peuk

Working Paper No. 4

TAX COMPLIANCE AS A SOCIAL NORM AND THE DETERRENT EFFECT OF INVESTIGATIONS

Marisa Ratto
University of Bristol

Richard Thomas
HM Revenue and Customs

David Ulph
HM Revenue and Customs

Tax compliance as a social norm and the deterrent effect of investigations.*

Marisa Ratto† Richard Thomas‡ David Ulph§

July 2005

Abstract

In this paper we focus on the effects of investigations on tax compliance. In a very general model we explain the direct and indirect effects of investigations and analyse taxpayers' response to an increase in the probability of audit when tax compliance is a social norm. We define the different elements that determine the impact of audits on compliance and show that if tax compliance is a social norm in the relevant community there is an additional effect arising because of social norm considerations. The behavioural response of taxpayers to an increase in the audit rate is stronger.

Our findings help explaining seemingly contradictory results that emerge from the empirical evidence.

Keywords: tax evasion, social norm, opportunities to evade, optimal audit rule.

JEL: D81, H26, H30, K42.

---

*This paper benefitted from the contributions of participants at the IFS Public Economics Working Group (Warwick, 2004), the International Conference on Public Economic Theory PET04 (Beijing), the Royal Economic Society annual conference 2005 (Nottingham) and the CMPO Workshop (Bristol). All remaining errors are ours.

†Centre for Market and Public Organisation, University of Bristol.

‡KAI: Analysis, HM Revenue & Customs, Somerset House, London.

§KAI: Analysis, HM Revenue & Customs, 1 Parliament Street, London.

1 Introduction

Different aspects affect taxpayers' willingness to evade: attitudes towards tax compliance are influenced by opportunities to evade and personal circumstances that may impact on the cost of compliance, by people's perceptions of how they are treated by the tax authority[1] and their perceptions of the enforcement system, and also by interactions with other taxpayers. An important decision that the tax authority has to make is how to allocate investigation resources among different groups of taxpayers. For this it is crucial for the tax authority to know how taxpayers are affected by audits.

Recent empirical studies have attempted to estimate the impact of audits on tax compliance. These studies, mostly based on US data, distinguish between the direct and indirect effects of investigations. The direct effect is described as the additional revenues in the form of unpaid taxes and fines collected by the tax authority through investigations. The indirect effect relates instead to the behavioural response of taxpayers to a change in the audit policy and measures the increase in tax compliance induced in the whole community of taxpayers, not only within those taxpayers being investigated.

The reported estimates on the effects of audits on compliance are quite different, but difficult to compare, as these studies use different data sets. When estimates for the direct and indirect effects are provided, a common result is that indirect effect tend to be much higher than the direct effect. A clear understanding of the direction and magnitude of these effects is crucial to inform the decision of the tax authority on how to optimally allocate investigation resources among different groups of taxpayers. However, the methodology used in the empirical studies to calculate the indirect and direct effects does not allow to distinguish the determinants of these effects and hence very little can be said on how these effects vary across different groups of taxpayers.

In this paper we formalise the concepts of direct and indirect effects of investigations, which seem to have been neglected by the theoretical literature. The theoretical models on tax evasion have analysed the overall impact of investigations without distinguishing between direct and indirect effects. In the standard portfolio models[2] the tax authority sets the probability of detection, the tax rate and the fine rate independently from the taxpayer's decision. There is no interaction between the tax authority and the representative taxpayer and the tax parameters are fixed, chosen independently from taxpayers' behaviour. Later contributions have analysed, by use of game theoretical models, the interaction between taxpayers and the tax authority. The assumption made in those models is that the choice of the tax parameters depends on the extent of evasion, in that

[1] Kristina Murphy (2003) shows that the way taxpayers feel treated and whether they feel they are trusted by the tax authority is crucial in determining taxpayers' behaviour.

[2] Allingham and Sandmo (1972), Yitzhaki (1974).

taxpayers' decisions have an impact on the tax revenues raised by the Government[3]. However these models consider the overall response of taxpayers to the audit policy, without distinguishing between direct and indirect effects.

In a very general model, we derive the rule for the optimal allocation of resources across different groups of taxpayers and identify the different elements that determine the direct and indirect effects of investigations. We then derive a formula for the ratio between indirect and direct effects. If the ratio were constant across different groups of taxpayers, there wouldn't be any need to know the values of the different components. However, we show that there is no reason to assume the ratio to be constant across different groups of taxpayers. The ratio depends on how intensively a given group of taxpayers is audited, on the effectiveness of investigations and on the behavioural elasticity of taxpayers. The first two factors do indeed vary across taxpayers. For the elasticity of evasion this is not clear. However, even if the elasticity of evasion were equal across different groups of taxpayers the ratio between indirect and direct effects would still vary across different groups of taxpayers and fiscalities. Hence measuring the single components of the ratio is very important for identifying the optimal allocation of investigation resources. In the empirical studies we briefly review the estimates for the elasticities are quite different and it is not clear what is the underlying assumption on the motivations driving tax compliance. The difference in the significance and magnitude of the behavioural response to audits emerging from these studies may be due to the different aspects of individual behaviour that may be captured by the different datasets. In particular, individual behaviour might not only be affected by purely individualistic calculus but might also reflect some sort of group norms, i.e. the compliance decision of an individual may depend on the proportion of taxpayers within their reference group who are honest. If this is the case individual level data will not capture this link and hence the deterrent effect of investigations might be underestimated. With this in mind we compare three different settings: one where taxpayers are purely motivated by selfish attitudes, one where they also attach a moral dimension to tax compliance and a third context where tax compliance is a social norm and analyse the impact of a rise in the audit rate on aggregate voluntary compliance. We show that if tax compliance is a social norm in the relevant community, the importance attached to the social norm affects overall voluntary compliance as well as the response to an increase in the audit probability. In particular, the more people care about the social norm the greater the reduction in total evasion for a given increase in the audit rate. This implies that the greater the importance attached to the social norm, the greater the elasticity of evasion calculated at the community level.

[3] See Reinganum and Wilde (1984),(1991), Graetz et al. (1986) and Cremer et al. (1990), Greenberg (1984).

In the following section we discuss the available empirical evidence on the effects of investigations on tax compliance. In section 3 we derive a formalisation of the concepts of direct and indirect effect and separate out the different elements characterising the two concepts. In section 4 we consider three possible ways taxpayers may be motivated whether or not to comply and show the impact of a rise in the audit rate on aggregate evasion. Section 5 presents some simulations on how the optimal allocation of investigation resources varies within two groups of taxpayers who have different evasion elasticities and for different budgets of the enforcement agency. Finally section 6 concludes.

2 Evidence on the effects of investigations.

The empirical evidence on the impact of audit rates on taxpayers' compliance is quite recent and rather thin. These studies analyse the effects of audits on reported income and tax liabilities. A distinction is made between the direct and indirect effects of investigations and the behavioural response of taxpayers is measured in terms of the elasticity of reported income and reported tax to the audit rate. Estimates of these elasticities are quite different as the estimates of the ratio between indirect and direct effects. We present the results of six studies, two of which are based on individual level data.

Beron et al.(1992) analyse the effects of audits and socioeconomic variables on compliance. The authors distinguish two compliance decisions: income reports and reports of subtractions on the tax return. This for two main reasons. First the Internal Revenue Service assumes that the probability of detection is much higher for subtractions than income reports. Secondly the behaviour of taxpayers may differ for these two types of compliance activities. In particular the authors believe that non-compliance from underreports is often an act of omission, whereas non-compliance from overstatement of subtractions requires actual misstatements. They use US data from individual tax returns for 1969 filed in 1970. These include reported Adjusted Gross Income (AGI), reported total tax liability and number of returns filed. The information on audit is from a different database (the IRS's Project 778) and relates to audits performed in 1969. Socio-economic variables are measures of income and tastes and preferences from the 1970 Census of Population and Housing. All data are aggregated to the three-digit zip code to match the data available on audits.

Each equation is estimated for five different audit classes (low-income tax returns taking the standard deductions, low-income tax returns taking itemised deductions, low-income proprietor tax returns, middle income wage and salary workers and middle income proprietor tax returns).

Results suggest that increasing the odds of an audit significantly increases reported

AGI for two groups (low-income tax returns taking the standard deduction and low-income proprietor tax returns) and tax liabilities for three groups (low-income tax returns taking the standard deduction, low-income proprietor tax returns and middle income proprietor tax returns). The magnitudes of the effects of audits on tax compliance are in general modest and the elasticities for reported AGI are smaller than the elastiticies for reported tax liability for all five audit classes[4]. The authors do not calculate the direct and indirect effects of investigations.

Another study based on individual level data is by Erard (1992). The author describes the indirect effects of investigations as the additional future tax revenues that arise from an increase in the audit rate due to improved compliance by two different groups of taxpayers: those who have been audited and respond to the examination experience by improving their compliance behaviour and those who are not audited but nevertheless perceive a greater threat of examinations. In order to estimate the magnitude to these two sources of indirect revenue gains, Erard focuses on the effects of tax audits on subsequent years reporting behaviour. The author uses two US data sources from IRS on taxpayers who were the targets of an audit in one year and, purely by chance, were subject of a second audit two years later. The first data source is the 1982 IRS TCMP Survey, which contains results of thorough, line-by-line, audits of a large stratified random sample of 1982 federal income tax returns and information about the prior year return characteristics of all taxpayers in the TCMP sample, including whether their previous two tax returns were audited. The second data source is the 1985 TCMP Survey. Information on the prior audit history of taxpayers was obtained through a social security number match with IRS records on audits of 1983 and 1984 tax returns. This second data source contains complete information on prior audit assessments as well as the disposition of prior audit cases. Two approaches are used to examine the influence of a tax audit on subsequent year reporting. The author first examines if those taxpayers who experienced a large audit assessment improve their compliance in a subsequent year, but finds inconclusive results. In the second approach the author investigates whether those taxpayers who experienced a prior audit differ in their subsequent year reporting behaviour from taxpayers who did not experience a prior audit, after controlling for taxpayers characteristics and the prior year audit selection process. The estimates obtained from the regressions are not statistically significant and the author shows that the findings are highly sensitive to the assumptions and specifications imposed. One problem pointed out by the author is a sample selection bias: unobserved factors influencing prior year audits may be correlated with the unobserved factors influencing subsequent year tax non-compliance. Although the hypothesis of sample selection bias seems to be rejected by the data, if the assumption

[4]These latter vary between 0.19 and 0.31 when significant.

of sample selection bias is imposed and the model is re-estimated, the coefficients of the prior audit become significant.

We would like to remark that the underlying assumption of Erard is that taxpayers know about the audit rule of the tax authority and even if they are not audited they are able to perceive the change in the audit rate. In reality taxpayers may assume that audits are carried out at random and believe that getting an audit in the current period does not mean getting another one in the future. Hence they might not change their behaviour after being investigated or perceiving a higher threats of audit.

Tauchen et al.(1989) analyse the effects of audits and of the tax code on reported income. They use individual level data from the 1979 Tax Compliance Measurement Program and combine them with IRS administrative records for District Offices and 1980 Census data at the five-digit Zip code level. Audit data are only at IRS District level[5]. All data are aggregated at district level to match with the audit data. The authors estimate a reported income equation for four audit classes that differ by their total positive income and have non business source: low income (below $10,000) , middle income (between $10,000 and $25,000), middle income (between $25,000 and $50,000) and high income (above $50,000). Their findings suggest that audits stimulate higher income reports for all four groups but the effect is statistically significant only for the highest income group. According to the authors' calculations, the indirect yield from increasing the audit rate for high income wage and salary workers by one percentage point (from the 1979 level of 10.4 to 11.4), would be three times the direct revenue. However, there is no derivation of this result. The authors report an elasticity of declared income with respect to the audit rate equal to 0.19.

Dubin et al.(1990) investigate the overall role of audits in the federal revenue collection process using state level data for the period 1977-1986, when there was a sharp decline in the audit rate. In particualr the authors aim at estimating the *spillover* effects of investigations, which they define as the "...increase in collections from taxpayers, whether or not they are audited, who report more taxes due in response to an increase in the likelyhood of an audit". The authors estimate two models. One specifies reported taxes per return filed as a function of audit rates and a variety of socio-economic factors. The other model specifies returns filed per capita as a function of the same variables. Data on audits consist of the total individual income tax returns examined divided by total individual income tax returns filed. In order to calculate the spillover effects they repeat

[5]The IRS Districts are administrative units responsible for conducting audits. The boundaries of the districts coincide more or less with the states, except in the most populous states, where multiple districts are established. Until 1984 Califormia, Illinois, Ohio, Pennsylvania and Texas all had two districts and New York had four. In 1984 another district was added to Texas and three were added to California.

the analysis using reported tax plus additional tax and penalty recommended after an audit per return as a dependent variable in the first model. They use the following procedure. From the results obtained from their regressions, they use the estimated reported tax liability per return and total returns filed per capita to calculate, for each year, the predicted value of total reported tax from individual returns that would have been realised if the audit rate had remained constant at its 1977 level over the period 1977-1986.

