Tax Compliance by Firms and Audit Policy[*]

Ralph Bayer[†] and Frank Cowell[‡]

University of Adelaide and London School of Economics
September 2010

The Toyota Centre
Suntory and Toyota International
Centres for Economics and Related
Disciplines
London School of Economics
Houghton Street
London WC2A 2A

(+44 020) 7955 6674)

[†] Correspondence: Dr Ralph-C. Bayer, School of Economics, University of Adelaide, North Terrace, Adelaide, SA 5005, Australia. Email: ralph.bayer@adelaide.edu.au

[‡] Correspondence: Professor Frank Cowell, London School of Economics, Houghton Street, London WC2A 2AE. Email: f.cowell@lse.ac.uk

Abstract

Firms are usually better informed than tax authorities about market conditions and the potential profits of competitors. They may try to exploit this situation by underreporting their own taxable profits. The tax authority could offset firms' informational advantage by adopting "smarter" audit policies .that take into account the relationship between a firm's reported profits and reports for the industry as a whole. Such an audit policy will create an externality for the decision makers in the industry and this externality can be expected to affect not only firms' reporting policies but also their market decisions. If public policy takes into account wider economic issues than just revenue raising what is the appropriate way for a tax authority to run such an audit policy? We develop some clear policy rules in a standard model of an industry and show the effect of these rules using simulations.

Distributional Analysis Research Programme

The Distributional Analysis Research Programme was established in 1993 with funding from the Economic and Social Research Council. It is located within the Suntory and Toyota International Centres for Economics and Related Disciplines (STICERD) at the London School of Economics and Political Science. The programme is directed by Frank Cowell. The Discussion Paper series is available free of charge. To subscribe to the DARP paper series, or for further information on the work of the Programme, please contact our Research Secretary, Leila Alberici on:

1   Introduction

Should a tax authority take into account the "real-economy" effects of its compliance policy? The actions of tax authorities are often perceived in purely financial terms, perhaps as a kind of tax farmer that seeks to maximise the revenue for the government or as a fiscal police officer that seeks to ensure enforcement of the law as effectively as possible. However, just as a conventional police force may properly have objectives other than simple law enforcement (fostering good community relations for example) so the tax authority may be required to have concern for a broader range of economic objectives than simple revenue-raising and compliance. Although it is convenient as a modelling device to assume that an agency has a single financial target it would be unreasonable to insist that the government's different policy objectives were located in separate watertight compartments. In this paper we suppose that a sensible tax authority is concerned about issues of productive efficiency in the economy and about equitable treatment of taxpayers. We develop a model of tax compliance by firms and show how their activity in product markets is connected with the design and implementation of enforcement policy by a tax agency.

Some aspects of the real-economy issues associated with tax compliance are already well known. For example in the case of the personal income tax and decisions made in the labour market the conventional Allingham and Sandmo (1972) model can be extended to incorporate labour supply. The conventional welfare-economic analysis of deadweight loss as applied to income taxes and commodity taxes can be extended to take tax noncompliance into account (Cowell 1990). However the further considerations that apply in the case of the taxation of firms have not been worked out. The case of firms is special in terms of both the efficiency and equity objectives .

First, the efficiency considerations arise from the interaction among firms within an industry as well as interaction of firms with the tax authority. Bayer and Cowell (2009) have demonstrated that the effectiveness of compliance policy depends on whether there is effective competition or collusion among the firms in the industry.[1] The interrelation between market organisation and the design of compliance policy raises several policy questions. Should audit rules be designed in such a way that firms will be induced to

---

[1]See also Besfamille et al. (2009) who show that with imperfectly competitive firms increased enforcement of an output tax will reduce output and may reduce revenue.

act more efficiently in product markets? Should a change in industry competitiveness change the design of compliance policy?

Second, the equity considerations arise precisely from the tailored audit rules that the tax-authority might employ to induce the behaviour in product markets that might be desirable on efficiency grounds. A "smart" compliance policy may give the appearance of treating equals unequally in a way that does not arise in compliance models involving the personal income tax. This implies that in evaluating the desirability of compliance policy one needs to go beyond the conventional individualistic welfare model in order to deal with questions of tax equity.

The paper is organised as follows. Section 2 explains our approach and relates it to the literature and Section 3 sets out the formal model. Section 4 develops the simple welfare-analytics of this model examines its workings using a simulation; Section 5 discusses the special issues of equitable treatment that occur in the audit model with firms; Section 6 draws the policy implications from this. Section 7 concludes.

2 The approach

2.1 Setting

Before we specify the precise model that we shall use to establish results and to simulate behaviour let us describe the economic agents and their interrelationships.

Firms In some treatments of the economics of tax compliance firms are treated as no more than profit centres which can be tapped by the tax agency. More sophisticated approaches take some account of the firms' role as producers but in a naive fashion that does not yield much economic insight. The standard assumption is either that each firm is a price-taker without market power or that there is a perfectly-informed monopolist with almost complete market power. However, under conventional treatment of risk and taxation, each of these idealised market forms turns out to produce an analysis of tax evasion in which the firms' characteristics are effectively absent:[2] because

---

[2]What happens is that under these special market conditions the production decisions can effectively be separated out from the tax compliance decisions, reducing the tax-compliance problem to a minor elaboration of the Allingham and Sandmo (1972) model

each firm is assumed to be in a particularly simple market environment, particularly simple results emerge.

To make the analysis interesting we need to think of the firm also as an information processor. This involves analysing the behaviour of the firm under uncertainty. The uncertainty comes from three sources:

1. exogenous uncertainty, such as demand shocks, cost shocks and assessment errors

2. uncertainty as to whether the firm will be audited for tax purposes,

3. uncertainty about the behaviour of other firms in a similar position to itself.

All three types of uncertainty will be seen to have a role within our model. The third type makes it clear that it is important to consider the firm within the context of an industry where the behaviour of other firms is important in determining its own behaviour.

Industry What is an industry? Our model of an industry focuses not so much on the physical characteristics of the outputs of the member firms but on the relationship among them. In the light of the exogenous uncertainty mentioned as point 1 in the list above it makes sense to suppose that members of the industry are better-informed than other economic agents about market events that may affect their profits: they intimately know the economic conditions that apply to their industry and could, if they wanted to, make reasonable estimates of the performance of other industry members. In a sense the industry is an information network in which the insiders have an advantage over an outside observer such as the tax authority. If there were no information advantage then the tax authority could work out the profit-maximising decisions and the associated industry equilibrium for itself and audits would become virtually irrelevant. To keep the problem manageable we assume that the industry is assumed to have a fixed number of firms: we do not attempt to account for entry into or exit from the industry. In our formal model it is sufficient to let the number of firms be 2, although this simplification is not essential to the main point of the argument.

(Cowell 2004, Lee 1998).

Tax authority We suppose that tax policy is entrusted to an agency that has the responsibility for enforcement, control over audit policy, and, possibly, over tax design but not over the structure or level of penalties for illegal non-compliance (evasion). Its objectives may be wider than simple revenue raising: this is important in our discussion of efficiency and policy design in Section 4. The tax authority will expect to find that different types of audit policy will have different types of impact on the firms' behaviour. Once again the role of information is crucial because, although the tax authority will not have as good information about an industry as the insiders it will find that there is some information that can be used to refine and improve the audit policy.

Policy evaluation We assume that the tax authority is answerable to a government that cares about the well-being of its citizens. Accordingly policy can be evaluated in terms of welfare-economic criteria that are applied as standard to other problems of public policy evaluation such as cost-benefit analysis.

