

NBER WORKING PAPER SERIES

TAX COMPLIANCE AND ENFORCEMENT

Joel Slemrod

Working Paper 24799

<http://www.nber.org/papers/w24799>

NATIONAL BUREAU OF ECONOMIC RESEARCH

1050 Massachusetts Avenue

Cambridge, MA 02138

July 2018

I am grateful to Garrett Anstreicher and Tejaswi Velayudhan for excellent research assistance and to the students in my Ph.D. class on government revenues at the University of Michigan for stimulating conversations. Helpful comments on an earlier draft were received from Claudio Agostini, Miguel Almunia, Alan Auerbach, Anne Brockmeyer, Wei Cui, Brian Erard, Christian Gillitzer, John Guyton, Michael Hallsworth, Anders Jensen, Paul Michael Kindsgrab, Ben Meiselman, Gareth Myles, Alan Plumley, Daniel Reck, Carlos Scartiscini, Christian Traxler, David Weisbach, and Eleanor Wilking, as well as from the JEL editor and referees. The author has an unpaid contractual relationship with the Internal Revenue Service to make research use of de-identified tax-return data; the views expressed here are not necessarily held by the IRS, the U.S. Department of the Treasury, or the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2018 by Joel Slemrod. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Tax Compliance and Enforcement
Joel Slemrod
NBER Working Paper No. 24799
July 2018
JEL No. H20,H26

ABSTRACT

This paper reviews recent economic research in tax compliance and enforcement. After briefly laying out the economics of tax evasion, it focuses on recent empirical contributions. It first discusses what methodologies and data have facilitated these contributions, and then presents critical summaries of what has been learned. It discusses a promising new development, the analysis of randomized controlled trials mostly delivered via letters from the tax authority, and then reviews recent research using various methods about the impact of the principal enforcement tax policy instruments: audits, information reporting, and remittance regimes. I also explore several understudied issues worthy of more research attention. The paper closes by outlining a normative framework based on the behavioral response elasticities now being credibly estimated that allows one to assess whether a given enforcement intervention is worth doing.

Joel Slemrod
University of Michigan
701 Tappan Street
Room R5396
Ann Arbor, MI 48109-1234
and NBER
jslemrod@umich.edu

Policy attention in developed countries to tax evasion and enforcement ramped up significantly after the financial crisis of 2008, the Great Recession, and the large deficits that followed. Particular scrutiny to high-income individuals and corporations accompanied heightened attention to income and wealth inequality. In the United States this momentum led to a major initiative aimed at reducing income tax evasion via unreported foreign accounts, albeit in the context of a shrinking IRS budget. In the United Kingdom, former Chancellor of the Exchequer George Osborne announced an ongoing crackdown on tax avoidance, tax evasion, and “imbalances” in the tax system that would bring in £5 billion in additional revenue each year. In developing countries, tax evasion is a long-standing policy concern. The IMF (2015, p. 6) recently stated that

“Long a central concern in developing countries, strengthening compliance has become a greater priority in many advanced and emerging economies since the financial crisis...Improving compliance has long been a core development objective, both to enhance revenue and as essential to building strong, trusted public institutions.”

Academic research into the determinants of tax evasion and the effectiveness of enforcement initiatives has also vastly expanded in the past decade, perhaps inspired by the heightened policy interest and certainly facilitated by expanded access of academics to administrative tax-return data and an increased willingness of tax authorities to partner with researchers on randomized controlled trials that hold the promise of compelling identification of the causal impact of alternative enforcement strategies.

In this paper I review this new wave of empirical research. Most of this scholarship postdates the other broad surveys about tax evasion and enforcement, including one in this Journal,¹ and so

¹ Andreoni, Erard, and Feinstein (1998), Slemrod and Yitzhaki (2002). A much shorter, and more policy-oriented, version of this paper was published as Slemrod (2017). Chalfin and McCrary (2017) review the literature on criminal deterrence, but do not address tax evasion. Mascagni (2018) surveys the experimental literature on tax compliance.

a fresh look is appropriate. Because this new wave of empirical research shows no sign of abating, it is particularly timely to take stock of what has been learned in order to assess and guide the direction of future research. Although this survey is wide-ranging, it is somewhat selective in its coverage of both topic and method. It touches only tangentially on some substantive issues such as multinational companies' tax compliance and the role of bureaucratic incentives and the behavior of tax officials,² and it does not address some methodological approaches such as lab experiments and macroeconomic measures of the informal economy.³

The rest of the paper is organized as follows. After section 1 briefly lays out the economics of tax evasion, the paper turns to the recent empirical contributions, with section 2 discussing what methodologies and data have facilitated these contributions. Section 3 focuses on the magnitude and nature of evasion. Section 4 assesses one prominent new development, randomized controlled trials mostly delivered via letters from the tax authority. The next sections discuss the principal enforcement policy instruments, audits in section 5, information reporting in section 6, remittance regimes in section 7, with other enforcement initiatives addressed in section 8. Section 9 explores several understudied issues worthy of more research attention and outlines a normative framework based on the behavioral response elasticities now being credibly estimated that allows one to assess whether a given enforcement intervention is worth doing. Section 10 concludes.

1. THE ECONOMICS OF TAX EVASION

Discussion of tax administration, compliance, and enforcement fits naturally into what Slemrod and Gillitzer (2014) call a “tax-systems” framework. A tax system is defined as a set of rules, regulations, and procedures with three aspects. First, it defines what events or states of the

² On the former, see, for example, Clausing (2009); on the latter, see for example, Khan, Khwaja, and Olken (2016).

³ On the former, see Alm and Jacobson (2007); on the latter, see Slemrod and Weber (2012).

world trigger tax liability—for example the earning of income, the ownership of a residence that might be subject to property tax, or the sale of a capital asset—and the rate schedule to be applied to these bases. Although this first aspect, *tax bases and rates*, is the principal object of modern tax analysis, it is only one part. A tax system also must specify who or what entity must remit the tax liability and when, which we might call *remittance rules*. For example, under most income tax systems, it is the employer that remits—actually sends to the government—an approximation of what tax their employees owe on that income. Although standard analysis often dismisses this as irrelevant, recent research suggests that, at least in some circumstances, it matters a lot. Third, a tax system details procedures for ensuring compliance, including the provision of third-party information-reporting requirements and the consequences, including penalties, of not remitting legal liability; these are the *enforcement rules*. This review focuses on the second and, especially, the third aspect of a tax system, but clearly there is substantial overlap among the three issues: for example, the tax base and the remittance system can have a profound effect on the enforceability of taxation, and the enforcement regime can affect the impact of tax rates on taxpayer decisions that determine the tax base.

1.1 Why Tax Evasion Matters

Tax evasion⁴ is an important issue because it affects both the resource cost of raising taxes and the distribution of the tax burden—the bread-and-butter concerns of public economics. If tax evasion could somehow be costlessly eliminated, the additional money collected could be used to

⁴ I will use the terms evasion and noncompliance interchangeably, although in some countries evasion has a particular legal meaning. I will also conform to standard usage in which evasion refers to illegal actions to reduce tax liability, while avoidance refers to legal actions to reduce tax, recognizing that in many situations the dividing line is blurry. Slemrod and Yitzhaki (2002) distinguish further between real (and legal) behavioral responses to taxation, such as reducing labor supply, which they call *real* responses, and actions that reduce tax liability without substantively altering one's consumption basket, such as re-characterizing a given action as tax-preferred R&D or delaying an asset sale with a taxable capital gain by a day to a lower-tax-rate year, for which they reserve the term avoidance.

finance worthy government projects or cuts in tax rates that would benefit most compliant taxpayers. But expanding government programs could be financed in a number of other ways, such as by raising tax rates or by broadening the income tax base, and a tax reduction could be financed by cuts in overall spending. The real policy question is whether curbing evasion would improve the equity and efficiency implications of the public finances, given the cost of doing so, and if so how best to go about it.

If opportunities or predilections for evasion were systematically related to income, then the income tax rate schedule or the relative reliance on tax bases could be adjusted to move toward the degree of progressivity is deemed optimal. Of course, not everyone evades taxes by the same proportionate amount or by an amount strictly related to income, both because of differences in personal characteristics—like attitudes toward risk, the tax system, and honesty— and because of different opportunities and potential rewards for evasion. Evasion thus creates horizontal inequity because otherwise equally well-off people end up with different tax burdens. Attempts to reduce tax evasion can themselves raise vertical equity concerns, as when the Internal Revenue Service (henceforth IRS) is criticized for spending resources to reduce fraud related to the Earned Income Tax Credit, whose recipients are low-income households, instead of devoting those enforcement resources to the types of noncompliance more likely to be pursued by high-income households, such as the use of unreported foreign accounts.

Tax evasion also generates efficiency costs. The most obvious are the resources taxpayers expend to implement and camouflage noncompliance, that third parties incur to put into effect withholding and provide information reports, and that the tax authority uses to administer the system and combat noncompliance. In addition, tax evasion generally provides a socially inefficient incentive to engage in those activities that facilitate evasion. For example, because the

income from house painting can be done on a cash basis and is therefore difficult for the tax authority to detect, this occupation is more attractive than otherwise. Although a supply of eager and cheap housepainters undoubtedly is greeted warmly by prospective buyers of that service, the effort of the extra people drawn to house painting, or any activity that facilitates tax evasion, will have higher social value in some alternative occupation. The same argument applies to self-employment generally, as the enhanced opportunity for tax noncompliance inefficiently attracts people who would otherwise be employees. The opportunity for noncompliance can distort resource allocation in a variety of other ways, such as causing companies that otherwise would not find it attractive to do so to set up operations in a tax haven in order to facilitate or camouflage evasion.

Of course, curtailing tax evasion is not costless, and its costs must be considered in developing optimal policy. The mere presence of tax evasion does not imply a failure of policy. Just as it is not optimal to station a police officer at each street corner to eliminate robbery and jaywalking completely, it is not optimal to completely eliminate tax evasion. The recognition of tax evasion focuses our attention on the set of enforcement policy instruments, the appropriate setting of which can be illuminated by optimal tax reasoning and fleshed out with empirical analysis; this is developed in Section 9.8. For instance, what should be the extent of audit coverage, the strategy for choosing audit targets, and the penalty imposed on detected evasion? The presence of evasion also alters the optimal setting of tax rates, because it affects the marginal efficiency cost of so doing, and the choice of tax base, because different tax bases are more or less susceptible to evasion.

1.2 The Evasion Decision

Why would an individual or business evade taxes? To an economist, the natural starting point is to consider the private costs and benefits of evasion. And indeed the standard framework for considering whether and how much to evade taxes is a deterrence model, formulated by Allingham and Sandmo (1972), who adapted Becker's (1968) model of criminal behavior to tax evasion. In this model, a risk-averse taxpayer decides whether and how much to evade taxes in a similar way as she would approach any risky decision or gamble.⁵

The basic version of the model, for the case of a proportional income tax, can be written as follows:

$$(1) \text{Max}_e (1-p(e,a))U((y(1-t) + te) + p(e,a)U(y(1-t) - fte),$$

where e is understated tax liability, p is the probability that evasion is penalized, a represents enforcement intensity, $U(.)$ is a von Neumann-Morgenstern utility function, y is exogenous true income, t is a linear tax rate and f is the proportional penalty assessed on detected evasion. Note that in this model people are influenced by possible legal penalties no differently than any other contingent cost: there is nothing per se about the illegality of tax evasion that matters. Nor is there any intrinsic willingness to meet one's tax obligations, sometimes referred to as "tax morale", discussed in Section 1.4.

If dp/de is zero, the first-order condition for optimal evasion is as follows:

$$(2) U'_N/U'_A = fp/(1-p),$$

where the subscripts N and A refer to the audited and not audited states of the world, respectively.

This model predicts that a risk-averse agent will do some evasion as long it has a positive expected

⁵ Recently, the Allingham-Sandmo framework has often been replaced by a simpler deterrence model that dispenses with the expected utility framework, so that the maximand is just $y(1-t) + te - c(e,\alpha)$, where c represents the private cost of evasion, which includes the utility cost of bearing risk and the expected value of punishment. This private cost may depend on the amount of attempted evasion and certainly depends on the vector α of enforcement instruments, including the extent of auditing and third-party information reporting.

value; i.e., if $p(1+f) < 1$, and that an increase in either the probability of detection or the penalty if detected will reduce evasion.

Of course, the model does not pin down how big these effects are, which depends on the parameters and the shape of the utility function, so this is a task for empirical analysis. The effect on evasion of a change in the tax rate is less clear, as even its sign depends on the form of the penalty function, as shown by Yitzhaki (1974); as formulated here, where the penalty is proportional to the tax evaded, the tax rate does not enter the first-order condition, as both the gain in the audited state and the loss in the unaudited state are proportional to the tax rate, so that a change in t causes only an income effect. Thus, an increase in t will *decrease* evasion as long as there is decreasing relative risk aversion.

With a nonzero dp/de , the first-order condition becomes

$$(3) \quad U_N'/U_A' = fp/(1-p) + (dp/de)(U_N - U_A)/U_A'$$

Not surprisingly, to the extent that the audit probability rises with evasion, the optimal amount of evasion is lower.

Much of interest is contained within the $p(\cdot)$ function. What matters, of course, is the perception of taxpayers, which raises the question of how these perceptions are formed. Presumably, perceptions are linked in some way to the reality. For one thing, enforcement policy can change the reality itself—increase the actual p —by, for example, hiring and training more auditors or more efficiently using the third-party information it routinely receives—and trust that the word about the reality gets out. The tax authority can also selectively disseminate information about its activities.⁶ Much of the empirical work I discuss below tries to measure the effect on behavior of

⁶ The IRS certainly does try. Blank and Levin (2010) show that the government issues a disproportionately large number of tax enforcement press releases during the weeks immediately prior to Tax Day (generally April 15) compared to the rest of the year, and conclude that government officials deliberately use tax enforcement publicity to influence individual taxpayers' perceptions and knowledge of the audit probability, tax penalties and the

communications from the tax authority to taxpayers designed to alter perceptions of the enforcement environment.

The idea that p increases with e —the more evasion one does, the more likely that the evasion will be detected and punished—is a reduced-form notion, as the tax authority does not know e without further action. Most tax authorities have some kind of risk management system that is used to allocate resources toward cases where enforcement action is likely to be more effective. It may also be that p depends positively on e/y , evasion relative to true income, in which case there is an implicit subsidy to earning more y , because it reduces the chance of a given amount of evasion being detected.

The value of p depends on the type of evasion considered. For example, in most countries, because of employer information reporting the probability of detection is close to one for unreported employee income. It is generally much lower, but increasing with the magnitude, for underreported self-employment income. For someone with multiple sources of income, the probability of detection would rise with the total amount of income evaded, as one first underreports those sources with the lowest probability of detection, and then moves on to underreport those types of income with higher probability. There will also generally be interaction among multiple evasions, within a period and across periods.

The standard deterrence framework applies naturally to tax compliance decisions made by individuals and small, single-owner businesses, but its applicability to big business is less clear. Arguably, large public companies should act as if they are risk-neutral, rather than like the risk-

government's tax enforcement efficacy while taxpayers are preparing to file their annual individual tax returns. Whether taxpayers contemplating evasion are more influenced (or influenced at all) by news of a successful enforcement action they learn of in April compared to one they learn of in November has not yet been demonstrated.

averse decision maker of the standard model. If this is true, one must look elsewhere for what constrains positive-expected-value evasion. Some firms might be concerned that publicized tax aggressiveness turns off some potential customers who would prefer to deal with civic-minded companies. On the other hand, some investors might take tax aggressiveness as a signal that a company is optimally aggressive both in its dealings with the tax authority but also with suppliers and customers (but, ideally, not with investors themselves!).⁷

The basic Allingham-Sandmo model has been enriched and extended along many dimensions, a literature that is not the subject of this survey. One extension worth mentioning introduces a dynamic element to the decision, as in Lee and McCrary (2017), which implies that it is harder to deter an impatient individual using a prison sentence, as most of the disutility is borne in the future, and that people who are myopic and engage in hyperbolic discounting will be less deterred by the threat of punishment, some of which occurs far in the future. This is likely to be less important for the crime of tax evasion, as imprisonment is a relatively rare punishment, but it correctly emphasizes that the punishment should be considered in a present-value sense.

1.3 The Informal Economy

A related but distinct concept to evasion is the informal economy, also known as the underground, hidden, or black economy. Many definitions of the informal economy have been suggested, having in common the notion of small-scale economic activities that are unobserved by official authorities.⁸ The informal economy encompasses small firms that don't register with the relevant tax or labor regulation authorities, employees who are not on the payroll, freelancers who

⁷ On this issue, see Hanlon and Slemrod (2009), who examine the stock-market response to publicized tax aggressiveness to sort out empirically these potential two concerns of public corporations, finding that on average stock prices decline when news about involvement in tax shelters becomes public. Stock price falls tend to be larger for retail-sector firms, which may be due to a possible consumer/taxpayer backlash.

⁸ Kanbur and Keen (2014) develop an insightful model of the relationship between informality and tax compliance.

don't file tax returns, and so on. Many evasion activities are clearly not part of the informal economy, such as overstating deductible charitable contributions or setting up a foreign bank account and not reporting the taxable income it generates. Moreover, not all informal enterprises are evading; for example, taxable income may be legitimately below the filing threshold at the same time a firm is skirting labor and safety regulations. Whether illegal activities should be included in the definition of the informal economy is controversial; if they were, one would have to acknowledge that not all such operations are small, as witnessed by narcotrafficking.

In situations where labor income in the formal sector is routinely reported by the employer to the tax enforcement agency and can be successfully matched to the worker's income report, the only way to evade tax may be by "moonlighting"—working extra hours for oneself at a different job—or by switching completely to the informal sector. The standard deterrence model can be easily modified to address the choice between formal- and informal-sector work by supposing that the taxpayer receives a higher pre-tax wage rate for formal-sector work but the income is taxed at the statutory rate and cannot be evaded, while informal-sector income is untaxed unless detected by a random audit.⁹

1.4 Non-deterrence Considerations

Some social scientists have argued that the deterrence framework misses important elements of the tax evasion decision, and question some of its central assumptions, including that nothing per se about the illegality of evasion matters, and that everyone acts as a free rider, so that there is no issue of intrinsic willingness to pay, or "tax morale."¹⁰ Some have gone further to suggest that,

⁹ See Kesselman (1989).

¹⁰ The evidence about the importance of nonpecuniary motivations for tax compliance, or tax morale, is nicely reviewed in Luttmer and Singhal (2014). Steven Durlauf has pointed out to me that civic duty and morale can be addressed as modifications of the standard model of preferences in that they introduce in the first case, utility from actions rather than their consequences, and in the case of the latter, social interactions. I don't want to stake out a strong case on this semantic issue, but find it helpful to refer to these considerations as non-deterrence because

in thinking about tax evasion, it is necessary to abandon the expected utility maximization model and incorporate “behavioral” considerations.

One approach stresses that some people may fully comply with their legal obligation because of a sense of civic duty regardless of, or in addition to, the possible expected pecuniary gains and argue that the tendency to perform one’s duty is susceptible to aspects of the enforcement process. Indeed, Frey (1997) argues that imposing more punitive enforcement policies may crowd out the “intrinsic” motivation to comply by making people feel that they pay taxes because they *have* to, rather than because they *want* to. Another approach suggests that some individuals’ behavior depends on the process by which the tax and tax enforcement system are formulated and its features, holding constant the incentives the system provides. For example, taxpayers may be more willing to comply with a system whose formulation they had a part in through voting. Taxpayer attitudes toward authority may also influence compliance behavior. Tyler (2006) argues that citizens are more likely to be law-abiding if they view legal authorities as legitimate, and the degree of legitimacy may itself be a function of the level of enforcement. When explicit enforcement is weak (e.g., few audits), legitimacy may erode, undermining the intrinsic willingness of taxpayers to comply with the law. People may be willing to comply with a law because they perceive it to be just, quite aside from their beliefs regarding the authority government has to enforce it. Such individual judgments can be complex; for example, expenditures on warfare might contribute to a sense of fairness in a patriotic period, but rejected during another period characterized by antimilitarism. Levi (1989) stresses the role of “reciprocal altruism,” in which some taxpayers’ behavior depends on the behavior, motivations, and intentions not of any subset of particular individuals, but of the government itself: when citizens believe that the government will act in

they move the focus away from the probability of punishment and the extent and nature of the punishment for detected noncompliance.

their interests, that its procedures are fair, and that their trust of the state and others is reciprocated, then people are more likely to become “contingent consenters” who cooperate in remitting taxes even when their short-term material interest would make evading the best option.