They estimate that maintaining the audit rate at its 1977 value, by 1986 total reported tax would have increased by 15.6 billion dollars, which is 4% of total individual tax in 1986. This value is the indirect effect or spillover effect of investigations. In order to get the direct effect of investigations they add additional taxes and penalties resulting from investigations to reported tax, for each state in each year, and divide by the number of returns filed. This generates a dependent variable which includes the revenue produced by audits. They call this variable assessed liability per return. Repeating the same calculation as for total reported tax, they obtain a predicted value for the increase in total assessed liability for 1986 from holding the audit rate to its 1977 value of 18.2 billion dollars. The difference between this figure and the predicted value of total reported tax (15.6 billion dollars) gives a direct revenue effect of 2.6 billion. Hence the ratio between indirect and direct effect is 6:1, i.e. the indirect effects of audits produce six out of every seven dollars of additional revenue.

More recently Plumley (1996) presents an econometric analysis on the determinants of voluntary compliance, using a very rich dataset by state and year, from 1982 through 1991, including data on taxpayer behaviour and IRS actions. As measures of voluntary compliance Plumley adopts the ratio between how many returns were filed over the returns required to be filed and how much income and offsets were reported over income required to be reported and offsets allowed to be claimed. The explanatory variables were grouped in different sets: tax policy measures (e.g. filing threshold, allowed exemptions), burden/opportunity variables (e.g. hours needed to complete a tax return, type of income), IRS enforcement measures (audit rate at the start of the period, information return matching program, non-filer notices, refund offsets and criminal tax convictions), IRS responsiveness (telephone assistance, return preparation services) and other demographic and economic factors.

The findings suggest that audits have a significant compliance effect and the indirect effects of an audit outweigh the direct effect. Plumley obtains an estimate for the ratio between indirect effects and direct effects in a similar way than Dubin et al. and gets a value of 11:1.

It is worth noting that the author criticises models using micro-level data, e.g. 3 digit and 5 digit ZIP code level data, to estimate the general deterrent effects of investigations. According to the author, models that use 3-digit and 5-digit level data, "...implicitly assume that the general deterrent effect operates only within the strict confines of each unit of observation (e.g. a ZIP code boundary), and it seems obvious that people will develop their compliance perceptions and propensities based on the information they get from a wide variety of sources from many locations[6]."

The only UK study is by Mayston and Martin (1998), on the deterrent effects of VAT assurance visits on VAT non compliance. The authors use cross section data for 48,000 traders across the UK for the year 1996. They distinguish three types of effects of investigations: the deterrent effect, the total direct net additional liability effect and the preventive effect. The deterrent effect is the decrease in non compliance of traders who are not the subject of the assurance activity. The total direct net additional liability effect is the revenue collected through investigations. The preventive effect is the increase in compliance in subsequent years by those traders who do receive an assurance visit. In their study they calculate the *incremental* deterrent effect, which is the effect on the VAT return declarations that are made by traders who are not the subject of investigation of a one percentage change in investigations.

The measure of non-compliance adopted is the Net Additional Liability (NAL) which would have been discovered by an assurance visit. This is estimated from the NAL discovered by assurance visits, under the assumption that assurance visits represent a sampling process of all traders, so that the NAL of visited traders can be expected to be at the same level for all traders with the same risk characteristics and other parameters (e.g. time since last visit) that determine the level of their NAL. The authors estimate a log-linear model and regress the positive NAL on the probability of receiving an assurance audit, the time since last visit and risk factors characterising each trader. The coefficient on the probability of receiving an assurance visit represents the elasticity of evasion for those traders who are not subject to investigations. The obtained estimate is 0.55: a one percent increase in the probability of an audit to all traders will induce a 0.55 per cent decrease in non-compliance of traders who haven't been investigated. The authors also calculate the values of the incremental deterrent effect for different percentage changes in VAT assurance activity and find that investigations have diminishing returns: the absolute value of the incremental deterrent effect declines for successive equal increases in the assurance activity.

A comparison of the findings on the effects of audits on tax compliance from these studies is very difficult as they refer to different levels of data aggregation, to different

[6] Plumley (1996), p. 6-7.

time periods and different taxes. Some studies consider only one year time span, others look at 10 years time periods. Only three studies provide estimates for the ratio between indirect and direct effects of investigations, though only two explain the methodology used to get the result. It remains however unclear what is driving the direct and indirect effects of investigations. Hence the interpretation of the findings remains a challenge. In the next section we set a very general model on the optimal allocation of investigation resources and derive the expressions for the direct and indirect effects in order to identify the determinants of these effects. We will use the results to comment the seemingly contradictory findings of the empirical studies.

3 The indirect and direct effects of investigations

In this section we analyse the effects of an increase in the frequency of detection and derive an expression for the direct and indirect effects. We model the optimal allocation of investigation resources across different categories of taxpayers, when the enforcement agency has a fixed budget to carry out investigations. We assume that the enforcement agency aims at minimising the tax gap, defined as the amount of evasion taking place minus the amount recovered through investigations.

Let N_k be the number of taxpayers of type k, and E_k the average amount of evasion carried out by taxpayers of that type. The frequency with which the tax authority carries out investigations of taxpayers of type k is p_k. The total number of investigations carried out on taxpayers of type k is $I_k = p_k N_k$. We define the ratio of the average amount recovered per investigation of taxpayers of type k to the average amount of evasion per taxpayer of type k as θ_k[7]. The cost of carrying out such an investigation is c_k. We assume taxpayers differ in the frequency with which they are investigated and in each group their behaviour depends solely on the frequency of audit with which the group is targeted, via the function $E_k(p_k)$. There are m different types of taxpayer. We measure the responsiveness of taxpayers of type k to the audit rule in terms of the elasticity of evasion: $\varepsilon_k = -\frac{dE_k}{dp_k}\frac{p_k}{E_k}$.

We should note few points before proceeding.

[7] There are two factors bearing on the value of θ for any given group. First, for a variety of reasons the tax authority would not necessarily expect to recover in any particular investigation the full amount of tax that is actually evaded, which would suggest $\theta < 1$. On the other hand there may be a great deal of targeting of resources within group k so that investigations are devoted to the high end of the spectrum, in which case we could have $\theta > 1$. Also, if the enforcement agency is carrying out an investigation over multiple years, it might well be the case that $\theta > 1$. Which value of θ will apply in any circumstance depends on the heterogeneity of the group and the extent to which investigations are targeted on high yield or affect a long period of time. If the group is pretty homogeneous or if taxpayers are selected more or less at random, we would expect $\theta < 1$.

- In the behavioural relationship adopted above we are not assuming that taxpayers necessarily correctly perceive the true frequency with which their group is inspected, just that there is some relationship between the actual frequency of inspection, the perceived frequency of investigation and behaviour. We are not modelling these more fundamental relationships, but we just adopt a reduced form that relates behaviour ultimately to the actual frequency of inspection. So the elasticity defined above confounds two elasticities: the sensitivity of evasion behaviour to the perceived frequency of inspection and the sensitivity of the perceived probability of inspection to the actual frequency of inspection. It is important to distinguish between the two elasticities as they measure different aspects of the individual response to audits. They may also take very different values: an individual may be very sensitive to the perceived odds of being investigated, but the actual probability and perceived probability may be matched very poorly. Or the opposite might occur. Hence the weak response to investigations which emerges from the empirical studies using individual level data we considered above, could be compatible with a high sensitivity of evasion behaviour to perceived probability if perceive probabilities do not adjust precisely to a change in actual probabilities. Alternatively, a low response to audits could be due to a low sensitivity to perceived probability, even if the match between actual and perceived probability is perfect.

- It is also important to recognise that the elasticity defined above measures the average behavioural response of taxpayers in the same group. It is a population elasticity rather than an individual elasticity. This has two implications. First, this allows for considerable heterogeneity of individual sensitivity within the group. Secondly this is consistent with the possibility that taxpayer behaviour might not be based on a purely individualistic calculus but might be affected by the proportion of taxpayers within the group who are compliant, thus reflecting some kind of social norm at work. The advantage of very reduced form specification of individual behaviour that we have employed is that it is consistent with a wide range of deeper structural models.

- In principal behaviour will depend on many factors other than the probability of investigation. It will also depend on: the likelihood of the investigations being effective - and hence on θ_k; the likelihood that, if effective, a penalty will be imposed. Since here we are mainly interested in the allocation of investigation resources we do not consider these other behavioural factors, but recognise their presence through the fact that the elasticity can vary across groups.

- On the other hand we are not allowing for the possibility that the behaviour of taxpayers of type k depends on the frequency with which other groups are investigated - as might be the case if people's perceived probability of being investigated depends on what they hear from taxpayers in other groups about their experience.

- In this setting we focus purely on the number of investigations carried out. There is also an issue of the quality of investigations. What we might expect is that there is a quality continuum to investigations, and that higher quality investigations *(i)* require more resources C; *(ii)* recover a higher fraction of evasion θ, and, possibly, *(iii)* have a bigger impact on taxpayer behaviour, ε. It would be interesting to investigate what could be said about the optimal quality of investigations balancing off all these considerations. However, for the purposes of this paper, the assumption we make is that the quality of investigations - and hence C, θ, ε - is fixed, possibly because managers have already chosen the optimal quality. We are not ignoring quality and assuming that it is the cheapest, lowest cost type of investigation that should be pursued. So, once again, the reduced form model employed here is consistent with a deeper structural account of there being a spectrum of investigation technologies.

The total expected amount of evasion by taxpayers in group k is $N_k E_k$, while, if they are inspected with frequency p_k, then the total expected compliance yield from investigations will be $N_k p_k \theta_k E_k$. So the total tax gap from group k will be:

$$G_k = N_k E_k \{1 - p_k \theta_k\} \tag{1}$$

The cost of investigating taxpayers in group k is $c_k p_k N_k$. The objective of the enforcement agency is to select the frequency of an audit for each group of taxpayers in order to minimise the tax gap, subject to the constraint that only a limited amount of resources (C) can be devoted to investigations:

$$\min_{p_k} \sum_{k=1}^{m} \{N_k E_k [1 - p_k \theta_k]\} \qquad s.t. \sum_{k=1}^{m} c_k p_k N_k \leq C \tag{2}$$

The first order condition for an interior solution is:

$$N_k \left\{ \frac{dE_k}{dp_k} [1 - p_k \theta_k] - \theta_k E_k \right\} + \lambda N_k c_k = 0 \tag{3}$$

where λ is the Lagrange multiplier on the resource constraint, representing the marginal reduction in the tax gap that could be brought about by an additional unit of resources for investigations.

With a bit of re-arranging we can re-write (3) as:

$$\frac{E_k\theta_k - \varepsilon_k E_k\theta_k + \frac{\varepsilon_k E_k}{p_k}}{c_k} = \lambda \tag{4}$$

Equation (4) characterises the optimal allocation of investigation resources. The expression on the left hand side represents the ratio of the marginal reduction in the payment/tax gap brought about by a unit increase in the frequency of inspections for group k, to the marginal cost of a unit increase in the frequency of inspections for group k.

An optimal allocation of resources implies that the marginal benefit:cost ratio should be the same across groups of taxpayers. This common marginal benefit :cost ratio will measure the marginal benefit of increasing resources available for investigation by 1 unit. Equation (4) also tells us that the optimal allocation of resources is independent of the size of the population of taxpayers in group k. It is the average yield for taxpayers in group k that matters not total yield.

We are interested in the expression for the marginal reduction in the payment gap (MRPG), the numerator of equation (4). We can write this as:

$$MRPG_k = E_k\theta_k - \varepsilon_k E_k\theta_k + \frac{\varepsilon_k E_k}{p_k} \tag{5}$$

Equation (5) allows us to explain the direct and indirect effects of investigations. The expression shows that there are three effects to be considered.

The first term of equation (5) represents the immediate yield brought in from an extra investigation. If one extra investigation is carried out, since each investigation is expected to yield $E_k\theta_k$ on average, then this is what the enforcement agency expects to get from the extra investigation.

The second and third terms of equation (5) represent the behavioural response of all taxpayers to an increase in the frequency of audit. A reduction in the average amount of evasion by all taxpayers in group k , means that all investigations (and not just the additional one) will find that the amount brought in from each investigation is now a bit lower. This effect is measured by the term $\varepsilon_k E_k\theta_k$. However a reduction in the average amount of evasion also means that the tax gap is reduced. The term $\frac{\varepsilon_k E_k}{p_k}$ captures the (absolute) reduction in the average amount of evasion brought about by a unit increase in the frequency of investigations. It is inversely proportional to the frequency with which the group is investigated. The smaller the number of investigations that the enforcement agency currently carries out, the greater will be the percentage increase that one additional investigation will represent. This is the *compliance* effect of an increase on investigations[8].

[8] We should note that here we are ignoring any timing issue, which is likely to affect the direct and the effect in a different way, in that the behavioural response to an increase in the frequency of investigations

So which of these are the direct effect and which the indirect effect? There are two possibilities. We can interpret the direct effect as the effect on yield, and hence consider the first two terms in equation (5) as the direct effect, or, and we believe this is more appropriate to reflect what we mean by direct and indirect effects, we can consider the direct effect as the immediate yield from investigations and the indirect effect as the behavioural impact of the actions of the tax authority. In this case the first term of equation (5) represents the direct effect and the second and third terms the indirect effect.

In the next subsection we derive the ratio between indirect and direct effects following from each of these two interpretations and consider how this analysis can help in clarifying the apparently contradicting figures emerging from the empirical studies we mentioned above.

