Munich Personal RePEc Archive

Tax Evasion, Information Reporting, and the Regressive Bias Prediction

Pinje, Jori Veng and Boserup, Simon Halphen

University of Copenhagen

26 January 2011

Online at https://mpra.ub.uni-muenchen.de/35935/
MPRA Paper No. 35935, posted 14 Jan 2012 02:48 UTC

Tax Evasion, Information Reporting, and the Regressive Bias Prediction[1]

Simon Halphen Boserup[2] and Jori Veng Pinje

University of Copenhagen

January 2012

[1]We are grateful to the Danish tax administration, SKAT, for providing data access and to Peer Skov and the Rockwool Foundation Research Unit for providing data and assistance on black market activities. We thank Alan Auerbach, Jonathan Feinstein, Henrik Kleven, Claus Thustrup Kreiner, David Dreyer Lassen, Emmanuel Saez, Suzanne Scotchmer, and Joel Sobel as well as seminar participants at Nordic Workshop on Tax Policy and Public Economics, UC Berkeley, the 6th Royal Economic Society PhD Meeting at City University London and the EDGE Meeting at University College Dublin for helpful comments and suggestions.

[2]Corresponding author. Tel. +45 2649 2696. *E-mail address:* simon.h.boserup@econ.ku.dk. Department of Economics, University of Copenhagen, Øster Farimagsgade 5, building 26, DK-1353 København K, Denmark.

Abstract

A robust, but untested, prediction from the tax evasion literature is that optimal auditing induces a regressive bias in effective average tax rates compared to statutory rates, reducing the degree of redistribution in the tax system. Using Danish administrative data, we show that a calibrated structural model of rational tax evasion and tax enforcement can convincingly replicate the moments and correlations of tax evasion and probabilities of audit once we account for the presence of information reporting in the tax compliance game. We find that both reduced-form evidence and simulations are in accordance with the prediction of regressive bias when conditioning on information reporting. However, information reporting counteracts the regressive bias generated by optimal evasion and auditing behavior and, as a consequence, the bias vanishes when considering the degree of redistribution in the overall economy.

JEL: D82, H26, K42

Keywords: Information Reporting, Regressive Bias, Tax Enforcement, Tax Evasion

1 Introduction

In this article, we develop a structural model of tax evasion and enforcement in a population of taxpayers. Highly detailed Danish administrative data allows us to perform a meaningful calibration exercise to investigate the model's ability to explain tax evasion and the tax agency's enforcement strategy. We show that the model's predictions closely match key empirical relationships in the data and, in particular, we provide the first empirical evidence of the regressive bias prediction established in the theoretical literature on tax evasion and optimal enforcement (see for example Reinganum and Wilde, 1986; Cremer, Marchand, and Pestieau, 1990; Sanchez and Sobel, 1993; Erard and Feinstein, 1994).

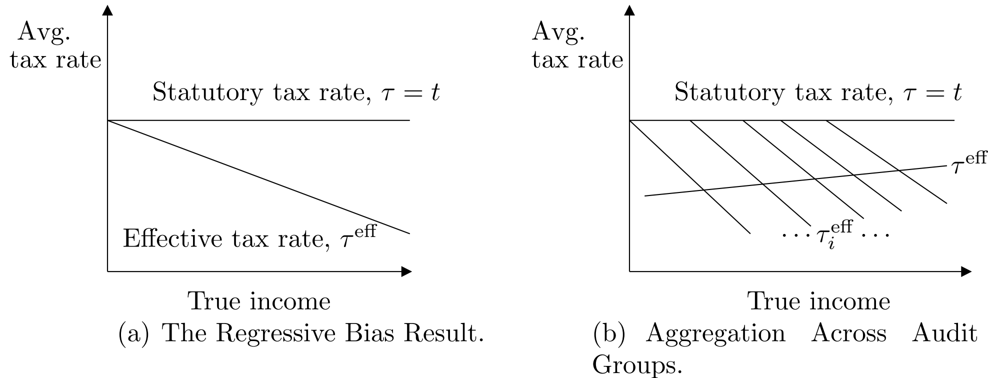
The potential for tax evasion requires a distinction between the statutory tax system and the *effective* tax system. Tax evaders pay less taxes than they should and this implies a wedge between statutory and effective average tax rates. The regressive bias prediction states that this wedge is larger for high-income taxpayers than for low-income taxpayers – even when the enforcement regime is revenue maximizing. Thus, the tax system may be substantially less redistributive than intended by the tax code. As shown by Scotchmer (1992), the prediction of regressive bias is theoretically robust. Model variations in the literature consistently arrive at regressively biased effective average tax rates.

The intuition behind this prediction is the following: The tax compliance game played by the tax agency and taxpayers is a screening problem in which high-income taxpayers can increase their expected payoff by imitating low-income taxpayers. If not all taxpayers can be audited, the tax agency should optimally prioritize tax returns reporting low income. Rather than eliminating tax evasion altogether, budget-constrained optimal enforcement primarily discourages very low reports by high-income individuals. Due to the optimal regressivity in tax enforcement, evading taxes on the margin subjects a low-income taxpayer to a greater risk of getting caught than a high-income taxpayer, which tends to make high-income taxpayers evade more. In equilibrium, the decreasing relationship between the probability of audit and reported income and the increasing relationship between evaded taxes and true income

lead to an increasing wedge between the statutory average tax rate and the effective average tax rate as a function of true income, i.e., a regressive bias. Figure 1(a) illustrates how the wedge between the effective average tax rate, τ^{eff}, and the average tax rate as implied by the statutory tax system, τ, is increasing in true income.

There is one important exception to the regressive bias result: when the tax agency uses *ex ante* observable population variables, such as gender, age, occupation, or employer-reported salaries, to predict true incomes, there may be no bias or even progressive bias in the population as a whole. How this plays out in particular economies will determine the appropriate way to account for the redistributional aspects of tax evasion and enforcement through economic policy. Scotchmer (1987) shows that when tax agencies facilitate prediction of taxpayers' true income by dividing taxpayers into *audit groups*, upon which the agency conditions its enforcement strategy, effective average tax rates remain regressively biased within audit groups but the direction of the bias between groups is ambiguous. The aggregate bias depends on the predictive power of the signals (i.e., the *ex ante* known population variables) and the allocation of audit resources across audit groups. Consequently, the regressive bias prediction should be interpreted as a *within-audit-group* phenomenon. Figure 1(b) illustrates the aggregate relationship between effective average tax rates, τ^{eff}, and true income, which is a composite of relationships within multiple audit groups, τ_i^{eff}. Whereas the regressive bias prediction remains valid within audit groups, effective tax rates may be *progressively biased* across audit groups.

The mechanism driving the result is that some low-income taxpayers benefit from being high-income individuals within their audit group while some high-income taxpayers instead are low-income taxpayers within their audit group. This reclassification changes the risk of being audited and, hence, the *ex ante* effective tax rate. In addition, the tax agency can more efficiently target high-income individuals by modifying the distribution of audit resources between audit groups. If the observable signal of true income is stronger or audits are more abundant among high-income taxpayers, progressive bias between groups may dominate in



(a) The Regressive Bias Result.

(b) Aggregation Across Audit Groups.

Figure 1. Correlation Structure of Effective Average Tax Rates.
Notes: τ is the statutory average tax rate (here, constant at $\tau = t$), τ_i^{eff} is the effective average tax rate within audit group i, and τ^{eff} is the aggregate effective average tax rate.

the aggregate.

We apply a specific theoretical structure to Danish administrative data on tax compliance/evasion and show that the empirical properties of tax evasion, tax enforcement, and effective tax rates are convincingly replicated by a screening game between a tax agency and taxpayers. To this end, we combine insights from two main sources, Kleven, Knudsen, Kreiner, Pedersen, and Saez (2011) and Erard and Feinstein (1994). In the former, the authors collect a uniquely detailed micro-data set based on a random sample of Danish taxpayers containing pre- and post-audit incomes and taxes, as well as reports on income, proxies for audit probabilities, etc. They show that third-party reported income is by far the best predictor of true income compared to other population variables. Since the Danish tax agency, SKAT, does in fact use these information reports extensively in its enforcement efforts, they are ideal for constructing audit groups.[1] Based on this insight, we generalize Erard and Feinstein's within-audit-group model to describe tax evasion and optimal enforcement both within and between audit groups. We calculate an internally consistent set of model parameters directly from data and calibrate the tax agency's budget to match the simulated level of tax evasion to data. We evaluate the model numerically and find that applying

[1] Other recent papers demonstrate the importance of explicitly considering information reporting. Phillips (2010) demonstrates the predictive power of an indirect measure of third-party reported information in US data and Pomeranz (2010) demonstrates the general importance of information as a deterrent of VAT evasion in a sample of Chilean firms.

structure to the data yields results in close correspondence with a minimal-assumptions reduced-form approach. This model convincingly replicates tax evasion behavior for both wage earners and the self-employed although these two groups differ markedly in terms of the propensity to evade taxes and the extent and distribution of third-party reported income. We conclude that (statically optimized) tax evasion and tax enforcement is sufficient to generate the observed structure of effective average tax rates.

Overall, our micro-data on Danish taxpayers suggests that there is a regressive bias within audit groups. Between audit groups, tax rates are progressively biased to such an extent that tax rates are actually progressively biased in total income. Thus, our findings support the regressive bias prediction at the theoretical level but not as an aggregate empirical outcome in Danish data; specifically, our results correspond closely to the structure of effective tax rates conjectured in Scotchmer (1987). Moreover, using information about the enforcement regime, we find evidence suggesting that the actual audit regime exhibits the key qualitative features of an optimal audit regime and that the correlation structure of effective average tax rates is, indeed, caused by the theorized combination of optimal enforcement and tax evasion.

In model simulations, the covariance structure of effective average tax rates is robust to parameter variations. In view of this, we predict that similar empirical relationships would be found in data from any tax agency that employs, as does the Danish tax agency, a strong signal in predicting true incomes. The model also suggests that enforcement regimes employing information reporting to a lesser extent may be substantially more regressive.

Our results have important implications for policy. Due to the theoretical robustness of the regressive bias prediction, it has been argued (e.g., in Scotchmer, 1992) that governments could increase the progressivity of the income tax schedule to counter regressive bias inherent in optimal tax enforcement. However, our results imply that such a policy adjustment is undesirable. In the first place, adjusting tax rates cannot eliminate the inequity between taxpayers that evade taxes and taxpayers that do not. Secondly, there may be no regressive

bias to correct in the aggregate due to tax agencies' use of third-party reported information in tax enforcement. If such is the case, the policy priority is correcting the horizontal inequity between evaders and non-evaders rather than the distortion of redistribution between high- and low-income taxpayers – for this purpose allocating more resources to the tax agency or collecting more information *ex ante* are superior approaches.

Our results illustrate the importance of including information reports in empirical analyses of tax evasion and enforcement. Neglecting to account for information reports may lead to counterintuitive comparative statics estimates such as for tax evasion with respect to total income or marginal taxes. This may partly explain the empirical literature's lack of consensus with respect to basic correlations between measures of tax evasion, tax rates, and income.[2]

We now proceed to the main body of the paper. Section 2 develops our model of the tax compliance/evasion game. Section 3 outlines the Danish tax system and describes the main features of the data. Section 4 describes the calibration of parameters, outlines the numerical strategy and establishes the correspondence of data and model-generated output. Section 5 concludes. The Appendix provides details of the numerical implementation and a description of black market activity in Denmark.

2 Theory: A Model of Income Tax Auditing Subject to Information Reporting

Several current theories are capable of analyzing behavior *within* audit groups, i.e., conditional on pre-defined groups based on *ex ante* observable information. However, as we wish to analyze *aggregate* reporting behavior as well as the tax agency's overall response, we need a model that can encompass a population of taxpayers, i.e., several audit groups. To this

[2]For example, Feinstein (1991) finds a negative effect of marginal tax rates on underreporting, whereas Clotfelter (1983) finds a positive effect. With respect to the effect of income on underreporting, Feinstein (1991) finds no effect, whereas Clotfelter (1983) finds a positive effect.

end we generalize the model in Erard and Feinstein (1994) to incorporate a population that is heterogeneous in third-party income reports.[3]

Erard and Feinstein (1994) introduce noise in taxpayer reports by incorporating the stylized fact that some taxpayers report their incomes honestly, even when they have ample opportunity to evade taxes. This is also the case in our data as we demonstrate in Section 3. As argued in Erard and Feinstein (1994), including inherently honest taxpayers increases the realism and usefulness of the model: it eliminates several potential equilibria and leaves them with a unique revenue maximizing equilibrium prediction. Further, it eliminates the unrealistic feature of earlier models that the tax agency in equilibrium would know the true incomes of all taxpayers *before* the actual audit.[4] Thus, for each tax return filed by a particular taxpayer, the tax agency decides whether or not to audit based on the expected reports of dishonest and honest taxpayers and the likelihood that any particular tax return is fraudulent.

To develop a model that we can apply to data, we extend the model in Erard and Feinstein (1994) to account for the tax agency's use of information reports. As shown by Kleven et al. (2011), in the Danish context, third-party reported income is by far the most powerful predictor available, making it an ideal candidate for defining audit groups. However, as this variable, like true income, is intuitively best understood as a continuous variable, we allow the tax agency to choose audit functions contingent on the third-party information of a particular taxpayer and interpret each *level* of third-party reported income as an audit group. Reflecting the very low evasion rates on third-party reported income in our data, we use the simplifying assumptions that these reports are always correct and are common knowledge to both taxpayer and tax agency. Overall, the probability that a particular taxpayer is

[3] We use a different specification for penalties in case of detected evasion compared to Erard and Feinstein (1994). We model penalties as proportional to evaded taxes rather than evaded income as this is also the structure of the actual Danish penalty system.

[4] A limitation of the modeling framework is that it does not explain why some taxpayers choose to report honestly. However, the model is well-suited for analyzing the behavior of rational tax evaders *given* that some taxpayers are, in fact, honest. Moreover, it provides a relatively simple framework for analyzing optimal enforcement in the face of this behavior and subject to the informational asymmetries inherent in the tax enforcement/compliance game.

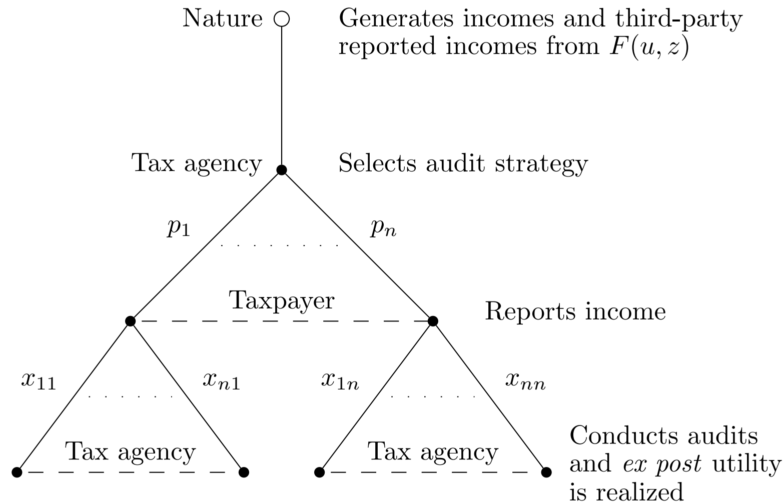


Figure 2. Game Tree.

audited depends both on the exogenous signal, i.e., third-party reported income, and the endogenously determined reported income.