Much of the evidence related to these nonstandard behaviors comes from how people react to other people’s behavior in lab experiments. But the psychological attitudes of individuals toward government might be fundamentally different than their attitudes toward other people, or even other organizations. Individuals might feel more dutiful and even obedient toward government. Invocation of the word obedience, though, invokes a darker side of the relationship between individuals and government as an authority figure. Indeed, notorious experiments conducted by Milgram (1963) showed that unwitting subjects were willing to deliver what they thought were substantial electric shocks when instructed to, and encouraged to, by authority figures.

An alternative strand of scholarship outside of the Allingham-Sandmo deterrence model tradition is rooted in behavioral economics. Early on, Benjamini and Maital (1985) built on the work of Kahneman and Tversky (1979) by exploring the implications for tax evasion of subjective probability bias, perception of other people’s behavior, and social stigma. Dhimi and al-Nowaihi (2007) argue that, compared to an Allingham-Sandmo model, a prospect-theory framework (with the addition of a stigma cost for discovered evasion) can more satisfactorily explain the level of observed evasion, the non-ubiquity of evasion, and the fact that tax rates negatively affect evasion. Durlauf and Nagin (2011) argue that, in a world in which the perceived probability of detection is very low, even small changes in that perceived probability can have a big effect.

A recent offshoot of behavioral public economic focuses on the notion that individuals, and perhaps also firms, are inattentive to certain aspects of a tax systems—aspects which are not *salient*, in the language of Chetty, Looney, and Kroft (2009). In their model, inattention is

measured by an under-reaction to a non-salient change in a tax rate relative to the response to a salient tax change. This notion of salience makes sense in the setting they address—grocery store items where the sales tax is not posted on the price tag—but does not generalize in a straightforward way to tax evasion, where it seems unlikely that an inattentive taxpayer would presume that, for example, there is *zero* chance of being caught. This is not to say that some taxpayers in some settings are not familiar with their tax obligations and the consequences of not meeting those obligations.

2. METHODOLOGY—HOW HAVE WE LEARNED NEW THINGS?

Measuring tax evasion is highly challenging due to tax evaders' incentive to conceal their behavior. Of course, measuring any sort of crime is subject to similar concerns, but tax evasion also has no identifiable victims, so no data can be obtained from victimization reports or victim surveys.¹¹ On the other hand, in contrast to most other crimes, one can observe many traces of changes in tax evasion, such as changes in reported tax bases.

The threat of punishment and perhaps social shame makes many taxpayers unwilling to respond accurately even to surveys done by organizations not associated with the tax authority.

¹¹ There are other important differences between tax evasion and alternative crimes. One is that, for other crimes, differentiating between a deterrent effect and an incapacitation effect is key, called a “first-order issue” by Chalfin and McCrary (2017). But incapacitation due to incarceration is not a major concern for tax evasion. To get a sense of the difference in magnitude, according to the U.S. Bureau of Justice Statistics, in 2013 2,220,300 adults were incarcerated in U.S. federal and state prisons, and county jails. In contrast, in fiscal year 2016 just 927 people were incarcerated for federal tax evasion not related to illegal activity or narcotics. Moreover, one can in principle evade taxes even while incarcerated, while this is more difficult for, say, auto theft. Indeed, in 2010 more than 91,000 inmate returns claimed \$758 million in fraudulent refunds (Ellis 2013). In addition, the stigma from being involved in tax evasion is qualitatively different from that associated with, say, murder. Finally, unlike many crimes, tax evasion is not a phenomenon of youth, which has many implications, especially when the criminal choice is considered in a life-cycle setting; for example, Anwar and Loughran (2011) argue that risk perceptions are more sensitive to experience early in a person's criminal career.

For this reason, almost all the empirical analyses of evasion, including the most credible ones discussed below, don't actually have a reliable measure of evasion. But scholars have risen to the challenge, and there are several promising developments in ascertaining the magnitude, nature and determinants of tax evasion as well as, most importantly, how different policies might affect tax evasion. I discuss these developments next. Because the econometric issues are addressed extensively elsewhere, here I only briefly address some of the tax-specific issues that arise with these methodologies.

2.1 Random Audits

Aggregating up from data on operational audits will not provide an accurate picture of tax evasion, as by design these audits focus on those returns that are most suspected of noncompliance; to the extent that this focus is successful, aggregation will overstate the magnitude of evasion. A more accurate picture can be obtained from random audits. In the United States, the IRS National Research Program (NRP)¹² provides a snapshot of compliance and evasion from annual stratified random samples of approximately 13,000 returns combined over several years with the primary objective of optimizing the IRS's audit strategy, but also to better understand the tax gap. Each return is classified according to one of three possible courses of action: accept the return as corroborated by third-party information, write to the taxpayer for additional information on up to three items that could not be corroborated, or conduct an in-person audit. The NRP oversamples returns from high-income taxpayers and individuals who report sole proprietorship income. Because the line-by-line audits can fail to uncover substantial amounts of noncompliance, the tax gap methodology includes adjustments for income that is not detected during audits using a technique known as detection-controlled estimation, discussed in Feinstein (1999).¹³

¹² The NRP replaced a similar program known as the Taxpayer Compliance Measurement Program (TCMP) in 2001.

¹³ See Internal Revenue Service (2016).

Given the scale of the effort required (and perhaps its political sensitivity), tax authorities in very few countries have done such studies.¹⁴ Even for the United States, where these studies have been done regularly since tax year 1963, it is difficult to draw inferences about changes in noncompliance over time, because the methodologies used have varied. The limited usefulness of these estimates for these purposes is not really a surprise, because their main objective is not to come up with precise aggregate “tax gap” measures of the magnitude of evasion or of the nature of evasion, but rather to help identify returns that are more likely to feature evasion, so as to guide the allocation of enforcement resources.

2.2 Randomized Controlled Trials

Randomized controlled trials¹⁵ (RCTs) have been part of the “credibility revolution” (Angrist and Pischke, 2010) in empirical economics because they enable identification of the causal impact of, for example, a policy intervention. When executed properly, the researcher need not worry about identifying an appropriate control group, because the control group is built into the randomization; in the simplest case, there are two statistically identical groups, one that receives the policy treatment of interest and the other that doesn’t.

Although governments are unlikely to allow tax rates and bases to be randomized in the field, for other tax-system instruments policy randomization is possible and, more important, has become a reality.¹⁶ Despite the unrivaled internal validity of well-designed randomized control

¹⁴ A notable exception is the United Kingdom, where the HMRC has calculated tax gap estimates for many British taxes; they use bottom-up estimates based on random audits to assess the magnitude of several aspects of noncompliance for individual income tax, as well as business tax for small- and medium-sized enterprises. See HMRC (2015) and IMF (2013). OECD (2015) discusses other countries’ tax gap studies.

¹⁵ Such studies are also referred to as randomized field experiments, but my experience with tax authorities has taught me that most, maybe all, governments do not want to *experiment* on their citizens. It is, though, arguably prudent to run *trials*, or *pilots*, of a contemplated policy reform before rolling it out to the whole population.

¹⁶ Hallsworth (2014) offers an insightful review of the early wave of these studies.

trials, they are not without issues.¹⁷ It is not always clear that the results can be “scaled up.” General equilibrium effects may matter, and without understanding the causal channels through which policy interventions affect taxpayers’ behavior, it may not be possible to predict the effect of variations in the policy intervention without running repeated experiments. In addition, some interventions that are credible in an experimental setting may not be credible in an economy-wide setting. For example, Kleven et al. (2011) sent treatment groups letters asserting they are subject to a 50 percent or 100 percent probability audit threat letters; savvy taxpayers would know that either policy would be prohibitively expensive were either of these treatments to be expanded to the entire population. Thus, it could be that taxpayers would react strongly to such a threat when explicitly made on a small scale, when it is plausibly credible, but would not react if the threat were made broadly, which would not be credible due to limited resources of the tax authority. Another issue is the difficulty of establishing external validity, discussed at more length below.

2.3 Wider Availability of Administrative Data

A very promising recent development is the wider availability for analysis of anonymized (also called de-identified) administrative tax data, including but not limited to tax returns, sometimes linked to other administrative records, often on the whole population of a country. These kinds of data first became available in Nordic countries, but now they are also available under varying protocols in Canada, in the United Kingdom, many other European countries, in the United States (in this case not generally linked to other administrative data) and most recently in numerous low- and middle-income countries including Brazil, Chile, China, Costa Rica, Ecuador, India, Pakistan, Rwanda, Tunisia, and Uganda. Compared to having small samples of tax-return data, when a researcher has *all* returns she has much more (statistical) power to reach reliable conclusions about

¹⁷ This is not the place to assess in detail the controversies about the internal and external validity of RCTs; the interested reader might want to consult Deaton (2010), List (2011), and Banerjee et al. (2017).

the effect of taxation and to investigate such issues as the heterogeneity among groups defined by geography or other demographic variables. Some current efforts in the U.S. have focused on linking tax data sets, such as linking firm records with worker records, in ways that have not heretofore been done, even by the IRS itself. The fact that tax-return data generally contain what the taxpayer reported, rather than the “truth” about, say, income, is a disadvantage if it is the truth one is hoping to learn about, but is a particular advantage when studying the elasticity of reported taxable income, which crucially *includes* evasion and avoidance responses to tax policy instruments.¹⁸

To be sure, some delicate issues arise with studies based on administrative data.¹⁹ The data must be anonymized to preserve taxpayer privacy, and other measures must be in place to prevent misuse of the data. These concerns have led to a situation where only a small group of academics get access to the data,²⁰ so that replication is not easy. One might also be concerned that results that shed an unflattering light on the operations or decisions of the tax authority are less likely to see the light of day.

2.4 Analysis of Administrative Data with Kink, Notch and Regression Discontinuity Research Designs

Empirical analysis outside of the randomized control paradigm is by no means dead, nor should it be, especially when making use of tax-return administrative data. But the bar for credible identification strategies in research using administrative data (and randomized field experiments) has gotten higher.

¹⁸ A survey of the literature on the elasticity of taxable income is presented in Saez, Slemrod, and Giertz (2012).

¹⁹ These issues are discussed in greater detail in Slemrod (2016).

²⁰ In the United States, the access has been formalized through the Joint Statistical Research Program of the IRS' Statistics of Income Division, which since 2012 has issued a call for proposals every two years, and accepts and facilitates projects that meet explicit criteria. In the 2016 round, 18 proposals were approved. For details see <https://www.irs.gov/pub/irs-soi/16jsrpprojects.pdf>.

Two non-RCT-based research designs hold particular promise and have been widely exploited recently in empirical tax analyses: regression discontinuity and “bunching” analysis of kinks and notches in policy. In a regression discontinuity design, there is a threshold above or below which a tax treatment applies. By comparing neighboring observations on either side of the threshold and therefore arguably quite similar, one can estimate the average effect of the treatment in that local area, even in environments where randomization is not feasible. In this setting it is crucial that the assignment of cases to treatment is random, and that it is impossible for the decision makers to manipulate their treatment status.

When policy introduces kinks in budget sets, so that the marginal (but not average) tax rate changes discontinuously around the kink, identification is aided because arguably in many cases the people on either side of the kink are on average fairly similar. How many people “bunch” at or very near the kink provides, under some assumptions, a measure of the average elasticity of choices with respect to the tax rate. Even more compelling is the analysis of behavior in the presence of policy-induced notches, where the budget set itself is discontinuous, so for instance reporting one additional dollar of income increases one’s tax liability by a few hundred dollars. Ever since their potential for identification was understood, an impressively high number of policy notches have been discovered and analyzed, some pertaining to tax enforcement policy. One particularly promising aspect of notches is that their presence often implies that there is some region of behavior that is always dominated by behavior in another region regardless of one’s preferences: earning one less dollar to get below the threshold of a tax-increasing notch in an income tax schedule saves money and requires less labor, a win-win for the majority of folks who prefer more leisure to less and more money to less. The so far ubiquitous finding that theoretically dominated

region is not empty sheds light on the constellation of reasons that might explain such behavior such as irrationality, cluelessness, and adjustment costs.²¹

2.5 Traces of True Income and Evasion

In a traces-of-income procedure, one compares an indicator of the true tax base, say income, to reported income.²² The classic paper is due to Pissarides and Weber (1989), who use food consumption as an income indicator. They assume, reasonably, that although how much food someone purchases is a function of income and other characteristics, it does not depend on what *kind* of income, in particular salary versus self-employment income, a person has. Next they examine the ratio of food purchases to reported income, separately for employees and self-employed people, and find that it is larger for the latter group, implying given their assumption that the self-employed understate their income. Thus, they infer (relative) true income from food purchases, and compare this “trace” of true income to (relative) reported income.

Under a traces-of-evasion technique, one seeks behavior that can reasonably only, or at least largely, be explained by tax evasion, for instance the hoarding of high-value currency, which facilitates both tax evasion and money laundering, or the maintenance of accounts in tax havens. This approach is related to the broader topic of “anomaly detection” that is used to assist, for example, in the detection of credit-card fraud, as discussed in Chandola, Banerjee, and Kumar (2009).

2.6 Other Promising Empirical Research Designs

²¹ Kleven (2016) surveys bunching analysis, focusing on tax issues; Kleven and Waseem (2013) is a key application.

²² More generally, this method focuses on traces of the true tax *base*, be it income, consumption, or some other base. There is also a macroeconomic version of this methodology that has analyzed traces of true aggregate income to shed light on the size of the informal economy across countries. For example, conditional on other determinants, a high ratio of aggregate electricity use (a trace of true income) or cash use (a trace of evasion) to formal income might indicate a relatively large informal sector. Slemrod and Weber (2012) critically discuss such estimates of the informal economy.

Some modern research designs that have provided insights into other questions have not yet been applied to tax compliance and enforcement questions. One is the so-called “judges’ model,” which takes advantage of random or quasi-random assignment of judges or other decision makers as an instrument to assess the impact of the severity of punishment on subsequent behavior. An excellent example of this approach is the study of Dobbie and Song (2015), who exploit the fact that U.S. bankruptcy courts use a blind rotation system to assign cases to judges who are of measurable heterogeneous “toughness” in order to study the impact of bankruptcy protection on debtor outcomes. To the extent that tax auditors are randomly or quasi-randomly assigned to cases, one might be able to learn about the impact on behavior of the intensity of evasion monitoring. Many other taxpayer interactions with the tax authority, such as offers in compromise—taxpayer proposals to settle tax liabilities for less than the full amount owed—or applications to enter an amnesty program, have this character.

3. THE MAGNITUDE AND NATURE OF EVASION

3.1 Evidence from Random Audits

The IRS’s most recent tax gap estimates were released in April, 2016 and pertain to an average over tax years 2008 to 2010.²³ The overall gross federal tax gap is estimated as \$458 billion, which amounts to 18.3 percent of the estimated actual (remitted plus unremitted) tax liability; the IRS calls the 81.7 percent of the total that is remitted the “voluntary compliance rate.”²⁴ The IRS estimates that it would eventually collect \$52 billion of the \$458 billion gross tax gap estimate,

²³ IRS (2016).

²⁴ Of note is the fact that the IRS’ 2014-2017 strategic plan states a target voluntary compliance rate of 87 percent by 2017, a substantial 23 percent decrease in the noncompliance rate. (IRS 2014).

which results in a “net tax gap” of \$406 billion, so that the “net compliance rate” is 83.7 percent--16.3 percent of the estimated tax liability will never be remitted.²⁵

More telling than the overall compliance rate is the sharp variation in the compliance rate by the extent of information reporting to the IRS. When there is little to no third-party-reported information (such as self-employment income), the estimated noncompliance rate is 63 percent, compared to 19 percent when there is “some” reporting, and 7 percent when there is “substantial” information reporting. Small businesses, which comprise a major part of the first category, represent a large portion of the tax gap in individual income; approximately 47 percent of underreporting of individual income tax comes from business income. When there is both withholding and substantial information reporting (such as for employee wages and salaries), the estimated noncompliance rate is just 1 percent. Kleven et al. (2011) find a similar disparity in, but lower overall, noncompliance rates in Denmark, with an estimated 14.9 percent for self-employment income but less than one percent for third-party reported items.²⁶

This variation is, in my view, stark and compelling evidence for the primary importance of deterrence as an explanation of tax evasion. When information reporting renders the likelihood of

²⁵ The 16.3 percent is slightly higher than the previous net tax gap estimate of 14.5 percent for tax year 2006, but the IRS attributes the difference to “improvements in the accuracy and comprehensiveness of the estimates” rather than a trend toward more noncompliance.

As a comparison, the HM Revenue & Customs (2015) recently calculated the overall tax gap in the United Kingdom as of 2013-2014 to be 6.4 percent of true liability: 5.0 percent for the individual income tax, 6.4 percent for the corporation income tax, and 11.1 percent for the value-added tax. Small and medium-sized enterprises account for over half of the overall tax gap. Whether the much lower estimated noncompliance rate for the British income tax compared to the U.S. income tax is real, or due to a different tax gap methodology, is not clear. Kleven et al. (2011) report the results of a smaller-scale random audit tax gap exercise in Denmark, and find overall evasion to be only 2.2 percent of net income.

²⁶ Unlike the US tax gap estimates, in the Danish study no correction was made to reflect that audits do not uncover all evasion, so the extent of detected evasion between the two countries is not as stark as the figure in the text suggest.

successful evasion to near zero (i.e., p is close to one), evasion is negligible; when that likelihood is much lower, evasion proliferates.

3.2 Traces

The wide divergence between the compliance rates of employees and the self-employed has attracted a lot of research interest and has focused attention on the latter group. Recall that Pissarides and Weber (1989) pioneered the “traces-of-income” approach using UK Family Expenditure Survey data on food consumption to estimate the extent of evasion, assuming that only the self-employed evade and that the relationship between food consumption and true income is independent of employment status.²⁷ With these assumptions, they are able to predict true income—and therefore underreporting—for the self-employed survey respondents. Assuming income reports in the survey match those given to the tax authority, they estimate that self-employed people in the United Kingdom on average underreported their income by about one-third.²⁸

Feldman and Slemrod (2007) pursue a similar methodology, but avoid the need to use survey data by instead using as the trace of true income charitable donations reported on U.S. income tax returns. They conclude that reported positive self-employment income of \$1 is associated with the

²⁷ A related traces approach is taken by Fisman and Wei (2004), who identify the response of evasion to product-specific import tariffs and value added tax in China. They measure evasion as the difference between exports to China from Hong Kong as reported in Hong Kong’s export data and in China’s import data, at the level of highly disaggregated products. In principle, these figures should be identical, but in practice imports are consistently underreported and the extent of underreporting is highly correlated with the return to evasion, the tax rate, such that a one percentage point increase in the tax rate is associated with 3 percent more misreporting, most of which seems to occur by mislabeling products as lower-taxed products. More recently, Handley and Moore (2017) show that reported (deductible) transport costs vary positively with tariff rates, which they attribute to misreporting.