3.1 The ratio between indirect and direct effects

For the calculation of the ratio between indirect and direct effects the simplest understanding would be to say that the direct effect is just the immediate yield brought in from an extra investigation, while the indirect effect is the effect of this activity on changing behaviour and hence compliance. So the ratio is :

$$R_{k1} = \frac{\varepsilon_k}{p_k \theta_k} \tag{6}$$

But this ignores the second term in equation (5). If we include the second term and say that the direct effect is about the effect of investigations on yield, whereas the indirect effect is about the effect on compliance, the ratio becomes:

$$R_{k2} = \frac{\frac{\varepsilon_k E_k}{p_k}}{E_k \theta_k - \varepsilon_k E_k \theta_k} = \frac{\varepsilon_k}{p_k \theta_k (1 - \varepsilon_k)} \tag{7}$$

Notice that because, compared to the first ratio, we have made the direct effect smaller and kept the indirect effect the same, we have $R_{k2} > R_{k1}$

On the other hand, if we say that the indirect effect is all about the behavioural impact of the enforcement agency actions, then we should consider the second and third terms of equation (5) as the indirect effect, in which case we get a third measure:

$$R_{k3} = \frac{-\varepsilon_k E_k \theta_k + \frac{\varepsilon_k E_k}{p_k}}{E_k \theta_k} = \varepsilon_k \left(\frac{1}{p_k \theta_k} - 1 \right) \tag{8}$$

is going to be observed later than the direct effect on the discovered evasion. For simplicity here we only consider one period, thereby modelling a steady state.

Because, relative to the first measure, we have now made the indirect effect smaller and kept the direct effect the same we have $R_{k1} > R_{k3}$[9]. As already anticipated our view is that the third ratio is the closest to capturing the spirit of what we mean by the direct and indirect effects.

The most important point is that, whatever the definition one adopts, there is absolutely no reason to think that this ratio is constant across fiscal areas. Indeed there is absolutely no reason to think that this ratio is going to be constant across different taxpayer groups within a fiscal area. For even if we thought that the behavioural elasticity, ε, was relatively constant across groups, the ratio depends on an operational/resource decision - how intensively to investigate taxpayers, p, and on the operational effectiveness of fraud investigations, θ, both of which will certainly vary both across and within fiscal areas.

It's worth noting that a high ratio is consistent with a low behavioural response, if the coverage rate is sufficiently low. This could explain the contradictory findings of Tauchen et al. (1989), who report a very high estimate for the ratio between indirect and direct effects but a very low elasticity of declared income with respect to the audit rate.

How do these results relate to the findings of the empirical studies we considered above?

First of all we should notice that the figures provided in the empirical studies on the ratio between indirect and direct effects are calculated using a very different approach to that we employed in our model. One difference with our approach is that these authors do not consider a marginal increase in the audit rate, but a 1 percentage point increase. Tauchen et al.(1989) consider a rise in the audit rate for high income wage and salaries workers from 10.4% to 11.4%, i.e. an increase by 110%. Dubin et al. (1990) calculate the indirect effects if the audit rate had remained to its 1977 value, a change from 1% to 1.88%, equivalent to an increase of 188%. Plumley (1996) considers the effects of raising the frequency of investigations from 0.65% to 1.65%, hence an increase by 254%. A second difference is that these authors do not use a formula to calculate the ratio, but directly calculate it from their estimated equations.

Both Plumley (1996) and Dubin et al. (1990) calculate it as the ratio between the change in total amount of evasion undertaken by taxpayers and the change in the compliance yield from investigations. Hence the ratio calculated is the analogue of the ratio

[9] We derive these ratios under the assumption that the tax authority does not have enough information on how taxpayers differ in one group and hence cannot target specific taxpayers. This implies that the audit is random within a given group. However, as we show in the Appendix, under some general conditions, the above expressions are not greatly affected, even if we allow for the possibility that the tax authority can target investigation resources and select, within a given group, those taxpayers with higher expected evasion.

R_2. The estimate for Plumley it is 11:1. This is an average value, as it considers average changes. Using 1991 parameters valued and considering a 1% change in the audit rate we derived a ratio indirect/direct effect of 8.54. For Tauchen et al. the average ratio is 30:1[10]. Dubin et al. report a ratio equal to 6:1. These estimates are quite different, but they all confirm the overwhelming role of the indirect effects. The figures may seem quite high, but, in the light of what we illustrated above, this is not surprising. The important point to bear in mind is that the direct effect operates on people who are investigated whereas the indirect effect operates on the whole population (or group). Given that only a small proportion is investigated, this factor in itself tends to make the indirect effect much larger than the direct effect, as our formulation shows. We can also explain the difference in values provided: as equation (7) shows, it is possible to observe different ratios across districts/states and across time if the elasticity of evasion, the probability and the effectiveness of audits vary across states or across time. We know that the probability of audit was different across the different studies. We do not have any estimate for the effectiveness of audits, apart from Plumley. For the elasticity of evasion we do not have an estimate for these models. When it is calculated, the elasticity is not the evasion elasticity as in our model, but it is the elasticity of reported income or reported tax liabilities. Tauchen et al. (1989) using data aggregated at district level (for the tax year 1979) report an elasticity of declared income with respect to the audit rate equal to 0.19. As the authors do not explain their results it is not possible to translate it in an evasion elasticity. From the details available in Plumley, who uses panel data at state level (from 1982 to 1991), we were able to calculate the elasticity of evasion. If the calculation is performed on average values the elasticity is about 0.7. A calculation based on 1991 parameters values gives a figure of 0.56.

An average elasticity of evasion of 0.7, as obtained from Plumley may seem quite a high value. However, the estimate is like our term ε_k: it is a population elasticity. This elasticity may capture quite different aspects of individual behaviour. One important aspect could be the role of social interactions in a community of taxpayers: individuals may be affected by the number of other people being compliant and this may translate in a greater response to investigations at the level of the whole community. In the next section we want to explore this idea in detail. This argument could conciliate the contrasting findings from micro level studies and more aggregate analysis on the value of the behavioural elasticities.

[10] The authors report an increase in tax revenues due to the improved compliance of taxpayers almost three times the direct revenue yield for an increase in the audit rate from 10.4% to 11.4%. Assuming a linear relationship between the audit probability and the indirect yield and ignoring second order effects, this would imply a ratio indirect/direct effects of approximately 30:1.

4 Tax compliance as a social norm and the impact of an increase in the frequency of audits

The idea that tax compliance does not only rely on deterrence and economic factors, but it is also affected by moral considerations and social interactions, has been explored in some recent developments of the theoretical literature on tax evasion and has been tested in recent empirical studies. The assumption made by the theoretical models and tested in the empirical studies is that individuals may be induced to be honest on the grounds of moral considerations and/or social pressure. This is typically modelled by introducing non-monetary factors as extra arguments in the utility function[11]. The findings from laboratory experiments and surveys seem all to converge to the view that there is indeed a social dimension in the subjects' decision whether or not to comply. There is evidence that many countries with similar fiscal systems have different compliance experiences. Torgler (2002) reviews some laboratory experiments and draws the following conclusions: "...(i) individuals who comply tend to view tax evasion as immoral, (ii) compliance is higher if moral appeals are made to the taxpayer, (iii) individuals with tax evaders as friends are more likely to be evaders themselves, and (iv) compliance is greater in societies with a stronger sense of social cohesion."(p. 664).

However, how individuals' attitudes actually translate into behaviour still remains to be explained. Results from a field experiment on the role of normative appeals on social conscience in deterring tax evasion, conducted on 60,000 US taxpayers, suggest that evaders are more likely to be individuals with higher opportunities to evade and these are the ones who seem to be less affected by normative appeals[12]. The crucial role of opportunities in determining the choice of evaders and also of non-filers is confirmed by other laboratory experiments and econometric studies[13].

In this section we want to analyse how the decision whether or not to evade is affected by the tax parameters and opportunities to evade and how their impact changes when tax compliance assumes the characteristics of a social norm. We assume tax compliance is a code of behaviour in the community of taxpayers, although we do not model how it may emerge and remain establish. Different factors could be responsible for the emergence of tax compliance as a social norm. The underlying idea is that taxpayers care about how many other individuals evade tax, and this could be because they care about their reputation within the community or because they are concerned that each member in the community pays their fair contribution to tax revenues. Following Myles and Naylor

[11] Benjamini and Maital (1985) and Gordon (1989) analyse the role of psychic costs in deterring tax evasion and Myles and Naylor (1996) analyse tax compliance as a social norm.

[12] See Blumenthal et al. (2001).

[13] See Slemrod et al. (2001), Crane and Nourzad (1993) and Erard and Ho (2001).

(1996), we assume that there exist an extra source of utility from not evading and this is increasing in the number of taxpayers behaving honestly. In terms of the previous analysis on the allocation of investigation resources, here we examine the behaviour of individuals belonging to the same group k. To keep notation simple we omit the subscript k, but when we refer to the community of taxpayers, where the social norm is at work, we mean those taxpayers in group k.

We are interested in analysing how the behavioural response to an increase in the probability of an audit is affected by the different assumptions we make on how taxpayers may be motivated to cheat.

We compare three settings: one in which tax compliance is simply an opportunistic behaviour, based on individual calculus. We distinguish the case of the standard portfolio model where the loss from being caught is simply the pecuniary fine charged by the tax authority, from the case where the fine has a broader interpretation and it also incorporates the psychic cost of being investigated, which can vary across taxpayers. In the second setting we allow for a warm glow from tax compliance. If the individual does not evade she/he benefits from an extra non-pecuniary utility gain from being honest. In the third setting we assume that tax compliance is a social norm, regarded as a code of behaviour in the relevant community: an individual, in making the choice whether or not to evade, is also influenced by the fraction of the taxpayers in the population who are compliant. We analyse the effect of an increase in the frequency of investigations on the overall compliance in these three cases.

We assume risk neutrality, as we want to abstract from any risk considerations and focus on the effects of non monetary considerations and social interactions on tax compliance. This assumption is not crucial for our qualitative results. We analyse the decision whether or not to evade rather than the decision how much to evade, hence we do not consider why people can evade different amounts, but rather how people decide to be compliant or not.

4.1 Selfish calculus.

We first consider the setting of the standard portfolio model, where the taxpayer decides whether or not to evade on the basis of a selfish and purely monetary calculus.

4.1.1 Individual behaviour

We first focus our analysis at the individual level.

We define the utility from non evading for an individual with income y and facing a tax rate t as:

$$U^{NE} = y(1-t) \tag{9}$$

Let e, $0 \leq e \leq y$, be the individual's opportunity to evade, i.e. the amount of income that can potentially be hidden and $\widetilde{e}$, $0 \leq \widetilde{e} \leq e$, the actual amount evaded. An individual is investigated with probability p, $0 < p < 1$, and in case of evasion he/she will need to pay back the taxes due and a monetary fine $F > 0$ on the amount of evaded income, $\widetilde{e}$ (as in Allingham and Sandmo, 1972). Hence the utility from evading is:

$$U^E = p[y(1-t) - F\widetilde{e}] + (1-p)[(y-\widetilde{e})(1-t) + \widetilde{e}] = y(1-t) + \widetilde{e}[t(1-p) - pF] \quad (10)$$

An individual is willing to evade if $U^E > U^{NE}$. Hence tax evasion will occur whenever if $\widetilde{e}[t(1-p) - pF] > 0$. If $t(1-p) - pF > 0$ the expected financial gain from evading one extra unit of income is positive and the taxpayer will always evade to the maximum amount, so $\widetilde{e} = e$[14]. The expression $t(1-p) - pF$ is decreasing in F, so that there will be an $\overline{F}$ such that $t(1-p) - p\overline{F} = 0$. For this particular value of the fine rate the individual will be indifferent between evasion and non evasion. Hence, $\overline{F} = \frac{t(1-p)}{p}$ defines the critical value above which an individual will opt for full compliance, as, above $\overline{F}$, $t(1-p) - p\overline{F} < 0$ and tax evasion is not profitable. If taxpayers face the same tax parameters there will be a unique value of $\overline{F}$ above which *everybody* will evade, even if their income differs. As long as taxpayers have an opportunity to hide their income ($e > 0$) a mixed equilibrium of evaders and non-evaders is not possible. Also, the only way to affect the decision whether or not to evade is to vary $\overline{F}$. In this case, in fact, opportunities to evade are exogenous in that they are not affected by any of the tax parameters: the critical level of opportunity for which $U^E = U^{NE}$ is $\underline{e} = 0$. Evasion will occur whenever $e > 0$[15].

We illustrate this in figure 1.

An increase in the probability of detection will affect the decision whether or not to evade only if the population of taxpayers is just indifferent between evasion and non evasion, i.e. if the value of the fine rate is exactly $\overline{F} = \frac{t(1-p)}{p}$. In this case, an increase in the probability of detection will make everybody to opt for full-compliance.