The structure of the model is illustrated in Figure 2. The tax agency selects the audit regime subject to a budget constraint without being able to commit to an audit strategy. The audit schedule for a particular audit group (i.e., conditional on a particular third-party reported income level) is a function of taxpayers' reported *residual* incomes, i.e., income in excess of third-party reported income, reflecting our assumption that third-party reported income is common knowledge. The tax agency allocates its resources across different strata of the population so as to equalize the shadow values of extending resources to auditing taxpayers with different amounts of third-party reported income. Whereas the distribution of true incomes, conditional on information reports, is known, actual true incomes of individual taxpayers are private information. Taxpayers choose income reports subject to their expectations about the audit regime. Finally, the actual returns and the audit schedule are realized, audits are conducted, and tax revenue and *ex post* utilities, as measured by income net of taxes and any penalty payments, are realized.

2.1 Individual Reporting Behavior

Individual taxpayers have true taxable incomes y and report taxable incomes, $\tilde{y}$. Part of true income, z, is reported by third parties and is known to all parties. Therefore, $y = z+u$, where u is residual income, which can be positive or negative as it includes both, e.g., wages and deductions not reported by third parties. u is *ex ante* unknown and can only be ascertained by the tax agency by conducting a costly audit, which we assume reveals all of "true" residual income.[5] We denote the reported residual x, such that $x = \tilde{y} - z$.

Erard and Feinstein (1994) split taxpayers into two broad groups, honest and dishonest taxpayers, and assume that these two types differ only in reporting behavior, and that honesty is uncorrelated with true income. However, empirically the ratio of compliant to noncompliant taxpayers is not constant on the domain of u due to a large mass of correct reports around $u = 0$. The reason for this is that third-party reported income is such a strong signal of true income that, for many taxpayers, it is, indeed, virtually a perfect signal.[6] However, this can be remedied by a minimal departure from the assumptions of Erard and Feinstein (1994) by letting the ratio of honest to dishonest taxpayers differ on the domain of u. We define the densities of true income conditional on third-party reports $f^h_{u|z}$ and $f^d_{u|z}$ for honest and dishonest taxpayers, respectively. In addition, we define the total density function as $f_{u|z} = f^h_{u|z} + f^d_{u|z}$ and $F_{u|z}$ the conditional distribution function associated with $f_{u|z}$.

[5] We follow Erard and Feinstein (1994) in assuming that taxpayers do not incur a cost from filing taxes (time costs, hiring of a tax accountant, concealment costs etc.). Such costs have welfare consequences in the form of deadweight losses. Cremer and Gahvari (1994) show that a concealment technology that allows taxpayers to lower the probability of detection at a cost can affect the effective progressivity of the tax system. This may result in more or less progressivity depending on the exact specification of the concealment technology. However, their model assumes a constant audit probability, whereas our model implies a non-increasing audit probability on the domain of reports of dishonest taxpayers. In any case, whether or not such costs are important, our results in Section 4 indicate that they are not necessary to explain the correlation structure of effective average tax rates.

[6] In principle, such taxpayers could still evade taxes by claiming unwarranted deductions. This type of reporting behavior is virtually non-existent in our data. A possible explanation is that the burden of proof is on the taxpayer in such cases. On the other hand, having negative residual income (i.e., some deductions not subject to third-party reporting) allows for tax evasion by overstating the value of otherwise legal deductions.

We follow Erard and Feinstein (1994) in assuming that taxes are linear in income.[7] Whereas honest taxpayers always report $x = u$, we assume that dishonest taxpayers are risk neutral and maximize expected utility given by expected income net of taxes and penalties

$$(1-t)\,z + p\,(x|z)\,[(1-t)\,u - \theta t\,(u-x)] + (1-p\,(x|z))\,[u - tx]\,,$$

where t is the tax rate, θ is the penalty rate on tax evasion, and $p(x|z)$ is the audit probability for report x given the level of third-party reporting z. The correct amount of taxes are paid with certainty on income reported by third parties, whereas taxes (and penalties) paid on residual income depends on both a taxpayer's evasion behavior and whether or not the taxpayer is audited.

In optimum, the taxpayer's choice must satisfy the first order condition

$$u = x + \frac{p\,(x|z) - \frac{1}{1+\theta}}{p'\,(x|z)}. \tag{1}$$

It is clear from Equation (1) that for $p\,(\cdot) = \frac{1}{1+\theta}$, $x = u$ and evasion is discouraged completely. However, $p \geq \frac{1}{1+\theta}$ is not compatible with equilibrium when the tax agency cannot commit to the audit regime: if evasion were completely discouraged, the tax agency would lower p for some x as a cost saving measure. Thus, in equilibrium $p\,(\cdot) \in \left[0, \frac{1}{1+\theta}\right)$. Furthermore, the incentive compatibility constraints on the tax agency's optimization problem implies that audit functions are decreasing on the domain of income reports (see Erard and Feinstein (1994) for a detailed demonstration of this point).

Given that $p'\,(x|z)$ is negative and $p\,(x|z) < \frac{1}{1+\theta}$, increasing the audit probability will, *ceteris paribus*, lower tax evasion as the risk of getting caught is higher. Lowering $p'\,(x|z)$

[7] Clearly, this an abstraction but not an extreme one. Although the income tax schedule has three brackets, the average tax rates are much smoother. It would also be possible to perform the analyses using a full, nonlinear specification of taxes. We do not expect that the conclusions of this paper would be substantially affected by this change. Moreover, to accomodate the progressiveness of marginal income taxes as much as possible, in the empirial application of the model we allow the model's constant marginal tax rate to vary in z.

(increasing its absolute value) also reduces tax evasion by increasing the risk of audit from taxes evaded on the margin.[8]

2.2 Optimal Audit Response

The tax agency chooses a continuum of audit schedules $p(x|z)$ and a budget allocation $B(z)$ for all z. In this way, the informational aspect of using third-party reported incomes to predict true income is incorporated into the population-wide equilibrium.[9] The audit schedule is chosen to maximize expected revenue (taxes plus fines)[10]

$$\int \left(\int_{\underline{x}}^{\bar{u}} \left[p(x|z) \left(t\mathrm{E}(y|x,z) + \theta t \left(\mathrm{E}(y|x,z) - \tilde{y} \right) \right) + \left(1 - p(x|z) \right) t\tilde{y} \right] \mathrm{d}F_{x|z} \right) \mathrm{d}F_z$$

subject to the budget constraint

$$c \int \left(\int_{\underline{x}}^{\bar{u}} p(x|z) \, \mathrm{d}F_{x|z} \right) \mathrm{d}F_z \leq \int B(z) \, \mathrm{d}F_z \equiv \mathbf{B}, \tag{2}$$

where $F_{x|z}$ is the induced conditional distribution function for reported residual income, x, given third-party reported income, z; F_z is the marginal distribution function for z; and $B(z)$ is the proportion or density of the overall audit budget, $\mathbf{B}$, allocated to income reports

[8] Taxpayers' income returns must also satisfy the second order condition, $p''(x|z)(x-u) + 2p'(x|z) \leq 0$.

[9] In principle, the tax agency could also condition audit schedules on other population variables such as gender, age, occupation, etc. However, as Kleven et al. (2011) show, these variables are less powerful as predictors. Conditioning on whether the taxpayer was audited in previous years would complicate matters as it would introduce a dynamic aspect to reporting decisions. However, as observations on past audits are not employed in SKAT's actual audit scheme, this limitation is unlikely to affect the fit of our model. In addition, the statute of limitations for retrospective audits is limited to 14 months.

[10] Scotchmer (1992) shows that maximizing some measure of social welfare instead of expected revenue does not change the qualitative prediction that (within an audit group) there will be regressive bias, although it may change the distribution of resources across audit groups. The similarity of the observed and simulated distribution of resources, cf. Section 4, suggests that revenue maximation is not an inappropriate simplification in this context.

with third-party reported income, z. For each (x, z), the tax agency must choose p to solve

$$\max_p \{p [t\mathrm{E}(y|x,z) + \theta t (\mathrm{E}(y|x,z) - \tilde{y})] + (1-p)t\tilde{y}\} \,\mathrm{d}F_{x|z}\mathrm{d}F_z$$
$$-\lambda(z)\, c \left[p \,\mathrm{d}F_{x|z} - B(z)\right] \mathrm{d}F_z,$$

where $\lambda(z)$ is the Langrangian multiplier on the budget constraint. This implies a point-wise first order condition

$$tE(y|x,z) + \theta t E(y|x,z) - \theta t\tilde{y} - t\tilde{y} - \lambda(z)\, c \gtreqless 0, \tag{3}$$

which is greater than, equal to, or less than zero as $p = \frac{1}{1+\theta}$, $p \in \left(0, \frac{1}{1+\theta}\right)$, or $p = 0$. We look for equilibria in which the tax agency chooses a mixed strategy such that (3) holds with equality.[11]

As mentioned, our model is a generalization of the model in Erard and Feinstein (1994). Specifically, our model simplifies to theirs if $i)$ z is zero for all individuals, such that $F_{u|z} = F_u = F_y$, and $ii)$ the ratio of honest to dishonest taxpayers, $\frac{f_u^h(u)}{f_u^d(u)}$, is constant on $[\underline{u}, \overline{u}]$. In this case, the problem becomes that of a partial optimization for a fixed $B(z)$ within an audit group. In this simpler version of the model, Erard and Feinstein (1994) show that the equilibrium audit and evasion functions have a number of useful properties. Due to the incentive constraints on reporting for high-income taxpayers, the audit function $p(x|z)$ is decreasing and continuous in reported income. The reporting function, $x(u|z)$ is strictly increasing in an upper region of the income domain and constant in a lower region as some taxpayers pool at the lowest possible report. As the audit and reporting functions are continuous and differentiable on the interior of the reporting domain, it is possible to solve for the equilibrium using methods of differential equations. In addition, as pooling occurs only at the lowest report, where the differential equation is undefined, sufficient conditions for equilibrium can be obtained by checking that the solution to the differential equation

[11] The second order condition is $\frac{\partial \mathrm{E}(y|x,z)}{\partial p(x|z)} \geq 0$. In our simulations the solutions always satisfy this criterion.

also satisfies the tax agency's first order condition for the lowest report, equivalent to (5) below. In the same way, we can leverage these properties to solve for the population-wide equilibrium as a range of within-audit-group equilibria coupled with the optimal budget distribution, $B(z)$.

The unique revenue maximizing equilibrium of the model is described by the collection of functions, $u(x|z)$ and $p(x|z)$, and the budget distribution, $B(z)$. Once $p(x|z)$ is determined, $u(x|z)$ is implicitly defined as the solution to the taxpayers' first order condition, and the tax agency chooses $p(x|z)$ such that (3) holds with equality. The two equations are connected by the tax agency's conditional expectation of taxpayers' true income given the reported income and third-party reports, $\mathrm{E}(y|x,z)$, which is

$$\mathrm{E}(y|x,z) = z + \frac{f_{u|z}^{h}(x)\,x + f_{u|z}^{d}(u(x|z))\,\frac{\partial u(x|z)}{\partial x}u(x)}{f_{u|z}^{h}(x) + f_{u|z}^{d}(u(x|z))\,\frac{\partial u(x|z)}{\partial x}}, \tag{4}$$

where the derivative $\frac{\partial u(x|z)}{\partial x}$ is derived from (1) by differentiating implicitly to get $\frac{\partial u}{\partial x} = 2 + \frac{p''(x)(x-u)}{p'(x)}$.[12]

We can then derive a second order differential equation, (A.1) in the Appendix, which determines the optimal equilibrium responses $p(x|z)$ and $x(u|z)$ in audit group z using the expressions for $\mathrm{E}(y|x,z)$, $u(x|z)$, $\frac{\partial u}{\partial x}$, and the tax agency's first order condition. However, as some taxpayers pool at the lowest report, to obtain sufficient conditions for equilibrium, we must check the tax agency's first order condition at $x = \underline{u}$ separately as

$$\mathrm{E}(u|x=\underline{u},z) = \frac{f_{u|z}^{h}(x)\,x + \int_{\underline{u}}^{u^{\text{pool}}} u \cdot f_{u|z}^{d}(u)\,\mathrm{d}u}{f_{u|z}^{h}(x) + \int_{\underline{u}}^{u^{\text{pool}}} f_{u|z}^{d}(u)\,\mathrm{d}u} = \frac{\lambda(z)\,c}{t+\theta t} + \underline{u}, \tag{5}$$

where u^{pool} is the residual income at which taxpayers (in this audit group) begin to pool at the lowest possible report.

Thus, given the equilibrium $\lambda(z)$, we can characterize the unique within-group equilib-

[12] Notice that $f_{x|z}(x(u)) = f_{u|z}(u(x))\left|\frac{\partial u(x,z)}{\partial x}\right| = f_{u|z}(u(x))\,\frac{\partial u(x,z)}{\partial x}$ since the SOC implies that $\frac{\partial u}{\partial x} \geq 0$ in interior optimum.

rium from Equations (5) and (A.1). By Equation (2), each $\lambda(z)$ corresponds to a required budget allocation, $B(z)$. Finally, the budget allocation across different z is pinned down by the requirement that the shadow value of increasing the budget, $\lambda(z)$, must be the same for all z, i.e., $\lambda(z) = \boldsymbol{\lambda}$, $\forall z$, for an interior solution. The shadow value, $\boldsymbol{\lambda}$, is pinned down by the requirement that the tax agency's overall budget, $\mathbf{B}$, may not be exceeded.

As mentioned above, the model contains Erard and Feinstein (1994) as a special case when attention is limited to a single audit group in which taxpayers without third-party income reports and the ratio of honest to dishonest taxpayers is constant on the domain of u. To illustrate, Figure 3 depicts the equilibrium for B at 10 percent, $\log(u) \sim \mathcal{N}(3.42, 0.3^2)$ truncated on $[20, 44]$, $Q = 0.4$, and $t = 0.5$.

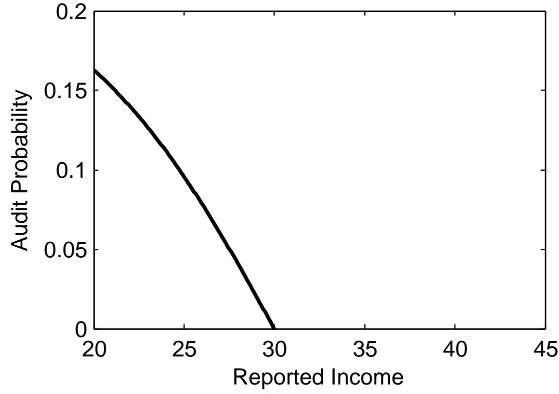
Figure 3(a) shows the audit schedule, $p(x)$: it starts in $\underline{u}$, is downward sloping, and terminates in $p(\bar{x}) = 0$. This form balances the need to audit in order to raise revenue with the cost of doing so. The negative slope reflects the need to discourage high-income taxpayers from reporting too low incomes.

Figure 3(b) shows the amount of evasion as a function of true income. The linear increase in the first part of the graph reflects pooling of dishonest taxpayers: for a given audit schedule, there will be some level of residual income, u^{pool} in $[\underline{u}, \overline{u}]$, for which the most profitable report is $\underline{u}$. Consequently, all taxpayers with residual incomes $u < u^{\text{pool}}$ also report $x = \underline{u}$. Therefore, there will be a point mass in the induced distribution of reports, $f_x(x)$. After this pooling point, evasion falls rapidly in income until evasion again becomes increasing in income as the probability of detection becomes sufficiently low.