²⁸ Tax evasion estimates for other countries using this method include Schuetze (2002) for Canada, Paulus (2015) for Estonia, Johansson (2005) for Finland, Martinez-Lopez (2013) for Spain, Engström and Holmlund (2009) and Engström and Hagen (2015) for Sweden, Cabral and Gemmell (2018) for New Zealand, Cabral, Kotsogiannis, and Myles (2014) for the United Kingdom, and Hurst, Li, and Pugsley (2014) for the United States.

same level of contributions as \$1.54 of wage and salary income, which implies—assuming a negligible wage and salary income noncompliance rate and that the self-employed are not inherently more charitable than others—a self-employment noncompliance rate of 35 percent ($0.54/1.54$), very similar to the Pissarides-Weber estimate but below the NRP figure discussed above; for positive farm net income, the implied noncompliance rate is 74 percent. Intriguingly, other things equal, *negative* reported values for self-employment income are associated with more contributions than is zero reported self-employment income, suggesting that on average these reported losses are associated with higher true incomes. Relative to Pissarides and Weber (1989), two aspects of this study are particularly worthy of note, one good and one not so good for the credibility of the estimates. It's good that the method does not require the researcher to classify a taxpayer as either an employee or self-employed, as discussed further below. It's not so good that the key assumption—that the conditional charity-income ratio does not vary by employment status—is stronger than the comparable assumption about food; indeed, Glazer and Konrad (1996) argue that some people donate to charities to signal wealth (or integrity), a motive that is arguably more relevant for some self-employed people.

Artavanis, Morse, and Tsoutsoura (2016) develop an ingenious research design that analyzes household microdata from one of ten large banks in Greece to estimate the relative extent of underreported income for self-employed Greek workers by type of occupation. They rely on the fact that in an environment with widespread evasion the banks have had to become skilled at inferring true income from reported income and other traces of income in order to remain competitive. Assuming that income is accurately reported for wage and salary earners, they estimate a credit supply equation for wage earners using reported income and information such as credit history, borrower characteristics, and local economic growth available to the bank. If this

credit supply equation to be valid for wage earners, they can infer the multipliers that the bank implicitly applies to reported self-employment income to assess the credit risk of those seeking loans. The authors estimate multipliers in excess of two for doctors, lawyers, engineers, scientists, accountants and financial service agents, indicating that for these professions reported income is less than half of true income. They note that even this may be an underestimate if the bank applies a discount for any additional income or collection risk assessed, or if the credit supply equation is biased because the bank applies a multiplier to wage earners' reported income because of suspected employee tax evasion.

In summary, these studies using a traces-of-income approach corroborate the finding from the tax gap analysis that evasion of self-employment income substantially exceeds that of employee income. This is a first-order demonstration of the explanatory power of the deterrence model.

Some of the underlying assumptions of this class of analyses have recently been re-examined. For example, income reports in household surveys have usually been treated as free of systematic misreporting, based on the reasoning that underreporting income on a survey does not reduce tax liability. However, Hurst, Li, and Pugsley (2014) argue that the benefit to a noncompliant individual household of reporting accurately in a survey setting is so small that even a slight probability that their report is not confidential could result in underreporting. Using a Pissarides-Weber-style methodology, they estimate that the self-employed underreport income in both the U.S. Consumer Expenditure Survey and the Panel Study of Income Dynamics by about 30 percent. This finding suggests that evidence on traces of evasion from survey data can be informative just as can what one observes in tax returns, as taxpayer reporting behavior may be similar in both settings.

All of these studies have specifically addressed the behavior of self-employed individuals relative to employees, who are assumed to not underreport their income. Clearly, a self-employed

individual with the same amount of income as an employee faces more opportunities and stronger incentives to evade taxes, as their income is not subject to third-party reporting. Nevertheless, some recent research has questioned the assumption that employees do not underreport income to the tax authorities. This could be particularly relevant for studies that do not identify taxpayers as self-employed by their source of income.²⁹ Individuals who might be full-time or part-time employees of a firm may also hold other jobs or perform freelance work and may self-identify as employees in a household survey while also receiving income from self-employment. If they fail to report this additional income, studies that assume no evasion among “employed” workers would underestimate evasion by those identified as “self-employed.”

Dunbar and Fu (2015) examine this issue by comparing imputations of households’ total annual expenditure to their reported annual income, using data from the Canadian Survey of Financial Security, the Survey of Household Spending, and income tax data reported by both employers and employees. They find evidence that income underreporting is not confined only to households that report some self-employment income; 30-40 percent of households that report only salaried income under-reported income, although the underreported income might not be wages and salaries. Paulus (2015) finds that Estonian private-sector employees underreport salary income. Using a dataset that links information from tax forms to a more comprehensive household survey, he estimates the extent of underreporting of income by employees whose income is subject to third-party reporting, arguing that both the employee and the employer have an incentive to coordinate and underreport income--the employer gains from owing lower payroll taxes and can also credibly lower reported revenue to save on value-added-tax liability. Making the slightly less

²⁹ The papers vary in their definition of a self-employed individual. Pissarides and Weber (1989) treat anyone as self-employed who reports more than 25 percent of their income as due to self-employment. Other research, such as Cabral, Kotsogiannis, and Myles (2015), uses taxpayers’ own categorizations of themselves to determine employment status.

restrictive assumption that (only) public-sector employees must (and do) report truthfully, he uses the correlation between income information in survey data and administrative data for these employees as a benchmark to compare to private-sector employees. He estimates that about 20 percent of private-sector employees in Estonia underreport income. Finally, Best (2014) analyzes matched tax administrative data on firms and salaried workers in Pakistan, and concludes that 19 percent of workers misreport their income.

These studies consistently show that evasion is substantially higher for income that is not subject to third-party reporting. The *level* of evasion, however, is more difficult to pin down. We can consider estimates from studies like Pissarides and Weber (1989) and Feldman and Slemrod (2007) as a lower bound of evasion by the self-employed. As Dunbar and Fu (2015) suggest, their crucial assumption that employees do not underreport may not always hold: individuals identified as “employees” in survey data underreport income earned outside of their primary job. Studies that estimate evasion by source of income instead of the employment status of the worker can come closer to a true estimate of the level of evasion. However, as Paulus (2015) finds, in some settings employees may underreport some portion of their income that is subject to third-party reporting.

The research into the level and nature of evasion informs policy because it illuminates “where the money is.” More crucial to policy is knowing the impact (and cost) of enforcement initiatives to reduce the level of evasion. From the perspective of the Allingham-Sandmo model, one might conclude that means learning about how p and f affect evasion. Next I address what has been learned recently about the impact of the principal linchpins of tax enforcement: audits, information reporting, and remittance rules. I begin by discussing a recent wave of research that has focused on communication from the tax authority to the taxpayer, usually, but not always, in the form of a letter.

4. LETTER, AND OTHER NOTIFICATION, INTERVENTIONS

4.1 Effects of Letter Interventions

The most prominent example of tax enforcement RCTs is letter interventions, where the tax authority sends one or more types of letters to randomly selected groups of taxpayers, with a control group receiving no letter. Because sending letters is a relatively inexpensive policy instrument, tax authorities have been particularly willing to partner with academics to design RCTs that measure their impact.

Before discussing what these RCTs have taught us, it is worthwhile to consider what a letter might say, what the words might convey to a taxpayer, and how this relates to the deterrence model of tax evasion laid out above. The content loosely falls within four categories, although often a particular intervention treatment combines aspects of more than one category.

In the first category are messages that provide information about the enforcement environment generally, such as the penalties for detected evasion (f). In principle a message in this category could convey the chances of detection (p) or of audit, but information about this is very difficult to convey precisely, especially as the probability depends on how much evasion is attempted, what kind of evasion is attempted, and so on.

In the second category are explicit audit threats—*your* return will be closely examined, or will be examined with some stated probability. The behavioral response to receiving such messages has the potential to shed light on the impact of changing p , but the interpretation is fairly subtle, because the reaction will depend on how the message *changes* the taxpayer's perceived $p(\cdot)$ function.

The third category includes messages that convey that the tax authority possesses personalized information that suggests that the taxpayer is, or has been, noncompliant. The information might

come from third-party-provided forms, information known to another jurisdiction, or from filings of other connected taxpayers. Possibly into this category would go the conveyance of apparently unthreatening information—“just checking in”—because the taxpayer may interpret this as a reminder that the tax authority knows of her existence and (potentially) her true tax liability.

The fourth category appeals to factors outside of the deterrence model, appealing to a sense of civic duty and norms of compliance, or else reminds the taxpayer of what programs tax revenue supports. Such messages would presumably be irrelevant to the free-riding, amoral taxpayer of the Allingham-Sandmo model, but might matter to the taxpayer who behaves in ways described in Section 1.4.

The studies track the reporting behavior (evasion is not directly observable), often relative to previous reporting, of those in these treatment groups compared to a control group that received no contact from the tax authority. With a few exceptions discussed below of studies that investigate network spillover effects of such interventions, these are all cases of what might be called specific dissuasion—where the intervention is intended to discourage noncompliance of the specific individual or business receiving the letter. This term is a modification of the concept of specific deterrence, which refers to a punishment intended to discourage criminal behavior in the specific individual receiving the punishment, and contrasts to general deterrence, which refers to inducing others tempted to commit this offence not to do so by changing the perceived chance of detection of noncompliance or the consequences of being detected evading.

In the first application to tax compliance of an RCT design, Slemrod, Blumenthal, and Christian (2001) analyzed the results of a randomized controlled experiment conducted by the State of Minnesota Department of Revenue (MDOR) and designed in cooperation with the

authors.³⁰ Randomly selected taxpayers within six distinct categories (low, middle, and high income, each with and without income not subject to employer withholding and information reporting) who filed a return for tax year 1993 were sent a letter from the MDOR in January of 1995. A treatment group was informed that their not-yet-filed 1994 federal and state income tax returns would be “closely examined”³¹ for any irregularities. The letters were sent after the tax year had ended, so that individuals generally could only respond by changing their reporting behavior. The authors compared the change in reported income (compared to tax year 1993) of this treatment group to that of a control group that did not receive any communication from the MDOR. Low- and middle- income taxpayers who received a letter promising a certain audit reported slightly, but statistically significantly, more, income than those who did not receive such a letter, and the difference was larger for those with better opportunities to evade in the form of income not subject to withholding and information reporting.

What do we learn from this result? First of all, if taxpayers surmised from this letter that any and all evasion would be discovered and penalized, then the conditional change in reported income is arguably a measure of the previous year’s evasion, as no rational taxpayer would evade if they believed the probability of detection to be 100 percent. More generally, the effect of receiving this letter on the beliefs of the treatment group depends on their prior beliefs about the probability of an audit. If the individual expected that their returns were routinely examined every year, then there might be little or no change in behavior. Others may have revised their probability of detection upward, but not to one. Thus, the measured increase in compliance does not reveal anything definitive about de/dp , even just in Minnesota in the mid-1990s, because the perceived

³⁰ Blumenthal, Christian, and Slemrod (2001) describe an appeal-to-conscience treatment arm of the same RCT, discussed in Section 8.3.

³¹ This language was chosen after careful consideration. Given the very limited auditing resources of the state, to say that returns would be “audited”, as the term is commonly understood, would probably have been deceptive.

value of dp is not known. It does, though, suggest that this type of intervention can reduce noncompliance, at least in the short run. It does not immediately inform about the longer-term impact for two reasons. First, the study did not track subsequent behavior. Second, even if it did, the effect in later periods would depend on whether changed perceptions persisted, and the promised interventions were perceived to be credible; an appropriate research design could in principle measure the latter, but the result might not be externally valid for other apparently similar interventions whose perceived credibility differed.

One other striking finding of Slemrod, Blumenthal, and Christian (2001) is worth mentioning: high-income taxpayers with self-employment or farm income receiving this audit-threat letter on average reported *lower* income, a seemingly perverse response from the perspective of the deterrence model. The authors speculate that more sophisticated taxpayers (and their accountants) understand an audit to be a negotiation, and view reported taxable income as the opening (low) bid in a negotiation that does not necessarily result in finding and penalizing all noncompliance; thus, the initial lower reported tax liability report might not indicate that the eventual tax remittance was lower, as well.³² This result offers a caveat that the compliance response to enforcement initiatives for very high-income individuals may be different than for others, and in particular may be very intervention-specific.

Kleven et al. (2011) conduct a similar audit experiment in Denmark. Randomly chosen individuals, some of whom had been (randomly) audited the previous year, received letters announcing either a 100 percent probability of audit or a 50 percent probability of audit, while a

³² In addition, if the taxpayer believes that there will be an audit regardless of what is on the return, then dp/de is zero rather than positive, and this reduces the optimal report, *ceteris paribus*. Conceivably, the letter could have induced people to seek out professional tax preparers, who alerted the taxpayer to legal means to reduce tax liability. However, receiving an audit-threat letter did not increase the fraction of these taxpayers that hired professional help (it was already above 90 percent), but perhaps the content and tone of the interaction changed.

control group received no letter. As the authors note, this is not the same as exposing taxpayers to a varying probability of *detection*, which is likely to be lower for self-reported income. The threat-of-audit experiment was conducted only on a sample of employees (i.e., it excluded any self-employed individuals). The individuals received the letter shortly after they received their pre-populated returns from the government with all relevant information known to the government from third-party information reports, and had one month to make adjustments to the return. Individuals who were threatened with a 50 percent probability of audit were about 1.1 percentage points more likely (14.5 percent versus 13.4 percent) to adjust net income upwards than those who received no letter, and those who received the 100 percent probability of audit were 0.9 percentage points more likely than the 50 percent threat of audit to do so.³³

Meiselman (forthcoming) uses an RCT to understand what messages from the tax authority are effective for eliciting a return from suspected resident nonfilers of the Detroit local income tax, assessing the efficacy of messages related to penalty salience, punishment probability, compliance cost, and civic pride. He finds that a penalty salience message that stated the statutory penalty for failing to file a return tripled response rates from 3 percent to 10 percent and raised the fraction of filed returns that admitted tax due from 39 percent to 52 percent. Compliance cost mailings that enclosed a blank tax return and punishment probability mailings that stated the recipient's federal income also raised response rates relative to the basic mailing, but civic pride mailings did not.

In an ambitious recent attempt to more closely tie the results of a letter intervention to the Allingham-Sandmo deterrence theory, Bergolo et al. (2018) carried out a large-scale field experiment involving value-added-tax compliance of over 20,000 Uruguayan small- and medium-

³³ The impact did not vary significantly depending on whether the taxpayers had received a random audit the year before, so no evidence of intertemporal cross-effects was discovered.

sized firms. Firms in a control group received a letter from the tax authority with generic information about taxes. Firms in the main treatment arm received the same letter with an added paragraph, conveying information about past audit rates and the penalty levels for tax evasion, based on historical means from a firm-specific sub-sample. Although in the main body of the letters these statistics are referred to as the “likelihood” of being audited and the likely penalty if evasion is determined to have occurred, elsewhere in the letter exactly how the figures were calculated is described in detail. In separate treatment arms, they measure the effect of varying the actual audit probability communicated to a firm between 25 percent and 50 percent, a letter stating that evasion increases the chance of being audited, and a public-goods letter listing a set of government services that could be provided if evasion in Uruguay ceased. They supplement these treatments with information on taxpayers’ subsequent perceptions about audits, measured with survey data, as well as on the actual taxes paid.

Bergolo et al (2018) find that adding to the baseline letter a paragraph with statistics about the probability of being audited and the penalty rates increased tax compliance by about 6.3 percent (about one-fourth of its current level), and adding a paragraph that informs firms that evading taxes increased the probability of being audited increases tax compliance by about 7.4 percent. In contrast, the message about public goods did not have a statistically significant and robust effect on tax compliance. Intriguingly, among firms who were sent the audit-statistics letter, those who received higher signals of the audit probability or penalty rates did not remit significantly higher taxes, nor did the variation in actual audit probabilities induce a significant change in reporting behavior; this contrasts to the result in Kleven et al. (2011) discussed just above. The authors conclude that these results cast doubt on whether firms are making an optimal cost-benefit calculation as in Allingham and Sandmo (1972); perhaps instead, they argue, taxpayers are

reacting to the information because it makes the cost of evading more *salient*, without any changes in beliefs about $p(\cdot)$. They argue that the results from the survey also favor the salience channel, because they suggest that on average the audit-statistics letter *reduced* the perceived probability of being audited and, thus, rational Allingham-Sandmo taxpayers would have reduced, rather than increased, their tax compliance, which did not happen.

These studies consistently indicate that letter interventions can substantially reduce noncompliance in the short term. They are less definitive regarding exactly what the mechanism is, with one recent study arguing that the letter affected the salience rather than the perception of the enforcement environment.

4.2 Mode of Intervention

One natural question that arises is the extent to which the delivery mechanism matters for a given message or, in other words, how the delivery mode affects what message is received. Ortega and Scartascini (2017) investigate this question by conducting a field experiment in Colombia that varies the way the National Tax Agency contacts taxpayers with tax delinquencies to deliver information about the account balance, the type of tax, the year or month it had not been paid, plus information on methods of payment and the cost that the taxpayer was incurring by not paying as well as a short moral suasion message. Taxpayers were randomly assigned to a control group, or to one of three possible delivery mechanisms: letter, email, or personalized visit by a tax inspector. The authors find sizable differences across delivery methods. Personal visits by a tax inspector are more effective than the impersonal methods; they are, alas, also much more expensive. In a follow-up study, Ortega and Scartascini (2015) find that the effect of phone calls falls between those of the impersonal methods and the personal visits.

Boning et al. (2018) investigate, using an RCT design, how U.S. businesses at risk of being delinquent in their tax remittances for IRS Form 941 (for withholding and social insurance taxes) respond to letters versus Revenue Officer visits. They find very large and immediate positive effects on remittance of the Revenue Officer visits that persist for at least four quarters thereafter, and a positive but substantially smaller and less persistent effect of the letter treatment. This study and the Ortega-Scartascini one speak with one voice that in-person contact has a larger effect on behavior than a letter. It is not obvious whether this differential response is due to a personal visit convincing recipients that the tax authority has become more serious about enforcement, and thereby increasing the perceived value of p , or whether it is a more successful in getting taxpayers to pay closer attention to an unchanged set of enforcement parameters.

4.3 Other than Income Taxes

Some recent research involving randomized controlled trials addresses the enforcement of taxes other than income tax. Castro and Scartascini (2015) focus on remittance of a municipal property tax in Argentina that differs from an income tax in some important ways relevant for enforcement. For one, revenue is directly linked to visible provision of public goods like street lights and trash collection, so that taxpayers may be more likely to connect their remittances to provision of these public goods. Second, payments are calculated on the basis of length of the property facing the street, number of street lights and trash collection services received at the property, and there is arguably little room for successfully misreporting these measures. The authors send three types of letters to test the effect of appeals to fairness, equity, and deterrence. They find that the deterrence messages have the strongest effect; informing taxpayers of the penalties of nonpayment increases the probability of remittance by 5 percentage points from a base of 40 percent.

In a field experiment in Austria, Fellner, Sausgruber, and Traxler (2013) use a randomized design to test the effect on compliance with Austrian television and radio licensing fees of various mailings to potential non-compliers. Austrian households owning a radio or television are required to remit a licensing fee; collection of the annual fee relies on self-reporting and individuals can access public broadcasting channels without paying the fee. In 2005, 94 percent of households were registered and paid a licensing fee, but only 1 percent of households owned neither a TV nor a radio, suggesting the presence of evasion. The authors sent letters emphasizing different messages to five treatment groups within a population identified as potential evaders by the enforcement authority. One publicized the threat of detection and sanction, another was a moral appeal equating compliance with fairness, and a third variant provided social information on the overall high level of compliance. Two others interacted the threat of detection with the moral appeal treatment and with the provision of social information. Those receiving any type of mailing were significantly more likely to make a payment within 50 days of receiving the letter, but only the variant emphasizing the threat of punishment induced an additional increase in compliance. The authors interpret the generic effect of the mailing as an “alert effect” signaling that nonpayment had been noticed, with the consequences of noncompliance amplified by the threat variant. The fact that *any* contact from the tax authority might affect compliance, at least in the short term, is a common finding in recent compliance research.