Pecuniary loss for being investigated. The act of being caught evading may imply some loss in reputation or some psychic cost for feeling guilty or ashamed. This non pecuniary cost is very likely to differ across individuals: the loss of reputation for being caught evading may be higher for a person with a high public profile, or the feeling of guilt or shame may be quite personal and differ across individuals, regardless of their occupation. In what follows we assume that the fine rate also includes a non pecuniary

[14]We should note that here both the probability of detection p and the fine rate F are fixed and do not depend on the amount of evasion. In reality both the frequency of an audit and the fine rate are positively related to the amount of concealed income. This may imply that tax evaders do not evade to the maximum extent of their possibilities.

[15]Whenever $e > 0$, $U^E > U^{NE}$, and an invidual will engage in tax evasion.

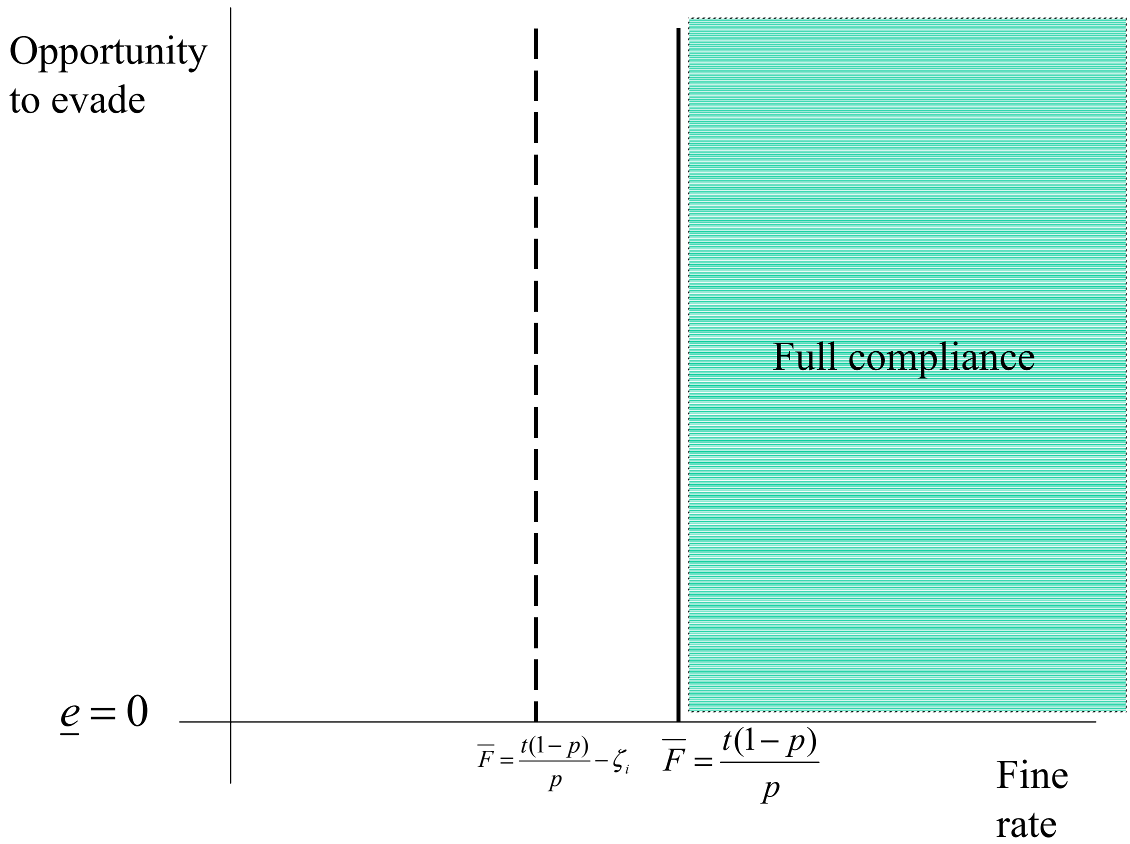

Figure 1: Selfish Calculus: the individual compliance decision.

cost, ζ_i which varies across taxpayers. We represent this broader concept of the fine rate as $f = F + \zeta_i$. The expected financial gain from evading one extra unit of income will therefore be $t(1-p) - p(F+\zeta_i)$, which is decreasing in both F and ζ_i. Like before, there will be a threshold level $\overline{f}$ such that $t(1-p) - p\overline{f} = 0$,above which an individual will opt for full compliance. This threshold level corresponds to $\overline{F} = \frac{t(1-p)}{p} - \zeta_i$ and will therefore vary across individuals depending on ζ_i. This implies that it will be possible to observe some individuals opting for evasion and others being fully compliant even if they face the same tax parameters. A mixed equilibrium of evaders and non-evaders will be possible.

In figure 1 we represent the threshold level of the fine above which individual i, who suffers a non pecuniary loss ζ_i for being audited, will not evade.

4.1.2 The Community

In the case of a monetary fine, applied to the amount of evaded income, the behaviour of a single individual also represents the behaviour of the whole community. In fact, if individuals face the same tax parameters, there will be a unique value of the fine rate above which everybody will be fully compliant and below which everybody will evade. The analysis is more interesting when the fine rate also includes a non pecuniary cost, which may vary across individuals, as we have noted above. Here we analyse overall evasion in a community where individuals differ in the psychic cost of being investigated.

We assume taxpayers differ in the opportunity to conceal income, e, the non pecuniary fine for being audited ζ_i, and income y. The density function for e is $g(e)$ and for f is $h(f)$. Since y doesn't affect any decision, its distribution is irrelevant. We define $\overline{f}$ as the threshold level of the fine below which tax evasion is profitable. This corresponds to $\overline{F} = \frac{t(1-p)}{p} - \zeta_i$. Some individuals will have ζ_i such that they will be above $\overline{f}$ and some will be below $\overline{f}$. The distribution of f will determine how many individuals evade. Total evasion is:

$$E = \int_0^{\overline{f}(p.t)} \left[\int_0^{\infty} eg(e)de \right] h(f)df. \tag{11}$$

Effect of an increase in the probability of detection on overall evasion The effect of an increase in the probability of detection on total evasion can be represented as:

$$\frac{\delta E}{\delta p} = -\int_0^{\infty} e\ g(e)\ de\ h(\overline{f})\frac{t}{p^2} \tag{12}$$

A change in the probability of detection will affect those on the margin, who were indifferent between evasion and non evasion, $h(\overline{f})$.

The magnitude of the impact of a rise in the audit rate on evasion depends on the distribution of the psychic cost. The greater the number of people who were previously indifferent between evading and being fully compliant, the greater the impact of a rise in the audit rate on tax evasion. Hence, for a given value of p and E, the greater the number of people who before the change in the audit rate were indifferent between evasion and non evasion, the higher the value of the elasticity of evasion.

4.2 Non-selfish considerations and tax compliance

We now consider a setting where the compliance decision is not merely based on selfish calculus, but an individual recognises the importance of paying taxes, for example because he/she values the provision of public goods, and/or regards honest behaviour as morally right. We represent this in terms of a warm glow for behaving honestly, modelled as a constant parameter in the utility from non evading. We keep the assumption that the fine rate also include a non-monetary cost for being investigated.

4.2.1 Individual behaviour

Let ω be the warm glow from being honest. The utility from non-evading is:

$$U^{NE} = y(1-t) + \omega \tag{13}$$

As before, the utility from evading is:

$$U^{E} = p[y(1-t) - f\widetilde{e}] + (1-p)[(y-\widetilde{e})(1-t) + \widetilde{e}] = y(1-t) + \widetilde{e}[t(1-p) - pf] \tag{14}$$

An individual will evade tax if the utility from evading is greater than the utility from non evading, i.e. if $\widetilde{e}[t(1-p)-pf] > \omega$ and will evade to the maximum of his opportunities ($\widetilde{e} = e$) if the fine is below the threshold $\overline{f}$[16].If there is a warm glow from behaving honestly, the entry condition for tax evasion is more restrictive: opportunities must exceed a threshold level, which is defined by the tax parameters and the warm glow for inducing an individuals to evade:

$$\underline{e}(f) = \frac{\omega}{t - p(f+t)} \tag{15}$$

Full compliance will occur if $e \leq \underline{e}(f)$. Two channels affect now the decision whether or not to evade: $\underline{e}$, which depends on p,f, and t, and $\overline{f}$.

An increase in p will affect both $\overline{f}$ and $\underline{e}$. In figure 2 we represent the decision whether or not to evade for an individual with a non pecuniary cost of being caught equal to ζ_i. $\underline{e}$ is an increasing function of the fine rate and tends to infinity when $f = \overline{f}$ .The individual will engage in tax evasion whenever his opportunities to evade are above the curve $\underline{e}$. Note that in the area below the curve $\underline{e}$ and to the left of $\overline{f}$ the expected gain from one extra unit of evaded income is positive but the warm glow from being compliant outweigh this gain and the individual is not willing to evade.

[16] This is the threshold level for which $t(1-p)-pf = 0$. As $f = F + \zeta_i$, $\overline{f}$ corresponds to $F = \frac{t(1-p)}{p} - \zeta_i$.

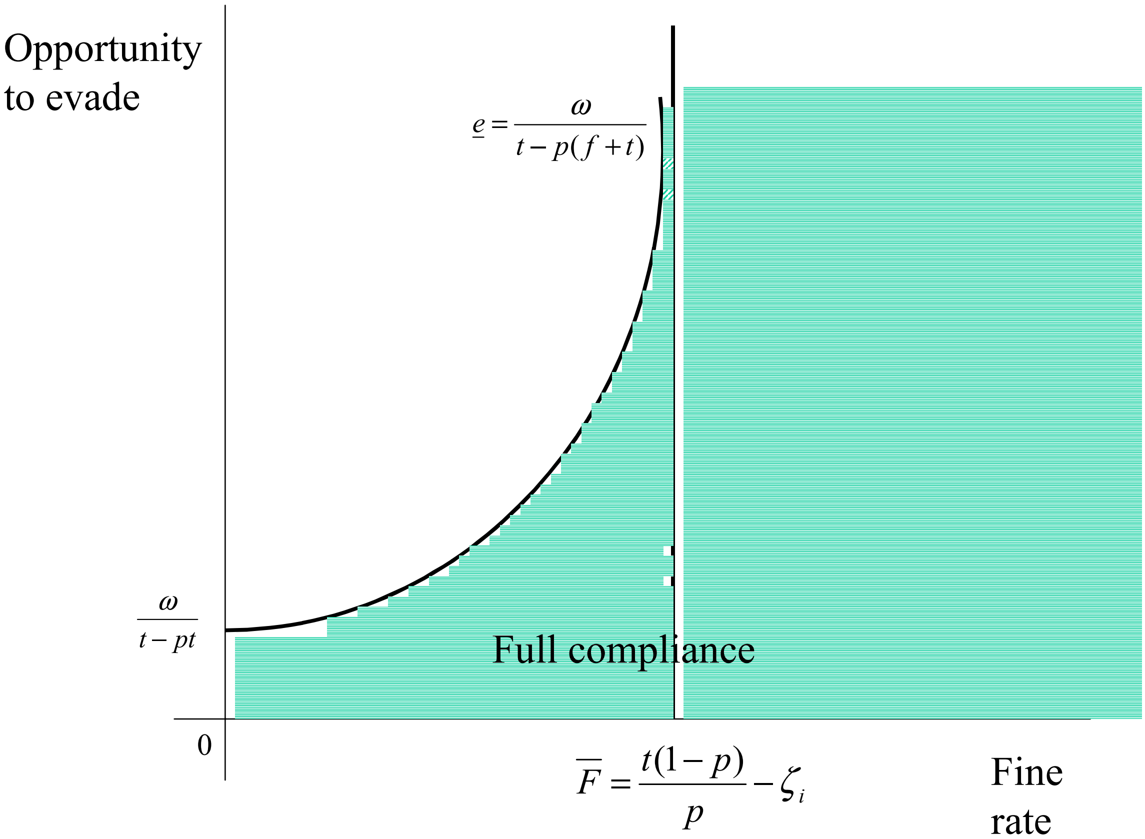

Figure 2: Moral considerations and the compliance decision.

4.2.2 The Community.

Within the community all those taxpayers for whom $f < \overline{f}$ and $e > \underline{e}(f)$ will evade. So total evasion will be:

$$E = \int_{0}^{\overline{f}} \left[\int_{\underline{e}(f)}^{\infty} eg(e)de \right] h(f)df \tag{16}$$

Effect of an increase in the probability of detection on overall evasion As we anticipated above there are now two channels through which an increase in the probability of detection will affect the tax compliance decision. In fact a change in the probability of detection will modify both the critical value of the fine rate and the critical value of opportunities:

$$\frac{\delta E}{\delta p} = -\int_{\underline{e}(\overline{f})}^{\infty} e\ g(e)\ h(\overline{f})\frac{t}{p^2} - \int_{0}^{\overline{f}} \underline{e}(f)\ g(\underline{e})\ \underline{e}(f)\frac{(f+t)}{t-p(f+t)} \tag{17}$$

The first term on the right hand side of equation (17) is the effect of a change in the probability of detection on $\overline{f}$: an increase in the probability of detection will lower $\overline{f}$, making tax evasion less profitable on the margin. The second term is the effect on the critical value of opportunities to evade: an increase in the probability of detection will increase $\underline{e}$, making the entry condition for evasion more restrictive. We illustrate the effect of an increase in the probability of detection on an individual's willingness to evade in figure 3. An increase from p to p' will shift the curve $\underline{e} = \frac{\omega}{t-p(f+t)}$ upwards and $\overline{f}$ leftwards and total evasion will decrease.