3 Model

We begin with the factors that determine the firms' taxable capacity. We represent the industry as a duopoly. The essential insights can easily be extended to an arbitrary number of heterogeneous firms: the two-firm model just requires some interpretation: if we focus on the behaviour of firm 1 then "firm 2" can be considered as a proxy for the rest of the industry in the eyes of firm 1's decision makers.

3.1 Firms and industry

We assume that firms make decisions about quantities of a good to produce and sell in a market. Each firm's market opportunities are given by a linear inverse demand schedule:

$$p(q_1 + q_2) := 1 - q_1 - q_2, \tag{1}$$

where p is the the market price of the industry's output given that firm 1 and firm 2 supply quantities q_1 and q_2 to the market. If $K_1(\cdot)$ and $K_2(\cdot)$

are the production-cost functions of firms 1 and 2 respectively, then pre-tax profits are

$$\Pi_1 (q_1, q_2) = p(q_1 + q_2)q_1 - K_1 (q_1), \tag{2}$$

$$\Pi_2 (q_1, q_2) = p(q_1 + q_2)q_2 - K_2 (q_2). \tag{3}$$

In the standard industrial-organisation model this is almost the end of the story. There remains a type of endogenous uncertainty for each firm about the output decisions of the other; this is usually resolved within a standard game-theoretic framework to capture the type of relationship between the economic agents in the industry; here we take each of two apparently standard cases:

- *Cournot*: each firm takes the other's output as fixed while solving its own profit-maximisation problem.

- *Collusion*: the firms act jointly in their decision making.

However, this is not almost the end of the story and, in the present context, these two cases are not quite the standard ones of the industrial-organisation literature. As we will discuss in Section 3.2, the introduction of taxation and the possibility of non-compliance introduce new elements to the profit-maximisation problem.

3.2   Tax regimes and (non-)compliance

Assume that the firms' profits Π_1, Π_2 are the basis for taxation. The tax authority is aware that firms may perceive that their information about the profits that they make in a given year is better than the tax authority's information and that this may give them to under-report or to conceal. If the tax is proportional at rate t and there is full compliance by the firm then the firms' profits. net of taxes are simply

$$[1 - t] \Pi_1 (q_1, q_2), \tag{4}$$

$$[1 - t] \Pi_2 (q_1, q_2). \tag{5}$$

However, these profits are not directly observable by the tax authority without incurring the cost. Instead it receives declarations d_1, d_2 from each of the two firms and it may choose to undertake a costly audit in order to check the

truthfulness of the report. If a firm is found to have under-reported, it is required to make up the shortfall of the tax and also to pay a fine F. The size of the fine is assumed to be outside the jurisdiction of the tax authority – we will assume it to be a fixed proportion of the under-reported profit, $\Pi_i - d_i$ for firm i.

Even with the presence of the fine, less than complete compliance may still be an attractive option for a firm, as discussed in section 3.3. How the firm may be expected to react will depend on the type of audit policy in place and the consequent probability of being subjected to a fine. We assume that the firms are well informed about the audit strategy being used by the tax authority although not about how it will be applied in their own case. In other words all in the industry know how the probability of auditing individual firms is determined but no firm knows for sure that it will be audited. Clearly there is a wide range of possibilities for the structure of audits in the light of firms' behaviour. However, we will focus on just two types of audit policy that are, perhaps, useful caricatures of actual practice and that enable us to analyse the role of information.

Fixed audit rule   The simplest type of audit rule is one where it is common knowledge that there is a given probability β_i^0 that firm i will be audited during the year: the probability does not depend on the reports d_1, d_2. We will use this primitive type of policy as a benchmark.

The relative rule   If the tax authority wants to make use of the imperfect information it has about the industry it could use this to tailor the audit rule for each individual firm in the light of that firm's declaration relative to declarations generally in the industry. The reports from each firm are free information and we can imagine the situation where an intelligent tax authority would use this to flag suspicious behaviour. If there were many similar firms in the industry the tax authority might well concentrate its investigations on individual firms reporting substantially below the industry average. In our two-firm case this translates into a rule where, *ceteris paribus*, one always assigns a higher audit probability to the firm reporting the lower profit. In the case where firms 1 and 2 are indeed similar it is instructive to look at the linear relative audit rule that generates detection probabilities

$$\beta_1(d_1, d_2) = a + b\left[d_2 - d_1\right], \tag{6}$$
$$\beta_2(d_2, d_1) = a + b\left[d_1 - d_2\right]. \tag{7}$$

where a and b are policy parameters. Parameter a reflects the total audit effort by the authority and is determined by its budget: it is the average detection probability for any pattern of declarations by the two firms. Parameter b captures the authority's reactivity: the higher is b the higher is the probability penalty for declaring low profits. To ensure that β_i is not negative and is not greater than 1 we require $b \leq a/\Pi_{\max}$, where $\Pi_{\max}$ is the Cournot profit. Clearly the special case of the fixed audit rule where $\beta_1^0 = \beta_2^0 = a$ can be taken as the limiting form of the relative rule.

3.3 The firm and its behaviour

We assume that firms are concerned just about expected net profits. If the firms declare $d_1 \leq \Pi_1$ and $d_2 \leq \Pi_2$ but are not audited, then their after-tax profits are, respectively,

$$\overline{\pi}_1 (d_1, q_1, q_2) := \Pi_1 (q_1, q_2) - td_1, \tag{8}$$

$$\overline{\pi}_2 (d_2, q_1, q_2) := \Pi_2 (q_1, q_2) - td_2. \tag{9}$$

If they are audited then, on the assumption that the audit immediately uncovers the true value of taxable profits, the firms profits after the tax and fine are, respectively,

$$\underline{\pi}_1 (d_1, q_1, q_2) := \Pi_1 (q_1, q_2) - td_1 - f \left[\Pi_1 (q_1, q_2) - d_1 \right], \tag{10}$$

$$\underline{\pi}_2 (d_2, q_1, q_2) := \Pi_2 (q_1, q_2) - td_2 - f \left[\Pi_2 (q_1, q_2) - d_2 \right]. \tag{11}$$

where f is the proportionate fine rate, which is assumed to exceed t, such that the payment after a successful audit at least covers the evaded taxes.

There is one other element to the problem for which we have not yet allowed. Effective under-reporting, that is, firms' activities leave trails in the product market and elsewhere. Simply reporting profits that are manifestly inconsistent with these evidence trails is not credible so that some sort of explicit concealment activity needs to be involved. This activity is costly and we may reasonably suppose the marginal concealment cost to be increasing in the amount being concealed. Accordingly we let the cost of concealment be represented as

$$C_i := \left[\Pi_i (q_1, q_2) - d_i \right]^2 \tag{12}$$

for $i = 1, 2$. Drawing together expressions (6) to (12) this means that the expected payoffs for the two firms, net of concealment costs, are

$$\beta_1(d_1, d_2)\underline{\pi}_1(d_1, q_1, q_2) + [1 - \beta_1(d_1, d_2)]\overline{\pi}_1(d_1, q_1, q_2) - C_1(\Pi_1(q_1, q_2) - d_1),$$
$$(13)$$
$$\beta_2(d_1, d_2)\underline{\pi}_2(d_1, q_1, q_2) + [1 - \beta_2(d_1, d_2)]\overline{\pi}_2(d_1, q_1, q_2) - C_2(\Pi_2(q_1, q_2) - d_2)$$
$$(14)$$

3.4 Workings of the model

As we noted above, there is rather more to this analysis than a conventional quantity-setting oligopoly. Each firm has two control variables q_1 and d_1 for firm 1, q_2 and d_2 for firm 2. Each firm is directly affected by the choices made by the other. The firm's activities are carried out in two stages:

- *production stage*: This covers the generation of taxable profit and includes production and sales of the product. In the model firms choose q_1, q_2.