Using a randomized controlled trial in Philadelphia, Chirico et al. (2016) study the extent to which property tax delinquency can be attributed to lack of salience, deterrence or tax morale. The study compares seven treatment arms of 19,000 newly delinquent taxpayers delivered through letters. These treatments range from a simple reminder of the liability, including accrued interest and penalties, to threats of a sanction and moral appeals. They find that moral appeals have no

more effect on payment than a simple reminder. Like other papers in this literature, Chirico et al. find the threat of sanctions to be most effective, raising more than \$65 for each additional dollar of administrative costs at the margin. Although the simple reminders increased taxpayer compliance in the immediate tax cycle, they did not in subsequent cycles, suggesting that the salience effect is short-lived.

5. AUDITS

Distinct from the impact on evasion of a change in the perceived probability of detection of noncompliance is the effect of audits on the audited--the specific deterrence effect. Why might an audit of this year's income tax return affect future reporting behavior? In the deterrence framework, it would matter only to the extent it affects the audited taxpayers' perceptions about $p(\cdot)$ or f , and exactly how they would be affected is ambiguous. A taxpayer may presume that the probability of getting audited a second time is low, sometimes referred to as the "crater effect," which would mean that taxpayers are less likely to comply in the years following an audit. Or, a taxpayer may revise upwards her prior on the probability of an audit, and would therefore be more likely to comply in the years following an audit. Of course, what matters is taxpayers' perceptions of the chance of evasion being detected and punished, not the chance of an audit per se, and an audit may change one's perceptions of what an auditor knows and can reasonably find out, whether an audit is akin to a negotiation, and so on; some taxpayers may be impressed with the information and skill of the auditor, and so revise upward their perceived p , while others may realize that at least for some sources of income the auditor knows little or nothing and does not have the resources to learn much more, and thus revise their perceived p downward.

Three recent studies have examined this issue.³⁴ DeBacker et al. (2015) use IRS data from the National Research Program (NRP) to study the behavior of audited individuals in the years following an audit. They construct a control group by randomly selecting (unaudited) returns from the same sampling pool as the NRP and who thus have similar characteristics to the randomly audited NRP sample. They find that an audit increases reported wage income over three years after the audit by 0.45 percent and increases self-employment (Schedule C) income by 7.51 percent. However, this large estimated effect on reported self-employment income is short-lived; indeed, five or six years following the audit, the treated group actually reports *lower* Schedule C income as compared to the control group. When they compare the reporting behavior of the same individual pre- and post-audit, they find the same positive effect of an audit. Kleven et al. (2011) also found, in Denmark, a positive deterrence effect of audits on subsequent evasion, with the effect being entirely composed of changes in self-reported income, with no effect found on third-party-reported income.

The external validity of these results is an issue. Because taxpayers audited under the NRP are informed that they have been randomly selected for research purposes, these audits may not have the same impact as an operational audit on the perceived probability of a future audit—the taxpayer learns nothing about whether the tax authority is “on to them.” Moreover, these taxpayers are not representative of those who are typically subject to audit, and their behavioral response behavior may not be representative of those who are normally targeted for operational audits.³⁵

Advani, Elming, and Shaw (2017) pursue a similar research strategy using data from the United Kingdom. The HM Revenue & Customs conducts a random audit each year similar to the NRP.

³⁴ Earlier studies include Long and Schwartz (1987) and Erard (1992), both of which did not find significant results.

³⁵ This raises a methodological Catch-22, of course, because the regular, operational, audits are not randomly chosen, raising difficult sample selection problems.

Instead of resampling the subject pool as DeBacker et al. (2015) do, the authors use individuals who appear in the treatment group in future years as the control group for those audited in the current year. For instance, individuals who were selected for random audit in 2006 and beyond are used as the control group for those audited in 2005. Consistent with the DeBacker et al. findings, those who are audited increase their reported tax liability more than the control group in years following the audit. Also based on UK audit data, Gemmell and Ratto (2012) found the response to depend on the results of the audit. In particular, those who were found to be compliant, reduced their subsequent compliance, while those found to be noncompliant increased their subsequent compliance.

There is one important difference between these two sets of studies of the specific effect of audits that makes the results difficult to compare. An American (or Danish) taxpayer who receives one of the random audits is informed that the return was chosen randomly—nothing about the return as filed triggered the audit, or increased its likelihood. In contrast, in the UK study the taxpayer was not told whether it was a random or targeted audit (Advani, Elming, and Shaw, 2017, p. 9). It follows that the US and UK audits would likely trigger very different re-evaluations of how likely a future audit is, and therefore induce different behavioral changes.

That audits matter other than for their general deterrent effect is also suggested by taxpayers' efforts to sidestep them. Almunia and Lopez-Rodriguez (2018) study the behavior of Spanish firms in response to a notch in enforcement intensity that arises because the Spanish Large Taxpayer's Unit (LTU) monitors firms with revenues above €6 million for all major taxes. Even though the compliance requirements and tax rates are the same above and below this threshold, enforcement changes discontinuously because the LTU has the resources to conduct more audits and to effectively utilize technology to cross-check reported information. Revenues,

and certainly reported revenues, are subject to firm choice, so that in the absence of prohibitive costs to changing firm size, one would expect a hole in the distribution of firms that report revenues just above the €6 million cutoff and bunching just below it: firms can earn the same expected after-tax income if they remain smaller and escape the intensive monitoring by the LTU. Sure enough, there is significant bunching of firms just below the threshold. The bunching is more pronounced for firms that produce intermediate goods, which makes sense because their transactions create more of a paper trail than firms that sell to final consumers and thus the discontinuous increase in enforcement intensity affects these firms more than retailers.

To combat sales and profit tax evasion by small firms and the self-employed, many developing countries have adopted some form of “reverse withholding,” where large firms withhold and remit to the tax authority a fixed share of their purchases from small firms (and individuals), who can then apply the withheld amount as a credit against their self-reported tax liability. While withholding does not change the firms’ true tax liability, there is typically a discontinuity in the audit probability at the withholding rate; firms seeking tax refunds (because self-reported tax liability is lower than the withheld amount) are audited at a higher rate than firms making additional tax remittances. Examining data from Ecuador, Carrillo, Emran, and Rivadeneira (2012) find evidence of bunching in reported tax liability just above the withholding threshold, suggesting that some firms manipulate their self-reported tax liability and possibly real economic choices to minimize expected tax payments subject to the discontinuity in the audit probability. The pattern of bunching changed dramatically in 2007 only for firms subject to a change in the required withholding rate, ruling out the possibility that the withholding rate had been chosen to match the distribution of true tax liabilities. Third-party data on sales and intermediate input costs reported by large firms who act as

withholding agents indicate bunching is indeed associated with tax evasion: self-reported sales are smaller than third-party reports for more than 10 percent of firms.

In principle, quantifying the extent to which taxpayers will alter their behavior to lower the probability of an audit can help reveal the expected cost of an audit, which according to expression (1) is $(1+ft)e$, and this is proportional to the amount of evasion as well as the penalty for detected evasion. This exercise has not yet been attempted.

6. INFORMATION REPORTING

In Section 4 one objective was to understand the effect of changing the perceived enforcement parameters via a direct communication (usually a letter) from the tax authority to the taxpayer. Another set of studies makes explicit why the probability has gone up. I next review research where the reason that $p(\cdot)$ increases is increased third-party information reporting, where entities (usually firms) report to the tax authority information relevant for someone else's tax liability. Information reporting by employers for employee wages and salaries is pervasive, and is also widespread for dividends, interest, share sales and real estate sales. Information reporting is also built into some invoice-credit value-added tax (VAT) systems, when credits for purchases from other firms are allowed only if accompanied by information on the seller, which can in principle be checked against the VAT returns of those sellers.

6.1 Firms

—Field evidence on Chilean firms' compliance with the VAT highlights the connection between information reports received by the tax authority and levels of evasion. Because firms can only claim tax credits for inputs bought from tax-compliant suppliers, the invoice-credit VAT system has a built-in (albeit imperfect) self-enforcement mechanism for firm-to-firm transactions.

Pomeranz (2015) tests this hypothesis by mailing increased-audit-threat letters to over 100,000 randomly selected Chilean firms, using a sample of over 300,000 firms receiving no letter as the control group. Consistent with theoretical predictions on the self-enforcement mechanism, the increase in VAT receipts (and therefore the inferred level of evasion) induced by the letters is concentrated at the level of sales from firms to final consumers, for which there is no paper trail.

Carrillo, Pomeranz, and Singhal (2017) examine the effect of a change in the tax authority's use of third-party information on reported corporate tax revenues. In Ecuador, the government has a few sources it can use to check firms' self-reports of revenue, including other firms' reports of purchases from the firm in question, credit-card sales from credit-card companies, as well as export and import information from the Ecuadorian customs authority. Although for a few years the tax authority had collected such third-party reports of firm revenues, it had not utilized this information. In the episode the authors study, the Ecuadorian tax authority informed some firms of the discrepancy between the two reports and offered them the opportunity to file an amended return. The authors compare the reporting behavior of firms before and after notification, and find that 24 percent of firms underreported revenue in the years when the government did not use the third-party-verified information. In the three rounds of the experiment, between 11 and 19 percent of notified firms filed an amended return.

One striking finding of this study is that, in amended returns, firms correctly report their revenues but also increased their reported costs almost one-for-one with the increase in reported revenues (96 cents for each dollar!), nearly eliminating any impact on apparent tax liability. The offset of reported expenses is similar to the finding of Slemrod et al. (2017) regarding the US Form

1099-K requiring credit-card companies to report business receipts, discussed below.³⁶ Because the tax authority does not have a comprehensive picture of firms' costs and revenues through third-party information, many firms apparently continued to reduce their tax liability through channels not covered by third-party reporting. The experience in Ecuador suggests that the comprehensiveness of third-party information is crucial and that, in the context of many developing countries where such comprehensive information is not available, one might not observe a sharp fall in overall evasion from using an additional source of third-party information. This study also reveals that reported costs were lower than third-party information on costs, which at first blush seems to be at odds with a model of firms that seek to maximize after-tax profits. The authors propose that this behavior is consistent with firms who may believe that the probability of an audit is a function of firm size and profits. In order to appear small, and lower the odds of being audited, some firms may underreport both revenues and costs.

6.2 Involving Charitable Contributions and Other Deductions

Two recent papers examine the impact on charitable contributions of altering the information reporting system involving charities. Fack and Landais (2016) do so by exploiting a change in the French tax treatment of charitable donations. Since the early 1970s, charities in France had been required to issue standardized receipts to donors but, starting in 1983, the reporting rules began to require taxpayers to attach these receipts to their tax filing when they claim deductions, lowering the cost to the tax authority of verifying claimed deductions. Fack and Landais find that reported donations fell by a stunning 75 percent after the introduction of this change, and argue that this decline resulted from a sharp drop in over-reporting rather than a decline in actual donations. Because the new rule only required donors to attach receipts they were presumably already

³⁶ In the same vein, Asatryan and Peichl (2017) find that for every dollar of audit-driven increase in reported revenue, firms in Armenia increase reported deductions by 70 to 80 cents.

receiving, it is not plausible that the decline was due to increased compliance costs. The authors also calculate the net-of-tax price elasticity of contributions before and after this increase in enforcement intensity, and conclude that the estimated elasticity before the enforcement change is about three times larger than the estimate after 1983. This is an illustration of the endogeneity of the elasticity of (one component of) taxable income to the vector of tax policy instruments, as stressed by Slemrod and Kopczuk (2002) and investigated empirically by Kopczuk (2005): a more effective enforcement regime apparently reduced the tax-price elasticity substantially.

Gillitzer and Skov (2016) examine the effect of a different kind of change in third-party reporting on claims of charitable contributions in Denmark. Starting in 2008, charities in Denmark were required to report contributions to the tax authority, which would then pre-populate individual tax returns with the information; individual taxpayers could either accept this information or amend it. This reduced compliance costs to individual donors, while also increasing the probability that a false claim by an individual would be detected. Gillitzer and Skov find that the number of claims actually *increased* substantially after 2008. Apparently, the effect of lower compliance costs far outweighed the effect of increased enforcement from third-party reporting, inducing people on average to report some tax-deductible contributions that they otherwise would not have bothered to claim. On the surface, this result appears to contradict the findings of Fack and Landais (2016), who found that requiring receipts from donors greatly reduced claims. It is possible that over-reporting of charitable contributions as a means of tax evasion is not as pervasive in Denmark as in France due to the difference in the cost of this channel of evasion.

Research by Kotakorpi and Laamanen (2016) sheds light on the response of deductions other than charity to a change in information reporting. They evaluate the impact of a policy experiment in Finland in the 1990s under which a proportion of taxpayers received a pre-filled income tax

return, whereas other taxpayers had to file a full return. They concluded that receiving a pre-filled income tax return led to a significant *reduction* in the number of individuals claiming deductions. Although which return one received did not affect the actual chance of evasion being detected, the authors argue that considerations related to tax evasion might have been in play if the reform affected individuals' perceived probability of detection: individuals receiving the pre-filled form might have become worried that the authorities also had information on other items that were not printed on the form. If this is the mechanism behind the results, it would imply that in the new system some individuals who would otherwise have claimed some unwarranted deductions did not do so upon receiving a pre-filled return.

6.3 Involving Credit-Card Companies

In an effort to reduce understatement of revenues, the US Housing Assistance Act of 2008 required that, as of January 1, 2011, payment settlement entities such as credit-card companies and other third-party payment organizations had to report on a new Form 1099-K electronic payments received by businesses. Analyzing administrative data on the universe of individual income tax returns that report sole proprietor income, Slemrod et al. (2017) find a large increase in the number of businesses reporting income that is exactly equal to the amount in the 1099-K report, consistent with a simple model of optimal reporting behavior. The new reporting requirement increased reported receipts of this relatively small group by up to 24 percent, although this was offset somewhat by a 13 percent increase in reported expenses. They also find that more than 20 percent of the affected group were induced by the introduction of Form 1099-K to file a Schedule C for self-employment income. Thus, information reporting seems to have had the intended effect of inducing more accurate reporting of business receipts, but the overall effect on evasion was

dampened by increased reported expenses, which are not directly observable to the tax authority;³⁷ the revenue effect may also have been offset by firms moving to cash receipts that are not covered by this information-reporting regime.

The phenomenon of one type of misreporting increasing when another declines due to increased enforcement has now been demonstrated multiple times, and calls out for more theoretical and empirical investigation—under what conditions will it happen, and under what conditions will legal avoidance displace evasion when it becomes less attractive? Yang (2008) documents another example of this phenomenon, where in the Philippines increased enforcement against a specific method of evading import duties reduced this type of behavior, but resulted in substantial displacement to an alternative method that had not been targeted, such that a zero change in total duty evasion could not be ruled out.

6.4 Involving Consumers and Workers

On the grounds of administrative efficiency, modern tax systems have for the most part largely excluded people in their role as consumers from remittance responsibility or even information reporting, relying on tax remittance from firms even for consumption taxes; neither retail sales tax nor value-added taxes generally involve consumer participation. Modern systems have also de-emphasized the role in remittance of people in their role as employees, looking to employers via withholding for the bulk of tax collection, and in exact withholding systems for *all* of collection. In general it is more efficient to rely on firms, especially larger firms that can take advantage of economies of scale and accounting systems already in place for non-tax reasons, to take the lead role in remittance. This system precludes using consumers and workers as a check on firm

³⁷ Note, though, that in the United States in an audit the burden of proof for expenses rests on the taxpayer, while it rests on the IRS for receipts.

compliance, however, and some recent research explores the possibility of making use of such checks.

Beginning in 2007, in an effort to reduce VAT evasion by retail firms, the São Paulo tax authority provided monetary incentives to customers to report evasion by retail firms in a program called the Nota Fiscal Paulista (NFP). Consumers receive tax rebates and are entered into lotteries in exchange for requesting receipts, and can also check online firms' reports of their transactions with the consumer and report any discrepancies. Programs with some similar features exist in China, Malaysia, the Philippines, Portugal, Puerto Rico, Slovakia, and Taiwan.³⁸ Naritomi (2016) finds that over a four-year period after the introduction of the NFP, retail firms increased their reported revenues by 22 percent compared to a control group of wholesale firms (who are not affected by the NFP). In addition, on average a firm's reported receipts go up by 14 percent right after they receive their first consumer complaint.³⁹ Consumers are encouraged to utilize what they know to check against what firms reported, facilitated by the fact that each retail purchase receipt contains the Social Security number of the purchaser.

Kumler, Verhoogen, and Frias (2013) study the effects on evasion of a 1997 pension reform in Mexico that tied younger workers' retirement benefits more closely to the wage reported by employers. This reform provided a new incentive for this group of workers to ensure that their employers accurately reported their wages, which in turn might lower payroll tax evasion by firms. To examine the impact of this initiative, the authors combine two sources of data on wages: administrative data from the Mexican Social Security agency (IMSS) and household survey data from the Encuesta Nacional de Empleo Urbano (ENEU). The authors find that, consistent with a

³⁸ See Marchese (2009).

³⁹ Note that this policy involves changes both to the information-reporting and remittance regimes, as discussed in Section 7.

decline in evasion, the gap in median or mean wage within a cell between the ENEU data and the IMSS data fell for younger age groups after the pension reform but, as predicted, for older workers not affected by this reform there was no decrease in the gap between the two income reports.

6.5. Involving Other Countries

The US Foreign Account Tax Compliance Act (FATCA) initiative takes information reporting to another level, by inducing⁴⁰ foreign financial institutions to report to the IRS, directly or through their home government, about the foreign accounts and income accruing of US citizens. Much anecdotal evidence suggests that the compliance costs of this effort (mostly borne initially by foreign institutions) are substantial, so (at least from a global perspective) the hurdle should be high for its compliance impact. However, because this crackdown on noncompliance applies mostly to high-income households, due to distributional concerns the dollar-against-dollar hurdle might be somewhat lower than otherwise. Notably, much of the world has committed to implement something similar to FATCA, called the Common Reporting Standard, by 2018.

7. THE REMITTANCE SYSTEM

Although public finance textbooks downplay its importance and assert, usually without any proof, that who or what entity remits a given tax—say the buyer versus the seller of a good—has no effect, the remittance system can in some settings be of first-order importance in efficiently enforcing and administering a tax system. Getting the money from what Logue and Slemrod (2009) call the “low-cost remitters” can matter. The textbook remittance invariance result depends, *inter alia*, on an assumption that evasion opportunities are essentially the same for all remitters, which is often not true.⁴¹ That this lesson has been widely recognized is exhibited most clearly in the

⁴⁰ As of July 1, 2014, failure to comply can trigger a 30 percent withholding tax on U.S.-source payments.