The distribution of f and e are crucial in determining the magnitude of the impact of the audit probability on evasion. The greater the number of taxpayers on the margin ($g(\underline{e})$ and $h(\overline{f})$), the greater the impact of the audit rate on evasion. For given values of tax evasion and of the audit rate, a higher number of taxpayers on the margin implies a higher elasticity of evasion.

4.3 Tax compliance as a social norm

We now relax the assumption that tax compliance is an individualistic choice, taken with no considerations for the behaviour of other taxpayers. We assume tax compliance is a social norm and individuals benefit from conforming to it. In particular, the number of the other taxpayers who are honest becomes a crucial element in the decision whether or not to evade.

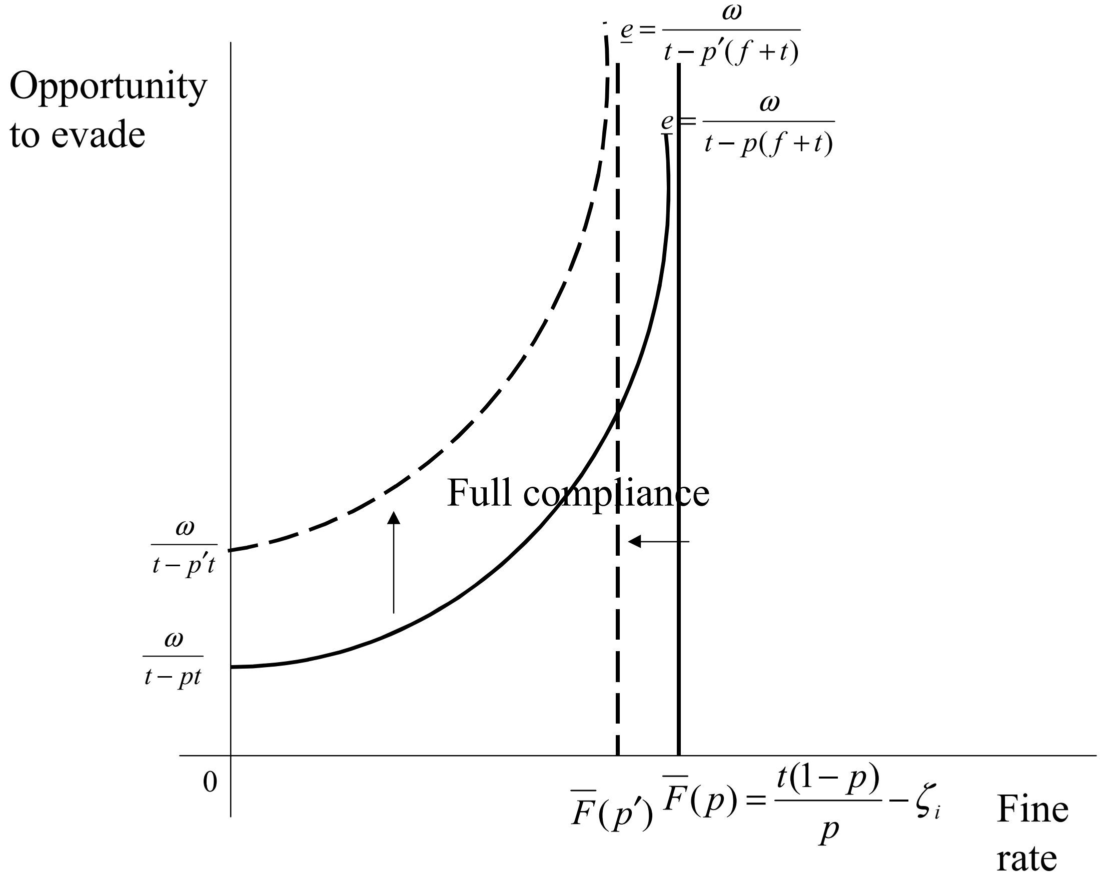

Figure 3: Moral considerations: the effect of an increase in the probability of detection on compliance.

4.3.1 Individual behaviour

Following Myles and Naylor (1996) we assume that when an individual pays his taxes honestly he gets two extra sources of utility, a social custom utility and a warm glow. For an individual with income y and facing a tax rate t, the utility from non evading is represented by:

$$U^{NE} = y(1-t) + b(1-\mu) + \omega \tag{18}$$

where μ, $0 \leq \mu \leq 1$, is the proportion of the population who evades, $b(1-\mu)$ is the utility of conforming with the group of honest taxpayers: we can think of it in terms of a non-pecuniary gain from enhancing one's reputation within the community by being honest. In line with the social custom approach[17] , reputation depends on the proportion of individuals who believe in a given code of behaviour, so that the larger the number of believers the more reputation is lost by disobedience of the code. $b \geq 0$ is the weight attached to the social norm and $\omega \geq 0$ is the warm glow from behaving honestly. It is important to keep the two effects separate as they imply quite different mechanisms of enforcement. In the case of a warm glow the mechanism of enforcement relies on the inhibitory power of personal conscience and civic responsibility. Individuals hold a personal conviction towards non evasion and are prepared to adhere to it, even if the other members of the community cannot observe any cheating activity. In the case of a loss in reputation for evading, the mechanism of enforcement relies instead on the community's behaviour and the greater the number of individuals complying with their tax duties the greater the loss for stepping out of the social norm. This reflects the distinction emphasised by sociological theories of social control (Grasmick-Green (1980) and Wrong (1961)) between the inhibitory power of moral commitment to the law and the inhibitory power of the threat of social disapproval. The utility from evasion is as before:

$$U^{E} = p[y(1-t) - f\widetilde{e}] + (1-p)[(y-\widetilde{e})(1-t) + \widetilde{e}] = y(1-t) + \widetilde{e}[t(1-p) - pf] \tag{19}$$

If $f < \overline{f}$, i.e. $t(1-p) > pf$, the individual will evade to the maximum of his possibilities, so $\widetilde{e} = e$. If this inequality were not true, then penalties alone would be enough to prevent evasion and we would have no need to invoke social norms. Hence the utility from evading is[18]

$$U^{E} = y(1-t) + e[t(1-p) - pf] \tag{20}$$

[17]See Akerlof (1980) and Cowell (1990)

[18]In this model we assume that once the individual has opted for evasion, he chooses the amount of evasion as in the standard model, on the basis of the expected financial gain. The model could be enriched by allowing evaders to decide on the amount of evasion also on the basis of how much evasion is carried out by other individuals in their peer group. In this case the relative amount of evasion would be an additional argument in the utility from evasion. However this modification of the model wouldn't qualitatively affect our results.

So an individual will evade tax if the utility from evading is greater than the utility from non evading. This occurs if the opportunity to evade e exceeds a threshold value $\underline{e}$:

$$e > \underline{e} = \frac{b(1-\mu)+\omega}{t-p(f+t)} \tag{21}$$

Notice that the entry condition for tax evasion is more restrictive than in the previous settings. The greater the importance attached to the social custom, the greater will be $\underline{e}$. Whereas a greater proportion of evaders (μ) is associated with a lower $\underline{e}$ which implies that the entry condition for evasion becomes less restrictive as the number of dishonest taxpayers increases.

In figure 4 we represent the compliance decision of two different individuals who face the same tax parameters, have the same opportunities to evade, attach the same importance to the social norm and get the same warm glow for being honest but suffer different non pecuniary costs in case they are caught cheating. In particular, taxpayer 1 bears a non pecuniary cost if caught evading $\zeta_1$which is less than the cost suffered by taxpayer 2, ζ_2. For an equal monetary fine imposed by the tax authority, taxpayer 2 will bear a higher cost for being caught cheating, i.e. $f_2 = F + \zeta_2 > f_1 = F + \zeta_1$ and hence, for each level of the fine rate in the interval $0 < f < \overline{f}$ the threshold level of opportunities above which taxpayer 2 will consider to evade will be higher, i.e. the $\underline{e}_2$curve lies to the left of $\underline{e}_1$. The shaded area in figure 4 represents the greater willingness to evade by taxpayer 1, which is due to the lower psychic cost of being caught. Figure 4 could also represent the same taxpayer who has to decide two different strategies of non-compliance: whether or not to overstate expenses or whether or not to understate income. The two strategies may imply different costs of being caught cheating, but if the individual decides to be honest in both cases he gets the same utility gains from reputational concerns and conforming with honest behaviour. If the stigma from being caught overstating one's expenses is lower than the stigma for understating one's income, the individual will be more willing to overstating expenses. And if his opportunities lie in the shaded, he will be willing to overstate his expenses but not to underdeclare his income.

4.3.2 The community

All those taxpayers for whom $f < \overline{f}$ and $e > \underline{e}$ will evade, so that total evasion in the community will be:

$$E = \int_0^{\overline{f}} \left[\int_{\underline{e}(f)}^{\infty} eg(e)de \right] h(f)df \tag{22}$$

Let

$$m(\mu; f, p, t) = \int_0^{\overline{f}(t,p)} \left[\int_{\underline{e}(f,\mu,t,p)}^{\infty} g(e)de \right] h(f)df \tag{23}$$

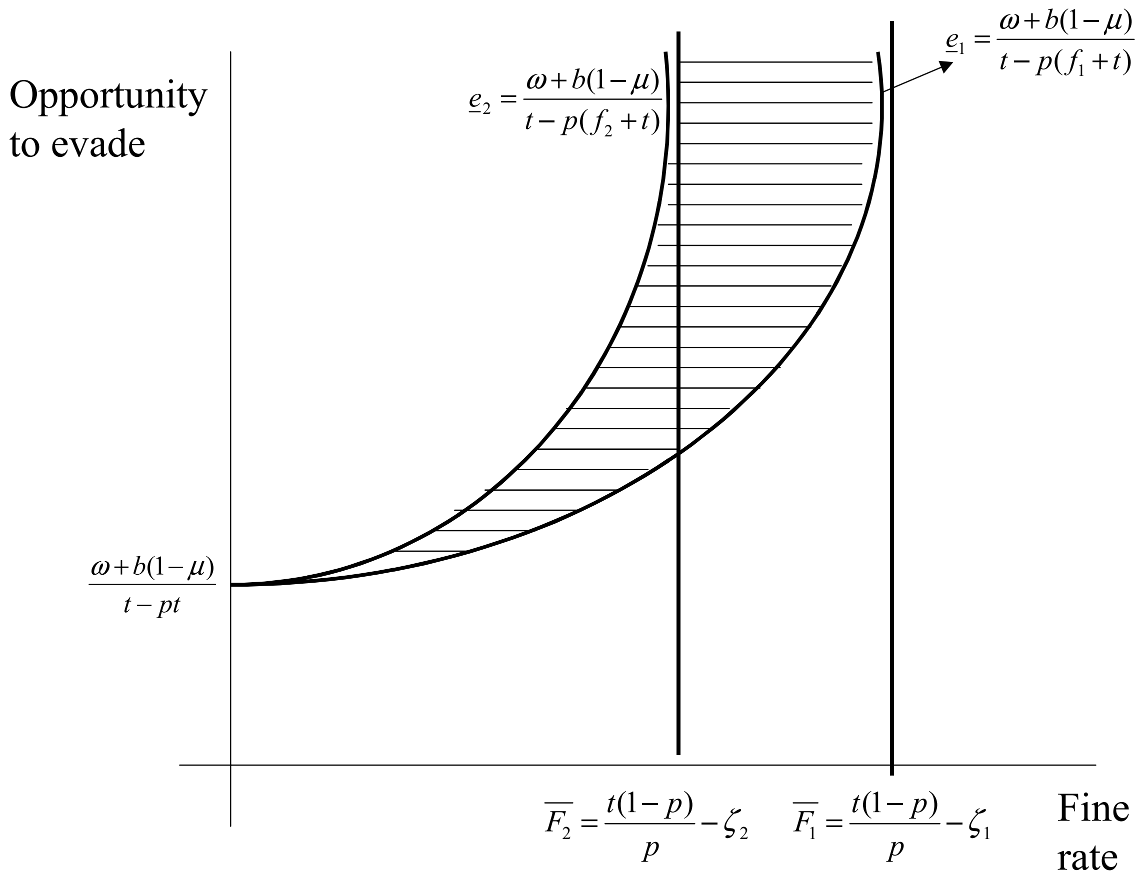

Figure 4: The compliance decision for different non pecuniary costs of being caught.

be the proportion of taxpayers in the community that are willing to evade. Notice that

$$\frac{\delta m}{\delta \mu} = \int_{0}^{\overline{f}(t,p)} \left[\frac{b}{t-p(f+t)} g(\underline{e})\right] h(f)df \geq 0 \tag{24}$$

which, for given $\underline{e}$, is strictly increasing in b. Notice that

$$m(1; f, p, t) = \int_{0}^{\overline{f}(t,p)} \left[\int_{\frac{\omega}{t-p(f+t)}}^{\infty} g(e)de\right] h(f)df \tag{25}$$

which is independent of b, though, for $\mu < 1$ an increase in b will increase $\underline{e}$ which will lower m. The equilibrium value of μ, $\widehat{\mu}$, is given by

$$\widehat{\mu} = m(\widehat{\mu}; f, p, t) \tag{26}$$

It occurs when the distribution of evasion opportunities is such that, if every individual faces the same proportion of evaders $\widehat{\mu}$, the actual proportion of evaders in the whole economy, $m(\widehat{\mu}; f, p, t)$, will be just $\widehat{\mu}$, i.e. $\widehat{\mu}$ is a fixed point for $m(\widehat{\mu}; \cdot)$. In other words, an equilibrium in the whole community occurs when, given the actual proportion of tax evaders in the population, no one has an incentive to switch from evasion to non-evasion, or vice versa.

