0.007***
	(0.000)	(0.001)	(0.001)	(0.000)
Peer same sector audited at t-1		-0.024***		
		(0.002)		
Peer same size audited at t-1			-0.026***	
			(0.001)	
Peer same age audited at t-1				-0.030***
				(0.002)
Controls at t				
Woman entrepreneur	0.004***	0.004***	0.004***	0.004***
	(0.000)	(0.000)	(0.000)	(0.000)
Married entrepreneur	-0.003***	-0.002***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Age of entrepreneur	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Years of activity	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Size of the firm	0.000***	0.000***	-0.000**	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Accountant: n. of clients/1000	-0.005***	-0.005***	-0.005***	-0.004**
	(0.002)	(0.002)	(0.002)	(0.002)
Accountant: n. of provinces	0.001***	0.002***	0.002***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
Year of move FE	yes	yes	yes	yes
Sector FE	yes	yes	yes	yes
Province FE	yes	yes	yes	yes
Audit policy controls peer	yes	yes	yes	yes
Audit policy controls own	yes	yes	yes	yes
Pseudo R-squared	0.046	0.051	0.052	0.049
N.Obs.	14,697,629	14,697,629	14,697,629	14,697,629

Notes: Marginal effects of probit models estimates with standard errors clustered at the tax accountant level (in parentheses). The sample includes all tax filings. Audit policy controls for peer and own audit include the average characteristics listed in Table 3 of the tax filings audited in the previous year. *, **,*** denote statistical significance at the 10, 5 and 1 percent level.

Table 11: Dynamic effects of audit on accountant switches

Dep. Var.: Moved to new accountant at t	(1)	(2)	(3)	(4)
Peer audit at t-1	-0.585***	-0.125***	-0.082***	-0.073***
	(0.022)	(0.010)	(0.013)	(0.011)
Peer audit at t-2		-0.128***	-0.094***	-0.072***
		(0.011)	(0.014)	(0.009)
Peer audit at t-3		-0.192***	-0.158***	-0.129***
		(0.011)	(0.012)	(0.008)
Own audit at t-1	0.093***	0.102***	0.088***	0.090***
	(0.004)	(0.005)	(0.005)	(0.005)
Own audit at t-2	0.070***	0.078***	0.074***	0.078***
	(0.004)	(0.004)	(0.005)	(0.005)
Own audit at t-3	0.036***	0.044***	0.045***	0.046***
	(0.004)	(0.004)	(0.005)	(0.005)
Controls at t				
Married entrepreneur	-0.058***	-0.066***	-0.065***	-0.077***
	(0.002)	(0.002)	(0.002)	(0.002)
Size of the firm	0.002***	0.003***	0.003***	0.004***
	(0.001)	(0.000)	(0.000)	(0.000)
Accountant: n. of clients/1000	-0.053*	-0.064**	-0.062**	-0.074***
	(0.031)	(0.030)	(0.030)	(0.023)
Accountant: n. of provinces	0.012***	0.013***	0.014***	0.015***
	(0.002)	(0.002)	(0.002)	(0.002)
Year of move FE			yes	yes
Province FE	yes	yes	yes	yes
Audit policy controls peer at t-1, t-2, t-3			yes	yes
Audit policy controls own at t-1, t-2, t-3			yes	yes
Sample: Old accountant still active				yes
Pseudo R-squared	0.044	0.018	0.025	0.022
N.Obs.	7,902,444	7,462,868	7,462,868	7,419,000

Notes: Marginal effects of probit models estimates with standard errors clustered at the tax accountant level (in parentheses). In column 1 the sample includes taxpayers filing in two consecutive years; in columns 2 to 4 taxpayers filing in four consecutive years. Time-varying characteristics in the year of filing are added. Audit policy controls for peer and own audit include the average characteristics listed in Table 3 of the tax filings audited in the previous years. *, **,*** denote statistical significance at the 10, 5 and 1 percent level.

C A model of tax professionals and their clients

We present here a simple theoretical framework to study the relationship between the tax authority, the tax accountants and their clients. The goal is to rationalize the social spillovers among the clients of the same tax accountant and thus to formally derive the testable hypotheses that are studied in the paper. In the model, accountants optimally choose the evasion of the taxpayers conditional on the information that, thanks to their role, they can aggregate by observing several realizations of tax audits; and they use it to anticipate the IRA auditing probabilities. In turn, the tax authority chooses these probabilities optimally to maximize tax revenues net of auditing cost. The model predicts a positive relationship between individual tax evasion and average tax evasion of the other clients of the tax accountant, highlighting two mechanisms: self-selection of tax payers into accountants with heterogeneous attitudes about tax evasion, and informational externalities generated in the tax accountant's activities.