- *declaration stage*: This concerns the financial decisions made by the firms: they choose d_1, d_2.

Although we reasonably imagine the declaration stage as being after the production stage, of course the decisions and expected outcomes in the second stage will feed back into decisions made in the first stage. Therefore we can expect that policy instruments that focus on the second-stage financial decisions may have repercussions also on the "real economy" decisions in the first stage. Furthermore, between the production stage and the declaration stage each firm may experience a profit shock, which is observable to the firms in the industry but unobservable to the tax authority. Because the quantities q_1, q_2 have already been chosen at the point when the shock occurs, profit shocks that result from fixed-cost shocks, marginal-cost shocks, demand shocks or observation errors by the tax authority can all be expressed in the same way. We assume that the firms are essentially identical except for the profit shock; in particular they are perceived ex ante as identical by the tax authority when determining its audit rule.

The two stages and the two contrasting market assumptions, Cournot competition or collusion, lead us to consider four possible cases, which we will briefly consider in turn.

Case 1: Cournot competition at both first and second stages.

Here the tax authority has a nice opportunity. Consider the standard model of a symmetric duopoly illustrated in Figure 1 where the two straight lines represent the reaction functions of the two firms. If firms were perfectly compliant on principle, or if the tax authority could perfectly observe market events so that there were no possibility of evasion, then equilibrium would be at point $\left(q_1^C, q_2^C\right)$ (we have $q_1^C = q_2^C$ by symmetry). If the tax authority cannot observe events perfectly and just uses a fixed audit rule then this does not affect the product market so that the reaction functions and equilibrium remain unaltered (Marrelli and Martina 1988). In our model this is the case when b, the *reactivity* of the rule in (6, 7), is zero. Then firms face a declaration-independent detection probability of a. However, if the authority switches to a relative rule it creates an informational externality: each firm knows that its probability of audit is going to depend on its declaration relative to the average declaration in the industry. Two things then happen. First, the switch to the relative rule causes each firm to increase the declaration for any given level of output, for reasons that are straightforward to see intuitively. The reactivity of the relative rule is of special importance here.

Proposition 1 *In a symmetric equilibrium the rule's reactivity b decreases the amount of taxes evaded.*

Proof. See appendix. ∎

Second, there is an effect on the first-stage reaction curves in Figure 1. To see this note that an increase in firm one's quantity q_1 in generally cause the profits of firm 2 to fall, which in turn reduces the optimal declaration of firm 2. Therefore, firm one can indirectly decrease its audit probability by increasing its production quantity. A firm wants to do this up to the point where the own gross profit reduction of a further increase in q_1 outweighs the improved scope for evasion. By this reasoning we can see that the switch in the audit regime will move firm 1's reaction function out to the right as shown. Of course the same effect works for firm 2 and so it is clear that equilibrium output must increase.[3]

We can state this positive effect of a relative rule on the quantity choices more formally.

---

[3]The result is general and does not depend on either the assumption that there are only two firms or the assumption of linear reaction functions (Bayer and Cowell 2009).
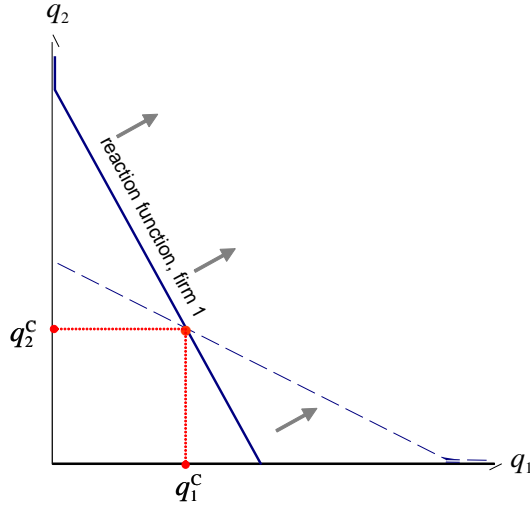
Figure 1: Effect on reaction function of informational externality

Proposition 2 *Under a fixed rule* $(b = 0)$ *in equilibrium firms produce the Cournot quantity, while under a relative rule they produce more than the Cournot quantity.*

Proof. See appendix. ∎

Case 2: Cournot competition at the first stage, collusion at the second stage.

If firms are able to cooperate on tax returns then it is clear that they will aim at eliminating the externality introduced by a relative audit rule. By coordinating their declarations they can avoid the dilemma that both firms have an incentive to increase their declarations in order to reduce the audit probability. Consequently, in the case of collusion at the declaration stage, a relative rule loses its positive effect on declared profits and declarations become the same as under a fixed rule (i.e. the reactivity of the rule b is zero).[4] Consequently, increasing the reactivity of the rule does not help to reduce evasion.

---

[4]This is true in a symmetric environment, where both firms have the same production and evasion cost. In asymmetric situations the declarations may differ from the fixed rule outcome, as collusion provides an additional incentive to minimise aggregate evasion cost.

Proposition 3 *If firms collude on the declaration stage than in a symmetric equilibrium the reactivity of the rule b has no impact on the evaded tax.*

Proof. See appendix ∎

Obviously, this raises the question how collusion at the declaration stage impacts on production decisions. The impact is not obvious. One might expect that eliminating the externality of the relative rule on the second stage also takes away any incentive to produce more than the Cournot quantity. Or even worse, one could conjecture that the collusion at the second stage might spill over to the production stage leading to quantities even smaller than those under Cournot competition. Luckily, these fears are unsubstantiated. Cooperation among firms when they file their tax returns does not fully eliminate the externality on production quantities created by a relative rule. The intuition is subtle. When firms individually decide on their production quantities they foresee already that they will collude on the declaration stage later on. The jointly optimal declarations will depend on the gross profits. As in the case without collusion firm i's optimal declaration (now the one that maximises joint ex-post expected profit) increases with firm $j's$ gross profit. For this reason – with the ultimate outcome in mind – a firm wants to reduce the profit of the competitor by increasing production even when it knows that they will cooperate when filing the tax returns.

Proposition 4 *If firms collude on the declaration stage then a relative rule $(b > 0)$ still leads to quantities greater than the Cournot quantity.*

Proof. See appendix. ∎

A relative rule loses its beneficial effect on evasion behaviour in the presence of collusive tax declarations but still delivers welfare gains in the product market through production quantities beyond the Cournot outcome.

Case 3: Collusion at the first stage, competition at the second stage.

Suppose the firms can agree on total output and some allocation of output and profit between them. Since firms are identical ex ante, assume further that they can only agree on quotas that lead to the same gross profit for both.[5] Once the gross profits are realised the firms independently declare

---

[5]This could be the outcome of Nash bargaining with identical bargaining power and wihtout side payments.

profits. Here the relative audit rule is obviously still effective in reducing the amount of taxes evaded compared to a fixed rule, as the externality stemming from the relative rule is still on operation in the declaration stage.

Proposition 5 *In a symmetric equilibrium with collusion at the declaration stage the rule's reactivity b decreases the amount of taxes evaded.*

Proof. Analogous to the proof of Proposition 1. ∎

It remains to be determined, which production quantities the firms will agree upon on the first stage. Intuition suggests that the relative rule operating on the second stage does not play a role. Since ex-post expected net profits should increase with the gross profits firms should be able to agree on a joint monopoly production plan. The following Proposition confirms this intuition.