Hence an equilibrium with zero evasion will occur when, given that nobody evades, nobody will ever consider to evade. Formally:

Condition 1: $\widehat{\mu} = 0$ is an equilibrium if, and only if, for all e and f with $h(f) > 0$ and $g(e) > 0$

$$U^E = y(1-t) + \widetilde{e}[t(1-p) - pf] \leq y(1-t) + b(1) + \omega \tag{27}$$

If $F < \frac{t(1-p)}{p} - \zeta_i$.for all taxpayers then $\widetilde{e} = e$ and this condition becomes:

$$e \leq \frac{b(1) + \omega}{t - p(f+t)}$$

Even if the monetary fine for evading is so low that the expected net gain from an extra unit of evaded income is positive for all taxpayers, opportunities to evade are too low to profitably engage in tax evasion and nobody will ever evade.

Similarly an equilibrium with full evasion will occur if nobody will ever consider not to evade when the observed behaviour is full evasion. Formally:

Condition 2: $\widehat{\mu} = 1$ is an equilibrium if, and only if, for all e and f with $h(f) > 0$ and $g(e) > 0$,

$$U^E = y(1-t) + \widetilde{e}[t(1-p) - pf] > y(1-t) + b(0) + \omega \tag{28}$$

If $F < \frac{t(1-p)}{p} - \zeta_i$.for all taxpayers then $\widetilde{e} = e$ and this condition becomes:

$$e > \frac{b(0) + \omega}{t - p(f + t)}$$

Full evasion will occur when the monetary fine for evading is so low and opportunity to evade are so high that everybody will find it profitable to evade.

Which of these two conditions is satisfied depends on how opportunities to evade the non-monetary cost of being caught are distributed across the population. For given distributions of opportunities and non-monetary costs of being caught, the zero evasion equilibrium is more likely the higher the probability of detection and the monetary fine.

The structure of the equilibrium depends on whether only one, both of neither of these conditions hold. There are four different cases:

- If only condition (27) holds then $\widehat{\mu} = 0$ will be a unique equilibrium if $m(\mu) < \mu$ for all $\mu > 0$. Otherwise there is also at least one other interior equilibrium since $m(\mu)$ must cross the 45^0line before $\mu = 1$.
- If only condition (28) holds then $\widehat{\mu} = 1$ will be a unique equilibrium if $m(\mu) > \mu$ for all $\mu < 1$. Otherwise there must be at lest one other interior equilibrium.
- If both (27) and (28) hold then two possibilities can arise: either a) $\widehat{\mu} = 0$, $\widehat{\mu} = 1$ constitute the unique set of equilibria, or b) $\widehat{\mu} = 0$, $\widehat{\mu} = 1$ are equilibria and there are interior equilibria between these points. It is sufficient for b) that either $m'(0) \geq 1$ and $m'(1) \geq 1$ or that $m'(0) \leq 1$ and $m'(1) \leq 1$.
- If neither conditions hold there are only interior equilibria and these will be in odd number.

The more spread are the distributions of opportunities to evade and of non-monetary costs for being caught, the more likely is that neither conditions hold and there are only interior equilibria.

In what follows we focus on a unique interior equilibrium and consider the comparative statics for a change in the audit rate. We make the assumption that neither (27) nor (28) hold and that $m' > 0$ and $m'' = 0$. Results are however valid for any locally stable interior equilibrium. The equilibrium is represented in figure 5.

The function $m(\mu)$ is represented by AB. The equilibrium occurs at the intersection of AB with the 45° line $\widehat{(\mu)}$.

It is easy to see that $\frac{\delta\widehat{\mu}}{\delta b} < 0$ since an increase in b pivots the line AB down through the point B. In fact, an increase in b implies, for $\mu \neq 1$, a greater $\underline{e}$, i.e. for a given observed proportion of evaders, there will be less taxpayers willing to evade than before. In figure

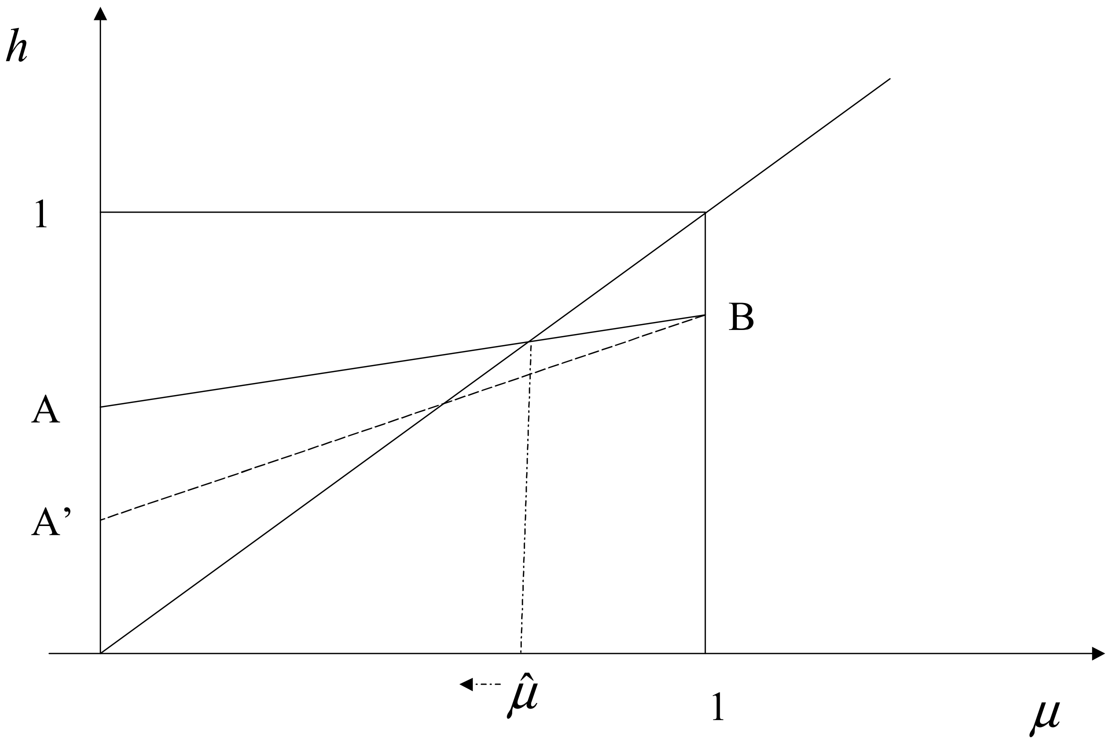

Figure 5: The equilibrium proportion of tax evaders

5 the line $A'B$ represents the function $m(\mu)$ for a higher b. The new equilibrium implies a lower proportion of evaders. So the more people care about the social custom, the fewer evade tax.

Effect of an increase in the probability of detection on overall evasion. We want to know how the amount of evasion is affected by the probability of detection.

Notice first of all that an increase in p raises $\underline{e}$ for all values of μ and so shifts the schedule AB down, thus lowering $\widehat{\mu}$, i.e. $\frac{\delta\widehat{\mu}}{\delta p} < 0$. Formally we have:

$$\frac{\delta\widehat{\mu}}{\delta p} = \frac{\frac{\delta m}{\delta p}}{1 - \frac{\delta m}{\delta\widehat{\mu}}} \tag{29}$$

Notice that the qualitative prediction does not depend on the presence of social norms and would be true even if $b = 0$. However, the magnitude of the effect does depend on the presence of social norm considerations, since, as we saw, $\frac{\delta m}{\delta\widehat{\mu}}$ is increasing in b.

The reason is clear. An increase in p causes the marginal individuals to stop evading; this lowers the proportion who evade, which in turns reduces evasion - and so on. The larger is b the more this effect ratchets up.

By differentiating (22) with respect to the p, we get that the impact of a change in the probability of audit is:

$$\frac{\delta E}{\delta p} = -\int_{\widehat{\underline{e}}(\overline{f})}^{\infty} e\ g(e)\ de\ h(\overline{f})\frac{t}{p^2} - \int_{0}^{\overline{f}} \widehat{\underline{e}}\ g(\widehat{\underline{e}})\ \widehat{\underline{e}}\frac{(f+t)}{t-p(f+t)} - \int_{0}^{\overline{f}} g(\widehat{\underline{e}})\ \widehat{\underline{e}}\ \left(-\frac{b}{t-p(f+t)}\ \frac{\delta\widehat{\mu}}{\delta p}\right) \tag{30}$$

Notice that the first two effects are the effects that would arise in the absence of social norms, if individuals only got a warm glow from non evading. The third argument on the right hand side is the additional effect arising because of social norm considerations. It would be zero if $b = 0$, and is strictly increasing in b. This is for two reasons: an increase in b raises the coefficient on $\frac{\delta\widehat{\mu}}{\delta p}$ as well as raising the magnitude of this response.

In figure 6 we decompose the effect of a rise in the audit rate in the case where tax compliance assumes the characteristics of a social norm. The line $A_{sc}B_{sc}$represents the proportion of those taxpayers who are willing to evade as a function of those who are actually evading. The initial equilibrium is at point $\widehat{\mu}$. If individuals do not care about the social custom, i.e. $b = 0$, and they only get a warm glow from tax compliance, the number of those who are willing to evade is not related to those who are actually evading and the relevant schedule for m is $A_{wg}B_{wg}$. If the initial equilibrium is $\widehat{\mu}$, a rise in the audit rate will shift both lines parallel downwards. The movement from $\widehat{\mu}$ to $\widehat{\mu}'_{wg}$ represents the first two effects of equation (30) and is the decrease in the number of evaders in the case of just a warm glow mechanism. The movement from $\widehat{\mu}'_{wg}$ to $\widehat{\mu}'_{sc}$ represents the third term in the right hand side of equation (30) and is the extra effect of a rise in the audit rate due to the social norm. Notice that the magnitude of this third effect depends on the slope of the line $A_{sc}B_{sc}$, which is determined by b, the importance attached to the social norm.

In conclusion, if people care about the social norm then an increase in p will cause the fraction of the population who evade to fall, and this will give an extra reason for people to stop evading over and above the normal deterrence effects. Moreover, the more people care about the social norm, the greater will be the fall in the proportion who evade. But, in addition, the more weight that people give to this effect, the more this will cause people on the margin to stop evading[19]. For a given value of p and E the elasticity of evasion will be greater the greater the importance attached to the social custom.

In our model we assume that the b parameter is equal across all taxpayers, but more realistically the importance attached to the social norm is likely to vary across different

[19]We should note that this analysis applies when there is a unique, locally stable equilibrium. In the presence of multiple equilibria we wouldn't be able to use our comparative statics as an increase in the probability of detection would shift the density function down and some initial equilibria might disappear, causing a jump to a different equilibrium.

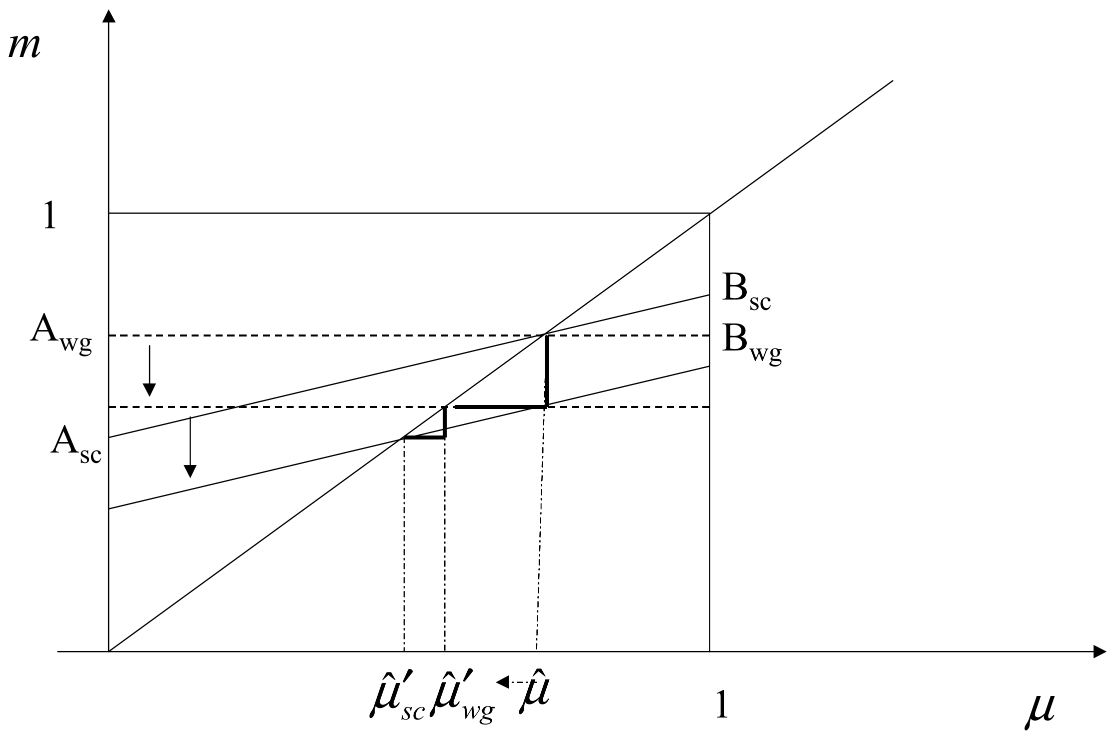

Figure 6: Effect of an increase in the audit rate on the number of evaders. Comparison between the social norm case and the warm glow case.

groups of taxpayers and hence we should expect a heterogeneous response of taxpayers to a rise in the audit rate, even if all other parameters are the same. This could explain the different findings of the empirical studies we consider in section 2: by looking at different time periods and different sections of the population of taxpayers, estimates are likely to capture a different magnitude of response. In particular, the value of the elasticity of evasion is very likely to be different across different groups of taxpayers considered in the studies.