Setup Assume there is a continuum of taxpayers with mass one. Taxpayer i chooses whether to report all income or to evade e_i dollars. Taxpayer i's utility of choosing a level of tax evasion $e_i \geq 0$ is assumed to be:

$$u(e_i, m_i) = [(1 - p_i)(e_i) - p_i (Te_i)] - F(e_i) - e_i^2/(2m_i). \tag{2}$$

where p_i is the probability of being audited and discovered and Te_i is the cost of being audited with evasion e_i. The term in brackets is the net expected benefit of the tax evasion; the second term, $F(e_i)$, is the fee to be paid to a tax accountant who prepares the tax returns, that may be a function of e_i (in the remainder we assume all taxpayers need a tax accountant).[21] The last term, $e_i^2/(2m_i)$, is the "ethical" cost of violating the law, where m_i is the taxpayer's type: the lower is m_i, the more costly it is to violate the law. We assume that the types m_i are uniformly distributed in $[\underline{m}, \overline{m}]$.

We assume that there is a finite number J of tax accountants with heterogeneous dispositions to allow their customers to violate the law. Specifically, the utility of an accountant who chooses a level of evasion e_i for a customer i is:

$$U(F, e_i, d_j) = F(e_i) - \frac{(e_i)^2}{2d_j} \tag{3}$$

where the term $\frac{(e_i)^2}{2d_j}$ is again the "ethical" cost of allowing evasion, where d_j is the accoun-

[21] In our dataset, only 3.4% of taxpayers choose to file without a tax accountant.

tant's type. Accountants are ordered according to their disposition to violate the law with $d_l > d_k$ if $l > k$, $d_j \in \left[\underline{d}, \overline{d}\right]$. With a slight abuse of notation, we denote the set of accountants as J. For simplicity, we assume that the d_j's are equidistant (i.e. $|d_j - d_k| = \Delta d$ for some $\Delta d > 0$) and taxpayers are uniformly distributed around them (as in Figure 9 below). If we denote as S_j the set of taxpayers m_i who are closest to d_j, then this assumption implies $|S_l| = |S_k|$ for all k, l.[22] Taxpayers and accountants share the net expected monetary benefit of the tax evasion, with the accountant receiving a fraction α of it, implying $F(e_j) = \alpha (1 - (1 + T) p_j) e_j$, which can be positive or negative.

We assume that neither the tax authority nor the tax accountant can observe the taxpayers' types. The tax authority may observe the accountant type and target its auditing effort to specific types of accountants.[23] The tax authority chooses the auditing rate to maximize expected revenue collection net of the cost of the auditing. If z_j dollars are spent in auditing a taxpayer assisted by accountant j and a level of evasion e_j, the probability of discovery is equal to $p(z_j) = \sqrt{z_j}$. The expected benefit for the tax authority is $p(z_j) \cdot (Te_j + \xi_j)$. The variable ξ_j is an i.i.d. realization reflecting idiosyncratic factors concerning j that may affect the tax authority's decision.[24] We assume the distribution of ξ_j is a truncated normal that takes only non-negative values, with mean $\overline{\xi} > 0$ and variance $1/r$. For simplicity, we assume r is sufficiently large that for all practical purposes ξ_j can be assumed to be normal with $\overline{\xi} > 0$ and variance $1/r$, which allows us to simplify the analysis. The cost of the audit is λz_j, where λ is the shadow cost of public funds. Naturally it must be that $z_j \leq 1$ for all j (or else the probabilities of a discovery will be higher than one). In the following, we assume that, as is natural, λ is sufficiently large such that this is always true. For simplicity, we will therefore ignore the constraint $z_j \leq 1$ going forward.

The timing of the game is as follows. In the first stage the tax authority chooses the probability of auditing a taxpayer. As mentioned, this probability can be contingent on the accountant type but not on the taxpayer type m_i, which is unobservable (taxpayers are otherwise identical). The model can be easily generalized to allow for different observable classes of taxpayers (say different regions, etc.), and we will discuss this extension in greater detail below. In the second stage, the taxpayers choose a tax accountant without observing the tax authority's auditing strategy. In the third stage, each tax accountant j observes L informative signals $\mathbf{s_j} = (s_{j,l})_{l=1}^{L}$ on the auditing probability p_j and chooses the level of tax

[22] That is $|m_i - d_j| \leq |m_i - d_k|$ for all $k \in J$.