Proposition 6 *Duopolists that can enforce a cartel agreement with identical quotas produce half the monopoly quantity each.*

Proof. See appendix. ∎

Case 4: Collusion at both stages.

It is clear that this combination results effectively in monopoly behaviour throughout; the distinction between the stages becomes artificial as does the distinction between the two types of audit rule. Under fairly weak conditions (e.g. symmetric cartel agreements) we know that output and declaration decisions become independent (Lee 1998).[6] If firms behave like one large profit maximising entity with respect to both production and declaration decisions then a relative rule loses all bite. It is worth noting that a relative rule at least does no harm in this highly collusive environment.

4 Audit policy

In the light of the diverse behaviour that will arise from auditor-firm interaction under various competitive regimes there are some important policy implications to be investigated. We will do this in two stages in order to

---

[6]The proofs to Propositions 3 and 6 showing independence of declarations and quantities from b if there is collusion on the respective stage still go through.

separate out pure efficiency objectives from equity considerations: first we will examine the case where the profit shock is vanishingly small so that the firms necessarily appear to the tax authority as identical if they make identical choices; then, in section 5 we will consider the impact of the profit shock.

In what follows we use a simulation to analyse how relative audit rules affect revenue, quantities and evasion cost.

4.1   Case 1 Cournot competition at both stages

We first look at the non-collusion scenario, where we have established that a relative rule increases tax declarations for given profits, but also reduces profits by inducing higher production quantities, which in turn reduces declarations and revenue. So the total effect of a relative audit rule on revenue has two conflicting components: a positive declaration effect and a negative profit effect. The declaration effect is a first-order effect, while the profit effect is only of second order. The declaration effect is anticipated by firms at the quantity-choice stage and production quantities are adapted accordingly. Thus, we can expect that a more flexible rule provides a "double dividend," which consists of an increased production quantity and an increased revenue. To investigate this we simulate the two-stage game using the model of equations (1)-(14) and the assumption that marginal production cost is a constant c.

First, if the authority uses a relative audit rule, what happens to output and tax revenue as the sensitivity of the rule and the tax rate change? Figure 2 shows contour plots of the simulated equilibrium quantity and revenue for $c = 0.1$, $a = 0.25$ and $f = 0.5$.[7] A lighter shading indicates higher values of the quantities and government revenue, respectively.

It is apparent that a more reactive rule increases production quantities and revenue for a given tax rate. The marginal quantity effect of an increase in the sensitivity is decreasing. The marginal revenue effect of the audit sensitivity increases with the tax rate. We see that a more reactive audit rule might lead to higher welfare, as a higher reactivity does not lead to an apparent conflict between the revenue and industry output. The tax rate has an influence on the quantity only if the detection rule is relative. The

---

[7]The   calculation   of   the   equilibrium   is   tedious   and   is   available   on http://darp.lse.ac.uk/taxcompliancefirms
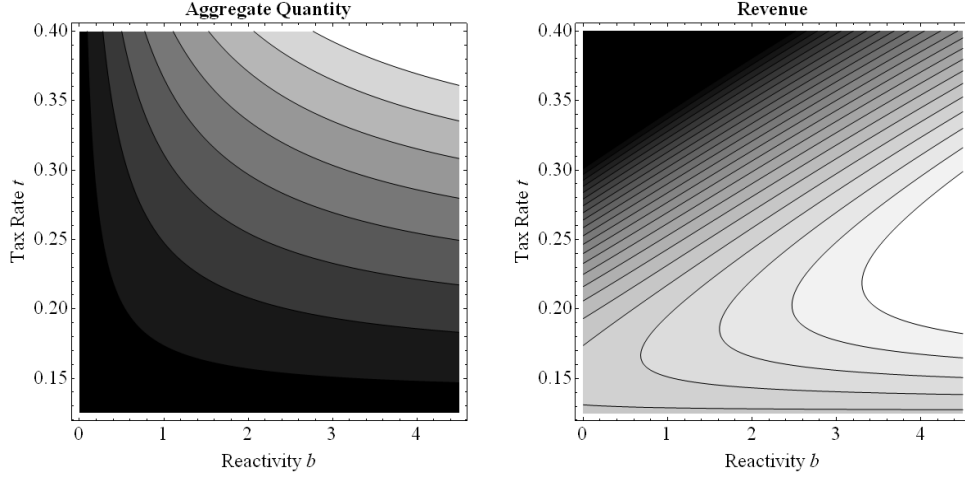
Figure 2: Quantities and revenue in equilibrium for different t and b

tax rate and the sensitivity are substitutes for generating higher outputs if $b > 0$.

For a partial-equilibrium welfare analysis we assume that the social welfare function places equal weight on consumer surplus CS, producer surplus and the revenue available for producing public goods. Then total surplus can be expressed as:

$$\text{total surplus} = CS + \pi_1(d_1^*, d_2^*, q_1^*, q_2^*) + \pi_2(d_1^*, d_2^*, q_1^*, q_2^*) + \text{revenue}$$
$$= \int_0^{q_1^* + q_2^*} [p(x) - c]\, dx - \underbrace{[C_1(d_1^* - \Pi_1(q_1^*, q_2^*)) + C_2(d_1^* - \Pi_2(q_1^*, q_2^*))]}_{\text{waste}}.$$

Figure 3 shows how the wasted resources (due to evasion activity) and the overall surplus depend on the tax rate and on the reactivity of the rule. Lighter shading again represents higher values. We see that a higher tax rate increases the waste, as it provides a higher evasion incentive. Since the evasion incentive is reduced by the reactivity of the rule, waste is reduced by an increase in b.

With respect to the surplus it is apparent that again a high reactivity is helpful. Overall welfare increases for a given tax rate with b. However, the impact of the tax rate is non-monotonic. On the one hand a relative
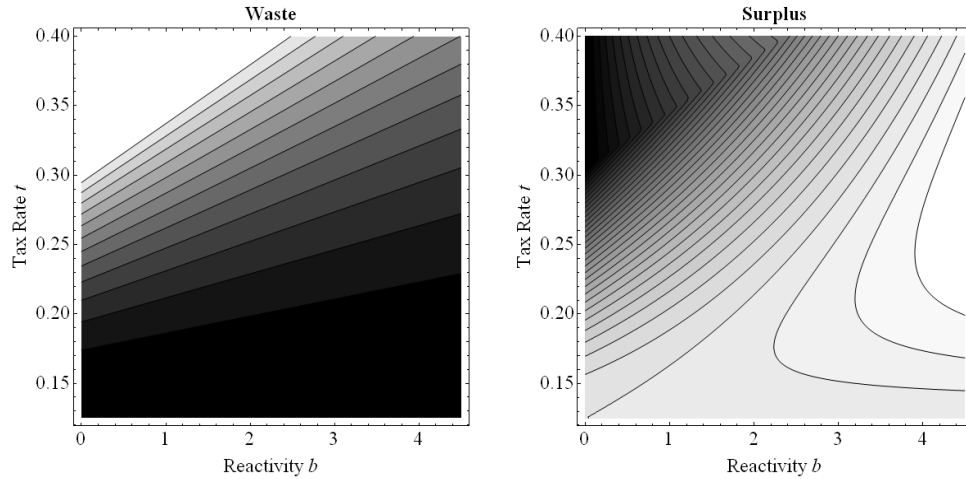
14

Figure 3: Waste and surplus

rule works best at increasing output if the tax is high. The externality on production is the higher the more severe the consequences of evasion are, since the tax rate increases the stakes. On the other hand, a high tax rate increases evasion and therefore the evasion cost: the increased concealment investment represents a waste for the economy. We conclude that for every level of reactivity, there is a tax rate that maximises welfare. Note that the upper left corner of the graph (with the highest waste and low surplus but downward-sloping contours) is the region where firms go underground and declare zero profit.