If the importance attached to the social norm varies across taxpayers, an important issue is how this would affect the allocation of investigation resources. How could the enforcement agency take advantage of the greater behavioural response, i.e. the greater indirect effect, of taxpayers sensitive to the social norm? Would this necessarily imply to put more resources where the social norm argument is more compelling?

In the next section we consider two groups of taxpayers, where tax compliance is regarded as a social norm to a different extent and, by use of simulations, show the implications for the optimal allocation of investigation resources.

5 Allocating investigation resources between two groups with different evasion elasticities: a simulation.

We consider two groups of taxpayers, 1 and 2. Individuals in group 1 regard tax compliance as a social norm, $b, \omega \neq 0$. Whereas individuals in group 2 do not attach any importance to the social custom, $b = \omega = 0$. From our previous analysis we know that the entry condition for evasion for group 1 is more restrictive than for group 2, hence we should expected a greater average evasion in group 2, $E_2 > E_1$.

In this setting the programme of the enforcement agency is:

$$\min_{p_1, p_2} N_1 E_1(p_1)(1 - p_1\theta_1) + N_2 E_2(p_2)(1 - p_2\theta_2)$$

$$s.t \quad C_1 N_1 p_1 + C_2 N_2 p_2 \leq C$$

The first order conditions are:

$$\frac{E_1\theta_1 - \varepsilon_1 E_1\theta_1 + \frac{\varepsilon_1 E_1}{p_1}}{C_1} = \lambda \tag{31}$$

$$\frac{E_2\theta_2 - \varepsilon_2 E_2\theta_2 + \frac{\varepsilon_2 E_2}{p_2}}{C_2} = \lambda \tag{32}$$

The solution depends on the values of the parameters. In what follows we run two simulations, each with three scenarios with different parameter values. Table 1 presents our first simulation. For convenience we define group 1 as "individuals" and group 2 as "corporations". In the first scenario we assume that both groups have the same elasticity of evasion, ε, and the average amount of evasion in both groups, E, is observed and is substantially higher for corporations (100 times higher than for individuals). We derive the optimal probability of an audit, p, under the assumption that the functional form for the average evasion is $E_i = \alpha_i p_i^{-\varepsilon_i}$. The optimal coverage rate for individuals is 0.6% and for corporations 18.9%. This gives an aggregate enforcement resource allocation of £56m to individuals and £944m to corporations.

In scenario 2 we assume a much smaller value of the elasticity of evasion for corporations (0.1 instead of 0.5). This could be due to the fact that companies attach less importance to the social norm. All other parameters are the same as in scenario 1. The change to the allocation compared to scenario 1 is slight. The values for the audit probabilities are 0.7% for individuals and 18.7% for corporations and the allocation is £66m and £934m.

In scenario 3, where the evasion elasticity for corporations is reduced to 0.01, there is still little change. Over 93% of total resources are still allocated to corporations. This is because average evasion for corporations is so high relative to individuals that the direct effect of investigations always dominates the combined direct and indirect effects for individuals and most resources are optimally allocated to investigate corporations.

Simulation 1							
		Scenario1		Scenario2		Scenario3	
		Same epsilon		Corp less elastic		Corp much less elastic	
		Indivs	Corps	Indivs	Corps	Indivs	Corps
Assumptions	**N (million)**	9	1	9	1	9	1
	c	1000	5000	1000	5000	1000	5000
	theta	5	3	5	3	5	3
	elasticity of evasion	0.5	0.5	0.5	0.1	0.5	0.01
Choice var	**p**	**0.006214**	**0.188815**	**0.007355**	**0.18676**	**0.007708**	**0.1861248**
Dependent var	**E**	500	50000	460	50055	449	50007
Lagrangian term	**direct**	2500	150000	2298	150164	2245	150022
	2nd term	-1250	-75000	-1149	-15016	-1122	-1500
	3rd term	40231	132405	31241	26802	29119	2687
	indirect	38981	57405.02	30092	11785	27997	1187
	total	41481	207405	32390	161949	30242	151208
	lambda	41.481	41.481	32.38987	32.38987	30.24161	30.241613
Constraint	**Resource all. (m)**	**56**	**944**	**66**	**934**	**69**	**931**
	budget (m)	1000		1000		1000	

Table 1 - Simulation 1: optimal allocation of investigation resources within two groups of taxpayers with different elasticities of evasion.

In the second simulation, which is represented in table 2, the value of average evasion for corporations is reduced from 50,000 to 5,000 (only 10 times higher than for individuals). The elasticities and all the other parameter values in each scenario are the same as in our first simulation. In this case the share of the total resources allocated to corporations decreases from 54% to 38% from scenario 1 to scenario 2, and further to 23% in scenario 3. The greater indirect effect of investigations for individuals plays now a major role for the marginal reduction in the payment gap and the optimal allocation of resources is such that more resources are devoted to investigate individuals than corporations.

Simulation 2									
		Scenario1			Scenario2			Scenario3	
		Same epsilon			Corp less elastic			Corp much less elastic	
		Indivs	Corps		Indivs	Corps		Indivs	Corps
Assumptions	**N (million)**	9	1		9	1		9	1
	c	1000	5000		1000	5000		1000	5000
	theta	5	3		5	3		5	3
	epsilon	0.5	0.5		0.5	0.1		0.5	0.01
Choice var	**p**	**0.051179**	**0.107878**		**0.069283**	**0.075291**		**0.085776**	**0.0456034**
Dependent var	**E**	500	5000		430	5183		386	5043
Lagrangian ter	**direct**	2500	15000		2149	15549		1931	15130
	2nd term	-1250	-7500		-1074	-1555		-966	-151
	3rd term	4885	23174		3101	6884		2251	1106
	indirect	3635	15674		2027	5329		1286	955
	total	6135	30674		4176	20878		3217	16084
	lambda	6.13485	6.13485		4.175678	4.175678		3.216861	3.2168609
Constraint	**Resource all. (m)**	**461**	**539**		**624**	**376**		**772**	**228**
	budget (m)	1000			1000			1000	

Table 2 - Simulation 2: optimal allocation of investigation resources between two groups of taxpayers with different elasticitity of evasion.

In table 3 we consider how the optimal allocation of resources varies with the budget available to conduct investigations. The analysis is based on the parameters for simulation 2, scenario 2, gradually increasing the enforcement agency's budget constraint, from £1m to £5,000m. Initially most of the enforcement resources are put into individuals. For a budget of £1m, 95% of the resources are devoted to individuals and for a budget of £100m, still 87.4% of the resources go into individuals. However, owing to the strong deterrent effect of investigations, the average level of evasion for the individuals quickly decreases and it becomes advantageous to switch resources to companies, primarily in order to obtain the direct yield. For a budget of £ 2,500m there is a switch in the audit probability and resources from individuals to corporations.

Simulation 3							
	budget	p-indiv	p-corp	Res all.-indiv	Res all-corp	E-indiv	E-corp
	m	%	%	%	%	m	m
Scenario 2	*5000*	11.37	*79.5*	*20.5*	*79.5*	*335*	*4095*
Budget growth path	1	0.01	0.001	95	5	11010	12655
	10	0.11	0.01	**94.1**	5.9	3498	9887
	100	0.97	0.25	**87.4**	12.6	1148	7282
	1000	6.93	7.53	**62.4**	37.6	430	5183
	2500	9.94	32.1	**35.8**	**64.2**	359	4483
	5000	11.37	79.54	20.5	79.5	335	4095

Table 3 - Optimal allocation of investigation resources for different values of the enforcement agency's budget constraint.

We have considered the optimal allocation of investigation resources between two groups of taxpayers, one group where tax compliance assumes the characteristics of a social norm and the other group with no such considerations. The findings from our simulations suggest that the decision how to optimally allocate investigation resources depends on the average evasion and on the elasticity of evasion in each group. The higher the average evasion in one group, the more substantial the direct effect of investigations and the more resources should be allocated to that group. In fact the direct effect tends to outweigh the combined direct and indirect effect in the other group. In this case the elasticity of evasion has less of a role in the decision on how to target different groups of taxpayers. As the difference in the direct effect across groups of taxpayers gets smaller, evasion elasticities assume a more important role. The results also show that there are diminishing returns from investigations. If the audit probability is very low, the indirect effect is very high and the decrease in tax evasion induced by a rise in the audit rate is quite substantial. But as the audit rate keeps on increasing the marginal reduction in tax evasion gets smaller and smaller. This implies that as more resources become available for investigations, they shouldn't be constantly focused on the group where initially the marginal reduction in the payment gap was greater.

6 Conclusion

In this paper we have focused on the effects of investigations on tax compliance. Recent empirical studies have attempted to estimate the impact of audits on tax compliance by measuring the elasticity of reported income and reported tax with respect to audits. Some of these studies also provide estimates for the direct and indirect effects of investigations. Results however are not very easily compared and the methodology used to obtain the estimates does not allow to distinguish the factors that determine the direct and indirect effects.

We decompose the elements of the direct and indirect effects in quite a general model and show that the ratio of the two depends on the behavioural elasticity, on how intensively a given group of taxpayers is investigated and on the operational effectiveness of investigations. The intensity and the operational effectiveness of investigations do indeed vary across fiscal areas and also across different groups of taxpayers within the same fiscal area. Hence there is no reason to expect the ratio to be constant across different groups of taxpayers or different fiscalities. It then becomes important to estimate the different components of the ratio between the indirect and direct effects to have some insights on how the impact of audits differ across different groups of taxpayers. The estimation of the elasticity of evasion is particularly problematic. The elasticity of evasion we define is at the level of a group of taxpayers. One important question is whether we should assume it to be constant across different groups of taxpayers. But this question can only be addressed empirically. We are however interested in understanding how the behavioural response to an increase in the audit rate may differ if considered at individual level or at the level of a group of taxpayers. This is important for choosing the level of data aggregation to calculate the elasticity of evasion. We investigate how different assumptions on the motivations driving taxpayers' behaviour may affect the elasticity of evasion. In particular we compare a setting where taxpayers decide whether or not to be honest with no regard for the behaviour of other taxpayers to a situation where the decision is interdependent and tax compliance is a social norm.

Our results suggest that if tax compliance is a social norm in the relevant community this has important implications on the impact of an increase in the coverage rate on voluntary compliance. At the aggregate level of the community of taxpayers, we can expect a higher response to a change in the allocation of audit resources than in the absence of a social norm. Essentially, social norms introduce a multiplier effect: generating greater compliance through a deterrent effect causes even more people to become compliant through the social norm. The magnitude of the impact of audits on aggregate behaviour will therefore be higher the greater the importance attached to the social norm. Hence, when tax compliance is a social norm, the mechanism of enforcement is richer and also more effective. In fact, if the social norm is sustained by moral commitment to the law and by the desire to conform with the behaviour of honest taxpayers, it is possible to use informal sanctions as an additional tool to enforce compliance. These rely on the inhibitory power of moral commitment to the law and of the threat of social disapproval. But at the same time, as we show, formal sanctions are also more effective.

The importance attached to the social norm is likely to vary across taxpayers, so that we should expect a heterogeneous response to a rise in the probability of detection by different taxpayers. Also, how compelling the social norm argument is depends on the type of non-compliance: for example, taxpayers have different attitudes towards tax evasion and tax avoidance. Kirchler et al. (2003), conduct a survey among 252 fiscal officers, business students, business lawyers and small bunisness owners and report that taxpayers discriminate between tax avoidance, tax evasion and tax flight (defined as the relocation of business only in order to save taxes), and they also perceive them as unequally fair, with tax evasion being the least positively regarded. Tax evasion can also be undertaken in different forms: late filing, false benefit claims, income underreports, expenses overstatements, VAT evasion, frauds linked to organised crime, and taxpayers may have different attitudes and opinions concerning these types of non-compliance. For example, Orviska and Hudson (2002) conduct a survey on moral attitudes to tax evasion and find that attitudes are more hostile to the evasion involving benefits and least hostile to the one involving VAT.

We also show that it is not necessarily optimal to put more resources in the groups where the social norm is more compelling. The optimal allocation will depend also on the average evasion taking place in the different groups. The greater the average evasion the greater will be the direct effect of investigations. And the greater the difference in the direct effect across different groups of taxpayers the less important are the elasticities in determining the optimal allocation of resources.