⁴¹ See Slemrod (2008) for further discussion of remittance invariance.

system of employer withholding almost all countries use for income and payroll taxes: it is more cost-effective to deal with collecting tax liability from a small number of organizations with relatively efficient bookkeeping done for non-tax reasons. It is also evident in the near-universal (the United States being the prominent holdout) adoption of the value-added tax, which diffuses remittance responsibility of a consumption tax to all firms—rather than just retailers—but almost always induces self-enforcement because business inputs are only creditable if purchased from verifiably tax-compliant firms. Some recent research has sought to quantify the effects of changes to a country's remittance regime.⁴²

Brockmeyer and Hernandez (2017) study the impact on compliance of tax withholding on business sales done by credit-card companies in Costa Rica. Credit card companies report all card transactions to the Costa Rican tax authority, as with the Form 1099-K in the United States, but also withhold (i.e., remit to the tax authority) a fraction of the sales, which can be credited against firms' sales tax liability in the same or future fiscal periods; firms with zero tax liability may request a refund, but this process is slow and onerous. A reform in August 2011 changed the withholding rate, on average doubling it but altering the rate with considerable heterogeneity. The study analyzes the change in the amount of tax remitted by firms that experienced an increase in the predicted withholding rate, compared to the change in tax remitted by firms that either experienced no change or were not subject to withholding. If fully and immediately credited by all firms, the change in remittance system should not change total tax collections. But it did, such that doubling the rate of withholding increased sales tax collection from those subject to withholding

⁴² Dusek and Bagchi (2017) examine the staggered adoption of employer withholding for US state income taxes in the 1940s through 1970s, and conclude that on average it led to an immediate and permanent increase in income tax collections of about 25 percent, holding tax rates constant, but they do not pursue the role of evasion in this response.

by 33 percent. The authors investigate two possible reasons for this result. The first is that reclaiming the withheld tax either as a credit against their tax liability or as a refund is costly; for some firms (especially small ones) these costs outweigh the benefits, and so they ended up over-remitting. These small firms were also more likely to under-remmit or fail to remit their reported tax liabilities prior to the reform, and thus the introduction of withholding raised compliance as well. In the second channel, firms whose tax liability was not withheld previously but would be after the reform now see an added column in their credit card statements for tax withheld. Firms whose withholding rate increases from zero exhibit a larger increase in their reported tax liability than firms that experience an equivalent withholding rate increase at higher initial values, suggesting that noticing the application of withholding for the first time may raise the salience of possible enforcement interventions, and thus lead firms to increase compliance. Notably, the analysis found no evidence that firms suddenly stopped accepting credit cards, although firms with low sales volume that accept credit cards had a small decline in the value of credit card transactions.

Another case where the remittance system apparently matters is diesel fuel taxation. Kopczuk et al. (2016) present empirical evidence that the identity of the remitting party in the U.S. diesel fuel market affects both evasion and therefore collections as well as the pass-through of taxes. Retail diesel prices are higher, and diesel taxes are passed through to retail prices to a greater extent, in states where the point of collection is at the distributor or prime supplier level rather than at the retail level, suggesting that these collection regimes reduce evasion compared to collection at the retailer stage. Note that some of the information-reporting initiatives discussed in this paper also shift the remittance pattern. The São Paulo system of involving final consumers in the VAT offers rewards to those who participate; in essence, the corroborative information comes with a negative remittance from consumers, which must ultimately be offset by larger remittance

elsewhere in the VAT chain, in other taxes, or lower expenditures. Another example is the dual landlord-tenant remittance system of the Italian TASI property tax (now mostly abolished), which assigned the remittance responsibility to both landlords and tenants and thereby generated information from two sources that could be compared. A crucial difference between classifying, say, Uber drivers as employees or as independent contractors is that only in the former case would Uber be responsible for withholding an approximation of the income tax liability for the drivers. Recently, Airbnb, Inc. has entered into agreements with certain cities to remit the hotel tax liability rather than relying on the compliance of the property hosts. Wilking (2017) examines the impact of these agreements on the prices of Airbnb properties, and finds that shifting the remittance obligation from the property host to the platform increases after-tax prices, suggesting that the change in remittance regime increased compliance.

Thus, a number of studies have shown that the tax remittance regime can in some settings affect the extent of evasion. Indeed, the remittance regime is an aspect of the enforcement regime. To be sure, in many settings variations in the remittance regime will not be of first-order importance, and future research should aim to clarify the conditions when remittance details matter.

8. OTHER ENFORCEMENT TOOLS

8.1 Public Disclosure

In Sections 5, 6, and 7 I have discussed the most important components of tax enforcement. But many other aspects of the enforcement environment can impact compliance, and I discuss a couple next, beginning with public disclosure of tax information. Public disclosure could complement deterrence by encouraging people with relevant information about others' true tax

liability to come forward as whistleblowers, and the fear of this occurring and subsequent tax noncompliance penalties (monetary as well as shaming) may constrain evasion. Opponents cite privacy concerns for individuals, and some businesses complain that complicated tax situations may be misunderstood or that proprietary information may be revealed. In principle, disclosure could backfire if some taxpayers reduce reported taxable income in order to minimize the attention of the press and those seeking to take advantage of their economic situation. On the other hand, some people might derive utility (bragging rights, if you will) from public appreciation of their apparent level of affluence, and may be willing to pay for it in the form of a higher tax liability.⁴³

Which incentives predominate is an empirical question. Hasegawa et al. (2013) study the impact of the Japanese income tax disclosure system that was abolished in 2004/2005 on tax reports of individuals and businesses. They focus on the fact that the disclosure regime applied only to taxable incomes above ¥40 million (about \$400,000), and find strong evidence based on “missing returns” with reported taxable income just above the disclosure threshold that, on average, individuals and businesses prefer to avoid disclosure. For businesses, this is consistent with the local characterization of so-called “39 companies,” whose reported taxable income was kept just below the ¥40 million disclosure cutoff so as not to provide evidence about their profitability, which might (so the story went) negatively affect the deals they could make with other companies. However, this study uncovered no evidence that disclosure increased reported taxable business income generally, and thus did not demonstrate its effectiveness as a compliance policy instrument.

Bø, Slemrod, and Thoresen (2015) explore the effect of public tax-return disclosure in Norway, which has a long history of disclosing tax filings, but in 2001 experienced a drastic expansion of

⁴³ The pros and cons of public disclosure are discussed at length in Lenter, Shackelford, and Slemrod (2003) and Blank (2014b).

effective public disclosure when anyone with access to the Internet could obtain information on other Norwegians' taxable income and income tax liability, taxable wealth and wealth tax liability. The authors exploit this change in the accessibility of the tax return data to measure the effects of public disclosure on income reporting. However, because this change happened in all of Norway at the same time, its aggregate effect may be conflated with the effect of other events happening contemporaneously. In this paper, identification of the deterrence effects of increased public access is facilitated by the fact that, prior to the shift to the Internet in 2001, in some municipalities easy access already existed because tax information was distributed widely through paper catalogues that were locally produced and disseminated as a fund raiser. The authors observe reported income changes that are consistent with public disclosure deterring tax evasion: an approximately 3 percent higher average increase in reported income is found among business owners living in (catalogue-free) areas where the switch to Internet disclosure represented a relatively large change in access. Thus, increased public disclosure was apparently an evasion deterrent to some taxpayers.

Malik (2017) studies the impact of the Pakistani government's decision to publish income tax records of all legislators several months after the tax returns had been filed. Using a differences-in-difference research, Malik finds evidence that the pressure to respond by reducing tax evasion was highest for competitively elected legislators and directly elected legislators.

Hoopes, Robinson, and Slemrod (forthcoming) investigate the impacts of public disclosure of information from corporate tax returns filed in Australia in 2014 and 2015 on consumers, investors, and the corporations themselves that were subject to disclosure. The results show that large private companies are likely to experience consumer backlash and are also, perhaps as a consequence, more likely to act to avoid disclosure. But the analysis does not reveal any material increase in tax remittances, one objective of legislating the disclosure regime. Investors, however,

apparently did react negatively to anticipated and actual disclosure of tax information, most likely due to anticipated policy backlash than the revelation of negative tax information.

In sum, the compliance-promoting effectiveness of public disclosure has not yet been widely demonstrated, and has not been shown in any setting for business tax compliance. There is some indication that it can be particularly effective for taxpayer groups that are especially sensitive to shaming.

8.2 Simplified Tax Regimes

It is often argued that the complexity of the tax system contributes to tax noncompliance.⁴⁴ Many countries offer smaller firms some form of a minimum alternative tax with a more easily measurable tax base, where the tax regime changes at a revenue or profit rate threshold. In Pakistan, corporations either pay a tax on profits or on turnover depending on which liability is greater. This effectively implies that at a profit rate lower than the ratio of the turnover tax rate to the profit tax rate, firms cannot deduct costs. Because in Pakistan a large portion of evasion is through misreporting of costs, this tax regime trades off loss in production efficiency (as firms move from a neutral profit tax to a distorting output tax regime) for a gain in revenue collection. Best et al. (2015) use administrative data on the universe of corporations in Pakistan to estimate the elasticity of taxable income using the bunching of firms below the threshold profit rate. They find clear evidence of such bunching, whose location shifts along with changes in tax rates that move the threshold. Using the analysis-of-bunching methodology, they estimate that turnover taxes reduce evasion by between 60 and 70 percent.

The introduction of the value-added tax in Japan in 1989 allowed firms with sales below ¥500 million in sales to opt for a simplified filing regime under which they could claim a fixed portion

⁴⁴ Ulph (2015) discusses the aspects of tax complexity, and addresses this claim.

of their sales (usually 80 percent) as input costs. Firms whose true input costs were below this threshold had an incentive to manipulate their size or structure to be eligible for this simplified filing. Firms could accomplish this through tax-motivated “splitting”—either transferring a portion of their operations to an existing small firm, or spinning off a portion of their firms as a new firm below the size threshold—or simply misreporting their sales. Onji (2009) studies the behavioral response to this system by constructing a counterfactual density of firms that allows one to separate changes in the density due to changes in the distribution of characteristics of firms from those due to the introduction of the tax benefit threshold. He finds that there is a bunching of firms below the threshold as well as a missing mass of firms right above the threshold, implying that Japanese firms did respond to the new tax incentive by, in part, masquerading as smaller firms.

Asatryan and Peichl (2017) study firms’ responses to size-dependent notches in the tax system of Armenia, a country with high rates of tax noncompliance. They study firms’ response by analyzing bunching in the distribution of firm size with respect to three notches: (1) the VAT registration threshold, (2) the threshold between quarterly and monthly filing and remittance requirements, and 3) the size determining whether international financial accounting standards or simplified rules are required. The authors find little responsiveness to the VAT registration notch, more to the filing and remittance frequency, and a large response to the accounting regime notch. Because they also have data on tax audit results, they can estimate that most of this response (at least 60 percent) occurs through under-reporting of income; firms escape the compliance costs of stricter accounting rules by understating their income.

8.3 Non-Deterrence Policies to Reduce Noncompliance

Enforcement policy could also attempt to enhance tax morale, and several studies have addressed how a government might try to achieve this. Many of the threat-of-audit letter RCTs

discussed earlier also contained a non-deterrence treatment. In the first tax RCT discussed earlier, Blumenthal, Christian, and Slemrod (2001) find no evidence that either of two written appeals to taxpayers' consciences had a significant effect on compliance; one letter stressed the benefits of tax-funded projects, while the other conveyed the message that most taxpayers were compliant. Based on a controlled field experiment in Switzerland, Torgler (2004) also found that moral suasion had hardly any effect on taxpayers' compliance behavior, nor did Fellner, Sausgruber, and Traxler's (2013) study of compliance with Austrian television and radio licensing fees. Pomeranz (2015) and Bergolo et al. (2017) found that a letter appealing to tax morale, but promising no increased enforcement, had little effect on VAT remittances. In Castro and Scartascini (2015), messages that emphasized reciprocity (taxes are used to pay for public services benefiting individuals) or peer effects (most citizens fulfill their tax obligations) did not have a significant effect. Bhargava and Manoli (2015) also find that a message about social stigma did not affect the take-up rate of the earned income tax credit.

In sum, a plethora of studies have failed to find evidence that appeals to tax morale, defined broadly, affect taxpayer behavior in the short run when delivered via a one-time mailing. Part of the reason could be the wording of these appeal-to-conscience letters; psychological research suggests the standard letter-based appeal-to-conscience wording such as "the entire community suffers" might affect behavior less than using terms wording such as "cheater," as in "Please don't be a cheater:" whether any government would be willing to employ such loaded terms remains to be seen. It is also important to keep in mind that, as Luttmer and Singhal (2014) argue, finding no effect is consistent with either tax morale not mattering, or these messages not effectively changing it. This point applies more generally to the broad class of interventions, as discussed more below.

The failure of such letters to affect compliance on the margin is not inconsistent with the existence of a substantial amount of “pathological honesty,” where taxpayers comply against their apparent self-interest. Two recent studies shed light on this phenomenon. LaLumia and Sallee (2013) examine panel data of tax returns before and after the United States required that dependent exemption claims be accompanied by a Social Security number, which resulted in a fall of between five and seven million dependent exemption claims. They focus on the vast majority of people who apparently did *not* claim a bogus exemption, and conclude that those that did not cheat were less likely to be heads of household and more likely to be married filing jointly; surprisingly, cheaters and non-cheaters faced similar benefits from falsely claiming a dependent. Dwenger et al. (2016) study motivations for tax compliance in the context of a legally binding, but unenforced, local church tax in Germany. Based on a randomized field experiment that introduces either positive deterrence or the provision of recognition and other non-pecuniary incentives, they find that about 20 percent of individuals remitted their true taxes tax liability in the baseline case with no deterrence in place. Recognition through social rewards for compliance caused some people to further increase their payments, but the provision of information on social norms or moral appeal had no impact.

Recently, a few studies have broken the previously solid set of field-experimental evidence finding no effect of such appeals. Bott et al. (2017) reports the results from a randomized field experiment in Norway conducted with more than 15,000 taxpayers who the tax authority deemed were likely to have misreported their foreign income. Shortly after sending the pre-populated tax returns for 2012, the tax administration in Norway mailed a letter to these taxpayers with information about how to report foreign income that randomly included two types of moral appeal—a fairness argument and a societal benefits argument. They find that including either

moral appeal in this letter almost doubled the average foreign income reported compared to a base letter without such an appeal, an effect similar in size to the effect of including a sentence that increased the perceived probability of detection by stating that the tax administration had received information that the taxpayer had income or assets abroad in previous years. The moral appeals mainly worked on the intensive margin, by increasing the amount reported of those who report any foreign income while the probability of detection mainly worked on the extensive margin, by increasing the share of tax subjects who report any foreign income. Neither moral appeal treatments had a statistically significant effect in the subsequent year, while the detection treatment still increased the share of taxpayers self-reporting some foreign income; the authors speculate that this might be due to the moral treatments affecting short-term salience but not causing a fundamental change in preferences, while the detection treatment caused a sustainable long-term updating of beliefs about detection probabilities.

Hallsworth et al. (2017) investigate whether letters that appeal to individuals' sense of social norms and public goods induce individuals to remit in a timely way taxes already declared. They run two large field experiments using administrative data from more than 200,000 individuals in the United Kingdom, and conclude that including social norms and public goods messages in standard tax payment reminder letters can considerably enhance tax compliance. The authors found wording that emphasized that the individual was in the minority of people who have not paid yet was the most effective in getting individuals to remit their taxes. They also find that mentioning financial penalties and remittance plans significantly increased the likelihood of compliance. Of note is that the outcome variable is the timing of liabilities that have been admitted, but not yet remitted, while most letter-based interventions look at the effect on initial reporting behavior. It might be that a taxpayer who is simply procrastinating on paying their taxes is more

likely to be persuaded by social norms to accelerate remittance than a firm or individual who is contemplating evasion. Finally, there are important differences in how the treatment might have been perceived by the recipients. In Hallsworth et al. (2017), the letter informs the taxpayer that the UK tax authority is aware of their delinquency. It says, “Nine out of ten people in the U.K. remit their tax on time. You are currently in the very small minority of people who have not paid us yet.” In contrast, the letter in the Blumenthal, Christian, and Slemrod (2001) Minnesota experiment says, “people who file tax returns report correctly and pay voluntarily 93 percent of income taxes they owe [...]; a small number of tax payers who deliberately cheat owe the bulk of unpaid taxes.” In this case, unlike in the Hallsworth et al. experiment, the letter does not convey to the taxpayer that the tax authority is aware of any wrong-doing by the individual. Thus the difference in results could be due to the distinction between informing an individual that the government has evidence pertaining to their actual evasion and appealing to their sense of duty without revealing any information it has regarding their avoidance behavior.

Del Carpio (2014) examines the role of norms and enforcement perceptions on tax compliance through a field experiment on property taxes in Peru. Randomly chosen subsets of residents in two municipalities in the Lima province were informed, through an official letter from the municipality, about the average rate of compliance, the average level of municipal enforcement, or both, while a third group was only reminded of the payment deadline. Analysis of the administrative data reveals that disclosing information on the level of compliance had a large positive impact on compliance (20 percent relative to the control group), while the payment reminder also raised compliance by 10 percent. Notably, the enforcement treatment did not have a significant effect on compliance net of the reminder effect, corroborating other evidence that *any*

contact from the tax authority to the taxpayer increases compliance, and additional treatments may or may not.

This set of results has somewhat moved my pre-2013 assessment that the evidence overwhelmingly supported that deterrence inhibits noncompliance but that manipulation of norms has no measurable effect. In some settings, norm-directed letter interventions do seem to matter. It now behooves us to understand better why this works in some, but not most, settings.

In the standard deterrence model, one's political affiliation should not affect evasion, conditional on one's perceptions of p , f , and one's risk aversion. And yet Cullen, Turner, and Washington (2018) find evidence in a quasi-experimental setting that political alignment with the party of the presidential administration in the United States has a positive impact on compliance. They analyze data from 1999, 2001, 2007 and 2009 – years surrounding the turnover elections of 2000 and 2008 to study how counties that consistently vote either Democrat or Republican change their evasion behavior following a change in the party of the President. They measure evasion at the county level using information from the IRS on reported income by category of income, including some types of income subject to third-party reporting (such as wages) and others such as business income that are not. The authors examine changes in reported income at the county level by source of income as a function of “political alignment” of the county, controlling for predictors of county-level income, where political alignment is measured as the average vote share in the county for the party of the sitting president over the past several elections. They find that moving into alignment results in a 0.4 percent increase in reported AGI (i.e., evasion declines), coming mostly from a 3.5 percent increase in business income. The authors interpret these results as consistent with taxpayers' approval of government activity affecting evasion, which is clearly inconsistent with the simple deterrence model that assumes free-riding behavior. As the authors

note, they cannot rule out or investigate the possibility that some taxpayers perceive that the probability of or cost of an audit also changes with alignment, although a simple story would suggest that it varies inversely, rendering their estimates lower bound on the impact of improved tax morale.

One recent paper, Perez-Truglia and Troiano (2015), compares the effect on tax debt repayment behavior of shaming versus financial penalties in a field experiment in three of the twenty-three U.S. states that publish online the names and amount of some tax debtors. The authors sent letters to delinquents that were worded to emphasize either shaming or the penalties. While all delinquents are informed by the states that their names will appear in the online list of delinquents, in the shaming treatment some individuals were randomly chosen to receive a letter giving details about how to access the online delinquent list along with a list of 10 delinquents in the neighborhood, including themselves. Some of these individuals are additionally told that their neighbors will receive a similar letter, to increase the salience of shaming. The authors investigate whether people will be more likely to pay off their debt if their perceived shaming adversely impacts their social capital, and if shaming might shift their motivation to pay from an intrinsic to an extrinsic one and therefore decrease the likelihood that they will remit. The authors find that both shaming and financial penalties increase the likelihood of payment within ten weeks of receiving the letter. Shaming mattered most for those with small amounts of debt, increasing the likelihood of payment by 2.1 percentage points, suggesting that there is a limit to the value of preventing social stigma. Because the tax authorities in these states generally give individuals an opportunity to clear their debts before publishing their names online, one may consider this effect as a lower bound of the total effect on tax debt payments.

9. UNDERSTUDIED EMPIRICAL ISSUES

Clearly during the last 15 years economists have produced much exciting empirical research about tax compliance and enforcement, using a variety of credible research designs. As is natural, there is a flavor of searching for one's lost keys at night under the few working lampposts. For example, we have learned much more about the short-term compliance effect of various letter interventions than the role of penalties in tax enforcement because tax authorities have been willing to undertake them, in part because they are inexpensive and non-disruptive. On the plus side, for the same reason the research can have a real and fairly immediate effect on policy. It is also worth noting that the lamppost technology is improving rapidly, so that as time passes we will be able to search in previously dark areas. In the spirit of providing some guidance for that search, in what follows I discuss a few topics that I think deserve some more attention.