4.2   Cases 2 and 3: Partial collusion

We have seen so far that welfare increases with the reactivity of the rule if we assume that firms do not collude. Figure 4 shows the total surplus for symmetric equilibria, when the firms collude on the declaration or the production stage.[8]   The chosen parameters are the same as above. It is apparent that increasing the reactivity of the relative rule is still welfare enhancing despite of the presence of collusion at one of the stages. In both

---

[8]Graphs for quantities, revenues, etc. for the collusion scenarios can be found on http://darp.lse.ac.uk/taxcompliancefirms
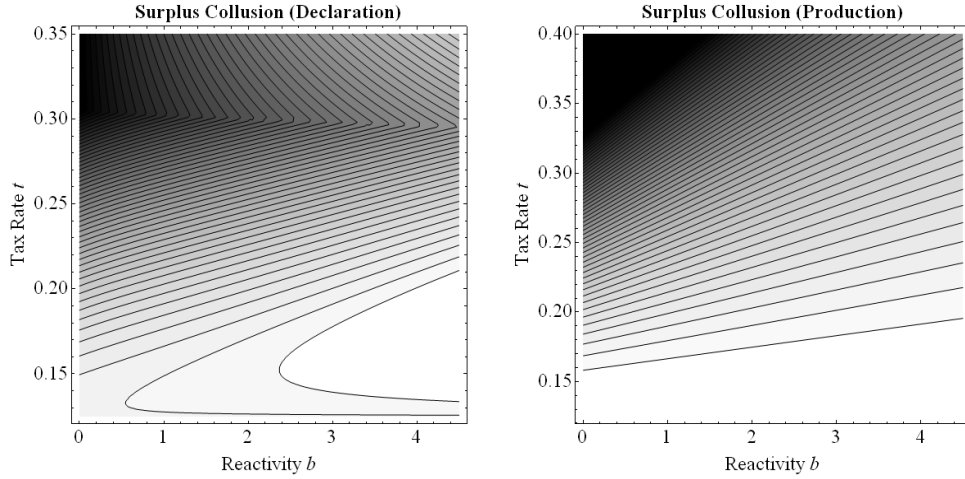
Figure 4: Surplus under collusion

cases, collusion removes one beneficial element of a relative rule. If firms collude on the declaration stage then the externality that reduces evasion (and thus waste) is internalised, while the incentive to produce more than under a fixed rule is preserved, as increasing the quantity improves a firm's position in the following collusive declaration stage. In the case of collusion at the production stage the relative rule has no impact on the quantities produced, as the firms choose jointly gross-profit maximising quantities. The remaining beneficial consequence of a relative rule is the reduction in evasion and wasted resources.[9]

5 Reactivity and inequity

We have seen that increasing the reactivity of the relative rule has a positive effect on efficiency. In this section we investigate if there are adverse distributive effects related to a high reactivity. In what follows we will set up an environment where ex-ante identical firms may end up with different gross profits due to some external shocks, which cannot be observed by the

---

[9]In the panel for collusion on the declaration stage we again have full evasion in the area where the contours are downwards sloping.

authority. The profit shocks result in the ex-ante identical firms becoming different observationally. The tax authority does not observe the shock and still believes that firms are essentially identical. Consequently, the authority attributes any declaration differences between firms to differences in the firms evasion activities. Applying a relative rule then is intended to punish the firm that tries to evade more. However, in the case when differences in declaration actually result from good or bad luck (i.e. the realisation of the profit shock) then the relative rule punishes the firm who had a bad draw, while it rewards the firm that was lucky. The reactivity of the rule might be positively related to the degree of unfairness created by the relative rule.

To develop this argument, we first set up a simple version of shocks within our model and outline the resulting equilibrium (section 5.1). We then investigate how the reactivity of the rule influences allocations and examine whether the reactivity of the rule appears to produce an inequitable outcome (sections 5.2 and 5.3). We make the background assumption that the nominal tax system reflects fairness an investigate the impact of the relative rule on different fairness criteria. Showing that our result holds for multiple measures provides a robustness check. The fairness criteria we use are linked to different distributional measures such as relative tax burden, relative profit after taxes, relative monetary expected net profit after audits and the relative total expected net profit including evasion cost. While we show that the reactivity of the rule has a negative impact on fairness, it is important to note that the positive effect of an increased reactivity on welfare survives the modification to our basic model. An efficiency-equity trade-off arises.

5.1 Shocks and unequal profits

Denote the *interim gross profit* of the firms by:

$$\tilde{\Pi}_i(q_i, q_j, \kappa_i) := \Pi_i(q_i, q_j) + \kappa_i, \ i, j \in \{1, 2\}, i \neq j. \tag{15}$$

where (κ_1, κ_2) are random profit shocks for the two firms with known probability density function φ. Because firms are identical ex ante we require the shock distribution to be symmetrical so that

$$\varphi(a, b) = \varphi(b, a).$$

This restriction ensures that if say a represents a large negative shock, while b is small that then both firms have the same probability of being the firm that suffers the large shock.

For each possible combination of interim gross profits of the two firms there starts a declaration subgame: the subgame-perfect continuation is a pair of declarations for each possible pair of interim profits. The rule for translating interim profits into optimal declarations can be derived from the following maximisation problem.

$$d_i^* \left(\tilde{\Pi}_i, \tilde{\Pi}_j \right) = \arg \max_{d_i} EU_i(d_i, d_j, \tilde{\Pi}_i, \tilde{\Pi}_j), \ \forall i, j \in \{1, 2\}, i \neq j.$$

In the case of the model in Section 3, solving for the optimal declarations depending on the interim profits we have:

$$d_i^* \left(\tilde{\Pi}_i, \tilde{\Pi}_j \right) = \frac{af - t}{2b + f} + \frac{2(1 + bf)}{2 + 3bf} \tilde{\Pi}_i + \frac{bf}{2 + 3bf} \tilde{\Pi}_j. \tag{16}$$

Through this decision rule a firm can foresee its expected net profit for any interim profit pair. Denote the function that maps the interim profits into an expected equilibrium profit by $R_i(\tilde{\Pi}_i, \tilde{\Pi}_j)$ which, in view of (15), is a function of q_1, q_2, κ_1 and κ_2. We now can compactly write down the first-stage maximisation problem of a firm:

$$q_i^* = \arg \max_{q_i} \int \int f(\kappa_i, \kappa_j) \cdot R_i(q_i, q_j, \kappa_i, \kappa_j) d\kappa_i d\kappa_j, \ \forall i, j \in \{1, 2\}, i \neq j.$$

So, anticipating the declaration decision, a firm will choose the quantity that maximises the expected net profit for a given quantity of the other firm. The expectation is taken over the possible realisations of the fixed-cost shocks for both firms. The explicit solution to this maximisation problem depends on the distribution of the shocks. It is important to note that the decision problem for both firms is identical, since the firms are identical ex ante (including the symmetry of the shocks). Consequently, if there exists at least one pure-strategy equilibrium then there is always a symmetrical equilibrium. Furthermore, if we have a unique pure-strategy equilibrium than this is the symmetric one. For simplicity we concentrate on symmetric pure-strategy equilibria. In such an equilibrium both firms choose the same production quantity, i.e. $q_1^* = q_2^* = q^*$.