[23] Intuitively, the tax authority can observe the tax accountant's activity with many clients over time, thus it can collect more accurate information on the tax accountant's type.

[24] As discussed above, many factors affect the decision to audit a taxpayer, including the business cycle and the sector in which the taxpayer operates. To these factors, we can add other unobserved factors such as the availability of tax inspectors and general guidelines periodically sent by the Treasury.

evasion for each customer. We assume that each signal $s_{j,l} = p_j^* + \varepsilon_l$, $l = 1, ..., L$ where ε_l is an i.i.d. normal random variable with mean zero and variance $1/k$, and p_j^* is the actual auditing probability. The idea is that the tax accountant can infer this probability by observing a small sample from his/her audited clients.

We study the perfect Bayesian equilibria in pure strategies of this game. A strategy for a taxpayer is a function $\varphi(m_i)$ mapping the taxpayer's type to a tax accountant j. A strategy for the tax authority is an allocation of available resources $\mathbf{z} = (z_j)_{j=1}^J$ such that $z_j \geq 0$ given the observed vector of shocks ξ_j.[25] A strategy for a tax accountant is given by a pair of functions $e(d_j, \mathbf{s}_j)$ and $\mu(d_j, \mathbf{s}_j)$. The function $e(d_j, \mathbf{s}_j)$ maps the accountant's type and the observed vector of signals $\mathbf{s}_j = (s_{j,1}, ..., s_{j,L})$ to a level of tax evasion in $[0, \infty]$. The function $\mu(d_j, \mathbf{s}_j)$ maps the observed vector of signals to a posterior distribution on the level of auditing chosen by the tax authority. This belief is part of the equilibrium because it depends on the tax accountants beliefs on the tax authority's auditing strategy, given the observed signals.

Equilibrium behavior We solve the game by backward induction. In the last stage, the accountant of type d_j chooses e_j to maximize (3) given $F(e_j)$. From the first order condition we obtain:

$$e(d_j, \mathbf{s}_j) = \begin{cases} \alpha d_j \left(1 - (1+T) E\left[p(z_j, \xi_j); \mathbf{s}_j\right]\right) & E\left[p(z_j, \xi_j); \mathbf{s}_j\right] < 1/(1+T) \\ 0 & else \end{cases} \tag{4}$$

where $E\left[p(z_j, \xi_j); \mathbf{s}_j\right]$ is the expected level of auditing. The accountant chooses a positive level of tax evasion only if the expected probability of auditing or the penalty T are sufficiently small. In this case, the level of tax evasion is decreasing in $E\left[p(z_j, \xi_j); \mathbf{s}_j\right]$.

To find the equilibrium, we follow a guess-and-verify approach where we first assume that the equilibrium belief $\mu(d_j, \mathbf{s}_j)$ is that $p(z_j, \xi_j)$ follows a normal distribution with mean equal to A_j and precision B_j. We will then verify that this expectation is correct in equilibrium.

When the accountant believes that $p(z_j, \xi_j)$ is a $N(A_j, B_j)$ random variable, Bayes' rule implies that the posterior probability of the probability of auditing, conditional on the sample $\mathbf{s}_j$, is normally distributed with mean and variance:

$$E\left[p(z_j, \xi_j); \mathbf{s}_j\right] = \Phi(\overline{s}_{j,L}) = \frac{A_j B_j + Lk \cdot \overline{s}_{j,L}}{B_j + Lk} \tag{5}$$

$$Var\left[p(z_j, \xi_j); \mathbf{s}_j\right] = \left(B_j^{-1} + Lk\right)^{-1} \tag{6}$$

[25] Implicitly, the tax authority has a budget R such that $\sum z_j = R$. This is captured by the fact that the cost of choosing z_j is λ, a parameter that can be interpreted as the Lagrangian multiplier associated with the budget.

where $\overline{s}_{j,L}$ is the sample mean of the L signals.[26] Intuitively, the posterior belief is an average of the equilibrium belief on the strategy followed by the tax authority and the evidence collected in the field, i.e. the signals $\mathbf{s}_j$.