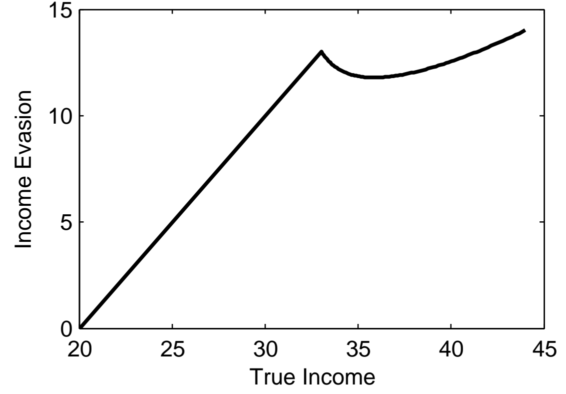
Figure 3(c) shows the effect of the optimal audit schedule on the *ex ante* effective tax rate, τ^{eff}, which is calculated as the ratio of expected payments (taxes and penalties) to true income

$$\tau^{\text{eff}} = \frac{p(x) \cdot (ty + \theta t(y - \tilde{y})) + (1 - p(x)) \cdot t\tilde{y}}{y}. \tag{6}$$

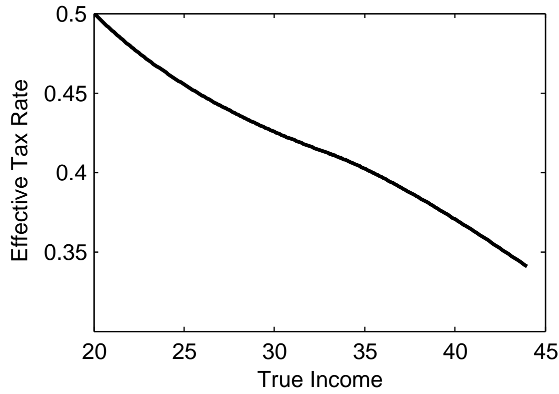
The declining profile of $p(x)$ together with the high propensity to evade taxes of high income taxpayers result in a negative relationship between the effective tax rate and income. There-



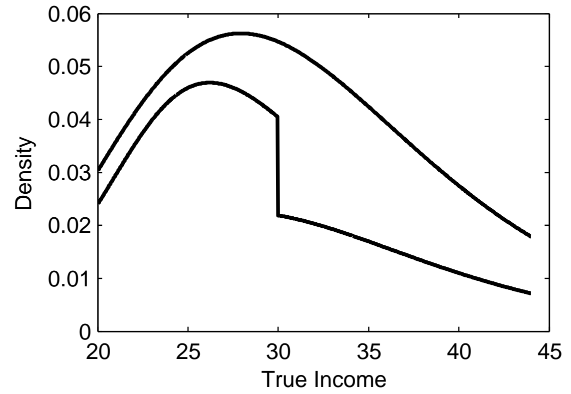
(a) The Optimal Audit Schedule, $p(x)$.



(b) Evaded Income, $u - x$, by True Income, y, for Dishonest Taxpayers.



(c) Regressive Bias, $\tau - \tau^{\text{eff}}$, for Dishonest Taxpayers.



(d) Induced Reporting Behaviour. The lower curve graphs the density of reports by dishonest taxpayers, excluding the mass point at $x = \underline{u}$, while the upper curve graphs the true income distribution.

Figure 3. Equilibrium Responses and Tax Bias.
Notes: All panels display an example of equilibrium functions from the Erard and Feinstein (1994) model without third-party reporting. Equivalently, this could be an example of the solution for a particular z in our model including third-party reporting. This example is produced assuming $B = 10$ percent, $\log(u) \sim \mathcal{N}\left(3.42, 0.3^2\right)$ truncated on $[20, 44]$, $Q = 0.4$, and $t = 0.5$.

fore, high-income taxpayers pay significantly less than the statutory tax rate, which, in the case of Figure 3(c), is $t = 0.5$, and we get regressively biased effective average tax rates.

Figure 3(d) shows the induced distribution of incomes and reports. The top graph is the original income distribution, which in this case is lognormal. The lower graph shows the distribution of induced reports, i.e., the equilibrium response of all taxpayers to the audit schedule. The right part of the graph is just a scaling of the original income distribution by Q while the left part is a weighted average of reports by honest and dishonest taxpayers. The whole graph is somewhat lower than the original income distribution as there is a mass point of dishonest taxpayers reporting at $\underline{u}$, the mass point being equal to the area between the graphs.

3 Data

SKAT's tax collection efforts extensively employ information reports by third parties. During some year t, incomes are earned and by the end of January in year $t + 1$, SKAT receives information reports from employers, banks, pension funds, and other entities, so-called third-party income reports. In general, all income received as salary, private/public pensions, honorarium, unemployment benefits, etc. is subject to third-party reporting as well as, e.g., mortgage interest payments and some capital income.[13] Self-employment income is rarely covered by information reporting except in cases where, e.g., remuneration is paid by a public institution. Third parties do not have discretion as to whether or not to supply SKAT with this information. The informational requirement is entirely related to the type of income.

By mid-March, SKAT sends out pre-populated tax returns based on third-party information and other information that they possess about the taxpayers, such as their residence and workplace for calculating commuting allowances. Subsequently, taxpayers have until May 1 to correct their tax return; in case of no corrections, the pre-populated tax return counts as final. After the deadline, SKAT's computerized system processes tax returns and attaches

[13]Dividends are reported by third parties, whereas capital gains were not reported in 2006/2007.

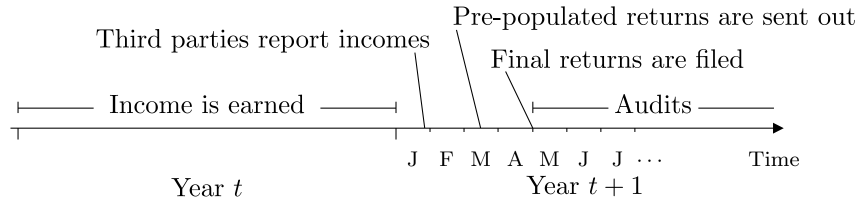


Figure 4. Tax Collection in Denmark – The Timing of Events.

audit flags to returns that the system finds likely to contain errors. The system does not as such assign a probability of audit or rank tax returns according to their likelihood of containing errors but assigns a recommended action, i.e., "audit" or "do not audit". Briefly, the audit flag system relies on third-party income reports and also a collection of auditing "best practices" that could be converted to algorithmic form, e.g., specific tax return compositions indicative of misreporting, cut-off rules based on expected incomes conditional on third-party reported income, etc. The flag system consists of a large number of flags, each of which is intended to signal the likelihood of tax evasion on particular line-items or combinations of line-items. Although, the flag system operates for both wage earners and the self-employed, in practice, it is only used for wage earners as the predictive power of the audit flags for self-employed has been judged too low by SKAT. For the self-employed, further information is gathered on a case-by-case basis. Predominantly, SKAT uses correlates of true income such as bank deposits, consumption of housing, cars, and other durables to signal of the likelihood non-reported income. They may also seek information exchange with known tax shelters about foreign deposits or uncover such deposits indirectly by tracking purchases with foreign credit cards, but such information is much harder to gather. All in all, the workings of the audit regime is very different for the self-employed and much more resource intensive.

After the tax returns have been processed, tax examiners assess the flagged returns and decide whether or not to initiate an audit based on the information available, local knowledge, and auditing resources. For wage earners, the information available is processed via the flag system and for the self-employed external information is gathered on an ad hoc basis. The process is depicted in Figure 4.

If an audit discovers underreporting, the taxpayer may pay the taxes owed immediately or postpone the payment at an interest. If the tax examiner views the underreporting as deliberate, the tax agency may impose a fine according to a fining scheme depending on the assessed intentionality of the misreporting.

3.1 Experimental Design

The data originates from an experiment conducted by SKAT in the years 2006–2008, originally analyzed in Kleven et al. (2011), and is in many ways comparable to the US Taxpayer Compliance Measurement Program. The experiment involved a stratified random sample of 17,764 self-employed individuals and 25,020 wage earners and recipients of public transfers in Denmark. In the present study, we use a sample of non-treated wage earners and recipients of public transfers (referred to as "wage earners") and a sample of non-treated self-employed for the fiscal year 2006.[14] The sample of wage earners is a stratified random sample of 10,740 Danish taxpayers, and the sample of the self-employed is a random sample (non-stratified) of 8,890 taxpayers.[15] The full populations of wage earners and self-employed, respectively, where approximately 4.2 million and 400,000 in 2006. For each taxpayer, SKAT conducted an unannounced audit after the deadline for changing the tax return (May 1, 2007). The tax audits were comprehensive in the sense that SKAT examined all items on the tax return, demanding documentation for all items on which SKAT did not possess information. Moreover, SKAT made a significant effort to have tax examiners perform homogeneous audits by, e.g., organizing training workshops and distributing detailed audit manuals. The audits took up 21 percent of the resources devoted to tax audits in 2007.

Of course, it is unlikely that tax examiners find all hidden income, such as that stemming

[14] In the original study in Kleven et al. (2011), some taxpayers were subject to treatments. These taxpayers received notifications prior to filing their final tax returns, indicating that they would be audited with either 50 or 100 percent probability.

[15] Note the randomness of our sample as opposed to tax compliance data obtained from the regular audits that is heavily biased by over-sampling taxpayers who are likely to have misreported their income in either direction. The sampling strategy for wage earners involved a stratification on tax return complexity. For the self-employed no stratification scheme was employed.

from cash-only businesses and other black market activities. We focus our attention on the detectable part of tax evasion given the methods available to SKAT and thus denote our empirical counterpart of true income "detectable income". In what follows, we will write true income when in fact we mean detectable income. In Section 4.3.4 we discuss the implications of this for our results.

For each taxpayer, we have income and tax records as reported by third parties, the final return as potentially changed by the taxpayer, and the post-audit return. In addition, the data contains information on the generated audit flags that would normally constitute a basis for selecting taxpayers for audits as well as a "compliance rating" reflecting the auditor's assessment of the degree to which discovered misreporting reflected deliberate fraud or accidental under/over-reporting.

3.2 The Tax System and Tax Compliance in Denmark

The Danish income tax system (in 2006) operates with many different measures of income. Here, we will provide the headlines; see Table 1 for details. Labor market income, i.e., salary, fringe benefits and other earned income, are taxed proportionally by a labor market tax of 8 percent while an earned income tax credit (EITC) of 2.5 percent is provided for labor market income up to 292,000 DKK.[16] Capital income is a net concept, and different tax rates apply depending on whether net capital income is positive or negative. For most taxpayers, net capital income is negative due to interest payments on mortgages. Central government taxes (bottom, middle and top tax) are levied on the so-called "personal income", which, in addition to positive net capital income, consists of labor market income plus social transfers and pensions less labor market taxes and some pension contributions. Central government taxes constitute a progressive tax scheme with a personal allowance and three brackets. Local taxes (county and municipality) are levied on "taxable income", which is similar to the central government tax base except that it allows for negative net capital income deductions

[16]Approx. 49,000 USD (1 USD $\approx$ 6 DKK in 2006).

Table 1. An Overview of the Danish Tax System, 2006.

Tax	Tax base	Bracket (DKK)[a]	Rate (pct.)
Labor market tax	Labor inc.	none	8.0
EITC	Labor inc.	up to 292,000	2.5
Bottom tax	Personal inc.+ max(cap.inc., 0)	38,500–	5.5
Middle tax	— // —	265,500–	6.0
Top tax	— // —	318,700–	15.0[b]
Local taxes	Taxable inc. (=pers.inc.+cap.inc.−deductions)	38,500–	33.3[c]
Stock income tax	Stock inc.	0–44,300; 44,300–	28.0; 43.0

[a]1 USD ≈ 6 DKK (in 2006).
[b]The top tax rate may be lowered by the "tax ceiling" that limits the sum of state taxes (bottom, middle and top) and local taxes (excl. church taxes) to 59 percent. In the average municipality the tax ceiling lowers the top rate by 0.08 percentage points.
[c]In the avg. municipality and county incl. optional church tax of on avg. 0.74.

and other deductions such as transport allowances. In this way, Denmark has a version of the Nordic dual income tax;[17] negative capital income is taxed at a flat rate, whereas positive capital income is taxed progressively just as regular income. Stock income (dividends and capital gains) is subject to a two-rate scheme with the high rate setting in at 44,300 DKK.

Table 2 presents some descriptive statistics on major income components for the two samples of wage earners and self-employed, respectively. The table shows sample means with standard errors of means in parentheses – all numbers for wage earners are calculated accounting for the stratification scheme. Column (1) presents pre-audit figures measured at the deadline, May 1, and column (5) shows figures reported by third-parties. Self-reported figures (the difference between (1) and (5)) are shown in column (6). Negative figures mean that taxpayers on average adjust the number downwards to less than what third-parties have reported. Columns (2)–(4) describe how the figures in (1) were adjusted by the tax examiners during the audits. Columns (3) and (4) split the audit adjustments into positive (meaning underreporting) and negative (meaning overreporting) adjustments while column (2) holds the average net adjustment, i.e., the sum of (3) and (4).

Panel A of Table 2 shows figures on total income and total taxes for wage earners. The former is defined as the sum of personal income, capital income, stock income, self-

[17]For a discussion of the Nordic dual income tax., see e.g. Nielsen and Sørensen (1997).

Table 2. Tax Compliance in Denmark, Income Year 2006.

	Reported income	Net audit adjustment	Under-reporting	Over-reporting	Third-party rep. inc.	Self-rep. inc.
	(1)	(2)	(3)	(4)	(5)	(6)
A.			Wage earners			
Total Income	193,277	1,664	1,825	-161	195,618	-2,341
	(1,906)	(480)	(479)	(22)	(1,844)	(584)
Total Tax	63,178	636	695	-59		
	(841)	(246)	(246)	(9)		
B.			Income components			
Earnings	156,127	672	683	-11	155,987	140
	(2,275)	(203)	(203)	(6)	(2,217)	(559)
Personal inc.	209,232	1,137	1,195	-58	209,726	-494
	(1,950)	(480)	(479)	(17)	(1,886)	(573)
Capital inc.	-10,884	142	198	-56	-11,308	424
	(272)	(27)	(24)	(11)	(266)	(81)
Deductions	-9,264	143	213	-70	-5,605	-3,659
	(178)	(28)	(26)	(11)	(85)	(144)
Stock inc.	3,612	239	262	-24	2,797	815
	(546)	(40)	(39)	(10)	(502)	(188)
Self-empl. inc.	103	21	23	-2	8	95
	(60)	(8)	(8)	(1)	(4)	(60)
Foreign inc.	479	-18	6	-25	0	479
	(92)	(19)	(4)	(19)	.	(92)
C.			Self-employed			
Total Income	298,388	21,480	22,697	-1,217	157,285	141,103
	(8,321)	(1,912)	(1,905)	(145)	(6,445)	(5,534)
Total Tax	124,392	8,719	9,089	-371		
	(4,423)	(609)	(606)	(50)		
D.			Wage earners and self-employed			
Total Income	202,310	3,367	3,619	-252	192,324	9,987
	(1,883)	(469)	(467)	(24)	(1,774)	(715)
Total Tax	68,439	1,331	1,416	-86		
	(858)	(231)	(231)	(9)		

Notes: Panels A and B show descriptive statistics for a stratified random sample of 10,740 taxpayers denoted as wage earners (incl. unemployed, pensioners, etc.). Due to the stratification strategy employed by SKAT, the sample contains 74.6 percent "heavy" taxpayers (i.e., with high-complexity tax returns) and 25.4 percent "light" taxpayers, whereas the population has 32.6 percent heavy taxpayers and 67.4 percent light taxpayers. In Panel C the sample consists of 8,890 randomly selected self-employed taxpayers. No stratification was employed. Panel D provides descriptive statistics for wage earners and self-employed combined using population weights.
Total income is defined as personal income + capital income – deductions + stock income + self-employment income + foreign income. The decomposition in Panel B is only available for the sample of wage earners. In the table, deductions are given as a negative amount. Reported income is the sum of third-party reported income and self-reported income. Standard errors of means in parentheses. All estimates for wage earners are population weighted.
All amounts in DKK (1 USD $\approx$ 6 DKK in 2006).

employment income, and foreign income less deductions. Pre-audit total income is on average a little less than 200,000 DKK with a significantly positive net adjustment from SKAT of almost 1,700 DKK. The positive net adjustment reflects an asymmetry in the reporting behavior with underreporting being more than ten times as high as the overreporting on average. Third-party reported total income is slightly higher than pre-audit total income mainly due to deductions not included in the third-party reports, implying a negative residual (i.e., self-reported) total income.