As a consequence firms' interim profits only differ with respect to the realisations of the fixed-cost shocks:

$$\Delta \tilde{\Pi} = \left| \tilde{\Pi}_i - \tilde{\Pi}_j \right| = |\kappa_i - \kappa_j| = \Delta \kappa.$$

In what follows we use the findings from above in order to judge if according to some criteria an increased reactivity has adverse effects on equity considerations.

18

5.2 Reactivity and the relative tax burden

A natural criterion for fairness is the relative tax burden. Taking the linear profit-tax scheme as the basis for fairness intentions by the legislator then a firm that has one dollar more profit than another should pay t dollar more in taxes. So suppose that a cost shock has led to the two firms having interim profits, which differ by $\Delta\kappa$. Note that we know from above that the gross profit excluding the shocks are the same for both firms. Then we can use (16) in order to calculate the difference in tax liabilities $t\Delta d$ after the declaration:

$$t\Delta d = t\frac{2+bf}{2+3bf}\Delta\kappa. \tag{17}$$

For a fixed rule with $b=0$ the difference between the tax bills is just equal to the gross profit difference multiplied by the tax rate. Therefore a fixed rule satisfies the criterion of fair relative tax burdens: a firm with one dollar more in profits pays t dollars more in taxes. The difference in (17) declines when b increases. For a given situation and a given profit difference an increase in the reactivity of the rule has a potentially unwanted distributional side-effect. The profit gap has a smaller impact on the differences in tax payments than intended by the tax law. This favours the firm that by chance ended up with a higher interim profit. It will be able to exploit the relative rule, as the authority directs more resources to the firm with the lower profit, since that firm looks more suspicious. Consequently, the firm will pay less than the intended $100t$ cents more per dollar of extra profit. The gap between the intended relative tax burden and the effective tax burden (before auditing takes place) widens with the reactivity of the rule; it is straightforward to extend this analysis to the difference in net profits before auditing. The difference is

$$\Delta\tilde{\Pi} - t\Delta d = \left(1 - t\frac{2+bf}{2+3bf}\right)\Delta\kappa,$$

where $(1-t)\Delta\kappa$ is intended by the legislator. The firm with the higher gross profit will enjoy a larger net profit gap to the firm with the lower profit than intended whenever the rule has positive reactivity.

This advantage enjoyed by the firm with higher gross profit suggests that the tax treatment of the two firms could be regarded as unfair. What is more, unlike the case with the taxation of personal incomes where individual taxpayers probably do not know the incomes and tax assessments of other individuals, in the case of firms it is reasonable to assume that firms know

19

each others' circumstances, know their competitors' profits and may know how other firms are being treated for tax purposes; the tax authority, although not able to observe the circumstances and profits of firms accurately will be aware that this information is common knowledge within the industry and will thus be aware that their audit policy may generate a perception of unfairness. Based on the difference in net profits (before audits take place) we can define an *unfairness measure* capturing the relative advantage of a firm earning more than the other firm. Our first measure ϕ_1 takes the net profit gap, normalises the underlying gross profit difference to one dollar and subtracts the fair gap of $1 - t$:

$$\phi_1 = t\left(1 - \frac{2 + bf}{2 + 3bf}\right)$$
$$= \frac{2bft}{2 + 3bf}.$$

Clearly ϕ_1 measures by how many dollars the resulting net profit gap exceeds the fair gap. More precisely ϕ_1 measures the gap between the two firms' after tax profits per dollar of gross profit difference over and above the gap intended by tax law, which is $1 - t$. Unfairness according to this measure increases with the reactivity of the rule. A fixed rule (i.e. $b = 0$) does not result in any unfairness, as ϕ_1 becomes zero in that case. For a given positive reactivity distributional unfairness increases with the tax rate.

5.3 Reactivity and relative expected net payoff after auditing

We have seen that the firm with a coincidentally higher gross profit can exploit the fact that the relative rule does not take this coincidence into account and therefore considers this firm as less suspicious. Does this unfair advantage prevail if we take into account that the lucky firm will have to pay higher fines if caught. The intended difference in expected net profit after auditing $\Delta\tilde{\Pi} - t\Delta d - \Delta EF$, where ΔEF denotes the expected fine, can be calculated as

$$\Delta\tilde{\Pi} - t\Delta d - \Delta EF = \left(1 - t\frac{2 + bf}{2 + 3bf} + \frac{2bf(2af - t)}{2 + 3bf}\right)\Delta\kappa$$
$$= \frac{bf(3 - 4af + t) + 2(1 - t)}{2 + 3bf}\Delta\kappa$$

20

Taking the same approach from above and expressing the gap per monetary unit and normalising by subtracting the fair gap we get a measure for the unfair advantage of the more profitable firm for the expected post audit profit:

$$\phi_2 = \frac{4b(t - af)}{2 + 3fb}.$$

The difference between ϕ_2 and ϕ_1 is that ϕ_2 takes into account two additional factors, the higher fine the richer firm has to pay if caught but also the reduced audit probability of the richer firm stemming from a higher declaration than that of the less profitable firm. This wider definition of fairness leaves our qualitative result from above unchanged. As for our first measure a fixed rule ($b = 0$) does not lead to unfairness, since then the expected profit gap (including auditing and fines) is exactly $1 - t$ as intended by the legislator; ϕ_2 is equal to zero. The profit gap increases with b.[10] So also if expected net profit after an audit is the criterion an authority adheres to when it comes to fairness then the reactivity of the rule has a negative impact on distributional fairness. The higher b the larger is the deviation from the intended net profit difference of $1 - t$ per dollar. As with ϕ_1 the tax rate is positively related to distributional unfairness if the rule is relative.

Reactivity and fairness when evasion cost are considered

Typically, one would expect that authorities are not too concerned about firms' costs that arise only from evasion activity. In the present case it might make sense though, since including the evasion cost provides a good robustness check of our qualitative results. A higher reactivity of the audit rule increases allocative efficiency but might cause some distributional concerns, as we have discussed. The distributional inequity comes from the fact that a relative rule does not take into account unobserved profit differences. Hence, a firm with a coincidentally higher profit has an advantage, since the rule treats the higher declaration as a sign of the firm evading less, whereas it is rather a sign of the firm having a higher profit. The firm can exploit this by evading more taxes. With evading more taxes the firm will also have to incur higher real resource cost arising from the evasion activity. These costs are not included in either of the two measures we investigated previously. Here

---

[10]To see this take $\partial \phi_2 / \partial b$ and note that this derivative is positive if $t > af$, which is the condition for evasion taking place.

21

we ask if the reactivity of the rule still leads to distributional concerns if we take into account evasion costs. So we are interested in the difference of total equilibrium net profits of two firms who accidentally ended up with different payoffs ΔEU. We construct the unfairness measure ϕ_3 that includes the evasion cost in the same way as above (normalising $\Delta\kappa$ to one and subtracting the fair gap $1 - t$):

$$\phi_3 = \frac{f(1+bf)}{(2+bf)} \frac{4b(t-af)}{(2+3bf)}.$$

Measure ϕ_3 is closely related to ϕ_2, as it only differs by a factor of $f(1 + bf)/(2 + bf)$. Not surprisingly, also for this measure a fixed audit rule does not lead to any distributional distortions, since $\phi_3 = 0$ for $b = 0$. However, even if one takes into account the evasion cost an increase in the reactivity leads to greater unfairness, since ϕ_3 increases in b

$$\frac{\partial\phi_3}{\partial b} = \frac{4f(4 + bf(8 + 5b)(t - af))}{(2+bf)^2(2+3bf)^2} > 0 \text{ for } t > af,$$

where $t > af$ is the condition for evasion to take place.[11] As for the other measures for a given positive b an increase in the tax rate increases unfairness.