In the second stage, the tax authority chooses the amount to spend on auditing j's customers, z_j, given the equilibrium strategy and beliefs of the tax accountants as described by (4) and (5). The tax authority's problem can be directly written as:

$$\max_{\mathbf{z}\geq 0} \sum\nolimits_j E\left\{\sqrt{z_j}\left[\alpha d_j T\left(1-(1+T)\,\Phi_j(\overline{s}_{j,L})\right)+\xi_j\right]-\lambda z_j\right\} \tag{7}$$

where the expectation reflects the fact that the tax authority does not know the actual sample of signals $\overline{s}_{j,L}$ observed by the consultant. Note that we have:

$$E\Phi(\overline{s}_{j,L}) = \frac{A_j B_j + Lk\cdot\left(E\overline{s}_{j,L}\right)}{B_j+Lk} = \frac{A_j B_j + Lk\cdot\left(\sqrt{z_j}\right)}{B_j+Lk} \tag{8}$$

Substituting (8) in (7), the authority's problem can be directly written in terms of the auditing probabilities $\mathbf{p}=(p_j)_{j=1}^{J}$:

$$\max_{\mathbf{p}\geq\mathbf{0}} \sum\nolimits_j \left[\begin{array}{c} p_j\left[\alpha d_j T\left(1-(1+T)\,\frac{A_j B_j}{B_j+Lk}\right)+\xi_j\right] \\ -\left(\frac{\alpha d_j T(1+T)Lk}{B_j+Lk}+\lambda\right)p_j^2 \end{array}\right]$$

Because we assumed above that ξ_j is positive with arbitrarily high probability, we will ignore for now the cases in which $\xi_j<0$, and so:

$$p(z_j^*,\xi_j) = \frac{\alpha d_j T\left(1-(1+T)\,\frac{A_j B_j}{B_j+Lk}\right)}{2\left(\frac{\alpha d_j T(1+T)Lk}{B_j+Lk}+\lambda\right)} + \frac{\xi_j}{2\left(\frac{\alpha d_j T(1+T)Lk}{B_j+Lk}+\lambda\right)} \tag{9}$$

The auditing probability $p(z_j^*,\xi_j)$ can be approximated by a normal random variable with mean equal to the expected value of the righthand side of (9) and variance equal to the variance of the second term in the righthand side of (9). In equilibrium, we need that the tax accountants' beliefs are correct, implying:

$$\begin{aligned} A_j &= \frac{\alpha d_j T\left(1-(1+T)\,\frac{A_j B_j}{B_j+Lk}\right)}{2\left(\frac{\alpha d_j T(1+T)Lk}{B_j+Lk}+\lambda\right)} + \frac{\overline{\xi}}{2\left(\frac{\alpha d_j T(1+T)Lk}{B_j+Lk}+\lambda\right)} \\ B_j &= \left[\left(2\left(\frac{\alpha d_j T\,(1+T)\,Lk}{B_j+Lk}+\lambda\right)\right)^2\cdot r - Lk\right]^{-1} \end{aligned} \tag{10}$$

[26]See, for instance Theorem 1 in DeGroot, 1970[ch. 9.5].

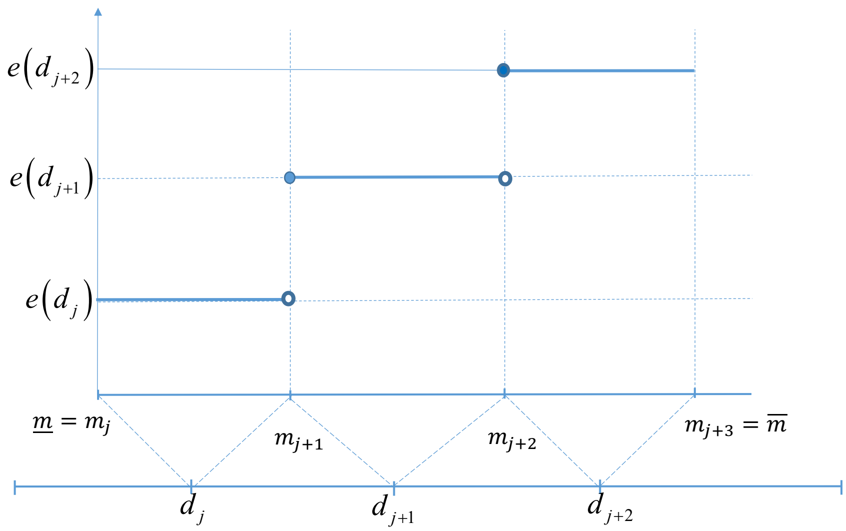


Figure 9: The equilibrium

for all j. The next result shows that a pure-strategy equilibrium of the subgame between the tax authority and the tax accountants exists and it characterizes the equilibrium level of auditing and tax evasion.