Panel B features a decomposition into main income components for wage earners. The asymmetry in the over- and underreporting found for total income is noticeable for all components.[18] Not surprisingly, the greatest relative amount of underreporting is found on items least subject to information reporting. Self-employment income tops the list with underreporting amounting to 18.5 percent of the mean post-audit self-employment income level followed by stock income (6.8 percent), deductions (2.3 percent), and the rest being less than 2 percent.

In Panel C we show descriptive statistics for the sample of self-employed taxpayers. As a decomposition into income components has not been possible, we only show numbers for total income and total taxes.[19] As with wage earners, we find a pronounced asymmetry in net audit adjustments corresponding to much higher underreporting compared to overreporting for the self-employed. The main difference compared to wage earners is spelled out in the average level of self-reported income. Income sources of self-employed are to a much lesser extent covered by the system of third-party reporting, resulting in an almost even split between income reported by third parties and self-reported income. This provides SKAT with a much greater challenge in discovering unreported income.

[18]Foreign income is the exception. Here, the average net adjustment is negative, corresponding to overreporting on average, yet, the adjustment is not significantly different from zero. The likely reason is that there are few cases of foreign income, and the variation in adjustments performed by SKAT is dominated by correction of mistakes.

[19]During the experiment, tax corrections concerning the interplay of the business and private side for the self-employed where not included in the data. This was remedied at the aggregate level for total income and total taxes but not for separate income components. Therefore, for the self-employed we only include descriptive statistics for aggregate income and tax measures.

We get a further idea as to where the opportunities to evade taxes are prevalent by looking at taxpayers' behavior and conditioning on the informational environment. In Table 3 we separate taxpayers according to whether or not their entire income was reported to the tax agency by a third party. Panel A shows the shares of under-/overreporting and correct reports for each sample (wage earners and self-employed, respectively). All figures in the table are calculated accounting for stratification whenever applicable. The overall population weighted share of compliers, given by wage earners not underreporting, amounts to approximately 94 percent for wage earners. For the self-employed, approximately 65 percent comply. To address taxpayers with ample opportunity to evade taxes, Panel B shows shares of particular groups conditional on whether or not their entire income is reported by a third-party (standard errors in parentheses). For example, less than 2 percent of wage earners with all income reported by third parties underreport taxes. For wage earners with some income not reported by third parties, this share is much higher, yet a substantial share of over 80 percent (depending on the definition of compliance) are found to comply with the tax laws despite having ample opportunity to evade.

Only few self-employed taxpayers (3.4 percent) have their entire income reported by third parties, underlining the tax agency's challenge in securing tax revenue from these taxpayers. Further, almost 35 percent are found to underreport their taxes. The share of self-employed who do not underreport their taxes is again high (95 percent) for those with all income reported by third-parties and much lower (64 percent) for those with some income not covered by the system of third-party reporting, albeit still a substantial share comply with given tax laws. Strikingly, wage earners and self-employed who have all their income reported by third parties do not differ significantly in reporting behavior, whereas this is not the case when some income is not reported by third parties.

Table 3. Reporting Behavior of Danish Wage Earners and the Self-Employed, 2006 Incomes.

	Wage earners		Self-employed	
Observations	10,740		8,890	
Entire income reported by third-parties?	Yes	No	Yes	No
A.	Share	Share	Share	Share
# underreported	0.010	0.049	0.002	0.346
# correct	0.653	0.269	0.032	0.570
# overreported	0.003	0.016	0.000	0.048
Total reports	0.665	0.335	0.034	0.966
B.	Share of sub-sample	Share of sub-sample	Share of sub-sample	Share of sub-sample
Correct reports	0.979	0.809	0.943	0.590
	(0.002)	(0.011)	(0.055)	(0.006)
Not underreporting	0.984	0.855	0.950	0.640
	(0.002)	(0.010)	(0.055)	(0.007)
"Honest" taxpayers*	0.988	0.901	0.957	0.690
	(0.002)	(0.008)	(0.055)	(0.007)

Notes: Standard errors of fractions in parentheses. The sample of wage earners is a stratified random sample. Fractions and standard errors are calculated subject to the stratification scheme. "Wage earners" also include recipients of benefits. The sample of self-employed is a non-stratified random sample.
*Calculated imposing the assumption that unintentional underreporting is as frequent as (unintentional) overreporting – i.e., symmetry in reporting errors. For example, for the self-employed (right-most column), the (unstratified) calculation is simply $(0.570 + 2 \cdot 0.048)/0.966 \approx 0.690$. For wage earners, we provide a population weighted estimate.

3.3 Effective Tax Rates

To address the distortion of tax evasion/enforcement of the effective tax system, we need a measure of *ex ante* effective average tax rates calculated directly from data. As with Equation (6), an appropriate measure of effective average tax rates must take into account the extent of tax evasion, the risk of detection, and the penalties paid in case of detection, all of which affect expected payments to the tax agency. As such, we must restrict our analysis of this phenomenon to the sample of wage earners for whom we have information about the enforcement regime. The *ex ante* effective average tax rate can be calculated from data as

$$\tau^{\text{eff}} = \frac{f \cdot \left(T + \Theta\left(T - \tilde{T}, I\right)\right) + (1-f) \cdot \tilde{T}}{Y}, \tag{7}$$

where f is the probability of getting caught, T and $\tilde{T}$ are taxes on true and reported income, respectively, Y is true income, and $\Theta(\cdot,\cdot)$ is a nonlinear function describing the penalty for underreporting taxes as a function of underreported taxes and the assessed intentionality of evasion, I. With a probability f, evasion is detected and the taxpayer pays the full taxes due plus a penalty that is proportional to the amount of taxes evaded. With a probability $1-f$, evasion goes undetected and the taxpayer only pays taxes on reported income. As tax evasion diminishes or as the risk of detection or the penalty increase, the effective average tax rate will increase, *ceteris paribus*.

We denote by τ the nominal average tax rate, defined in the usual way, $\tau = T/Y$. As a matter of convenience, we define the tax rate bias as $\tau - \tau^{\text{eff}}$. This allows us to compare how much statutory and effective tax rates differ when both vary across individuals in the sample. Although we focus on individuals for which the data reveals some underreporting, (7) is equally valid for taxpayers not engaging in tax evasion. For these individuals, taxes due on reported and actual income are the same, and the effective average tax rate is simply the average tax rate, $\tau^{\text{eff}} = \tau$.

Y, T, and $\tilde{T}$ are observed in the data as post-audit total income and taxes, and pre-audit taxes.[20] We use SKAT's audit flag system as a proxy for the probability of getting caught for wage earners. Not all taxpayers with flags are audited, so we assume that the probability is proportional to the number of flags assigned to a tax return.[21] Specifically, we calculate our proxy for the probability of detection simply as the ratio of flags assigned to a tax return to the maximal number of flags assigned to any tax return. With this approach, the audit rate among wage earners is 3.3 percent. This is slightly lower than the total population audit rate of 4.2 percent reported by Kleven et al. (2011). As this rate includes audits of the self-employed, who, presumably, are audited relatively more intensively, the average audit rate suggested by our proxy seems more or less reasonable.

[20] Recall the definition of total income as the sum of personal income, capital income, stock income, self-employment income, and foreign income less deductions.

[21] Alternatively, this can also be interpreted as an assumption that each part of the tax return, to which an audit flag corresponds, is audited with probability 1.

We specify the penalty function, $\Theta\left(T - \tilde{T}, I\right)$, using the actual rules for calculating penalties for tax evasion and the compliance rating system applied by the tax examiners during the audits. In Denmark, evasion penalties are calculated as a factor on taxes evaded; that factor, however, varies for the amount evaded and the intentionality of evasion as assessed by the auditor. In the case of intentional tax evasion, the fine is calculated as 1 times evaded taxes under 30,000 DKK and 2 times the evaded taxes exceeding 30,000 DKK. In the case of gross negligence, the rates are instead 0.5 times evaded taxes not exceeding 30,000 DKK and 1 times evaded taxes exceeding 30,000 DKK. Fortunately, the compliance ratings in the data are exactly intended to measure the degree of intentionality of uncovered tax evasion. Compliance ratings take on values in $\{0, 1, 2, \ldots, 6\}$ indicating decreasing degrees of intentionality of misreporting. According to this classification, compliance ratings of 0, 1, or 2 signify deliberate tax evasion, whereas $3, \ldots, 6$ signify gross negligence (approaching 3) or innocent mistakes (approaching 6). Using these classifications, we can accurately calculate the penalty rate applicable for each individual tax evader.[22]

4 Calibration and Results

Due to the considerable detail of our data, we can construct a set of parameters for the purpose of simulating the model that are internally consistent, i.e., they all derive from the same data set. Using the samples of wage earners and the self-employed we approximate penalty and tax rates from the actual tax system. The parametrized share of honest taxpayers we allow to differ between the groups of wage earners and the self-employed to account, in some measure, for self-selection into these employment categories. For the same reason, we also estimate the bivariate income distributions separately for the two groups. As we calculate below, the share of honest taxpayers is indeed much lower for the self-employed

[22] Assuming, e.g., that innocent mistakes (rated 6) are not penalized or that the threshold in compliance ratings between intentional evasion and gross negligence is between 1 and 2 or between 3 and 4, turns out not to affect the results we present in Section 4.

corresponding to the intuition that some people may self-select to exploit more ample evasion opportunities. Finally, for each group, we calibrate the audit budget to match simulated average tax evasion among evaders to observed average evasion. Without loss of generality, we can normalize the per-audit cost, c, to 1 such that overall budget parameters $\mathbf{B}$ can be interpreted as the share of the population subject to audit within the groups of wage earners and the self-employed.

4.1 Calibration

4.1.1 Income Distributions

We use the taxpayer data to construct the income distributions needed in the model. As income measure we use total income defined as the sum of personal income, capital income, stock income, self-employment income, and foreign income less deductions.

In principle, the densities of honest and dishonest taxpayers can be estimated separately but with the size of our data set this would introduce a large element of uncertainty in estimates of dishonest taxpayers. Instead, we follow Erard and Feinstein (1994) closely and estimate a common distribution for both honest and dishonest taxpayers with the only difference being a mass point of honest wage earners for whom true income is perfectly predicted by third-party reported income, i.e., $u = 0$, which allows this mass point to vary in z. This is important because richer wage earners are much more likely to have non-zero residual income than poorer wage earners. However, for the self-employed there are very few individuals without some residual income and we can estimate income distributions without accounting for a mass point. In practice, to fit the simultaneous distribution of z and u, we exclude any honest taxpayers in $u = 0$ and fit a mixed lognormal distribution.[23] The distribution of the mass point of wage earners at $u = 0$ across z is estimated separately.

[23] Our results do not appear to alter significantly if, instead, a kernel estimation is used. However, kernel densities are inconvenient as they allow for "troughs" of zero density in the interior of $[\underline{u}, \overline{u}]$ which may cause our algorithm to fail. By using a sufficient number of component distributions in the mixed lognormal distribution, the difference between this distribution and a bivariate kernel distribution becomes negligible.

The exact characteristics of this distribution is documented in the Appendix. Briefly, the variance of $u|z$ is generally increasing in z; however, the taxpayers with very low or negative z have relatively complicated income compositions resulting in high variance of $u|z$ and, for wage earners, a relatively small mass point at $u = 0$.

4.1.2 Honesty

With our simplified version of conditional densities, we can write $f^h_{u|z} = Qf_{u|z}(u) + \mathbf{1}_{(u=0)}M(z)$ and $f^d_{u|z} = (1-Q)f_{u|z}(u)$, where $\mathbf{1}_{(\cdot)}$ is the indicator function. Thus, for $u \neq 0$ the share of honest taxpayers is Q, whereas for $u = 0$ it is $Q + M(z)$, where $M(z) \geq 0$ is the mass point at $u = 0$ for some level of third-party reporting, z. To determine an appropriate value of the parameter, Q, we must account for the fact that, in reality, some taxpayers seem to make reporting mistakes. For example, in the data some reports are adjusted downward by the auditor, which means that, in the absense of an audit, the taxpayer would have payed more than intended by the statutory tax system.

We approach the problem in the following way. First, we assume that no taxpayer will try to evade taxes on income that is reported by a third party (this assumption is bourne out in the data for wage earners as shown in Table 3). Secondly, in keeping with the model, we disregard the fact that some taxpayers make reporting mistakes. A revenue maximizing tax agency is indifferent about the motivation for underreporting and about overreporting.[24] As a consequence, taxpayers reporting too large taxable incomes are treated as if they are exactly compliant and taxpayers that underreport taxable incomes by mistake are treated as tax evaders. Then we separate taxpayers by whether they underreported taxes (non-compliant taxpayers, $x < u$) or reported correctly/overreported taxes (compliant taxpayers, $x \geq u$). Compliant taxpayers are then decomposed into those with zero residual income and non-zero residual income. We define the parameter Q as the ratio of compliant taxpayers with non-zero residual income to the total number of taxpayers with non-zero residual income in the sample.

[24]We do not consider the, rather implausible, scenario that the tax agency might refrain from auditing certain groups because this would reveal overreporting by some taxpayers thus lowering collected revenue.

The idea is that having some income not subject to third-party reporting provides taxpayers with ample opportunity for evasion. By not seizing the opportunity, they reveal themselves as being honest in the present context. Table 3 shows this decomposition. First, note that among wage earners whose entire income is reported by third parties, the compliance rate is 97.9 percent. Among those wage earners that have some of their income not reported by third parties, the compliance rate is 80.9 percent. The number of honest taxpayers is the sum of those reporting correctly and those overreporting by mistake, which corresponds to $Q = 85.5$ percent.[25] The residual consists of both dishonest wage earners and wage earners underreporting by mistake whom we cannot distinguish. To partially control for self-selection into occupations according to a taxpayer's proclivity to evade taxes, we calculate Q separately for the self-employed as shown in Table 3. The resulting value, $Q = 64.0$ percent, is indeed substantially lower and suggests that this distinction is important.