One issue that we recognise but do not develop here is how people form their perceptions about the probability of being investigated. As we mentioned above the individual elasticity of evasion confounds two different elasticities: the sensitivity to the perceived frequency of inspections and the sensitivity of the perceived probability of inspection to the actual frequency of inspection. Distinguishing the factors which determine these two elasticity is very important to better understand how taxpayers are influenced by audits.

In modelling the optimal allocation of resources among different groups of taxpayers, we assumed that in each group taxpayers are only affected by the probability of being investigated in their group. But there may be some spillover effects: an increase in the coverage rate in one group could be observed in another group and this could alter the perception of being investigated also in this group and increase their voluntary compliance. These are possible extensions of our analysis.

References

[1] Akerlof, G. A., 1980, A theory of social custom of which unemployment may be one consequence. Quarterly Journal of Economics, 95, 749-775

[2] Allingham, M., and Sandmo, A. 1972, Income tax evasion: a theoretical analysis, Journal of Public Economics, 1, 323-338.

[3] Beron, K.J., H.V. Tauchen, and A.D. Witte, 1992, The Effect of Audits and Socioeconomic Variables on Compliance, in *Why People Pay Taxes: Tax Compliance and Enforcement*, Joel Slemrod, ed., (The University of Michigan Press, Ann Arbor), 67-89.

[4] Blumenthal, M., Christian, C., Slemrod, J., 2001, Do normative appeals affect tax compliance? Evidence from a controlled experiment in Minnesota, National Tax Journal, 54, 1,

[5] Cowell, F., 1990, Cheating the government, The MIT Press, Cambridge, MA.

[6] Crane, S., Nourzad, F., 1993, An empirical analysis of factors that distinguish those who evade on their tax return from those who choose not to file a return, Public Finance, 49, 106-118.

[7] Cremer, H., Marchand, M., Pestieau, P., 1990, Evading, auditing and taxing, Journal of Public Economics, 43, 67-92.

[8] Dubin, J., M. Graetz and L. Wilde, 1990, The Effect of Audit Rates on the Federal Individual Income Tax, 1977-1986, National Tax Journal 43, 395-409.

[9] Erard, B, 1992, The Influence of Tax Audits on Reporting Behavior, in *Why People Pay Taxes: Tax Compliance and Enforcement*, Joel Slemrod, ed., (The University of Michigan Press, Ann Arbor), 95-114.

[10] Erard, B., Ho, C., 2001, Searching for ghosts: who are the nonfilers and how much tax do they owe?, Journal of Public Economics, 81, 25-50.

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

[12] Grasmick, H., Green, D., 1980 Legal Punishment, Social Disapproval and Internalization as Inhibitors of Illegal Behavior. The Journal of Criminal Law and Criminology, 71, 3, 325-335.

[13] Greenberg, J., 1984, Avoiding tax avoidance: a (repeated) game-theoretic approach, Journal of Economic Theory, 32, 1-13.

[14] Kirchler, E., Maciejovsky, B., Schneider, F., 2003, Everyday representations of tax avoidance, tax evasion, and tax flight: do legal differences matter?, Journal of Economic Psychology, 24, 535-553.

[15] Myles, G., Naylor, R., 1996, A model of tax evasion with group conformity and social customs. European Journal of Political Economy, 12, 49-66.

[16] Kristina Murphy, 2003, Who me? I didn't do anything wrong: Trust, resistance and compliance among tax scheme investors. MIMEO.

[17] Orviska, M., Hudson, J., 2002, Tax evasion, civic duty and the law abiding citizen, European Journal of Political Economy, 19, 83-102.

[18] Plumley A., 1996, The Determinants of Individual Income Tax Compliance, Internal Revenue Service Publication 1916 (Rev. 11-96) Washington, DC: 1996

[19] Reinganum, J., Wilde, L., 1985, Income tax compliance in a Principal-Agent framework, Journal of Public Economics, 26, 1-18.

[20] Reinganum, J., Wilde, L., 1991, Equilibrium enforcement and compliance in the presence of tax practitioners, Journal of Law, Economics and Organization, 7, 1, 163-181.

[21] Slemrod, J., Blumenthal, M., Christian, C., 2001, Taxpayer response to an increased probability of audit: evidence from a controlled experiment in Minnesota, Journal of Public Economics, 79, 455-483.

[22] Tauchen, H.V., A.D. Witte , and K.J. Beron, 1989 Tax Compliance: An Investigation Using Individual TCMP Data, NBER Working Paper No. 3078, (National Bureau of Economic Research, Inc., Cambridge, MA).

[23] Torgler, B., 2002, Speaking to theorists and searching for facts: tax morale and tax compliance in experiments, Journal of Economic Surveys, 16(5), 657-683.

[24] Wrong, P., 1961, The Oversocialized Conception of Men in Modern Sociology, 26 Am. Soc. Rev. 183.

[25] Yitzhaki, S., 1974, A note on Income Tax Evasion: A Theoretical Analysis, Journal of Public Economics, 3, 201-202.

7 Appendix

Allocating investigation resources: the targeting case.

In section 3 we assumed a random audit within a specific group of taxpayers. The tax authority did not have any information of how evasion differs across taxpayers within the same group, so that was not able to distinguish marginal evasion from average evasion withing the same group of taxpayers.

We now suppose that the tax authority has some information about taxpayers that enables it to distinguish the *expected* amount of evasion that one type of taxpayer might be involved in relative to another.

More precisely, assume that the actual amount of evasion, e_{jk}, in which the j-*th* taxpayer in group k is involved given by:

$$e_{jk} = \phi_k(x_j, p_k) + \xi_k \tag{33}$$

where x_j is a vector of observable individual characteristics and $\xi_k$is a random variable with zero mean. We assume that the tax authorities have a risk-profiling model that enables to observe for any given taxpayer the expected amount of evasion

$$\overline{e}_{jk} = \phi_k(x_j, p_k) \tag{34}$$

For simplicity, assume that the effectiveness of investigations is independent of taxpayer type and is given once again by the constant θ_k.

Since behaviour depends solely on the fraction of people investigated - and not their identity - in order to minimise the tax gap the tax authority will obviously want to target those taxpayers with highest expected evasion.

To understand the implications of this, for expositional simplicity, we assume that x is a scalar and that the distribution of x in the k-*th* group is given by the density function $l_k(x)$. We also assume that $\phi(x, p)$ is a strictly increasing function of x.

This implies that if a fraction p_k of taxpayers in group k are investigated these will be all taxpayers for whom $x \geq \underline{x}_k$, where $\underline{x}_k$ is defined by

$$\int_{\underline{x}_k}^{\infty} l_k(x)dx = p_k \tag{35}$$

For later purposes notice that

$$-l(\underline{x}_k)\frac{d\underline{x}_k}{dp_k} = 1 \tag{36}$$

As before let $E_k(p_k) = \int_0^{\infty} \phi_k(x_j, p_k)l(x)dx$ be the average amount of evasion in group k, and let

$$E_k^t(p_k) = \frac{\int_{\underline{x}_k}^{\infty} \phi_k(x, p_k)l_k(x)dx}{p_k} \tag{37}$$

be the average amount of evasion amongst the sub-group of group k who are targeted for investigations. Obviously, $E_k < E_k^t$.

Also let $E_k^m = \phi_k(\underline{x}_k, p_k)$ be the expected evasion of the marginal taxpayer who is targeted for investigation. Obviously $E_k^m < E_k^t$.

Since tax authorities typically only investigate a very small fraction of taxpayers, we would normally expect that $E_k < E_k^m$, and so $E_k < E_k^t$.

Finally let $\varepsilon_k = -\frac{dE_k}{dp_k}\frac{p_k}{E_k}$ be the sensitivity of taxpayer behaviour to the probability of investigation for the k-th group of taxpayers as a whole, and $\varepsilon_k^t = -\frac{dE_k^t}{dp_k}\frac{p_k}{E_k^t}$ be the sensitivity of the sub-group of taxpayers who are targeted for investigation.

The payment/tax gap of the k-th group of taxpayers is

$$G_k = N_k \left[E_k - \theta_k \int_{\underline{x}_k}^{\infty} \phi_k(x, p_k) l_k(x) dx \right] \tag{38}$$

An alternative way of writing this is

$$G_k = N_k[E_k - \theta_k p_k E_k^t] \tag{39}$$

From (38) we get:

$$MRPG_k = -\frac{1}{N_k}\frac{dG_k}{dp_k} = \theta_k \phi_k(\underline{x}_k, p_k) + \frac{E_k \varepsilon_k}{p_k} - \theta_k \int_{\underline{x}_k}^{\infty} \left[-\frac{\delta \phi_k}{\delta p_k} \right] l_k(x) dx \tag{40}$$

From (37) it is straightforward to show that

$$\int_{\underline{x}_k}^{\infty} \left[-\frac{\delta \phi_k}{\delta p_k} \right] l_k(x) dx = E_k^t \varepsilon_k^t + (E_k^m - E_k^t) \tag{41}$$

Substitute (41) into (40) and we get:

$$MRPG_k = \theta_k E_k^m + \frac{E_k \varepsilon_k}{p_k} - \theta_k \left[E_k^t \varepsilon_k^t + (E_k^m - E_k^t) \right] \tag{42}$$

But notice that we can re-write this as:

$$MRPG_k = \theta_k E_k^t - \theta_k E_k^t \varepsilon_k^t + \frac{E_k \varepsilon_k}{p_k} \tag{43}$$

which is the formula one would get from (39).

The formula in (43) is very similar to that in (5).

The question is what we can say from this about the ratio of the indirect to the direct effect of investigation activity. This depends on what one means by the direct and indirect effect - the average or marginal effect of the tax authority activity.

If we defined the direct effect as $\theta_k E_k^t$ - i.e. the *average* yield form an investigation - then, from (43), the formula for the ratio of the indirect effect to the direct effect would be:

$$R_{3k} = \varepsilon_k \left[\frac{E_k}{E_k^t} \frac{1}{p_k \theta_k} - \frac{\varepsilon_k^t}{\varepsilon_k} \right] \tag{44}$$

If the tax authority were unable to target investigation resources then we would have $E_k = E_k^t$; $\varepsilon_k = \varepsilon_k^t$ and (44) would collapse to (8).

If the tax authority were able to target,but there were no reason to think that the behavioural response of targeted taxpayers was significantly different from non-targeted taxpayers, then we would have $E_k < E_k^t$; $\varepsilon_k \approx \varepsilon_k^t$ and so we would have

$$R_{3k} \approx \varepsilon_k \left[\frac{E_k}{E_k^t} \frac{1}{p_k \theta_k} - \frac{\varepsilon_k^t}{\varepsilon_k} \right] < \varepsilon_k \left[\frac{1}{p_k \theta_k} - 1 \right] \tag{45}$$

and so, as we might expect, targeting gives a lower ratio of the indirect to the direct effect.

However this approach would be very odd since the direct doesn't reflect the fact that the tax authority is targeting resources and so, if given extra resources, would deploy those on the marginal taxpayer. So if we define the direct effect as $\theta_k E_k^m$ - i.e. the *marginal* yield per investigation - then, from (42), the ratio of the indirect effect to the direct effect is:

$$R_{3k} = \varepsilon_k \frac{E_k^t}{E_k^m} \left[\frac{E_k}{E_k^t} \frac{1}{p_k \theta_k} - \frac{\varepsilon_k^t}{\varepsilon_k} \right] + \left[\frac{E_k^t}{E_k^m} - 1 \right] \tag{46}$$

If we compare (46) with (44) then we see that the expression on the RHS of (46) is larger than the expression on RHS of (44) - which is not surprising since the MRPG is the same and, in (46) we are using as denominator the marginal direct effect which is smaller than the average, which is denominator in (44). This shows up in two ways. First of all we have to multiply (44) by $\frac{E_k^t}{E_k^m} > 1$ - which is a re-scaling effect to reflect the different denominators - but then we have to add $\left(\frac{E_k^t}{E_k^m} - 1 \right) > 0$.

If the coverage rate is very low then we would expect the marginal and average values to be very similar, so, the values we get in (44) and (46) are likely to be very similar.

Once again, if the tax authority were unable to target investigation resources then we would have $E_k = E_k^t$; $\varepsilon_k = \varepsilon_k^t$ and (46) would collapse to (8).

If the tax authority were able to target, but there were no reason to think that the behavioural response of targeted taxpayers was significantly different from non-targeted taxpayers, then we would have $E_k < E_k^m < E_k^t$; $\varepsilon_k \approx \varepsilon_k^t$ and so we would have

$$R_{3k} \approx \varepsilon_k \frac{E_k^t}{E_k^m} \left[\frac{E_k}{E_k^t} \frac{1}{p_k \theta_k} - 1 \right] + \left[\frac{E_k^t}{E_k^m} - 1 \right] \tag{47}$$

What this suggests is that in many circumstances the ratio of the indirect effect to the direct effect could be well approximated by

$$R_{3k} \approx \varepsilon_k \left[\frac{E_k}{E_k^t} \frac{1}{p_k \theta_k} - 1 \right] \tag{48}$$

which is just a mild adjustment to the original formula in (8).