Proposition 1. *In equilibrium, the tax authority will monitor accountant j with probability $p(z_j, \xi_j)$ given by (9) with mean A_j and variance B_j given by the system of equations (10). Tax accountant j chooses a level of evasion $e(d_j, \overline{s}_{j,L})$ given by (4) and (5) that is monotonically decreasing in the $\overline{s}_{j,L}$.*

While the equilibrium level of evasion is not, in general, expressible in closed form, it is easy to characterize it in the limit case in which the accountant observes a large number of signals, or very precise signals. In the limit as $L \to \infty$, $e(d_j, \overline{s}_{j,L})$ converges to $e(d_j, \overline{s}_{j,\infty})$ with:

$$e(d_j, \overline{s}_{j,\infty}) = \begin{cases} \alpha d_j \left(1 - \frac{(1+T)[\alpha d_j T + \xi_j]}{2[\alpha d_j T(1+T)+\lambda]}\right) & \frac{(1+T)[\alpha d_j T+\xi_j]}{2[\alpha d_j T(1+T)+\lambda]} < 1 \\ 0 & else \end{cases}$$

which is increasing in d_j.

We now study the taxpayers' decision. It is natural that if ξ_j or T are very large and λ small, then tax evasion may be zero.[27] In general, however, when λ is sufficiently large, the probability of auditing will be sufficiently small so that $e(d_j, \overline{s}_{j,\infty}) > 0$. Moreover, as the following results show, for a sufficiently large λ, auditing is insufficient to equalize the level of evasion among tax accountants even if their types are observable. In this case, taxpayers will sort themselves among tax accountants with taxpayers with higher m_i choosing accountants

[27]Note that as $L \to \infty$, we have $s_{j,L} \to \xi_j$ by the law of large numbers, so it is as if the tax accountant could see ξ_j.

with higher d_j.

Proposition 2. *There is a λ^* such that for $\lambda \geq \lambda^*$ the equilibrium is characterized by a partition of taxpayers types* $\{\widehat{m}_j\}_{k=1}^{J}$ with $\widehat{m}_1 = \underline{m}$, $\widehat{m}_J = \overline{m}$ and $\widehat{m}_j < \widehat{m}_{j+1}$ *such that* ***a*** *taxpayer of type $m_i \in (\widehat{m}_j, \widehat{m}_{j+1}]$ evades $e(d_j, \overline{s}_{j,L})$ as defined in (4),(5) and (10).*

Figure 9 illustrates the equilibrium. Taxpayers with a higher propensity to evade (higher m_i) match with accountants that are more likely to allow them to do it (higher d_j), which causes the distribution of tax evasion to be systematically dependent on the identity of the accountant. This phenomenon is attenuated if the tax authority can see the accountants' types because the authority can then check accountants with higher d_j more intensively. For a sufficiently high cost of public funds λ, however, this is not sufficient to eliminate sorting.

This leads to the following observations. First, we have:

Observation 1. (Sorting Effect) *Except when the cost of auditing is zero, in equilibrium, taxpayers with a higher propensity to evade match with tax accountants who are more accommodating. This implies that the expected tax evasion of a client of an accountant is increasing in the share of other customers who are found evading taxes.*

A second important implication of the model is that the final level of tax evasion depends not only on the accountant type, but also on the information that the accountant acquires regarding the auditing strategy followed by the tax authority. The tax accountant fine-tunes the level of tax evasion based on the accountant's tolerance for evasion (i.e. the type d_j) and the observed signal s_j. When d_j's are positive (i.e. when there is some tolerance for evasion), we observe heterogeneity in behavior due to heterogeneous signals. This leads to the following observation:

Observation 2. (Informational Externality Effect I) *Even if there is no sorting because all tax accountants have the same type $d_j = d^*$, the expected tax evasion of a client of an accountant is increasing in the share of other customers who are found evading taxes.*

Without directly observing the tax accountant's types, it would be hard to separate the sorting effect vs. the informational spillover effect and thus test Observations 2 and 3. The the informational spillover effect, however, has two additional testable implications. The most likely signal used by the tax accountant to fine-tune his/her activities at time t is his/her direct experience with customers at time $t-1$ and perhaps the experience of nearby accountants if they can communicate. It follows that:

Observation 3. (Informational Externality Effect II) *The expected probability of receiving an audit at time t is increasing with the share of clients that are audited at $t-1$, or, if the accountant j is in communication with an accountant k, increasing with the share of k's* *clients who are audited. Hence, reported income (evasion) of a taxpayer at t is expected to correlate positively (negatively) with the number of other clients of the same tax practitioner audited at $t-1$.*