4.1.3 Penalty

The model has a fixed penalty factor, θ, as opposed to the more complicated penalty function, $\Theta(\cdot, \cdot)$, from Section 3. We approximate an appropriate value of θ by calculating the average penalty rate for the sample of tax evaders accounting for stratification between light and heavy taxpayers within the group of wage earners and for the relative shares of wage earners and self-employed in the population. We take a simple approach and use the OLS slope coefficient between calculated penalties, $\Theta(\cdot, \cdot)$, and underreported taxes as our value of θ. The resulting penalty rate on underreported taxes is 1.15.

4.1.4 Tax Rates

We estimate a marginal tax function, $t(z)$, using local mean smoothing of marginal tax rates on the entire sample of wage earners and self-employed accounting for stratification of light and heavy taxpayers in the group of wage earners and the relative shares of wage earners

[25] Of course, we also account the for the sample stratification in calculating Q.

and self-employed in the population. We allow the approximated tax rates to vary in z to partially account for the progressiveness of Danish income taxes. Because our data set contains all line items, we can calculate each taxpayer's marginal tax rate on all income components, such as earnings, capital income, stock income, etc. For each component, we calculate marginal taxes with respect to reported income. To obtain an average marginal effective tax rate, conditional on z, we then weight marginal taxes of different components according their relative prominence on a taxpayers tax return (i.e., before the taxpayer is audited).

4.2 Simulation Strategy

An individual solution, $\left(p, \frac{\partial p}{\partial x}\right)$, to Equation (A.1) in the Appendix that corresponds to a particular z is found numerically using methods of Ordinary Differential Equations (ODE). The solver is initialized using $p(\bar{x}) = 0$ and $p'(\bar{x}) = \left(\frac{1}{1+\theta}\right) / (\bar{u} - \bar{x})$, where $\bar{x} \equiv x(\bar{u})$. Thus, starting at the end-point of the equilibrium-path audit probabilities, a numerical solver finds values in steps until $\underline{u}$ is reached, ensuring that the taxpayers' as well as the tax agency's optimality conditions are met for reports $x \in (\underline{u}, \bar{x}]$. However, since a positive mass of taxpayers are pooling their reports at $x = \underline{u}$, the expectation $\mathrm{E}(u|x, z)$ is not differentiable in this point. Therefore, we check that the tax agency's FOC is met in the pooling point separately after finding some candidate solution, cf. (5).

The difficulty in identifying equilibria in this model stems from *a priori* indetermination of λ and $\bar{x}$: we must satisfy $\mathrm{E}(u|x = \underline{u}, z) - \underline{u} = \frac{\lambda c}{t+\theta t}$, which depends on both variables. Our solution method, the so-called shooting method for parametrized ODEs, searches the space of possible $(\lambda, \bar{x})$ for candidate solutions, for each checking that the tax agency's optimization constraints are satisfied on the entire domain of x, until satisfactory solutions are found. The optimal budget allocation, which in our simulations is always interior, equates shadow prices of increasing the budget density across levels of z.

While mathematically and intuitively z is naturally understood to be a continuous vari-

able described by the simultaneous distribution of u and z, we approximate the optimal allocation of the total audit budget on the domain of z by constructing a representative, evenly spaced grid. We provide detailed documentation of the numerical implementation in the Appendix.

We have estimated $t(z)$, θ, Q, and the income distribution from data. Thus, the remaining free parameter is the budget value, $\mathbf{B}$, which we do not know. Since the mean level of evasion is inversely proportional to total tax revenue, it is monotonically declining in $\mathbf{B}$. To calibrate $\mathbf{B}$, we use the estimated income distribution to simulate a population of taxpayers: we vary $\mathbf{B}$ until the average level of evasion for tax evaders matches the level observed in the data, approximately DKK 8,312 for wage earners and DKK 25,991 for the self-employed. The resulting budget values are $\mathbf{B} = 0.0412$ and $\mathbf{B} = 0.4565$, respectively.

4.3 Results

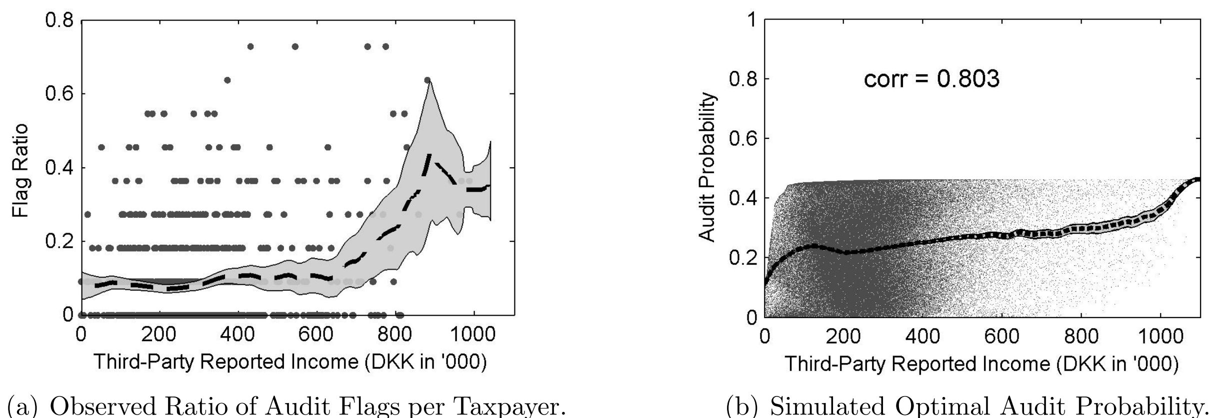
As mentioned, we calibrate the model to the average level of evasion among tax evaders in the data. The match between data and simulations may seem trivial as it is imposed by the calibration procedure. However, in the context of the economic literature on tax evasion, being able to match a structural model to moments of the data for reasonable parameter values is, to our knowledge, novel. For example, Alm, McClelland, and Schulze (1992) argue that observed evasion is too low to be explained by a model of actual audit and penalty regimes. Our analysis lends support to the argument of Andreoni, Erard, and Feinstein (1998) and Slemrod (2007) that third-party reporting and tax-return-dependent audits can explain a substantial part of observed evasion. However, in accordance with Feld and Frey (2002), our analysis also requires us to take into account the substantial number of taxpayers that report honestly despite incentives to evade.

4.3.1 Tax Evasion and Enforcement

In Figure 5 we compare the observed distribution of flags across third-party reported income with the optimal distribution obtained in the simulations. In panel (a) we show, for each individual, third-party reported income (in '000 DKK) and the ratio of flags to the maximally observed number of flags assigned to any return. In addition, we show the local average ratios and 95 percent confidence bounds using local mean smoothing. In panel (b) we show individual and average observations of audit probabilities from simulated data. Generally, the audit intensity is increasing in third-party reported income, reflecting the fact that higher-income taxpayers find it relatively easier to evade taxes since the conditional variance of true residual income is larger. As we do not know how the number of flags assigned to tax returns translates into the likelihood of being audited, it is not surprising that there is a level difference between the two graphs. This reflects the fact that our minimal assumptions proxy for the empirical audit probability suggests an audit rate among wage earners of 3.3 percent, whereas the audit rate required to calibrate the model is 4.1 percent. Nonetheless, the graphs have very similar profiles. Both are increasing in third-party reported income and the audit intensity is especially high in the right tail of the distribution. This is bourne out in a correlation coefficient between local averages of 0.803.

Similarly, Figure 6 shows the empirical and simulated covariation of reported residual incomes (x), denoted in DKK in '000, and the probability of audit, which in panel (a) is proxied by the ratio of the number of flags assigned to a tax return to the maximally observed number of flags assigned to any tax return. As in Figure 5, there is a level difference between the two graphs. However, under our minimal assumption that the number of flags is positively correlated with the actual likelihood of audit, Figure 6(a) does suggest that the actual likelihood of an audit is distributed across the distribution of reported residual income in a manner broadly consistent with a revenue maximizing tax agency.[26] This conclusion is

[26]Note that the increasing average probability of audit for $x > 0$ is perfectly consistent with audit probability functions being strictly descreasing, conditional on z. In the simulations, the average audit probability is increasing for $x > 0$ because the equilibrium audit intensity and the variance of $u|z$ are increasing in



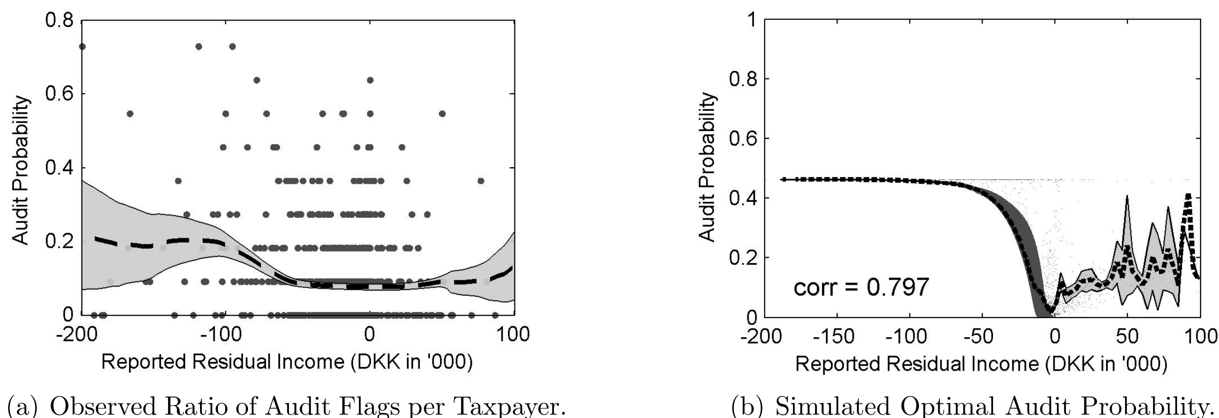
(a) Observed Ratio of Audit Flags per Taxpayer. (b) Simulated Optimal Audit Probability.

Figure 5. Observed and Simulated Optimal Audit Intensity Across the Distribution of Third-Party Reported Income.
Notes: Panel (a) shows, for the subsample of only tax evading wage earners and recipients of benefits (905 obs.), the number of flags per taxpayer as a share of the maximally observed number of flags across the distribution of third-party reported income, z. Panel (b) shows the simulated audit probability ($\sim$194,000 obs.) across the distribution of third-party reported income, z. In both panels, the dotted lines give the local average of the observations together with 95 percent confidence bands using local mean smoothing with the Epanechnikov kernel function and a rule-of-thumb bandwidth. The local mean smoothing in Panel (a) does not account for the stratification scheme. The simulated data in Panel (b) is not stratified. Income is defined as the sum of all income less deductions and is measured in '000 DKK. 1 USD $\approx$ 6 DKK (in 2006). In Panel (b), the budget is allocated such that approximately 4.1 percent of all wage earners and recipients of benefits are audited.

reinforced by the relatively high correlation (0.797) between local averages of the share of flags and simulated audit probabilities in the distribution of reported residual incomes.

The simulations accurately reproduce the covariance structure of tax evasion with respect to the composition of the tax return in terms of third-party reported income and residual income. Figure 7 shows empirical and simulated covariation of tax evasion ($u-x$) and residual income (u), denoted in DKK '000 for both wage earners and the self-employed. For each panel, we show individual data points and local averages and 95 percent confidence intervals across the domain of residual incomes using local mean smoothing. As shown in Panel (b) and (d), the local averages of simulated tax evasion are highly correlated (correlation coefficients 0.951 and 0.966, respectively) with the local averages of observed tax evasion for both wage earners and the self-employed. Moreover, except for a slight clustering of wage

z. Therefore, the higher is a taxpayer's z, there more likely it is, on average, that he is audited which, in equilibrium, lessens the degree to which he evades taxes, making it more likely that he reports a positive residual income.



(a) Observed Ratio of Audit Flags per Taxpayer.

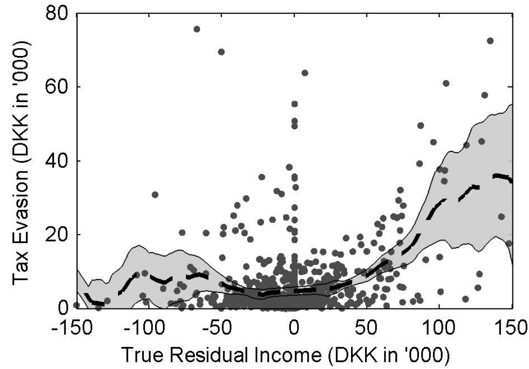
(b) Simulated Optimal Audit Probability.

Figure 6. Observed and Simulated Optimal Audit Intensity Across the Distribution of Reported Residual Income.

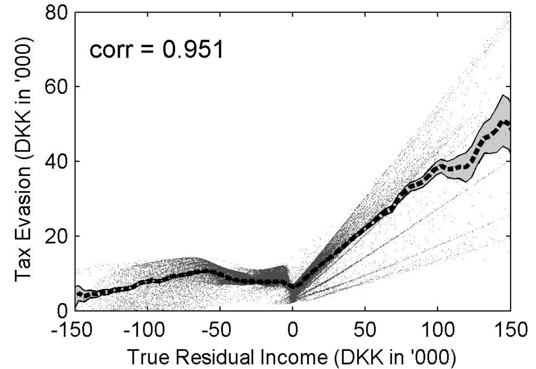
Notes: Panel (a) shows, for the subsample of only tax evading wage earners and recipients of benefits (905 obs.), the number of flags per taxpayer as a share of the maximally observed number of flags across the distribution of reported residual income, x. Panel (b) shows the simulated audit probability ($\sim$194,000 obs.) across the distribution of reported residual income, x. In both panels, the dotted lines give the local average of the observations together with 95 percent confidence bands using local mean smoothing with the Epanechnikov kernel function and a rule-of-thumb bandwidth. The local mean smoothing in Panel (a) does not account for the stratification scheme. The simulated data in Panel (b) is not stratified. Income is defined as the sum of all income less deductions and is measured in '000 DKK. 1 USD $\approx$ 6 DKK (in 2006). In Panel (b), the budget is allocated such that approximately 4.1 percent of all wage earners and recipients of benefits are audited.

earners with small negative residual incomes but relatively large degrees of evasion in the observed data, the distribution of individual data points also closely resembles that observed in the data. Although the self-employed evade more taxes on average, in neither data nor simulations do the self-employed appear to be more prone to evasion on the margin. Rather, the self-employed evade more taxes because they tend to have larger incomes and because less of that income is revealed by third parties. Finally, in equilibrium, tax evasion for the self-employed is curtailed to a large extent by intensive auditing.

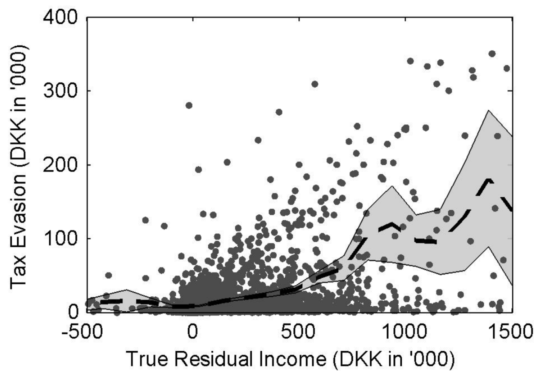
In conjunction, Figures 5-7 suggest the direct evidence of tax evasion and the indirect evidence on the Danish tax agency's enforcement strategy is consistent with our theory of rational tax evaders and a revenue maximizing tax agency. Moreover, since, as shown in Figure 5, the budget intensity increases with third-party reported income, we should expect Schotchmer's conjecture of progressive/regressive bias conditional on residual/third-party reported income to be bourne out in both data and simulations.