9.1. The Role of Tax Professionals

One understudied issue is the role of professional tax preparers in tax administration and enforcement. Their role is potentially important given their ubiquity. In the United States, 63 percent of individuals and 97 percent of corporations use some professional assistance. An earlier literature, notably Klepper, Mazur, and Nagin (1991), using summary tabulations of line-item tax return data from the IRS's 1982 randomized audit study, found evidence consistent with preparers discouraging noncompliance on legally unambiguous income sources, but encouraging noncompliance on ambiguous sources. Countries vary substantially in how the law treats professional preparers, from no official contact to significant regulation. In the United States, as part of the Professional Preparer Initiative, about 750,000 tax preparers registered with the IRS by 2011, but no evaluation has yet been made of registration's impact on tax compliance.

For the most part, modern empirical methods have not been brought to bear on the impact of preparers and how they might be regulated. One noteworthy exception is Zwick (2018), which examines the role of paid preparers in the take-up of a tax refund for corporate losses, in part to explain why only 37 percent of eligible firms claim their refund. He finds that, relative to preparers without a professional license, certified public accountants are 6.8 percentage points more likely to claim the carryback refund for their clients. The data do not support the possibility that firm selection—savvier firms hire savvier accountants—explains the observed preparer effect with a research design based on firms switching their preparer following the death or relocation of their previous one (arguably an exogenous event). However, they do not apply this method to compliance behavior.

9.2 Networks

The role of networks in tax evasion and enforcement is no longer unstudied, but given its importance economists have just begun to scratch the surface. These networks might involve families, as in Alstadsæter, Kopczuk, and Telle (2014)), who use detailed administrative data from Norway to identify family networks and describe how take-up of tax avoidance progresses within a network. As discussed above, networks might also involve tax preparers. They might involve the Internet. Hoopes et al. (2015) examine data on capital-gains-tax-related information search—on Google, Wikipedia, and IRS information platforms—to determine when and how taxpayers acquire information, and find increases in information search around tax deadlines, suggesting that taxpayers seek information to help them comply with the tax law. Positive correlations between stock market activity and search as well as year-end spikes in information search on capital losses

when the market performs poorly indicate that taxpayers seek information for tax-planning purposes.⁴⁵

Drago, Mengel, and Traxler (2015) and Rincke and Traxler (2011) study the spread of compliance behavior in neighborhood networks in the context of the Austrian TV license fee regime, discussed earlier. The first study involves a randomized field experiment run that varied the content of mailings sent to potential evaders. The authors first provide survey evidence showing that the communication intensity of neighbors in rural areas is strongly correlated with spatial distance, and then document that households who were not part of the experimental sample (and were therefore untreated) were more likely to switch from evasion to compliance in response to the mailings received by their neighbors in the same network. Rincke and Traxler, using snowfall as an instrument for local inspections, also find that compliance rises significantly among those who had no exposure to field inspections. In another Austrian setting, Paetzold and Winner (2016) investigate the effect of one's work environment on the improper claiming of commuter tax allowances in Austria, and find an asymmetric effect: once individuals learn from co-workers that over-reporting of the allowances goes undetected, they are more likely to start cheating, but being exposed to an environment of compliance does not reduce previous cheating behavior.

The Boning et al. (2018) RCT, mentioned briefly earlier, investigates how enforcement efforts might have spillover deterrence effects on firms within a network. They study three kinds of networks: firms in close geographic proximity (5- or 9-digit ZIP code endings), firms that share a tax preparation company or the same tax preparer, and firms that have a common ownership link. In the experiment, the IRS sent letters and made in-person visits to businesses whose individual

⁴⁵ Another aspect of networks is the type of chains that form in value-added tax regimes, wherein firms out of the tax net tend to sell to other firms outside of the tax net (and consumers), while firms in the tax net tend to sell to other firms in the tax net, because only purchases from tax-compliant firms can be credited against revenues by firms. See De Paula and Scheinkman (2010).

income tax withholding and payroll taxes had declined. They find that firms that share the same tax preparer with a firm that received an in-person visit show an increase in tax remittance relative to firms in tax preparer networks that did not receive a visit. However, they do not find a similar network effect among businesses that use the same tax preparation company as a treated firm, or among businesses in close geographic proximity to a treated firm.⁴⁶ Strikingly, subsidiaries of firms that get a Revenue Officer visit report *less* tax subsequently, due either to a cash-flow effect or a substitution of evasion toward a now perceived-to-be less likely target audit; this effect is, though not symmetric: the parent companies of firms that get a visit do not change their remittance pattern.

Understanding more about these networks has the potential to address an important gap in our understanding of the impact of tax enforcement initiatives. Nearly all of the research discussed in this survey attempts to estimate the “specific” or “direct” deterrent effect that operates solely on those subject to enforcement action. But these studies do not take account of the “general” deterrent effect in the population as a whole that operates on all taxpayers through the effect of a policy action on the perceived probability of detection and punishment, which is after all the effect that Allingham and Sandmo (1972) stress. General deterrence is arguably the most important channel through which tax enforcement initiatives work, as in principle it applies to the whole population, but it is difficult to measure because it works via the perceptions of potentially noncompliant (i.e., nearly all) taxpayers. Although the general deterrent effect is much more difficult to measure than the specific deterrent effect, networks may shed light on how particular enforcement actions

⁴⁶ Nor did Meiselman (forthcoming) find evidence of geographic spillovers in his RCT regarding nonfilers with the Detroit city income tax.

diffuse into the population and affect perceptions broadly.⁴⁷ Taxpayers connected in some way to those directly subject to enforcement action may learn of the enforcement action, which changes their perception of the enforcement environment, which may in turn spread to their connections. Thus, the network deterrent effect captures the word-of-mouth spread of information about a change in enforcement policy, a plausible mechanism that could result in general deterrence.

The presence of spillover effects of course has a potential downside, that it might pollute what otherwise would be a control group. On this issue, there is voluminous work in the development literature about how to design RCTs to limit the bias.

9.3 The Role of Firms

Information reporting and the remittance regime are crucial to tax enforcement, and firms are the linchpin of both of these functions. Recent work by OECD (2017) and Slemrod and Velayudhan (2018) reveals that, in both developed and in at least one developing country (India), firms remit on average about 85 percent of all taxes. The welfare implications of this issue are fascinating, because it implies that in the presence of taxes the equilibrium pattern of firm borders and size is not optimal, contrary to the suggestion of Coase (1937). Taxes can be collected with less cost when the tax authority can make use of information generated (and reported) by arms-length transactions between firms, and between firms and employees. Sole proprietorships and small businesses, especially family firms,⁴⁸ are difficult for the tax authority to penetrate,

⁴⁷ Sah (1991) formalizes the endogenous determination of the perception of p based on the available information, including from acquaintances, and the true probability; it addresses crime in general, and not tax evasion in particular.

⁴⁸ Kopczuk and Slemrod (2010) provide a sketch of how to model the taxation of family firms, stressing that in some developing countries the weakness of legal institutions encourages the formation of family firms, whose family bonds informally enforce against theft; these bonds have a social cost because they increase the opacity of firms, making tax enforcement more difficult.

providing an example of when production efficiency may not be desirable when taxes must be raised, contrary to the classic result of Diamond and Mirrlees (1971).

We should take a closer look at the relationship of tax compliance and self-employment. When we say that modern tax systems rely heavily on firms, we are referring to medium-sized and large firms, because small firms and the self-employed are ubiquitously problematic. Theory suggests why this might be so, but in an over-identified way (i.e., there are too many theories). Third-party information reporting for non-employee income is not easily done. Self-employed enterprises are by definition small, and the agency argument formalized by Kopczuk and Slemrod (2006) and Kleven, Kreiner, and Saez (2016) suggests that for this reason evasion is more sustainable. Self-employed people choose that status, and may be less risk-averse to all forms of uncertainty, including potentially costly detection of evasion. Future empirical analysis should aim at sorting out these issues, perhaps by leveraging the fact that some people, and some families, have both employee income and self-employment income.

A mountain of micro evidence, using multiple methodologies, documents a strong association between self-employment and noncompliance and between self-employment and the “flexibility” of reported taxable income locally to kinks and notches in tax schedules. Kleven (2014, p. 82) plots for over 80 countries the fraction of workers who are self-employed against the tax/GDP ratio, and documents a strong negative relationship: countries that have more self-employed people collect less tax. Although he rightly cautions that no causal inferences can be drawn from such a graph, I agree with his conclusion that the availability from employers of third-party information on employee income plays a key role in tax compliance and in explaining a country’s overall tax take. Consistent with this conclusion, Jensen (2016) shows that, as countries develop, their employment structure shifts from self-employment to employees and exemption thresholds for

income tax liability fall, a pattern that is consistent with tax authorities setting the threshold at a level that justifies enforcement costs.

We also need to focus more on the compliance by firm-withholders which, with just a few exceptions such as Boning et al. (2018), has nearly been ignored and often implicitly assumed to be perfect. Of interest is the fact that, in the United States, withheld income and employment taxes such as Social Security taxes are called trust fund taxes as recognition that legally the firm holds the employee's money in trust until it makes a federal tax deposit. Noncompliance can trigger a “trust fund recovery penalty” that pierces the corporate veil, and can be levied on any person who has the duty to perform and the power to direct the collecting, accounting, and paying of trust fund taxes, including but not restricted to officers or an employee of a corporation as well as a corporate director or shareholder. Whether this qualitatively different penalty feature has a qualitatively different deterrent effect is not known.

Finally, firms and workers may in some situations collude to facilitate evasion, as was explored theoretically by Yaniv (1988). Best (2014) finds that firms in Pakistan aggregate the preferences of workers and facilitate tax avoidance by bunching their salary offers around kinks in the tax schedule. If and under what circumstances they facilitate evasion is worth exploring. One setting in which this is suspected is firms’ reclassifying workers as independent contractors rather than employees. As discussed in Section 7, this eliminates their role as withholder/remitters, and may induce more evasion, which in turn could reduce the cost of labor to the firm.

9.4 The Distributional Impact of Evasion and Enforcement

Paying attention to the distributional implications of compliance and enforcement policies is another logical extension of the recent wave of empirical analysis. Johns and Slemrod (2010) assess the distributional consequences of income tax noncompliance in the US federal income tax

for the tax year 2001 using data from the National Research Program.⁴⁹ They find that, when taxpayers are arrayed by their estimated “true” income, defined as reported income adjusted for the underreporting estimated by the IRS, the ratio of aggregate misreported income to true income generally increases with income, although it peaks among taxpayers with adjusted gross income in the 99.0 to 99.5 percentile. In sharp contrast, the ratio of underreported tax to true tax is highest for the lowest-income taxpayers, reflecting the fact that a given percentage reduction in taxable income corresponds to a particularly high percentage reduction in tax liability for taxpayers with taxable income just above the taxpaying threshold. Much of the distributional pattern of noncompliance is associated with the fact that, on average, high-income taxpayers receive their income from sources, such as business income, that have higher noncompliance rates. But this is not the whole story because similar, although not identical, patterns apply to the misreporting percentages of given income sources.

Nygård, Slemrod, and Thoresen (forthcoming) examine the distributional implications of off-the-tax-books transactions between service providers and consumers in Norway, wherein the supplier reduces her tax burden by underreporting income and the consumer may gain from buying an untaxed and therefore perhaps lower-priced product. The distributional implications of such joint tax evasion depend on the amounts evaded, on where the evaders on both sides of the market are found in the income distribution, and how the financial gain is split between the suppliers and demanders. The authors use multiple data sources to identify tax evasion among sellers and buyers of goods and services, and conclude that the tax-evasion-adjusted estimate of disposable income inequality in Norway exhibits more dispersion than official estimates.

⁴⁹ Matsaganis and Flevotomou (2010) address this question for pre-crisis Greece by comparing a sample of income tax returns to data from a household budget survey, concluding that tax evasion produces higher income inequality and lower effective tax progressivity.

An old saw goes “the poor evade, and the rich avoid,” suggesting that there are enough legal ways for sophisticated, wealthy taxpayers to reduce their tax obligations that they need not resort to illegal evasion. Both the recent direction of policy focused on evasive foreign accounts, the U.S. components of which are summarized and analyzed in Johannesen et al. (2017), and some new empirical research, suggest that the old saw is wrong. Zucman (2014), relying on anomalies in global investment statistics caused by offshore fortunes (i.e., more liabilities than assets show up in global investment data), estimates that U.S. residents hold about \$1.2 trillion of wealth offshore, equal to about 4 percent of their financial wealth, resulting in an annual revenue loss of \$36 billion. Evasion amongst the very top of the income distribution is difficult to uncover through traditional means like random audits, as the auditor typically lacks the resources to trace the sophisticated means of evasion often involving layers of financial intermediaries. However, high-profile leaks from these intermediaries, such as the 2007 leak from HSBC Private Bank in Switzerland and the 2015 “Panama Papers” from the firm Mossack Fonseca, have recently allowed researchers to gain some insight into tax evasion by the richest. Alstadsæter, Johannesen and Zucman (2017) use data from these leaks along with administrative data on income and wealth from Norway, Sweden and Denmark to show that evasion rates rise across the income distribution, and conclude that the top 0.01 percent evade about 30 percent of the income and wealth taxes they owe. The authors link the account names from the HSBC leak with individual tax data, and find that 95 percent of these foreign account holders did not report the existence of the account on their tax forms, which they classify as evasion.

9.5 Penalties

The severity and nature of the punishment for detected evasion is the neglected sibling among the two central policy parameters of the Allingham-Sandmo deterrence model; little

empirical research has been devoted to this topic, even in the new wave of research covered in this article. The relative neglect arises in part because governments have not so far proven willing to experimentally vary the extent of punishment, although some have participated in RCTs that randomly “remind” taxpayers about existing penalties. One exception is Bergolo et al. (2017), which in an RCT informs firms about the average penalty imposed for detected evasion among a sample of similar firms. The effect of providing information about penalties cannot be discerned, however, because the relevant treatment informs VAT taxpayers about both the average penalty *and* the average chance of being audited.

The nature of punishment has received almost no attention. For example, although the literature on other types of crime has addressed the impact of the certainty or celerity (i.e., swiftness) of punishment, the tax evasion literature has not focused on it. Recently, Blank (2014a) and Kuchumova (2015, 2017) have discussed “collateral tax sanctions” such as revoking from individual tax evaders drivers’ licenses, professional licenses, and passports (as the IRS is now legally able to do) and from business evaders the ability to obtain government contracts. Blank argues that such sanctions may be especially effective if, for example, they are more salient or create greater economic costs than monetary penalties, but no empirical evidence yet exists to evaluate their impact.

The flip side of penalties for evasion is rewards for compliance. Dunning et al. (2016) report that as many as 25 percent of municipalities in Brazil, Colombia, Ecuador, and Uruguay, and many localities in Argentina, Peru, and Mexico offer reward programs such as prize lotteries to compliant taxpayers.⁵⁰ Since 2004, Montevideo has raffled a year free of tax payments to

⁵⁰ Of possible interest to some readers note is the fact that in 2010 Stockholm experimented with a “speed camera lottery” such that if your vehicle comes in at or below the speed limit, the driver was automatically entered into a lottery; see Hagerty (2010).

eligible “good” taxpayers who have been punctual compliers over the previous year. The authors use the random awarding of these tax holidays to investigate the impact of habit on tax remittance, and find that winning the lottery reduces the winners’ future tax payments for up to three years.

Not enough has been done to learn about the effect of this type of reward program on compliance, including whether the effect is symmetric to the impact of penalties for noncompliance. One notable exception is Carrillo, Castro, and Scartiscini (2017), who study the impact of a randomized natural experiment in an Argentine municipality, where 400 compliers with property tax were provided with a new sidewalk in front of their property and public recognition. They found that being selected as a lottery winner had a positive and persistent direct effect on future compliance as well as a geographic spillover effect on neighbors. However, they find no evidence that taxpayers with pending liabilities paid them off in order to become eligible for the lottery; thus, there is evidence of a specific, or direct, effect, but no evidence of a general reward (i.e., the opposite of a deterrent) effect. Ahmed et al. (2012) describe a first step in learning about the compliance impacts of a program of social recognition of tax compliance in Bangladesh.

9.6 Policies Aimed at Evasion-Facilitating Behavior

Countries use a wide range of policies to inhibit actions that facilitate tax evasion, but few have been evaluated rigorously. Many such policies are aimed at the use of cash. For example, some governments have introduced a ceiling for cash transactions: DKK 10,000 in Denmark, €1,000 in France and Italy, \$10,000 in Australia as of July 1, 2019, €5,000 in Belgium, €1,500 in Greece and, as of November, 2016, India banned its highest-value 500 and 1,000 rupee notes.⁵¹ Others have required point-of-sale terminals in, for example, taxis. Although it has been suggested that

⁵¹ See Williams (2014, p. 103).

(evasion-facilitating) cash be directly taxed,⁵² this has not been widely implemented. There was a tax on cash withdrawals in India from 2005 to 2009, designed primarily as an audit trigger.⁵³ Alternatively, one could consider providing incentives to use credit cards, as countries such as Argentina and South Korea do.⁵⁴ From this perspective, the U.S Form 1099-K initiative discussed earlier is counterproductive in that it cracks down on underreporting of credit-card sales while leaving cash sales untouched. In many European countries, certified cash registers are required, in part to counter the use of zappers, software installed on electronic cash registers or other electronic point of sale systems that allows users to erase recorded transactions.⁵⁵ All of these policies are costly, and future policy would be informed if their benefits could be credibly quantified.

9.7 External validity of tax enforcement RCTs

I will confess that about fifteen years ago, when it became clear that RCTs would play a leading role in the credibility revolution in empirical economics, I despaired that tax analysis would be left behind because no government would randomize tax rates levied on their citizens. That remains true, but I have been thrilled that many governments have proven willing to randomly vary the enforcement of their tax systems. Although there are now scores of such RCTs, this field is in its infancy compared to the use of RCTs in development economics, where there are several hundred.

But it is time for the infant inquiry to address the challenges of a mature inquiry. Although partner governments are usually quite happy to learn only about whether a given policy would achieve a valued goal in their own jurisdiction, as scholars it is important to design studies that

⁵² See Benshalom (2012). Macroeconomists are also interested in this notion as a way to facilitate a negative interest rate; see, for example, the discussion in Rogoff (2015) regarding the costs and benefits of phasing out paper currency.

⁵³ See Tax Administration Reform Commission (2014).

⁵⁴ Note that the United States has recently gone in the opposite direction, as witnessed by the 2017 Supreme Court ruling against New York's state anti-card-surcharge law.

⁵⁵ These policies are discussed in greater detail in Williams (2014, pp. 101-103). Zappers and appropriate policy responses are discussed by Ainsworth (2010).

will shed light on tax enforcement more generally, i.e., that have external as well as internal validity. That requires paying attention to the mechanisms that underlie observed effects, so we can learn why—ideally with a formal theory—some policies work better in some settings versus others. The follow-up survey methodology used by Bergolo et al. (2017) is promising as a way to learn about how a given policy intervention affected perceptions of the fundamental parameters of the Allingham-Sandmo model. We also know almost nothing about the special problems that might arise when scaling up tax enforcement from pilot to population policies.

Some of the threats to external validity that are concerning in other settings are probably not as important for tax enforcement. For example, the general equilibrium effects seem less serious, although some enforcement compliance might affect the tax professional industry. Randomization bias, the issue that the governments that facilitate research are likely to be special is a concern. A citizenry that tolerates such government experimentation and monitoring is undoubtedly different from other citizenries in ways that are relevant to the questions at hand: the magnitude and nature of noncompliance, the norms that matter, and the institutional environment. But this issue loomed larger at the beginning of this new wave when most of the research occurred in developed countries, and particularly in the Nordic countries. I am happy to report that this is no longer true, as much recent research is happening in developing countries, where the issues of compliance and enforcement are especially critical. But, governments in both developed and developing countries probably share the incentive to demonstrate the wisdom of existing policies and of carrying out pilots of contemplated new policies.