To see the second implication, it is useful to consider a simple generalization of the model. In the previous analysis we have assumed that the tax authority can only target auditing at the tax accountant level. It is, however, natural to assume that the tax authority can target with a finer grid that depends on observables: at the accountant/sector of the client (for example, it can treat clients of a tax accountant differently if they are, say, new or old businesses, lawyers, or plumbers). From a theoretical point of view, this scenario is exactly as the one studied above, except that now the tax accountant receives sets of signals that are contingent on the sector and/or demographics of the clients (say, the number of lawyers vs. the number of plumbers who have been audited).[28] If this is the case, then we should expect evasion of i to be especially sensitive to the share of similar customers who have been audited. Summarizing:

Observation 4. (Informational Externality Effect III) *The behavioral response of taxpayers to audits of other customers of their tax accountant is higher when the other audited taxpayers are similar to them in terms of observables, such as the business sector, size or gender.*

In practice, we can have four cases. When the heterogeneity in the accountants' types is sufficiently important and the accountants' signals are sufficiently precise, we should observe both self-selection and informational spillovers in the data. We might, however, have three other cases: if accountants' d_j's are not very heterogeneous, but signals are important, then we might observe only the informational spillover effect; when accountants' types are heterogeneous, but signals are uninformative, then we might observe only the sorting effect; if both types of heterogeneity are weak and signals are uninformative, then we might not observe either of the two effects.

C.1 Proof of Proposition 1

In equilibrium the mean and variance of the accountant's beliefs are a fix point $(A_j, B_j) = F(A_j, B_j)$, where $F(A_j, B_j)$ is defined by (10). Note that F is continuous in A_j, B_j and

[28] A way to see that the formalization presented above extends easily to this scenario is to consider the tax accountant/sector of client as independent entities in the model. A tax accountant with clients in, say, three sectors, would then be split into three different "tax accountants."

A_j, B_j must be in $[0,1] \times [0, \overline{B}]$ where

$$\overline{B} = \left[(2\lambda)^2 \cdot r - Lk\right]^{-1}$$

It follows by the Kakutani fix point theorem that A_j, B_j exists. By construction, the tax accountant choice of tax evasion is optimal given the beliefs and the beliefs are correct in equilibrium. Similarly, the tax authority chooses the optimal level of auditing given the correct beliefs of the accountant's evasion. ■

C.2 Proof of Proposition 2

The taxpayer does not know $\overline{s}_{j,L}$ when choosing consultant, so expects evasion:

$$Ee(d_j, \overline{s}_{j,L}) = e(d_j) = max\left(\alpha d_j\left(1-(1+T)\frac{A_jB_j + Lk \cdot A_j}{B_j^* + Lk}\right), 0\right)$$

where A_j^* and B_j^* are the equilibrium levels of A_j and B_j given by (10). Taxpayers choose j to maximize:

$$(1-\alpha)\left(1-(1+T)\, p(d_j)\right) e(d_j) - \frac{e(d_j)^2}{2m_i}$$

Note that there is a λ^* such that for $\lambda \geq \lambda^*$, $e(d_j)$ is monotonically increasing in d_j and $(1-\alpha)\left(1-(1+T)\, p(d_j)\right) > 0$. This implies that $(1-\alpha)\left(1-(1+T)\, p(d_j)\right) e(d_j)$ is increasing in d_j as well. Assume now that type m_i prefers d_j to d_k with $d_j > d_k$. Then we have:

$$m_l \geq m_i \geq \frac{1}{2}\left(\frac{e(d_j)^2 - e(d_k)^2}{(1-\alpha)\left[(1-(1+T)\, p(d_j))\, e(d_j) - (1-(1+T)\, p(d_k))\, e(d_k)\right]}\right)$$

for any $m_l \geq m_i$, implying that m_l also prefers d_j to d_k. Similarly, we can show that m_i prefers d_j to d_k with $d_j < d_k$, then m_l prefers d_j to d_k for any $m_l \leq m_i$. This implies that the set of m_is who chooses a consultant d_j is convex and increasing in d_j. That is there is a set of cut points $\{\widehat{m}_j\}_{j=1}^{J}$ with $\widehat{m}_j \leq \widehat{m}_{j+1}$ and $\widehat{m}_j \in [\underline{m}, \overline{m}]$, such that all types in $(\widehat{m}_j, \widehat{m}_{j+1}]$ find it optimal to choose a consultant of type d_j. ■