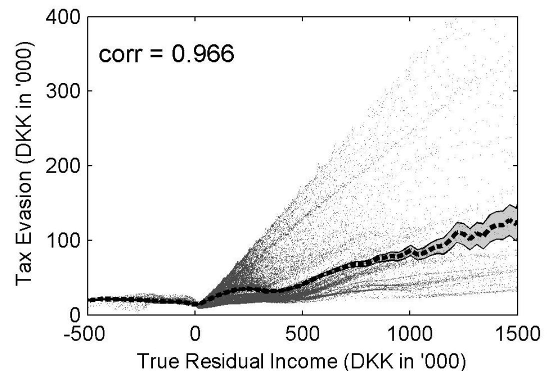
(a) Observed Tax Evasion, Wage Earners.



(b) Simulated Tax Evasion, Wage Earners.



(c) Observed Tax Evasion, Self-Employed.



(d) Simulated Tax Evasion, Self-Employed.

Figure 7. Observed and Simulated Tax Evasion Across the Distribution of True Residual Income.

Notes: Panels (a) and (c) show observed tax evasion across the distribution of true, i.e., post-audit, residual income, u, for wage earners (905 obs.) and self-employed (2,980 obs.), respectively. Panels (b) and (d) show simulated tax evasion across the distribution of true, i.e., post-audit, residual income, u, for wage earners ($\sim$194,000 obs.) and self-employed ($\sim$190,000 obs.), respectively. In both panels, the dotted lines give the local average of the observations together with 95 percent confidence bands using local mean smoothing with the Epanechnikov kernel function and a rule-of-thumb bandwidth. The local mean smoothing in Panels (a) and (c) does not account for the stratification scheme. The simulated data in Panels (b) and (d) is not stratified. Income is defined as the sum of all income less deductions and is measured in '000 DKK. 1 USD $\approx$ 6 DKK (in 2006). In Panel (b), the budget is allocated such that approximately 4.1 percent of all wage earners and recipients of benefits are audited. In Panel (d), the fraction of self-employed taxpayers audited is approximately 45.7 percent. Note, however, that the self-employed make up a much smaller group (approx. 400,000) compared to wage earners (approx. 4.2 million).

4.3.2 Effective Tax Rate Bias

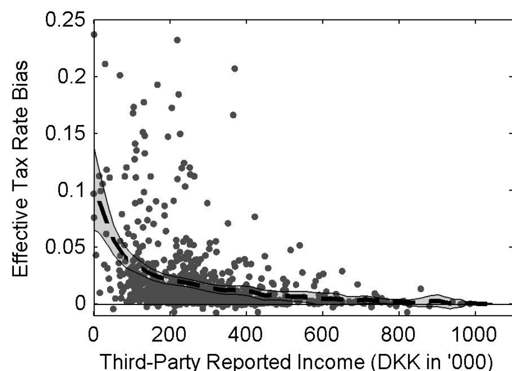
We calculate the bias of effective average tax rates as described in Section 3, $\tau - \tau^{\text{eff}}$, for data using the actual tax and penalty systems while for simulations using our approximations of a constant penalty rate, θ, and a set of constant marginal tax rates, t_z, that vary with third-party reported income.

In Figure 8 we display for each individual third-party reported income and our calcutation of effective tax rate bias. Panel (a) shows observations from the data set and Panel (b) shows simulated data. In each panel we also show local averages calculated using local mean smoothing. Both data and simulations exhibit effective average tax rates that are progressively biased with the bias decreasing towards 0 as third-party reported income increases. Moreover, the estimated local averages are highly correlated (correlation coefficient 0.974).

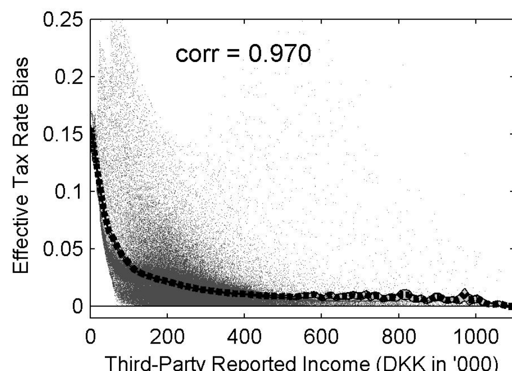
Figure 9, shows the corresponding figure of the data points and local averages of effective tax rates and residual income. Panel (a) and (b) share the same overall shape, namely, effective tax rates relatively unbiased (flat) in negative residual income but strongly biased and increasing in positive residual income. In both panels, the tax bias seems to decrease slightly at very high positive residual incomes. For the simulations, this reflects, similar to Figure 5(a) and 6(a), that high residual income is more common among taxpayers that also have large third-party reported incomes and that are audited relatively intensely. The structure of the data also seems consistent with this explanation. Again, local averages of effective tax bias in data and simulations in the distribution of residual income are highly correlated. The correlation, however, is somewhat smaller than for the progressive bias, reflecting the fact that regressive bias is generated partly by the allocation of audit probabilities within audit groups which, in the data, we observe imperfectly.

4.3.3 Regressions

Another way to assess the correlation structure in the data is to run reduced-form regressions, as we have done in Table 4. First, Panel A shows estimates from running a median regression



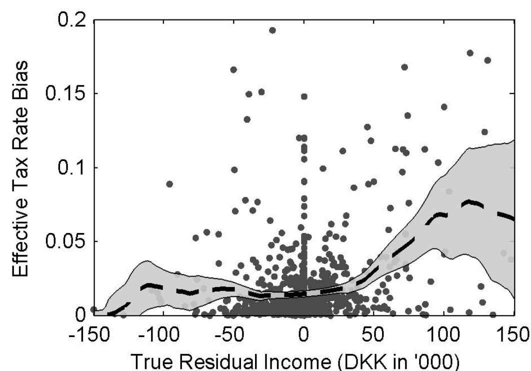
(a) Observed Tax Bias for Tax Evaders Across the Distribution of Third-Party Reported Income.



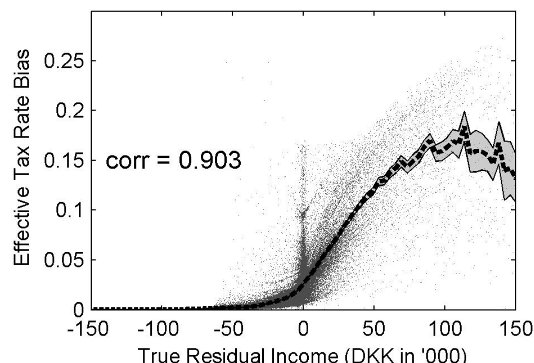
(b) Simulated Tax Bias for Tax Evaders Across the Distribution of Third-Party Reported Income.

Figure 8. Observed and Simulated Progressive Bias in Third-Party Reported Income.
Notes: The effective tax rate bias, $\tau - \tau^{\text{eff}}$, is the difference between the average statutory tax rate and the average effective tax rate as implied by the tax system, tax enforcement, and tax evasion behavior. Panel (a) shows the observed tax bias as a function of third-party reported income, z, for the subsample of tax evading wage earners and recipients of benefits (900 obs.). Tax rate bias is calculated as in (7). Panel (b) shows the simulated tax bias as a function of third-party reported income, z, for tax evading wage earners and recipients of benefits ($\sim$194,000 obs.). Tax rate bias is calculated as in (6). In both panels, the dotted lines give the local average of the observations together with 95 percent confidence bands using local mean smoothing with the Epanechnikov kernel function and a rule-of-thumb bandwidth. The local mean smoothing in Panel (a) does not account for the stratification scheme. The simulated data in Panel (b) is not stratified. Income is defined as the sum of all income less deductions and is measured in '000 DKK. 1 USD $\approx$ 6 DKK (in 2006). In Panel (b), the budget is allocated such that approximately 4.1 percent of all wage earners and recipients of benefits are audited.

on the sample of tax evaders of tax evasion on true residual income, u, (allowing slopes to differ depending on whether u is positive or negative) and third-party reported income, z, as well as a median regression of evasion on true total income, y. Whereas evasion does not appear to be increasing in total income y, it is, in fact, strongly increasing (0.381 [0.050] for wage earners 0.029 [0.008] for the self-employed) in positive residual income. As we can see from Figure 7, this is because positive residual income is much easier to disguise – tax evaders for whom $u > 0$ simply evade the entire amount of their residual income. For the tax agency, these taxpayers are indistinguishable from the many honest taxpayers reporting around $u = 0$, so this type of evasion is costly to uncover. For wage earners this is an especially attractive strategy due to the large mass of honest taxpayers reporting $x = 0$. For the self-employed there is virtually no excess mass of honest taxpayers reported $x = 0$,



(a) Observed Tax Bias for Tax Evaders Across the Distribution of True Residual Income.



(b) Simulated Tax Bias for Tax Evaders Across the Distribution of True Residual Income.

Figure 9. Observed and Simulated Regressive Bias in True Residual Income.
Notes: The effective tax rate bias, $\tau - \tau^{\text{eff}}$, is the difference between the average statutory tax rate and the average effective tax rate as implied by the tax system, tax enforcement, and tax evasion behavior. Panel (a) shows the observed tax bias as a function of true, i.e., post-audit, income, u, for the subsample of tax evading wage earners and recipients of benefits (900 obs.). Tax rate bias is calculated as in (7). Panel (b) shows the simulated tax bias as a function of true, i.e., post-audit, income, u, for tax evading wage earners and recipients of benefits ($\sim$194,000 obs.). Tax rate bias is calculated as in (6). In both panels, the dotted lines give the local average of the observations together with 95 percent confidence bands using local mean smoothing with the Epanechnikov kernel function and a rule-of-thumb bandwidth. The local mean smoothing in Panel (a) does not account for the stratification scheme. The simulated data in Panel (b) is not stratified. Income is defined as the sum of all income less deductions and is measured in '000 DKK. 1 USD $\approx$ 6 DKK (in 2006). In Panel (b), the budget is allocated such that approximately 4.1 percent of all wage earners and recipients of benefits are audited.

but it is still the case that the conditional distributions of residual income given third-party reported income is centered around $u = 0$ which makes such a reporting strategy attractive. As we also noted above, Table 4 suggests that the observed average marginal propensity to evade taxes is smaller for the self-employed than for wage earners. In our model, this is explained by the much higher audit rate for self-employed compared to wage earners. Because the self-employed on average have higher incomes and are subject to less third-party reporting, a self-employed taxpayer would tend to evade more than a wage earner for the same audit risk. Despite the high audit rate for self-employed, they nevertheless evade substantially more than wage earners.

Next, Panel B shows marginal effects, multiplied by a factor 100 for readability, from a Tobit regression of audit flag intensity (our empirical counterpart to the audit probability) on third-party reported income, z, and reported residual income, x, allowing slopes to differ

depending on whether x is positive or negative, and a Tobit regression of audit flag intensity on total reported income, $\tilde{y}$.[27] We find that the correlations exhibited by SKAT's audit flags are broadly consistent with Scotchmer's conjecture that a population-wide description of an optimal enforcement regime should entail decreasing audit probabilities within audit groups but increasing probability of audit between groups, exhibited by the negative coefficient on $x \cdot D_{x \leq 0}$ and a positive coefficient on z. This simple picture is slightly complicated by the positive and significant coefficient on $x \cdot D_{x \geq 0}$. However, as shown in Figure 6(b) this does not contradict the theoretical prediction of decreasing audit probability in u within an audit group. Rather, this coefficient is positive because only tax evaders with very high u report residual incomes substantially above 0 – these same tax evaders are predominantly found among taxpayers that also have large z and who are audited very intensively (c.f. Figure 5) leading to a positive correlation between $p(x|x > 0)$ and $x|x > 0$. We do not fully account for this complex relationship in our simple estimation specification.

The correlation of audit probability and income within and between audit groups translates into a significant positive correlation between audit probability and total reported income, $\tilde{y}$. Thus, despite the largely decreasing relationship within audit groups, third-party information reporting allows SKAT to audit taxpayers with high incomes more frequently.[28]

Lastly, in Panel C of Table 4 we run a median regression for the effective tax rate bias (in percentage points) either on true residual income, u, with slopes allowed to differ on the positive and negative domain of u, and third-party reported income, z, or on true total income, y. Since the data on tax evasion and audit flags seem consistent with the mechanisms driving the theoretical prediction of regressively biased effective average tax rates within audit groups, it is not surprising that we find a regressive bias within audit groups (i.e., a positive coefficient on $u \cdot D_{u \geq 0}$) and progressive bias between audit groups (i.e., a negative

[27] Left and right censoring of the audit flag intensity in the Tobit regressions is at 0 and 1, respectively.

[28] The intention to audit high-income taxpayers with higher probability is not a specific feature of Danish tax enforcement. Internal Revenue Service (2012) shows how, in 2011, 1.0 percent of taxpayers with incomes less than \$200,000 were audited, 3.9 percent of taxpayers with income in the range of \$200,000-1,000,000 were audited, and 12.5 percent of taxpayers with incomes over \$1,000,000 were audited.

coefficient on z). These effects combine to make tax rates *progressive* in total income, y. As shown above, the progressive bias between audit groups derives from the fact the SKAT intensitvely audits taxpayers with high z.

Overall, Table 4 suggests a correlation structure of effective tax rates as depicted in the stilized Figure 1(b). The data supports the theoretical prediction that effective tax rates are regressive within audit groups. Between audit groups, there is a progressive bias such that average tax rates are actually progressively biased in total total income.

4.3.4 Non-Detectable Income

A potential problem for the robustness and validity of our results concerns non-detected tax evasion. As we discuss in Section 3, some unreported income is almost certainly missing from our measures of tax evasion, despite SKAT's dilligent effort in making audits comprehensive. In particular, black market income is likely hard to detect. In Appendix B.1, we briefly present the best available evidence on the distribution of black market income in the Danish population based on survey data collected by the Rockwool Foundation Research Unit. This evidence suggests that black market income may be of a nonnegligible magnitude averaging approximately 3,143 DKK in the population. In comparison, the population weighted average underreported income is 3,619 DKK, cf. Table 2, Panel D.

However, assuming that black market income is completely non-detectable, the presence of such income will not affect neither the taxpayers' nor the tax agency's optimization criteria.[29] Consequently, the equilibrium of the model is unaffected, and the calibration exercise in this section remains valid because we fit the model to average tax evasion not including black market income.

Of course, even if black market income is completely non-detectable, it implies a measurement error in true residual income and translates into an underestimation of the effective average tax rate bias. Given that black market income is negatively correlated with reported

[29]For taxpayers, this hinges on the assumption of risk neutrality. For example, if taxpayers are risk averse, non-detectable black market income may interact with tax evasion behavior.