6   Policy

Some policy consequences are clear. In the absence of shocks a government that wants to maximise welfare should set the reactivity to its maximum level. This finding does not depend on our specific example and it does not depend on what we assume about competition or collaboration.

By contrast the tax rate that maximises welfare for the maximum b does depend on the specific example and on the competitiveness of choices at each of the the two stages. Figure 5 plots the highest optimal tax rate depending on the reactivity of the audit rule and the particular assumption on collusion for the parameter values from above. We consider the tax rate as optimal if it maximises the surplus given the parameters and the form of competition. Observe that given our parameter values a tax rate below 0.125 leads to truthful profit declarations in all three scenarios. Tax rates between zero and 0.125 yield the same surplus in all scenarios. In Figure 5 we only consider

---

[11] If the evasion condition is not satisfied then reactivity has no influence on efficiency nor on the distribution.

the highest optimal tax rate, which is the one that yields the highest revenue among those that maximise surplus.

In all three scenarios under a fixed rule ($b = 0$) a tax rate of 0.125 (or below) is optimal, as a fixed rule has no effect on allocative efficiency in the goods market. Therefore, welfare is maximised, when the wasted resources for covering evasion is minimised. The optimal tax rate then should prevent evasion. Now turn to a relative rule. The only scenario where the relative rule has no positive influence on the welfare created in the goods market is when there is collusion at the production stage. Consequently, the optimal tax rate in the presence of cartels should again minimise the wasted resources from evasion. Independent of the reactivity of the rule a tax rate of or below 0.125 is optimal here, since then evasion does not occur.

In the case of collusion at the declaration stage there is a positive effect of the relative rule on production quantities. This effect is the stronger the more reactive the rule and the higher the tax rate. The trade-off between the additional waste from increased evasion caused by a higher tax rate and the increased allocative efficiency in the goods market yields an increasing relationship between the reactivity of the rule and the optimal tax rate.

If firms collude neither on the declaration stage nor on the production stage then a more reactive rule does not only increase welfare in the goods market but also ceteris paribus decreases evasion and waste. It follows that the waste does not increase as rapidly with the tax rate as under collusion at the declaration stage. This implies that the optimal tax rate is higher and increases more strongly with the reactivity of the rule when competition rules on both stages.

To summarise our insights from the welfare simulations, we conclude that a relative audit rule has widespread advantages over a fixed rule. Furthermore, a more reactive rule is usually preferred, as it leads to welfare gains.

What are the implications for tax-enforcement policy of a change in competitiveness of the industry? As long as firms do not collude in the goods market, a higher reactivity allows for higher tax rates without damaging welfare. So, increasing the reactivity has another desirable effect for governments.

On the other hand it might be argued that conventional welfare analysis reveals a feature of the model that should be considered slightly unpleasant. Typically; if one assumes that firms are essentially identical, it is optimal to spend all the audit resources on the firm that reports the lowest profit ($b \rightarrow \infty$). It might be thought that this somewhat extreme position is like
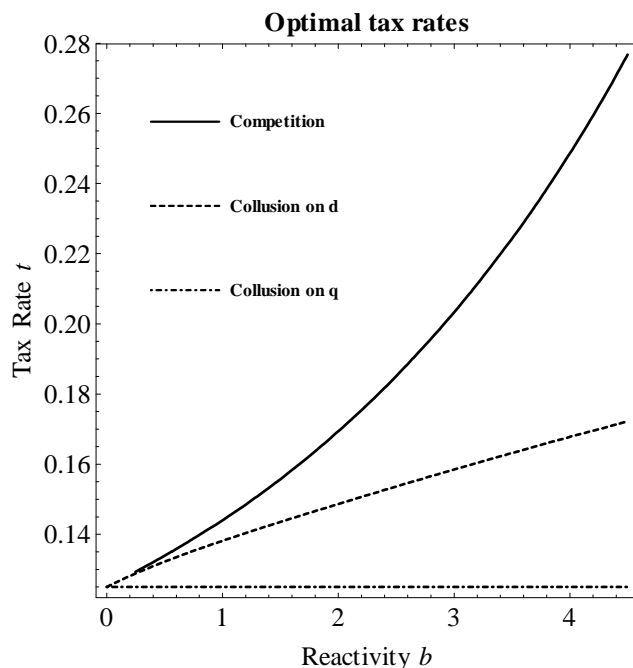
Figure 5: Optimal tax rates for different scenarios

the old question about severity of punishment versus probability of detection as a deterrent to a rational tax evader (because increasing the fine appears to have no resource cost whereas increasing the detection probability does, it appears to be optimal to require the death penalty for the slightest amount of tax evasion – Kolm 1973, Cowell 1989). However, the analogy is not appropriate. If there are errors in information or in administration then the "Death to Tax Evaders" policy produces awful outcomes for those who are innocent; but in our model such errors may just mean that the tax authority is focusing resources on small fry and letting some big corporate evaders off the hook. But the innocent are not suffering and no agent is faced with a threat of infinite penalty.

These points illustrate the limitations of the conventional individualistic welfare approach in the present case. Because we are not dealing with distributional outcomes for individual persons conventional welfare-based approaches to inequality are not applicable. Nevertheless it still makes sense to discuss fairness or equity issues in terms of differences in outcome (net

24

profits) and differences in treatment arising from the audit rule. What underlies this is the first form of uncertainty described in section 2; what can make it seem unfair are the second and third forms of uncertainty: external shocks make like entities appear unalike and may provide the basis for a relatively favoured firm to exploit the relative audit rule to its own advantage. Concern for this unfairness of treatment will impose a limit on the optimal choice of b.

7   Conclusions

This paper has shown that the design of a tax authority's audit policy can have important effects on production decisions by firms.   The   nature of those effects depends on whether firms compete or collude. Accordingly an appropriately designed audit policy may not only achieve greater compliance and higher net revenue for given output and resources spent on audit but may also have other effects that would be normally be considered desirable in a wider economic context. By a smart design of the audit policy the authorities can create an informational externality that partially offsets the informational advantages of industry insiders. Since decisions in the product market are taken in the light of the eventual outcome of net after-tax expected profits the audit policy can create a linkage to output decisions: specifically it may be possible to nudge firms in the direction of greater efficiency.

We have shown that, if firms are essentially similar then it pays the authority maximise "reactivity": to focus its audit resources on the firm making the lower declaration, even though non-compliance may be more widespread. Depending on whether there is competition or collusion in the product market then not only does high reactivity "pay" the tax authority in the narrow sense, it also enhances economic welfare.

References

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

Bayer, R.-C. and F. A. Cowell (2009). Tax compliance and firms' strategic interdependence. *Journal of Public Economics 93*, 1131–1143.

Besfamille, M., P. De Donder, and J.-M. Lozachmeur (2009). Tax enforcement may decrease government revenue", economics bulletin, vol. 29 no.4 pp. 2665-2672. *Martin Besfamille and Philippe De Donder and Jean Marie Lozachmeur, (2009) "Tax enforcement may decrease government revenue", Economics Bulletin, Vol. 29 no.4 pp. 2665-2672. 29,* 2665–2672.

Cowell, F. A. (1989). Honesty is sometimes the best policy. *European Economic Review 33*, 605–617.

Cowell, F. A. (1990). *Cheating the Government.* Cambridge, Massachusetts: The MIT Press.

Cowell, F. A. (2004). Carrots and sticks in enforcement. In H. J. Aaron and J. Slemrod (Eds.), *The Crisis in Tax Administration*, pp. 230–275. Washington DC: The Brookings Institution.

Kolm, S.-C. (1973). A note on optimum tax evasion. *Journal of Public Economics 2*, 265–270.