9.8 Optimal enforcement policy

Finally, it is time to better connect the new empirical literature on tax compliance and its more credible estimates of the effect of enforcement policies to the bread-and-butter normative issues

of taxation: efficiency and equity. Regarding efficiency, the focus on evasion may seem puzzling to those who are steeped in the idea of the elasticity of taxable income (or more generally, the elasticity of tax base), which holds that under some conditions this elasticity is a sufficient statistic for the marginal welfare cost of changing tax rates, and therefore understanding the anatomy of the behavioral response (e.g., labor supply versus evasion) is irrelevant. How evasion fits into this framework has been the subject of some controversy,⁵⁶ but in any event knowing how evasion contributes to the behavioral response helps focus policy discussions; in the extreme, if we were to discover that there is no evasion under any circumstances, pondering optimal enforcement would be a waste of time. Regarding equity, studies of the effect of taxation on income inequality must address the apparent disproportionate presence of certain kinds of tax evasion.

Integrating compliance and enforcement into optimal tax requires attention to the effect of evasion and enforcement on real decisions such as labor supply and business formation. Sometimes in policy debates this is ignored, when supporters imply that cracking down on evasion can raise revenue while avoiding the real behavioral responses we associate with raising tax rates. But this argument is logically flawed. Increased enforcement of, say, income taxes raises the expected tax rate (only for prospective evaders), and will trigger similar real responses as an explicit tax rate increase. Certainly many of the empirical papers discussed here investigate both real and compliance responses, but they generally do not focus on this interaction. This is a particular challenge in the case of labor supply, because most of the administrative data sets naturally contain information on reported taxable income, but do not match it with data on labor supply, although in some cases data on job flows exist. In the same vein, we need to know more about the substitutability between evasion and (legal) avoidance: if an enforcement policy cracks

⁵⁶ Compare Chetty (2009) with Gillitzer and Slemrod (2016).

down on evasion, to what extent will people respond by increasing (untaxed or lightly taxed) avoidance behavior?

Increased enforcement is just one way to raise revenue, with the obvious alternatives being to raise tax rates or broaden the tax base. In formulating optimal policy one needs to consider the marginal social costs of enforcement relative to the costs of alternative ways to raise revenue. Thus, the overall objective of this aspect of tax policy is not different than the objective of choosing tax rates, bases, and other elements of a tax system. The costs of increased enforcement include administrative costs (that show up in the IRS budget), compliance costs (that don't show up in the IRS budget), excess burden (due to behavioral responses of all kinds), and the extra uncertainty to taxpayers that the "tax audit lottery" creates.

Consideration of the social costs of tax evasion highlights the difference between the "recoverable" portion of the tax gap and the "economically recoverable" portion, borrowing language usually applied to oil reserves. The optimal tax gap is not zero any more than it is socially optimal to extract every last drop of oil beneath the ground, or to put a police officer at every corner to eliminate all street crime. For this reason ascertaining the size of the tax gap is not as helpful for policy design as is the susceptibility of the gap to alternative tax-system policies.

Two separate policy issues arise: (1) how big should the IRS enforcement budget be, and (2) how best to allocate a given budget. Slemrod and Yitzhaki (1987) show that one superficially intuitive rule for the former—increase the probability of detection until the marginal increase of revenue thus generated equals the marginal resource cost of so doing—is incorrect. Although the cost of hiring more auditors, buying better computers, and the like, is a true resource cost, the revenue brought in does not represent a net gain to the economy, but rather a transfer from private (noncompliant) citizens to the government. The correct rule equates the marginal social benefit of

reduced evasion (which is not well measured by the increased revenue) to the marginal resource cost. The social benefit includes the reduced risk-bearing that comes with reduced tax evasion, and any reduction in the inefficiencies discussed earlier. Cowell (1990, p. 136) suggests another complication: perhaps a specific social welfare discount should apply to the utility of those who are found to be guilty of tax evasion and thus “are known to be antisocial.”

For the allocation question, a useful rule of thumb is that all tax policies should equalize the marginal efficiency cost of funds, a simple expression that accounts for all the costs of raising revenue, which should in turn equal the marginal social benefit of raising revenue (Mayshar, 1991; Slemrod and Yitzhaki, 1996, 2002; Slemrod and Gillitzer 2014). Distributional considerations can be introduced into this framework in a straightforward way.

A framework for integrating empirically estimated parameters regarding tax enforcement into welfare analysis is provided by Keen and Slemrod (2017). In the model, the planner sets enforcement parameters and an income tax rate on reported income to maximize social welfare, equivalent to the welfare of a representative individual. In the simplest setting, enforcement is a single, continuous parameter, but the results extend to settings with multiple enforcement instruments and to enforcement actions that are inherently discrete, such as the existence of a large taxpayers’ unit. Given the tax rate and level of enforcement, individuals choose their labor supply and how much income to conceal subject to convex concealment costs. The model characterizes the answer to three core policy questions: (1) What information does the policy maker require about cost and effectiveness of administrative interventions to set optimal policy? (2) What is the optimal setting of enforcement instruments, for any given tax rate? (3) What is the optimal

compliance gap? Analogous to the now-standard elasticity of taxable income (with respect to the income tax rate)⁵⁷, it develops sufficient statistics for optimal administrative intervention.

At the optimal level of enforcement, $\phi = E(z, \alpha)$, where $E(z, \alpha)$ is the enforcement elasticity of tax revenue with respect to the administrative instrument α (the percentage change in tax revenue with respect to a percentage increase in enforcement resources), and ϕ is the adjusted marginal cost-revenue ratio, where the costs include both compliance costs and administrative costs, with a larger weight on the latter because these costs must first be raised by distorting taxes. This result suggests that tax authorities should identify the enforcement elasticity, which can be thought of as the goal of many of the empirical analyses discussed in this paper, and the marginal cost and revenue associated with a potential enforcement action rather than the average or total costs, even though the latter are easier to estimate and widely used as benchmarks.⁵⁸

The tax gap has traditionally been used a performance measure for tax administration, with little theoretical rationale. Keen and Slemrod (2017) show that the tax gap is not a sufficient statistic for the optimal level of enforcement, although there is a close relationship between the compliance gap and the enforcement elasticity of tax revenue. At an optimum, the compliance gap is characterized by $\frac{G}{1-G} = \frac{-\phi}{E(e, \alpha)}$, where G is the compliance gap expressed as a ratio of true liability and $E(e, \alpha)$ is the enforcement elasticity of evasion. The optimal compliance gap is thus characterized by a simple inverse elasticity rule, with the relevant elasticity being the enforcement

⁵⁷ See Saez, Slemrod, and Giertz (2012).

⁵⁸ The Detroit RCT in Meiselman (2018), discussed earlier, combines all the ingredients needed to apply the approach set out above. Employing a discrete version of this framework, it concludes that there was a negative direct welfare effect of the average experimental letter, driven by the large proportion of taxpayers for whom estimated compliance costs exceeded the incremental net tax remittance.

elasticity of evasion and the factor of proportionality being the adjusted marginal cost-revenue ratio.⁵⁹ The optimal compliance gap is inversely proportional to the enforcement elasticity of tax evasion. The higher the enforcement elasticity, the lower the optimal compliance gap. However, the behavioral impact of enforcement also depends on the adjusted marginal cost-revenue ratio. Expending effort to reduce the compliance gap is warranted if the gap exceeds the ratio of adjusted marginal cost-revenue ratio to enforcement elasticity of evasion. This condition shows how the kind of tax gap estimates discussed in Section 3 can be combined with information on cost-revenue ratios and behavioral responses to provide actionable advice: increasing enforcement to reduce the compliance gap is warranted if and only if the gap measure on the left-hand side of this expression exceeds the inverse elasticity on the right-hand side.

10. CLOSING THOUGHTS

The explosion of empirical research into tax compliance and enforcement is welcome news, as it was high time that the modern arsenal of empirical methods be applied to the second and third pillars of a tax system, remittance and enforcement, to complement the long-standing and continuing analysis of tax rates and bases.

What do we know? That raising the chance of getting caught high enough deters evasion is incontrovertible, as evidenced for example by the difference between the U.S. noncompliance rate of 63 percent for income not covered by third-party information reports (or withholding) compared to 1 percent for wages and salaries, covered by both. This evidence in favor of the Allingham-Sandmo deterrence model of evasion is as powerful as comes from the stark natural experiments concerning other crimes discussed in Nagin (2013) and Durlauf and Nagin (2011), such as the

⁵⁹ This expression holds under the simplifying assumption that enforcement does not affect labor supply, although it can be relaxed.

increase in crime in Nazi-occupied Denmark after the German officials dissolved the entire Danish police force (Andenaes 1974).

Finding compliance responses to stated changes in the probability of audit in an RCT has proven more elusive, with mostly positive but mixed results, perhaps because a given change in the chance of audit corresponds to vastly different changes in the chance of detection and because the prior perceptions may vary vastly depending on the setting. Little is known about the effect of changing penalties, mostly because of the difficulty of designing an RCT to measure this effect.

The leading alternative, but not mutually exclusive, paradigm to the Becker-Allingham-Sandmo model is one that relies on duty, conscience, and adherence to norms. A large number of RCTs have investigated whether appeals to these factors affects compliance. In most cases the answer is no, but in a few recent cases such appeals seem to have worked to reduce evasion or accelerate remittance of tax debt. We need to know more why this (only) occasionally seems to work.

What emerges more clearly is that contacts from the tax authority to the taxpayer can increase compliance in the short run. Sometimes even an anodyne, placebo contact has this effect, perhaps because it signals to the taxpayer that she is on the tax authority's "radar." In-person delivery of a given message is much more powerful, but it is also much more expensive, so the welfare superiority of that mode of delivery is not clear. Providing concrete evidence that the government has information indicating the presence of evasion consistently works to reduce evasion, apparently because it combines a you're-on-our-radar message with actionable information about noncompliance. As this literature moves from infancy to maturity, it should focus on clarifying the mechanism underlying effectiveness, so as to establish the external validity of claims about policy impact.

Knowing the effect on tax noncompliance of p (and f) is by no means sufficient information for guiding policy decisions. We also need to know what policies deliver effective deterrence, by obtaining verifiable information through third-party information reporting and information trails more broadly, and by optimizing the remittance regime to get the money from the low-cost remitters. We need to know much more about the social costs of such policies, including the administrative and compliance costs as well as intangible costs such as intrusions on privacy. This suggests that, as the empirical project matures, researchers should hew to the venerable tradition in public economics of drawing out the welfare implications of the empirical findings, for which a sufficient-statistics-based framework is now available. By so doing, this literature can inform the ongoing debates around the world about the design of tax-system policies: rates, bases, remittance rules, and the broad panoply of enforcement instruments. This is especially true because of the interaction between imperfect compliance and tax policy design, and specifically how compliance problems can justify a fundamentally different set of policy instruments than recommended by traditional public finance models.

References

- Advani, Arun, William Elming, and Jonathan Shaw. 2017. "The Dynamic Effects of Tax Audits?" IFS Working Paper W17/24.
- Ahmed, Nasiruddin, Raj Chetty, Mushfiq Mobarak, Aminur Rahman, and Monica Singhal. 2012. "Improving Tax Compliance in Developing Economies: Evidence from Bangladesh." Working paper.
- Ainsworth, Richard. 2010. "Zappers – Retail VAT Fraud." *International VAT Monitor*: 175-182.
- Allingham, Michael, and Agnar Sandmo. 1972. "Income Tax Evasion: A Theoretical Analysis." *Journal of Public Economics* 1 (3-4): 323-338.
- Alm, James, and Sarah Jacobson. 2007. "Using Laboratory Experiments in Public Economics." *National Tax Journal* 60 (1): 129-152.
- Almunia, Miguel, and David Lopez-Rodriguez. 2018. "Under the Radar: The Effects of Monitoring Firms on Tax Compliance." *American Economic Journal: Economic Policy* 10 (1): 1-38.
- Alstadsæter, Annette, Wojciech Kopczuk, and Kjetil Telle. 2014. "Social Networks and Tax Avoidance: Evidence from a Well-Defined Norwegian Tax Shelter." Working Paper, University of Oslo.
- Alstadsæter, Annette, Niels Johannesen, and Gabriel Zucman. 2017. "Tax Evasion and Inequality." NBER Working Paper.
- Andenaes, Johannes. 1974. *Punishment and Deterrence*. Ann Arbor, MI: University of Michigan Press.
- Andreoni, James, Brian Erard, and Jonathan Feinstein. 1998. "Tax Compliance." *Journal of Economic Literature* 36 (2): 818-860.

- Angrist, Joshua, and Jorn-Steffen Pischke. 2010. "The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics." *Journal of Economic Perspectives* 24 (2): 3-30.
- Anwar, Shamena, and Thomas A. Loughran. 2011. "Testing a Bayesian Learning Theory of Deterrence among Serious Juvenile Offenders." *Criminology* 49 (3): 667-698.
- Artavanis, Nikolaos, Adair Morse, and Margarita Tsoutsoura. 2016. "Measuring Income Tax Evasion using Bank Credit: Evidence from Greece." *Quarterly Journal of Economics* 131 (2): 739-798.
- Asatryan, Zareh, and Andreas Peichl. 2017. "Responses of Firms to Tax, Administrative and Accounting Rules: Evidence from Armenia." CESifo Working Paper Series No. 6754.
- Banerjee, Abhijit, Rukmini Banerji, James Berry, Esther Duflo, Harini Kannan, Shobhini Mukerji, Marc Shotland, and Michael Walton. 2017. "From Proof of Concept to Scalable Policies: Challenges and Solutions, With an Application." *Journal of Economic Perspectives* 31 (4): 73-102.
- Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76 (2): 169-217.
- Benjamini, Yael, and Shlomo Maital. 1985. "Optimal Tax Evasion & Optimal Tax Evasion Policy: Behavioral Aspects." In Wulf Gaertner and Alois Wenig (eds.), *The Economics of the Shadow Economy*. Studies in Contemporary Economics, vol. 15. Springer, Berlin, Heidelberg: 245-264.
- Benshalom, Ilan. 2012. "Taxing Cash." *Columbia Journal of Tax Law* 4 (1): 65-93.

Bergolo, Marcelo, Rodrigo Ceni, Guillermo Cruces, Matias Giacobasso, and Ricardo Perez-Truglia. 2017. "Tax Audits as Scarecrows: Evidence from a Large-Scale Field Experiment." Working Paper, University of California, Los Angeles.

Best, Michael C. 2014. "The Role of Firms in Workers' Earnings Responses to Taxes: Evidence from Pakistan." Working Paper, Stanford University.

Best, Michael C., Anne Brockmeyer, Henrik Jacobsen Kleven, Johannes Spinnewijn, and Mazhar Waseem. 2015. "Production vs. Revenue Efficiency with Limited Tax Capacity: Theory and Evidence from Pakistan." *Journal of Political Economy* 123 (6): 1311-1355.

Bhargava, Saurabh, and Dayanand Manoli. 2015. "Psychological Frictions and Incomplete Take-up of Social Benefits: Evidence from an IRS Field Experiment." *American Economic Review* 105 (11): 3489-3529.

Blank, Joshua D. 2014a. "Collateral Compliance." *University of Pennsylvania Law Review* 162: 719-800.

Blank, Joshua D. 2014b. "Reconsidering Corporate Tax Privacy." *New York University Journal of Law and Business* 11 (1): 31-121.

Blank, Joshua D., and Daniel Z. Levin. 2010. "When is Tax Enforcement Publicized?" *Virginia Tax Review* 30 (1): 1-38.

Blumenthal, Marsha, Charles Christian, and Joel Slemrod. 2001. "Do Normative Appeals Affect Tax Compliance? Evidence from a Controlled Experiment in Minnesota." *National Tax Journal* 54 (1): 125-138.

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

Boning, Will, John Guyton, Ronald Hodge, Joel Slemrod, and Ugo Troiano. 2018. "Heard it

through the Grapevine: A Randomized Experiment Assessing Direct and Network Effects of Tax Treatment Strategies.” NBER Working Paper 24305.

Bott, Kristina M., Alexander W. Cappelen, Erik Ø. Sørensen, and Bertil Tungodden. 2017. "You've Got Mail: A Randomised Field Experiment on Tax Evasion." NHH Department of Economics Discussion Paper 10/2017.

Brockmeyer, Anne, and Marco Hernandez. 2017. “Fiscal Capacity, Information and Taxation.” World Bank Working Paper.

Cabral, Ana Cinta G., and Norman Gemmell. 2018. "Estimating Self-Employment Income-Gaps from Register and Survey Data: Evidence for New Zealand." Victoria University of Wellington Working Paper No. 07/2018.

Cabral, Ana Cinta G., Christos Kotsogiannis, and Gareth Myles. 2015. “Self-Employment Underreporting in Great Britain: Who and How Much?” Tax Administration Research Centre, University of Exeter Discussion Paper No. 010-14.

Carrillo, Paul, Edgar Castro, and Carlos Scartascini. 2017. “Do Rewards Work? Evidence from the Randomization of Public Works.” Working paper, Inter-American Development Bank,

Carrillo, Paul, Dina Pomeranz, and Monica Singhal. 2017. “Dodging the Taxman: Firm Misreporting and Limits to Tax Enforcement.” *American Economic Journal: Applied Economics* 9 (2): 144-164.

Carrillo, Paul, M. Shahe Emran, and Anita Rivadeneira. 2012. “Do Cheaters Bunch Together? Profit Taxes, Withholding Rates and Tax Evasion.” Working Paper, George Washington University.

Castro, Lucio, and Carlos Scartascini. 2015. "Tax Compliance and Enforcement in the Pampas: Evidence from a Field Experiment." *Journal of Economic Behavior & Organization* 116: 65-82.

Chalfin, Aaron, and Justin McCrary. 2017. "Criminal Deterrence: A Review of the Literature." *Journal of Economic Literature* 55 (1): 5-48.

Chandola, Varun, Arindam Banerjee, and Vipin Kumar. 2009. "Anomaly Detection: A Survey." *ACM Computing Surveys* 41 (3): 1-58.

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

Chetty, Raj, Adam Looney, and Kory Kroft. 2009. "Salience and Taxation: Theory and Evidence." *American Economic Review* 99 (4): 1145-1177.

Chirico, Michael, Robert P. Inman, Charles Loeffler, John MacDonald, and Holger Sieg. 2016. "An Experimental Evaluation of Notification Strategies to Increase Property Tax Compliance: Free-Riding in the City of Brotherly Love." *Tax Policy and the Economy* 30 (1): 129-161.

Clausing, Kimberly A. 2009. "Multinational Firm Tax Avoidance and Tax Policy." *National Tax Journal* 62 (4): 703-725.

Coase, Ronald H. 1937. "The Nature of the Firm." *Economica* 4 (16): 386-405.

Cowell, Frank. 1990. *Cheating the Government*. Cambridge, MA: MIT Press.

Cullen, Julie Berry, Nicholas Turner, and Ebonya Washington. 2018. "Political Alignment, Attitudes toward Government and Tax Evasion." NBER Working Paper 24323.

Deaton, Angus. 2010. "Instruments, Randomization, and Learning about Development." *Journal of Economic Literature* 48 (2): 424-455.

DeBacker, Jason, Bradley T. Heim, Anh Tran, and Alexander Yuskavage. 2015. "Once Bitten, Twice Shy? The Lasting Impact of IRS Audits on Individual Tax Reporting." *Journal of Financial Economics* 117 (1): 122-138.

De Paula, Aureo, and Jose A. Scheinkman. 2010. "Value-added Taxes, Chain Effects, and Informality." *American Economic Journal: Macroeconomics* 2 (4): 195-221.

Del Carpio, Lucia. 2014. "Are the Neighbors Cheating? Evidence from a Social Norm Experiment on Property Taxes in Peru." Working Paper, Princeton University.

Dhami, Sanjit, and Ali al-Nowaihi. 2007. "Why Do People Pay Taxes? Prospect Theory versus Expected Utility Theory." *Journal of Economic Behavior & Organization* 64 (1): 171-192.