	Wage earners						Self-employed	
Dependent variable:	**A.** Evasion		**B.** Audit flag intensity		**C.** Tax bias (in pct. points)		**D.** Evasion	
$u \cdot D_{u \leq 0}$	-0.007	-	-	-	-0.004	-	0.004	-
	(0.008)				(0.004)		(0.005)	
$u \cdot D_{u > 0}$	0.381 ***	-	-	-	0.030 ***	-	0.029 ***	-
	(0.050)				(0.003)		(0.008)	
z	0.000	-	0.002 **	-	-0.003 ***	-	0.006 ***	-
	(0.001)		(0.001)		(0.000)		(0.003)	
y	-	0.002	-	-	-	-0.002 ***	-	0.015 ***
		(0.001)				(0.000)		(0.004)
$x \cdot D_{x \leq 0}$	-	-	-0.053 ***	-	-	-	-	-
			(0.013)					
$x \cdot D_{x > 0}$	-	-	0.002 **	-	-	-	-	-
			(0.001)					
$\tilde{y}$	-	-	-	0.003 ***	-	-	-	-
				(0.001)				
Estimation method	Median regression		Tobit regression		Median regression		Median regression	
Constant term included	x	x	x	x	x	x	x	x
$D_{u>0}$ or $D_{x>0}$, respectively, included	x		x		x		x	
Sample	Evaders	Evaders	Full	Full	Evaders	Evaders	Evaders	Evaders
Observations	905	905	10,584	10,584	900	900	2,980	2,980
Obs. left-censored	-	-	8,555	8,555	-	-	-	-
Obs. right-censored	-	-	2	2	-	-	-	-
Pseudo R-squared	0.19	0.00	0.12	0.02	0.16	0.02	0.04	0.02

Table 4. Evasion Behavior, Tax Enforcement, and Tax Bias – Regressions on Data Sample.

Notes: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.
All regressions are for wage earners and recipients of benefits except the rightmost two columns, which are regressions for self-employed. All samples are restricted to only contain taxpayers with strictly positive true total income.
Variable definitions: Monetary variables are in thousands ($u, z, y, \tilde{y}, x$, and evasion). Audit flag intensity is defined as no. of flags on the tax return divided by maximum no. of flags on any tax return (13) so that audit flag intensity is distributed on the unit interval. Tax bias is the percentage point difference between the true average tax rate and the effective average tax rate (reflecting both evasion behavior, audit probabilities, and penalty rates).
Evasion regressions (wage earners): Numbers in parentheses are stratified bootstrapped standard errors. The sample contains wage earners and benefit recipients with detected tax evasion and positive true income.
Audit flag regressions: Left and right censoring at 0 and 1, respectively. Numbers in parentheses are robust stratified standard errors. Estimates presented in the Tobit regressions are marginal effects. Marginal effects and standard errors in the Tobit regressions are multiplied by a factor of 100 for readability. The sample contains wage earners and benefit recipients with positive true income.
Tax bias regressions: Numbers in parentheses are stratified bootstrapped standard errors. The sample contains wage earners and benefit recipients with detected tax evasion and positive true income. There are five less evaders than in the evasion regressions. This is due to missing values in the compliance rating used to construct the tax bias measure.
Evasion regressions (self-employed): Numbers in parentheses are bootstrapped standard errors. The sample contains self-employed taxpayers with detected tax evasion and positive true income. As we do not have credible information on tax enforcement for the self-employed, there are no regressions for audit flag intensity and tax bias for the self-employed.

income (cf. Appendix B.1) and third-party reported income is a very large part of reported income (approximately 95 percent in the population), we can deduce that black market income is also negatively correlated with third-party reported income. Therefore, accounting for black market income implies a level shift in the effective tax rate bias as a function of third-party reported income and, in addition, that this level shift is largest for taxpayers with little third-party reported income. As a result, including black market income implies a stronger progressive bias.

With respect to the regressive bias, the effect of including black market income depends on how black market income and detectable residual income are correlated. For the group of "wage earners", we know that low-income earners and recipients of public transfers more frequently provide black market labor. These individuals have little to no detectable residual income, and, as black market income constitutes positive residual income and at the same time increases the effective tax rate bias, including black market income will tend to strengthen the positive relationship between tax rate bias and residual income (i.e., the regressive bias) for positive residual incomes depicted in Figure 9.

All in all, black market activities strenghten the distortions of the statutory tax system already generated by tax evasion and enforcement with respect to the formal economy.

4.3.5 Sensitivity Analysis

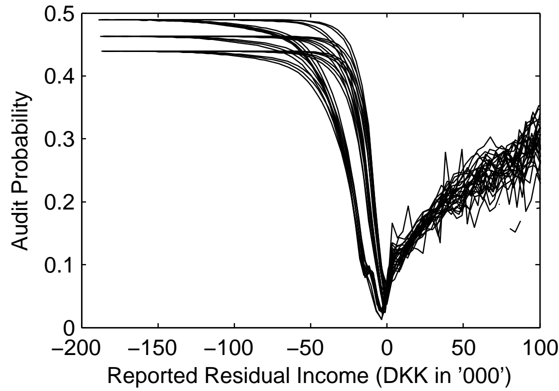
To show that our conclusions are robust to changes in parameters, we present in Figure 10 the simulated results of parameter changes for wage earners. We do this by changing the key parameters t_z, θ, and Q, and for each permutation letting $\mathbf{B}$ be calibrated to match simulated and observed average tax evasion among evaders. This we do for 27 permutations of the key parameters, i.e., all combinations of $-10\%, 0\%, +10\%$ changes to the set of parameters.[30] Focusing on the mechanism driving the regressive bias within audit groups, Panels (a) and (b) of Figure 10 show local averages of audit probabilities as a function of reported income

[30]For t_z the changes are implemented as across-the-board increases/decreases in the marginal tax rate.

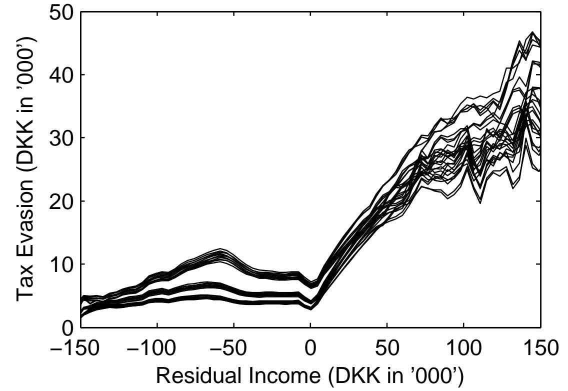
and local averages of tax evasion as a function of true residual income, respectively. For audit probabilities, the changes are relatively minor, the main effects being a level shift in the maximal audit probability corresponding to changes in θ. For tax evasion, the local averages are all qualitatively similar, although the impact of parameter changes are larger among tax payers with larger residual incomes.

The structure of tax rate bias within and between audit groups in the simulations is also highly robust. The progressive bias between audit groups, shown in Panel (d), is virtually unchanged as it is generated mainly by the distribution of audit resources in the population, which is more or less unchanged by the parameter changes. The impact of parameter changes on the regressive bias within audit groups, shown in Panel (c), is more substantial as it compounds the effects of parameter changes shown in Panels (a) and (b). However, in all cases the qualitative relation between effective tax rate bias is very similar to the baseline simulation.

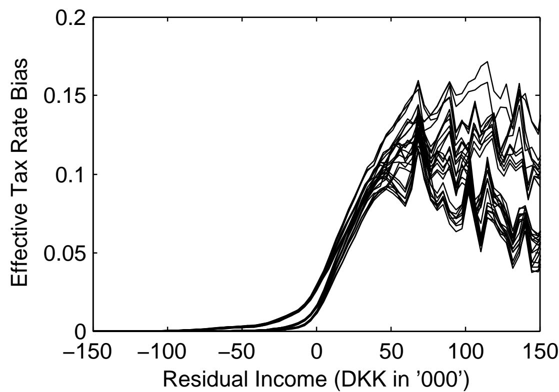
Varying the model parameters also affects the correlations of local averages in data and simulations, although not to a large extent. For the relationship between residual income and tax evasion, for example, the correlation coefficient lies between 0.875 and 0.952 compared to the baseline of 0.951. The most variable correlation the relationship between third-party reported income and the probability of audit which lies between 0.602 and 0.877 compared to the baseline of 0.803. However, this largely does not affect the correspondence of the progressive bias relationships between data and simulations – the correlation of local averages for third-party reported income and effective tax rate bias lies between 0.929 and 0.962. The regressive bias relationship varies more as it is affected by changes in both tax evasion and the audit probabilities within audit groups and lies between 0.666 and 0.907.



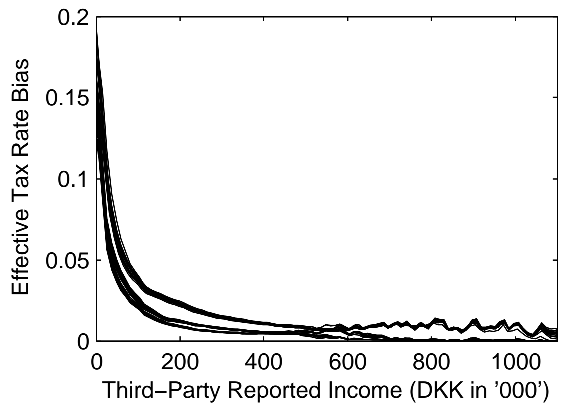
(a) Audit Probability, Reported Residual.



(b) Tax Evasion, True Residual Income.



(c) Effective Tax Bias, True Residual Income.



(d) Effective Tax Bias, Third-Party Reported Income.

Figure 10. Robustness Checks for Simulations of Wage Earners.
Notes: This figure checks the robustness of our simulation results graphically by plotting variations in estimated local average means on the basis of simulations with parameter permutations. We simulate the changes for wage earners in the four key relationships of the model, (a) audit probability as a function of reported residual income, (b) tax evasion as a function of true residual income, (c) regressive tax bias within audit groups as a function of true residual income, and (d) progressive tax bias between audit groups as a function of third-party reported income. The local averages depicted in the four panels are calculated in a similar manner to Figures 5–9, using local mean smoothing with the Epanechnikov kernel function and a rule-of-thumb bandwidth. We simulate the model for 10 percent parameters variations around the baseline estimates of t_z, θ, and Q, corresponding to 27 separate simulations. Thus, Panel (a) depicts variation in local means around the baseline simulation depicted in Figure 6(b) and similarly Panels (b)–(d) correspond to variations around the baseline simulated local means depicted in Figures 7(b), 9(b), and 8(b), respectively. All amounts in '000 DKK (1 USD ≈ 6 DKK in 2006).

5 Concluding Remarks

This paper highlights the importance of information in tax enforcement. We find evidence in favor of the regressive bias prediction and Scotchmer's (1987) conjecture that it is crucial to distinguish regressive bias *within* an audit group from *aggregate* or *between*-group variation. Using detailed administrative data, we find evidence suggesting that, whereas effective tax rates are regressively biased within audit groups as theory suggests, this relationship is negated by a progressive bias between audit groups induced by the distribution of audit resources and third-party information. The outcome is that tax rates are progressively biased in total income. However, the model also suggests that an enforcement regime with much less third-party reported information would be substantially more regressive. In Denmark, this is avoided by a large information collection effort.

However, as emphasized by the literature, distortions may be substantial in settings in which third-party reporting is less comprehensive. The standard optimal auditing literature seems to suggest that regressive bias can be countered simply by adjusting marginal tax rates across the board. However, once we allow for population heterogeneity of behavior and income composition, this is no longer feasible. Our results suggest an obvious policy to ameliorate these distortions: increasing the share of income reported by third parties will reduce both the extent of evasion and the regressive bias in tax enforcement.

From a theoretical point of view, including third-party reported information and the likelihood of honest reporting conditional on the income composition is crucial in understanding tax evasion. A large literature on compliance versus non-compliance emphasizes behavioral/social explanations such as guilt and shame (e.g., Grasmick and Bursick, 1990), fairness (e.g., Spicer and Becker, 1980), and trust in government (e.g., Slemrod, 2003; Torgler, 2003). We focus on the implications, rather than the explanations, of honest reporting by some taxpayers. We analyze whether the observed moments and correlation structures of data are consistent with that generated by an optimizing tax agency, a group of honest taxpayers, and a group of expected utility maximizing tax evaders. We find that our model

can replicate the extent of observed evasion as well as the subtle correlation structure of tax evasion, the probability of audit, and effective average tax rates with income. In addition, our results indicate that the Danish tax agency employs a distribution of resources across audit groups that is surprisingly similar in key respects to the optimal distribution generated by the model. All in all, there seems to be a role for both standard economic theory and behavioral/social extensions in explaining the behavior of tax evaders.

The correlation structure of effective tax rates seems robust: it is generated by our realistically complex model as well as in Scotchmer (1987). Furthermore, while variations in parameters change the level of average tax rate bias as well as the rate of progressivity between audit groups, in no variations is the correlation structure of effective tax rates qualitatively different from our baseline simulation. Thus, we are confident that similar empirical relationships would be found in data from any tax agency that employs, as does SKAT, a strong signal in predicting true incomes.

Based on data on the distribution of black market income in Denmark, we argue that our results are also robust to the lack of non-detectable income in our data. In fact, as discussed in Section 4.3.4, the data on the distribution of black market income suggests that our finding of regressive and progressive tax rate bias within and between audit groups, respectively, are lower bounds on the actual distortions of the statutory tax system. Under the assumption that black market income is non-detectable, the tax evasion and enforcement equilibrium is unaffected by the presence of this type of income, and the simulation results remain a valid description of optimal tax evasion and enforcement with respect to the formal economy. Moreover, the close fit between simulated and observed tax evasion and enforcement indicates that the assumption of complete non-detectability of black market income is an appropriate simplification.

A natural objection to the model we employ is the lack of general interactions with labor market choices. We accomodate to some extent the self-selection of taxpayers into employment categories by allowing the fraction of honest taxpayers to differ between wage earners

and self-employed. In addition, disregarding dynamic aspects is not likely to be important due to the limited retrospectivity of SKAT's actual audit scheme and the restrictive statute of limitations on retroactive penalties for tax evasion. However, we do not account for other effects, e.g., how tax enforcement affects labor supply on the intensive margin.

Despite these limitations, our paper advances the literature in the direction of developing a full-fledged structural model of tax evasion, which can be estimated directly with maximum likelihood or GMM methods. Moreover, our paper is an important next step towards an understanding of the comparative statics of tax evasion and enforcement. Such an avenue of research may be seen as a necessary complement to the literature on their behavioral and social determinants, which have been extensively explored in the literature.

Appendix

A.1 Numerical Implementation

The second order differential equation is obtained by combining (1), (3), (4), and the expression for $\frac{\partial u}{\partial x}$ to get

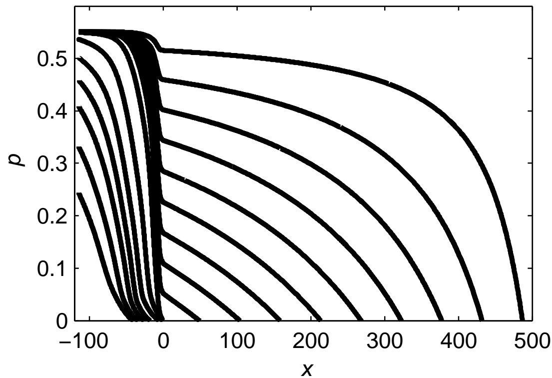
$$p''(x) = \left(\frac{f_u^h(x)\,\frac{\lambda c}{\theta t + t}}{f_u^d(u(x))\left[\frac{p(x) - \frac{1}{\theta+1}}{p'(x)} - \frac{\lambda c}{\theta t + t}\right]} - 2 \right) \cdot p'(x)^2 \left(\frac{1}{1+\theta} - p(x) \right)^{-1}, \tag{A.1}$$

suppressing z for convenience. Thus, sufficient conditions for equilibrium, given $B(z)$, are the two equations (5) and (A.1).