Lee, K. (1998). Tax evasion, monopoly and nonneutral profit taxes. *National Tax Journal 51*, 333–338.

Marrelli, M. and R. Martina (1988). Tax evasion and strategic behaviour of the firms. *Journal of Public Economics 37*, 55–69.

A   Proofs omitted in the main text

A.1   Proposition 1

Proof. The first-order condition for optimal declarations are

$$\frac{\partial}{\partial d_i} EU_i(d_i, d_j, \Pi_i, \Pi_j) = 0, \ \forall i, j \in \{1, 2\}, i \neq j$$

$$\Rightarrow d_i^*(d_j) = \frac{f(a + bd_j) - t + (2 + bf)\Pi_i}{2(1 + bf)}.$$

Best responses are linear with a slope smaller than 1, which ensures uniqueness of an interior declaration equilibrium for any admissible pair of gross profits. Solving for the optimal interior declaration we get:

$$d_i^* = \frac{af - t}{2b + f} + \frac{2(1 + bf)}{2 + 3bf}\Pi_i + \frac{bf}{2 + 3bf}\Pi_j. \tag{18}$$

As firms are ex ante identical we can use symmetry (i.e. $\Pi_j = \Pi_i$) on the quantity stage. The equilibrium declaration becomes.

$$d_i^* = \frac{af - t}{2b + f} + \Pi_i,$$

which implies that an interior solution requires $t > af$. The amount of tax evaded is

$$t \left(\Pi_i - d_i^* \right) = \frac{t \left(t - af \right)}{2 + bf}.$$

Taking the derivative with respect to b gives the desired negative effect of b on the taxes evaded.

$$\frac{d}{db} t \left(\Pi_i - d_i^* \right) = ft \left[\frac{af - t}{\left(2 + bf \right)^2} \right] < 0. \tag{19}$$

■

A.2 Proposition 2

Proof. Take the first order condition for the optimal quantity at the gross profit maximum of firm i (i.e. $\partial \Pi_i / \partial q_i = 0$):

$$\left. \frac{\partial}{\partial q_i} EU_i \right|_{\partial \Pi_i / \partial q_i = 0} = \frac{\partial \Pi_j}{\partial q_i} \left[bf \left(\Pi_i - d_i^* \right) \left(\frac{\partial d_i^*}{\partial \Pi_j} - \frac{\partial d_j^*}{\partial \Pi_j} \right) - t \frac{\partial d_i^*}{\partial \Pi_j} \right]. \tag{20}$$

From (18) we know that

$$0 < \frac{\partial d_i^*}{\partial \Pi_j} = \frac{bf}{2 + 3bf} < \frac{\partial d_j^*}{\partial \Pi_j} = \frac{2(1 + bf)}{2 + 3bf},$$

which together with $\partial \Pi_j / \partial q_i < 0$ implies that

$$\left. \frac{\partial}{\partial q_i} EU_i \right|_{\partial \Pi_i / \partial q_i = 0} \begin{cases} = 0 & if \quad b = 0 \\ > 0 & if \quad b > 0 \end{cases}.$$

Consequently, under $b = 0$ the duopolists set $\partial \Pi_i / \partial q_i^* = 0$, which yields the Cournot solution. Under a relative rule a duopolist's optimal quantity given the quantity produced by the opponent is greater than Cournot best-response to the opponent's quantity, i.e. $\partial \Pi_i / \partial q_i^* < 0$. Exploiting symmetry we know that the equilibrium quantity lies on the 45-degree in the q_i, q_j space. Since the best-response quantity is greater than the Cournot best response quantity for any quantity of the opponent, the best-response function crosses the 45-degree line on a point above and to the right of the Cournot solution.

■

27

A.3 Proposition 3

Proof. Joint profit maximisation on the declaration stage leads to declarations such that

$$(d_1^*, d_2^*) = \arg\max_{d_1, d_2} \left[EU_1(d_1, d_2, \Pi_1, \Pi_2) + EU_2(d_2, d_1, \Pi_2, \Pi_1) \right].$$

Taking first-order conditions with respect to d_1 and d_2 and solving gives

$$d_i^* = \frac{(1 + 2bf)(af - t) + (2 + 3bf)\Pi_i + bf\Pi_j}{2 + 4bf} \forall i, j \in \{1, 2\}, i \neq j \quad (21)$$

Invoking symmetry on the production stage (i.e. $\Pi_i = \Pi_j$) the optimal declaration under collusion at the declaration stage becomes

$$d_i^* = \frac{af - t}{2} - \Pi_i.$$

The evaded tax per firm is

$$t(\Pi_i - d_i^*) = \frac{t(t - af)}{2},$$

which does not depend on b. ∎

A.4 Proposition 4

Proof. Take the first-order condition for firm $i's$ quantity and evaluate it at it's Cournot best-response

$$\frac{\partial}{\partial q_i} EU_i \bigg|_{\partial \Pi_i / \partial q_i = 0} = \frac{\partial \Pi_j}{\partial q_i} \left[bf(\Pi_i - d_i^*) \left(\frac{\partial d_i^*}{\partial \Pi_j} - \frac{\partial d_j^*}{\partial \Pi_j} \right) - t \frac{\partial d_i^*}{\partial \Pi_j} \right],$$

which is the same as (20). However, the reaction of the optimal declarations to changes in the gross profits are different in the case of collusion. They can be taken from the jointly optimal declarations in (21):

$$0 < \frac{\partial d_i^*}{\partial \Pi_j} = \frac{bf}{2 + 4bf} < \frac{\partial d_j^*}{\partial \Pi_j} = \frac{2 + 3bf}{2 + 4bf}.$$

By the same logic as in the proof of Proposition 2 we have

$$\frac{\partial}{\partial q_i} EU_i \bigg|_{\partial \Pi_i / \partial q_i = 0} \begin{cases} = 0 & if \quad b = 0 \\ > 0 & if \quad b > 0 \end{cases}$$

and consequently equilibrium quantities are greater than the Cournot quantity if $b > 0$. ∎

A.5 Proposition 6

Proof. First observe that gross profits will be identical when we restrict the firms to the use of only symmetric cartel agreements. Anticipating this the two firms will know that they both will declare the same profit, which means that they will end up audit probability a. We can write the joint profit as

$$E\Pi_i(d^*(Q), Q) + E\Pi_j(d^*(Q), Q)$$
$$= 2\left[\Pi(Q) - td^*(Q) - af\left(\Pi(Q) - d^*(Q)\right) - \left(\Pi(Q) - d^*(Q)\right)^2\right],$$

where $\Pi(Q)$ is one firm's profit for total output $Q = 2q$. Taking the first-order condition and dividing by two gives

$$\frac{\partial \Pi}{\partial Q}\left[1 - af - 2\left(\Pi(Q) - d^*(Q)\right)\right] - \left(t - af - 2\left(\Pi(Q) - d^*(Q)\right)\right)\frac{\partial d^*}{\partial Q} = 0. \tag{22}$$

Now observe that according to the second-stage declarations form (18)

$$\frac{\partial d^*}{\partial Q} = \frac{2(1 + bf)}{2 + 3bf}\frac{\partial \Pi}{\partial Q} + \frac{bf}{2 + 3bf}\frac{\partial \Pi}{\partial Q} = \frac{\partial \Pi}{\partial Q},$$

which can be used to write (22) as

$$\frac{\partial \Pi}{\partial Q}\left[1 - t\right] = 0,$$

which implies that the joint monopoly quantity is chosen (i.e. where $\partial \Pi / \partial Q = 0$). ∎