Diamond, Peter A., and James A. Mirrlees. 1971. "Optimal Taxation and Public Production I: Production Efficiency." *American Economic Review* 61 (1): 8-27.

Dobbie, Will, and Jae Song. 2015. "Debt Relief and Debtor Outcomes: Measuring the Effects of Consumer Bankruptcy Protection." *American Economic Review* 105 (3): 1272-1311.

Drago, Francesco, Friederike Mengel, and Christian Traxler. 2015. "Compliance Behavior in Networks: Evidence from a Field Experiment." IZA Discussion Paper 9443.

Dunbar, Geoffrey R., and Chunling Fu. 2015. "Sheltered Income: Estimating Income Under-reporting in Canada, 1998 and 2004." Bank of Canada Working Paper No. 2015-22.

Dunning, Thad, Felipe Monestier, Rafael Pineiro, Fernando Rosenblatt, and Guadalupe Tunón. 2016. "Is Paying Taxes Habit Forming? Experimental Evidence from Uruguay." Working paper, University of California, Berkeley.

Durlauf, Steven N., and Daniel S. Nagin. 2011. "Imprisonment and Crime: Can Both Be Reduced?" *Criminology & Public Policy* 10 (1): 13-54.

Dusek, Libor, and Sutirtha Bagchi. 2017. "Third-party Reporting, Tax Collections, and the Size of Government: Evidence from Withholding." Working paper, Villanova University.

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

Ellis, Blake. 2013. "Prisoners Rake in Millions from Tax Fraud." *CNN Money*, January 30.

Engström, Per and Johannes Hagen. 2015. "Income Underreporting among the Self-Employed: A Permanent Income Approach." *European Economic Review* 92: 92-109.

Engström, Per, and Bertil Holmlund. 2009. "Tax Evasion and Self-Employment in a High-Tax Country: Evidence from Sweden." *Applied Economics* 41 (19): 2419-2430.

Erard, Brian. 1992. "The Influence of Tax Audits on Reporting Behavior." In Joel Slemrod (ed.), *Why People Pay Taxes: Tax Compliance and Enforcement*. Ann Arbor: University of Michigan Press.

Fack, Gabrielle, and Camille Landais. 2016. "The Effect of Tax Enforcement on Tax Elasticities: Evidence from Charitable Contributions in France." *Journal of Public Economics* 133: 23–40.

Feinstein, Jonathan S. 1999. "Approaches for Estimating Noncompliance: Examples from Federal Taxation in the United States." *Economic Journal* 109 (456): 360-369.

Feldman, Naomi, and Joel Slemrod. 2007. "Estimating Tax Noncompliance with Evidence from Unaudited Tax Returns." *Economic Journal* 117 (518): 327-352.

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

Fisman, Raymond, and Shang-Jin Wei. 2004. "Tax Rates and Tax Evasion: Evidence from 'Missing Imports' in China." *Journal of Political Economy* 112 (2): 471-496.

Frey, Bruno. 1997. "A Constitution for Knaves Crowds Out Civic Virtues." *Economic Journal* 107 (443): 1043-1053.

Gemmell, Norman, and Marisa Ratto. 2012. "Behavioral Responses to Taxpayer Audits: Evidence from Random Taxpayer Inquiries." *National Tax Journal* 65 (1): 33-57.

Gillitzer, Christian, and Peer Ebbesen Skov. 2018. "The Use of Third-Party Information Reporting for Tax Deductions: Evidence and Implications from Charitable Contributions in Denmark." *Oxford Economic Papers*.

Gillitzer, Christian, and Joel Slemrod. 2016. "Does Evasion Invalidate the Welfare Sufficiency of the ETI?" *B.E. Journal of Economic Analysis and Policy* 16 (4).

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

Hagerty, Elizabeth. 2010. "Speed Camera Lottery Pays Drivers for Slowing Down." *Toronto Star*, December 9.

Hallsworth, Michael. 2014. "The Use of Field Experiments to Increase Tax Compliance." *Oxford Review of Public Policy* 30 (4): 658-679.

Hallsworth, Michael, John A. List, Robert D. Metcalfe, and Ivo Vlaev. 2017. "The Behavioralist as Tax Collector: Using Natural Field Experiments to Enhance Tax Compliance." *Journal of Public Economics* 148: 14-31.

Handley, Kyle, and Timothy J. Moore. 2017. "Tariffs, Transport Costs, and Tax Evasion." Working Paper, University of Michigan.

Hanlon, Michelle, and Joel Slemrod. 2009. "What Does Tax Aggressiveness Signal? Evidence from Stock Price Reactions to News about Tax Shelter Involvement." *Journal of Public Economics* 93 (1): 126-141.

Hasegawa, Makoto, Jeffrey L. Hoopes, Ryo Ishida, and Joel Slemrod. 2013. "The Effect of Public Disclosure on Reported Taxable Income: Evidence from Individuals and Corporations in Japan." *National Tax Journal* 66 (3): 571-608.

HM Revenue & Customs. 2015. *Measuring Tax Gaps 2015 Edition: Tax Gap Estimates for 2013-14*.

Holtzblatt, Janet. 2007. "Implications of Return-free Tax Systems for the Structure of the Individual Income Tax." *FinanzArchiv/Public Finance Analysis* 63 (3): 327-349.

Hoopes, Jeffrey L., Daniel Reck, and Joel Slemrod. 2015. "Taxpayer Search for Information: Implications for Rational Attention." *American Economic Journal: Economic Policy* 7 (3): 177-208.

Hoopes, Jeffrey L., Leslie Robinson, and Joel Slemrod. Forthcoming. "Tax-Return Disclosure." *Journal of Accounting and Economics*.

Hurst, Eric, Geng Li, and Benjamin Pugsley. 2014. "Are Household Surveys Like Tax Forms? Evidence from Income Underreporting of the Self-Employed." *Review of Economics and Statistics* 96 (1): 19-33.

Internal Revenue Service. 2014. *Strategic Plan FY 2014-2017*. Washington, DC, IRS Publication 3744.

Internal Revenue Service, Research, Analytics & Statistics. 2016. *Federal Tax Compliance Research: Tax Gap Estimates for Tax Years 2008-2010*. Publication 1415 (Rev. 5-2016), Washington, DC.

International Monetary Fund, Fiscal Affairs Department. 2013. "United Kingdom Assessment of HMRC's Tax Gap Analysis." IMF Country Report No. 13/314, Washington, DC.

International Monetary Fund. 2015. *Current Challenges in Revenue Mobilization: Improving Tax Compliance*. Washington, DC.

Jensen, Anders. 2016. "Employment Structure and the Rise of the Modern Tax System." London School of Economics Working Paper.

Johannesen, Niels, Patrick Langetieg, Daniel Reck, Max Risch, and Joel Slemrod. 2017. "Taxing Hidden Wealth: The Consequences of U.S. Enforcement Initiatives on Evasive Foreign Accounts." Working Paper, University of Michigan.

Johansson, Edvard. 2005. "An Estimate of Self-Employment Income Underreporting in Finland." *Nordic Journal of Political Economy* 31 (1): 99-109.

Johns, Andrew, and Joel Slemrod. 2010. "The Distribution of Income Tax Noncompliance." *National Tax Journal* 63 (3): 397-418.

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47 (2): 263-292.

Kanbur, Ravi, and Michael Keen. 2014. "Thresholds, Informality, and Partitions of Compliance." *International Tax and Public Finance* 21 (4): 536-559.

Keen, Michael, and Joel Slemrod. 2017. "Optimal Tax Administration." *Journal of Public Economics* 152: 133-142.

Kesselman, Jonathan R. 1989. "Income Tax Evasion: An Intersectoral Analysis." *Journal of Public Economics* 38 (2): 137-182.

Khan, Adnan Q., Asim I. Khwaja, and Benjamin A. Olken. 2016. "Tax Farming Redux: Experimental Evidence on Performance Pay for Tax Collectors." *Quarterly Journal of Economics* 131 (1): 219-271.

Klepper, Steven, Mark Mazur, and Daniel Nagin. 1991. "Expert Intermediaries and Legal Compliance: The Case of Tax Preparers." *Journal of Law and Economics* 34 (1): 205-229.

Kleven, Henrik Jacobsen. 2014. "How Can Scandinavians Tax So Much?" *Journal of Economic Perspectives* 28 (4): 77-98.

Kleven, Henrik Jacobsen. 2016. "Bunching." *Annual Review of Economics* 8: 435-464.

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

Kleven, Henrik Jacobsen, Claus T. Kreiner, and Emmanuel Saez. 2016. "Why Can Modern Governments Tax So Much? An Agency Model of Firms as Fiscal Intermediaries." *Economica* 83 (330): 219-246.

Kleven, Henrik Jacobsen, and Mazhar Waseem. 2013. "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan." *Quarterly Journal of Economics* 128 (2): 669-723.

Kopczuk, Wojciech. 2005. "Tax Bases, Tax Rates and the Elasticity of Reported Income." *Journal of Public Economics* 89 (11-12): 2093-2119.

Kopczuk, Wojciech, Justin Marion, Erich Muehlegger, and Joel Slemrod. 2016. "Does Tax-Collection Invariance Hold? Evasion and the Pass-through of State Diesel Taxes." *American Economic Journal: Economic Policy* 8 (2): 251-286.

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

Kopczuk, Wojciech, and Joel Slemrod. 2010. "Taxation of Family Firms." Working Paper, University of Michigan.

Kotakorpi, Kaisa, and Jani-Petri Laamanen. 2016. "Prefilled Income Tax Returns and Tax Compliance: Evidence from a Natural Experiment." Working Paper, University of Turku.

Kuchumova, Yulia. 2015. "Tax Debt Collection Enforcement: When Does Suspension of a Driver's License Help?" Working Paper, University of Michigan.

Kuchumova, Yulia. 2017. "A Collateral Tax Sanction: When Does it Mimic a Welfare-Improving Tag?" Higher School of Economics Research Paper No. WP BRP 181/EC/2017.

Kumler, Todd, Eric Verhoogen, and Judith A. Frias. 2013. "Enlisting Employees in Improving Payroll-Tax Compliance: Evidence from Mexico." National Bureau of Economic Research Working Paper No. 19385.

LaLumia, Sara, and James M. Sallee. 2013. "The Value of Honesty: Empirical Estimates from the Case of the Missing Children." *International Tax and Public Finance* 20 (2): 192-224.

Lee, David S., and Justin McCrary. 2017. "The Deterrence Effect of Prison: Dynamic Theory and Evidence." *Advances in Econometrics*, Volume 38, Chapter 3: 73-146.

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

Levi, Margaret. 1989. *Of Rule and Revenue*. Berkeley and Los Angeles: University of California Press.

List, John A. 2011. "Why Economists Should Conduct Field Experiments and 14 Tips For Pulling One Off." *Journal of Economic Perspectives* 25 (3): 3-16.

Logue, Kyle D., and Joel Slemrod. 2009. "Of Coase, Calabresi, and Optimal Tax Liability." *Tax Law Review* 63: 797-866.

Long, Susan B., and Richard D. Schwartz. 1987. "The Impact of IRS Audits on Taxpayer Compliance: A Field Experiment in Specific Deterrence." *Annual Meeting of the Law and Society Association*.

Luttmer, Erzo F.P., and Monica Singhal. 2014. "Tax Morale." *Journal of Economic Perspectives* 28 (4): 149-168.

Malik, Rabia. 2017. "Transparency, Elections, and Pakistani Politicians' Tax Compliance." Working paper, New York University Abu Dhabi.

Marchese, Carla. 2009. "Rewarding the Consumer for Curbing the Evasion of Commodity Taxes?" *FinanzArchiv* 65 (4): 383-402.

Martinez-Lopez, Diego. 2013. "The Underreporting of Income by Self-employed Workers in Spain." *SERIEs* 4 (4): 353-371.

Mascagni, Giulia. 2018. "From the Lab to the Field: A Review of Tax Experiments." *Journal of Economic Surveys* 32 (2): 273-301.

Matsaganis, Manos, and Maria Flevotomou. 2010. "Distributional Implications of Tax Evasion in Greece." London School of Economics Hellenic Observatory Papers on Greece and Southeast Europe GreeSE Paper 31.

Mayshar, Joram. 1991. "Taxation with Costly Administration." *Scandinavian Journal of Economics* 93 (1): 75-88.

Meiselman, Ben. Forthcoming. "Ghostbusting in Detroit: Evidence on Nonfilers from a Controlled Field Experiment." *Journal of Public Economics*.

Milgram, Stanley. 1963. "Behavioral Study of Obedience." *The Journal of Abnormal and Social Psychology* 67 (4): 371-378.

Nagin, Daniel S. 2013. "Deterrence: A Review of the Evidence by a Criminologist for Economists." *Annual Review of Economics* 5 (1): 83-105.

Naritomi, Joana. 2016. "Consumers as Tax Auditors." Working Paper, London School of Economics.

Nygård, Odd E., Joel Slemrod, and Thor O. Thoresen. Forthcoming. "Distributional Implications of Joint Tax Evasion." *Economic Journal*.

OECD. 2015. *Tax Administration 2015: Comparative Information on OECD and Other Advanced and Emerging Economies*. OECD Publishing, Paris.

OECD. 2017. *Legal Tax Liability, Legal Remittance Responsibility & Tax Incidence: Three Dimensions of Business Taxation*. Centre for Tax Policy and Administration, Committee on Fiscal Affairs.

Onji, Kazuki. 2009. "The Response of Firms to Eligibility Thresholds: Evidence from the Japanese Value-Added Tax." *Journal of Public Economics* 93 (5): 766-775.

Ortega, Daniel, and Carlos Scartascini. 2015. "Who's Calling? The Effect of Phone Calls as a Deterrence Mechanism." Working Paper, Inter-American Development Bank.

Ortega, Daniel, and Carlos Scartascini. 2017. "Don't Blame the Messenger: A Field Experiment on Delivery Methods." Working Paper, Inter-American Development Bank.

Paetzold, Jörg, and Hannes Winner. 2016. "Taking the High Road? Compliance with Commuter Tax Allowances and the Role of Evasion Spillovers." *Journal of Public Economics* 143: 1-14.

Paulus, Alari. 2015. "Tax Evasion and Measurement Error: An Econometric Analysis of Survey Data Linked with Tax Records." Institute for Social and Economic Research, University of Essex Working Paper No. 2015-10.

Perez-Truglia, Ricardo, and Ugo Troiano. 2015. "Shaming Tax Delinquents: Evidence from a Field Experiment in the United States." Available at SSRN: <https://ssrn.com/abstract=2558115>.

Pissarides, Christopher A., and Guglielmo Weber. 1989. "An Expenditure-based Estimate of Britain's Black Economy." *Journal of Public Economics* 39 (1): 17-32.

Pomeranz, Dina. 2015. "No Taxation without Information: Deterrence and Self-Enforcement in the Value Added Tax." *American Economic Review* 105 (8): 2539-2569.

Rinke, Johannes, and Christian Traxler. 2011. "Enforcement Spillovers." *Review of Economics and Statistics* 93 (4): 1224-1234.

Rogoff, Kenneth S. 2015. "Costs and Benefits of Phasing Out Paper Currency." *NBER Macroeconomics Annual* 29 (1): 445-456.

Saez, Emmanuel, Joel Slemrod, and Seth Giertz, 2012. "The Elasticity of Taxable Income With Respect to Marginal Tax Rates: A Critical Review," *Journal of Economic Literature* 50 (1): 1-50.

Sah, Raaj K. 1991. "Social Osmosis and Patterns of Crime." *Journal of Political Economy* 99 (6): 1272-1295.

Schuetze, Herb J. 2002. "Profiles of Tax Non-Compliance among the Self-employed in Canada: 1969 to 1992." *Canadian Public Policy/Analyse de Politiques* 28 (2): 219-238.

Slemrod, Joel. 2008. "Does It Matter Who Writes the Check to the Government? The Economics of Tax Remittance." *National Tax Journal* 61 (2): 251-275.

Slemrod, Joel. 2016. "Caveats to the Research Use of Tax-Return Administrative Data." *National Tax Journal* 69 (4): 1003-1020.

Slemrod, Joel. 2017. "Tax Compliance and Enforcement: An Overview of New Research and its Policy Implications." In Alan Auerbach and Kent Smetters (eds.), *The Economics of Tax Policy*, Oxford University Press: 81-102.

Slemrod, Joel, Marsha Blumenthal, and Charles Christian. 2001. "Taxpayer Response to an Increased Probability of Audit: Evidence from a Controlled Experiment in Minnesota." *Journal of Public Economics* 79 (3): 429-453.

Slemrod, Joel, Brett Collins, Jeffrey Hoopes, Daniel Reck, and Michael Sebastiani. 2017. "Does Credit-Card Information Reporting Improve Small-Business Tax Compliance?" *Journal of Public Economics* 149: 1-19.

Slemrod, Joel, and Christian Gillitzer. 2014. *Tax Systems*. Cambridge, MA: MIT Press.

Slemrod, Joel, and Wojciech Kopczuk. 2002. "The Optimal Elasticity of Taxable Income." *Journal of Public Economics* 84 (1): 91-112.

Slemrod, Joel, and Tejaswi Velayudhan. 2018. "Do Firms Remit At Least 85% of Tax Everywhere? New Evidence from India." *Journal of Tax Administration* 4 (1): 124-137.

Slemrod, Joel, and Caroline Weber. 2012. "Evidence of the Invisible: Toward a Credibility Revolution in the Empirical Analysis of Tax Evasion and the Informal Economy." *International Tax and Public Finance* 19 (1): 25-53.

Slemrod, Joel, and Shlomo Yitzhaki. 1987. "The Optimal Size of a Tax Collection Agency." *Scandinavian Journal of Economics* 89 (2): 183-192.

Slemrod, Joel, and Shlomo Yitzhaki. 1996. "The Costs of Taxation and the Marginal Efficiency Cost of Funds." *International Monetary Fund Staff Papers* 43 (1): 172-198.

Slemrod, Joel, and Shlomo Yitzhaki. 2002. "Tax Avoidance, Evasion, and Administration." In Alan Auerbach and Martin Feldstein (eds.), *Handbook of Public Economics*, Vol. 3 (Amsterdam: North-Holland): 1423-1470.

Tax Administration Reform Commission. 2014. *Tax Administration Reform in India: Spirit, Purpose and Empowerment*. Government of India.

Torgler, Benno. 2004. "Moral Suasion: An Alternative Tax Policy Strategy? Evidence from a Controlled Field Experiment." *Economics of Governance* 5 (3): 235–253.

Tyler, Tom R. 2006. *Why People Obey the Law*. Princeton: Princeton University Press.

Ulph, David. 2015. "Measuring Tax Complexity." In Chris Evans, Rick Krever, and Peter Mellor (eds.) *Tax Simplification*, Kluwer Law International: 41-54.

Wiling, Eleanor. 2017. "Hotel Tax Incidence with Heterogeneous Firm Evasion: Evidence from Airbnb Remittance Agreements." Working Paper, University of Michigan.

Williams, Colin C. 2014. *Confronting the Shadow Economy: Evaluating Tax Compliance and Behaviour Policies*. Cheltenham, UK and Northampton, MA: Edward Elgar.

Yang, Dean. 2008. "Can Enforcement Backfire? Crime Displacement in the Context of Customs Reform in the Philippines." *Review of Economics and Statistics* 90 (1): 1-14.

Yaniv, Gideon. 1988. "Withholding and Non-withheld Tax Evasion." *Journal of Public Economics* 35 (2): 183-204.

Yitzhaki, Shlomo. 1974. "A Note on 'Income Tax Evasion: A Theoretical Analysis.'" *Journal of Public Economics* 3 (2): 201-202.

Zucman, Gabriel. 2014. "Taxing across Borders: Tracking Personal Wealth and Corporate Profits." *Journal of Economic Perspectives* 28 (4): 121-148.

Zwick, Eric. 2018. "Do Experts Help Firms Optimize?" NBER Working Paper No. 24382.