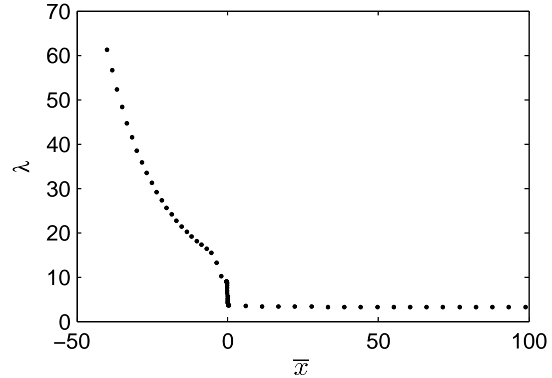
We approximate the equilibrium solution by discretizing z into an evenly spaced grid point vector of dimension 200.[31] Equilibrium functions for other values of z are approximated by interpolation. For each gridpoint, we solve the 2nd order ordinary differential equation (ODE) in (A.1) for many values of $\overline{x}$, where $\overline{x} \equiv x(\overline{u})$. The ODE algorithm is then initialized using $p(\overline{x}) = 0$ and $p'(\overline{x}) = \left(\frac{1}{1+\theta}\right) / (\overline{u} - \overline{x})$, cf. (1). For each value of $\overline{x}$ and z, we need a corresponding value of $\lambda(z)$, the shadow value of increasing the budget size. However, $\lambda(z)$ and $\overline{x}$ are not separately identified. Therefore, we must take a heuristic approach, solving for each $\overline{x}$ the ODE for many values of λ until one is found that satisfies the equilibrium conditions everywhere, in particular at $x = \overline{u}$. In practice, we do not merely guess repeatedly at $\lambda(z)$, but employ a search algorithm to find the $\lambda(z)$ that satisfies (5); this provides a candidate $\lambda(z)$ corresponding to a particular $\overline{x}$ that satisfy the FOC everywhere with a small error tolerance. Figures 1(a) and 1(b) illustrate an example of the set of solutions resulting from the algorithm.

When this algorithm has executed for all grid points of z, we can determine the optimal

[31] The model for the self-employed is substantially more computationally intensive so there we use only 100 grid points. Of course, this implies that interpolations will be less precise, but this does not appear to be important. Likewise, solutions using only 50 grid points are graphically indistinguishable in terms of Figure 7.



(a) Examples of Optimal Audit Functions: $p(x|z)$.



(b) Shadow Values with respect to $B(z)$.

Figure A.1. Solutions Examples
Notes: $\bar{x}$ (xbar) is defined as the lowest value of x that solves $p(x|\cdot) = 0$, i.e., the highest report of dishonest taxpayers. Reported residual, x, is measured in '000 DKK (1 USD $\approx$ 6 DKK in 2006).

budget allocation using the fact that in an interior equilibrium $\lambda(z)$ must be equalized across different levels of z.

Equation (A.1) can be solved by standard numerical methods. We employ a Runge-Kutta-type algorithm developed in Shampine (2009), which outperforms most standard ODE algorithms in terms of precision and robustness. However, two main numerical issues must be resolved.

First, due to point mass in $f^h_{u|z}$ at $u = 0$, $\mathrm{E}(u|x, z)$ is discontinuous at $x = 0$, which induces what is known as a "singularity" in the differential equation. We take a standard approach to this problem and approximate solutions for which $\overline{x} > 0$ by substituting the logical function $\mathbf{1}_{(x=0)}$ with a smooth, differentiable approximation. The resulting function displays a relatively smooth transition from 0 to 1 in a small band around $x = 0$. An alternative approach is to split the ODE algorithm in two, corresponding to the domains $[\underline{u}, 0)$ and $[0, \overline{x}]$, and identifying the discontinuous jump in $p'(x)$ from the equations characterizing the equilibrium and the measure of point mass at $x = 0$. However, as the size of this discontinuity cannot be identified analytically, this introduces an element of imprecision in the algorithm which, in our experience, may negatively affect the robustness of the algorithm.

Second, the ODE algorithm may fail to converge if we allow the conditional density

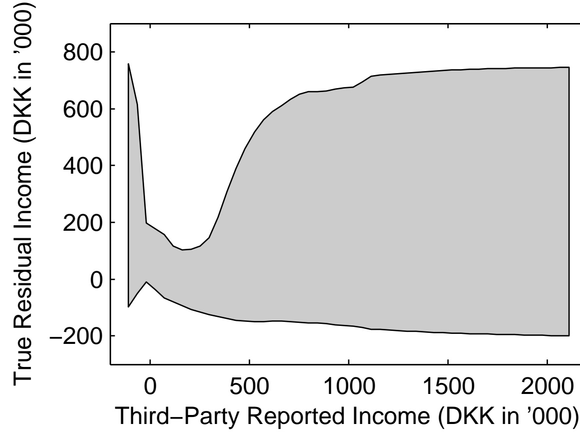


Figure A.2. The Support of u Across Audit Groups.
Notes: The estimated conditional densities of $u|z$ for wage earners and benefit recipients are truncated at the 0.25 and 99.75 percent fractiles of the unrestricted conditional distributions. Residual income, u, and third-party reported income, z, are measured in '000 DKK (1 USD $\approx$ 6 DKK in 2006).

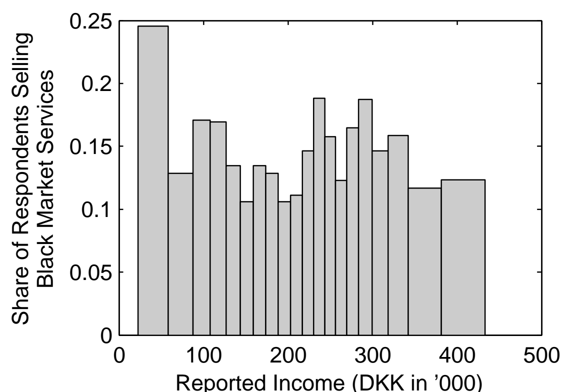
function to take values extremely close to 0 since the ratio $\frac{f^h_{u|z}(x)}{f^d_{u|z}(u(x))}$ may diverge toward infinity. Estimating the density f_{uz} as a bivariate kernel density is numerically inconvenient as it tends to result in "troughs" of zero density in the interior of the domain of some conditional distributions. Instead, as mentioned in the main body of the paper, we estimate f_{uz} as a lognormal mixture distribution.[32] Specifically, the mixture distribution consists of six component distributions. Increasing the number of component distributions allows a more flexible fit of the distribution but alters our results only very slightly. Lastly, we truncate the domain of the potential tax evaders' conditional true income distributions where the densities are very close to zero to keep the fraction in equation (A.1) from diverging to infinity. Specifically, we truncate the unrestricted conditional densities at the 0.25 percent and 99.75 percent fractiles. The resulting supports of the conditional distributions vary in z as illustrated in Figure A.2.

[32] Of course, the lognormal distribution is not defined on domains that include negative values. In practice, in estimating the mixture distribution, we create a simple additive mapping of the observations to a set of "virtual residual incomes" that are entirely positive, estimate the lognormal mixture distribution using six component distributions, and use the mapping to obtain the actual bivariate income distribution. The resulting distribution is indeed very close to that obtained by using a bivariate kernel density algorithm.

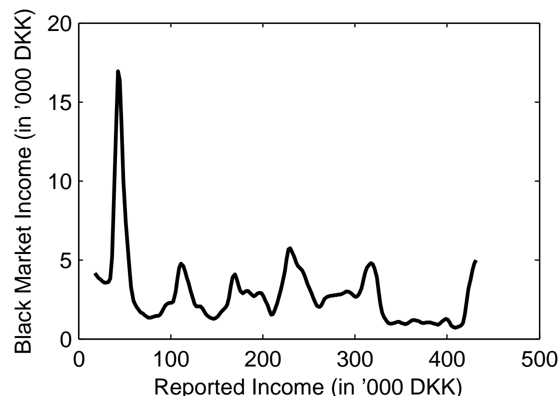
B.1 Black Market Activities

A potentially important avenue for tax evasion is black market income. This type of income is much harder to discover by tax auditors and, thus, less likely to be included in our data despite the intensive auditing of tax returns for the experiment. To quantify the extent of black market activities in the population at large, we utilize survey data collected by the Danish Rockwool Foundation Research Unit. They have since 1985 collected survey data for random samples of the population in an attempt to quantify the incidence of and return to black market activity. Although the surveys in principle collect identifiable information, such as social security numbers, the data set is anonymized. As such, no one has ever been prosecuted for having black market income due to participation in these surveys. Unfortunately, there were no surveys carried out in 2006 so we use two surveys from 2005 and two from 2007 instead (all amounts in 2006-prices). The surveys contain many variables, but we focus on measures of the incidence of and return to the supply of black market services (i.e., remuneration for black market labor). In Figure B.1, we show the incidence of black market work and the return to this activity across the distribution of reported income. Panel (a) shows the share of taxpayers having performed black market work for 20th fractiles of the reported income distribution in the sample. The figure indicates that black market work is more common among low-income taxpayers, whereas middle and top earners figure less prominently.[33] Panel (b) shows the average black market income across the distribution of reported income for the entire sample, unconditional on whether or not taxpayers have participated in black market work, using local mean smoothing. Again, mainly low-income taxpayers have black market labor income. This is consistent with a comprehensive study by the Danish Economic Council in 2011 (DØRS, 2011) using the same data as here but for all available years, which concluded that black market earnings were negatively correlated with total reported income. In addition, DØRS (2011) finds that the self-employed and low-

[33] In the survey samples from 2005 and 2007 that we use, there are very few top earners. But in DØRS (2011), which uses a larger sample spanning more waves of the survey, there is a clear picture that top earners supply black market work less frequently than the middle income earners.



(a) Share of Taxpayers Selling Black Market Services by 20th Fractiles of the Reported Income Distrition.



(b) Mean Black Market Income Across the Reported Income Distribution.

Figure B.1. Size and Distribution of the Black Market Economy in Denmark.
Notes: The data on black market activity stems from a survey collection effort undertaken by the Rockwool Foundation Research Unit, and income data stems from linked administrative data. The Rockwool Foundation Research Unit's surveys on black market income has been collected since 1985. For each wave in the collection effort, surveys are dispatched to individuals, both wage earners and the self-employed, with the understanding that their answers are kept anonymous. As such, no one has ever been prosecuted for acknowledging black market income unreported on their tax return in the surveys. Unfortunately, the surveys were not collected in 2006. Instead, we have obtained data for four surveys collected in 2005 and 2007 (in 2006 prices), for a total sample of surveyed individuals of 3,806. Only 10 individuals did not respond so the sample of responsive individuals is 3,796. Of these individuals, 560 responded that they had sold services on the black market during the last 12 months with an average income of 19,439 DKK. In the total sample of responsive individuals, this corresponds to an average black market income of 3,143 DKK. Panel (a) shows the share of taxpayers in the sample selling black market services by 20th fractiles of the distribution of total reported income. Panel (b) shows the average return to black market activity in the sample using local mean smoothing with the Epanechnikov kernel function and a rule-of-thumb bandwidth. All amounts are in '000 DKK (1 USD ≈ 6 DKK in 2006).
Source: Rockwool Foundation Research Unit and own calculations.

income wage earners most frequently supply labor on the black market, and black market wages are substantially higher for the self-employed, indicating that this group of taxpayers on average earns larger black market incomes than wage earners.

References

Allingham, M. G., Sandmo, A., 1972. Income tax evasion: A theoretical analysis. Journal of Public Economics 1, 323–338.

Alm, J., McClelland, G. H., Schulze, W. D., 1992. Why do people pay taxes? Journal of

Public Economics 48 (1), 21–38.

Andreoni, J., Erard, B., Feinstein, J. S., 1998. Tax compliance. Journal of Economic Literature 36 (2), 818–60.

Clotfelter, C., 1983. Tax evasion and tax rates: An analysis of individual returns. The Review of Economics and Statistics 65 (3), 363–373.

Cremer, H., Gahvari, F., 1994. Tax Evasion, Concealment and the Optimal Linear Income Tax. Scandinavian Journal of Economics 96 (gr), 219–239.

Cremer, H., Marchand, M., Pestieau, P., 1990. Evading, auditing and taxing: The equity-compliance tradeoff. Journal of Public Economics 43 (1), 67–92.

DØRS, 2011. Dansk økonomi, Forår 2011. www.dors.dk.

Erard, B., Feinstein, J. S., 1994. Honesty and evasion in the tax compliance game. RAND Journal of Economics 25 (1), 1–19.

Feinstein, J., 1991. An econometric analysis of income tax evasion and its detection. The RAND Journal of Economics 22 (1), 14–35.

Feld, L. P., Frey, B. S., 2002. Trust Breeds Trust: How Taxpayers Are Treated. Economics of Governance 3 (2), 87–99.

Grasmick, H. G., Bursick, Jr., R. J., 1990. Conscience, significant others, and rational choice: Extending the deterrence model. Law and Society Review 24 (3), 837–861.

Internal Revenue Service, 2012. Fiscal year 2011 enforcement and service results. http://www.irs.gov/newsroom/article/0,,id=251923,00.html.

Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S., Saez, E., 2011. Unwilling or unable to cheat? Evidence from a randomized tax audit experiment in Denmark. Econometrica 79 (3), 651–692.

Nielsen, S. B., Sørensen, P. B., 1997. On the optimality of the nordic system of dual income taxation. Journal of Public Economics 63, 311–329.

Phillips, M. D., 2010. Taxpayer Response to Targeted Audits. mimeo, University of Chicago.

Pomeranz, D., 2010. No taxation without information: Deterrence and self-enforcement in the value added tax.

Reinganum, J. F., Wilde, L. L., 1986. Equilibrium verification and reporting policies in a model of tax compliance. International Economic Review 27 (3), 739–60.

Sanchez, I., Sobel, J., 1993. Hierarchical design and enforcement of income tax policies. Journal of Public Economics 50 (3), 345–69.

Scotchmer, S., 1987. Audit classes and tax enforcement policy. American Economic Review 77 (2), 229–33.

Scotchmer, S., 1992. The regressive bias in tax enforcement. Public Finance 47, 367–371.

Shampine, L. F., 2009. Vectorized solution of ODEs in Matlab with control of residual and error. `http://faculty.smu.edu/shampine/`.

Slemrod, J., 2003. Public Finance and Public Policy in the New Century. Cambridge, MA: MIT Press, Ch. Trust in Public Finance, pp. 49–88.

Slemrod, J., 2007. Cheating ourselves: The economics of tax evasion. Journal of Economic Perspectives 21 (1), 25–48.

Spicer, M. W., Becker, L. A., 1980. Fiscal inequity and tax evasion: An experimental approach. National Tax Journal 33 (2), 171–175.

Torgler, B., 2003. Tax morale, rule-governed behavior, and trust. Constitutional Political Economy 14 (2), 119–140.