

An Evasive Topic: Theorizing about the Hidden Economy*

Abstract This paper reviews some central issues that arise in theorizing about tax evasion decisions and the hidden economy. It starts from the Allingham-Sandmo (1972) modelling of the tax evasion decision as a choice under uncertainty based on expected utility maximization and risk aversion. It goes on to discuss alternative specifications of the taxpayer's preferences with particular regard to the explanation of the extensive margin, i.e. the decision on whether or not to engage in tax evasion. It extends the model to the case of variable labour supply with work in both official and black labour markets. It then considers the application of the theory to taxes on wealth and income from capital, indirect tax evasion and smuggling. It also includes a consideration of general equilibrium effects and of the problems that evasion causes for the theory of optimal income and commodity taxes. It concludes with a brief discussion of the implications of tax evasion for economic policy in the welfare state.

Keywords Tax evasion, expected utility, the hidden economy, optimal taxation

JEL Classification: H26

*This paper formed the basis for a Keynote Lecture at the 66th Annual Congress of the International Institute of Public Finance at Uppsala University, August 2010. I am grateful to Friedrich Schneider and the referees for helpful comments and suggestions.

Agnar Sandmo, Department of Economics, Norwegian School of Economics, 5045 Bergen, Norway. E-mail: agnar.sandmo@nhh.no.

1 Introduction

Most of the literature on the theory of taxation is based on a simple theory of taxpayers' compliance: They do as they are told. If the size of the tax base as well as the tax rate is known, the revenue is simply the product of the two. But economists and administrators alike have known for a very long time that this is not strictly true. An early example of the awareness of the importance of tax evasion can be found in Adam Smith's *The Wealth of Nations* (Smith 1776). Among his four basic maxims or principles for the design of tax systems, the fourth may be of special interest in the present context:

“Every tax ought to be so contrived as both to take out and to keep out of the pockets of the people as little as possible, over and above what it brings into the publick treasury of the state.” (Smith 1776; 1976, p. 826.)

Modern readers may see in this statement an early understanding of the efficiency costs of taxation, and Smith does in fact argue that taxes “may obstruct the industry of the people,” which is a central concern of the modern literature on optimal taxation. In addition, however, two of his reasons have to do with the costs of tax administration in ensuring tax compliance and prevent evasion. On the one hand there is the cost incurred by the government: The tax “may require a great number of officers, whose salaries may eat up the greater part of the produce of the tax” (p. 826). On the other hand one must also take account of the cost to the taxpayer: “By subjecting the people to the frequent visits, and the odious examination of the tax-gatherers, it may expose them to much unnecessary trouble, vexation, and oppression” (p. 827). Of particular interest is his argument that

“by the forfeitures and other penalties which those unfortunate individuals incur who attempt unsuccessfully to evade the tax, it may frequently ruin them and thereby put an end to the benefit which the community might have received from the employment of their capitals” (p. 826).

Clearly, Smith was well aware of the problems raised by tax evasion and its implications for the costs of tax administration. It is of particular interest to note from the last quotation that he realized that not only evasion itself but also the policies adopted to control it may impose

costs on society by increasing the gap between the private burden of taxation and government revenue.

In spite of Smith's attention to the problem, the interest in tax evasion gradually faded out from the public economics literature. In his *Study in Public Finance* Pigou pointed out briefly that under adverse circumstances "an income tax would be rendered unworkable by evasion" (Pigou 1928; 1947, p. 115) but did not pursue the subject any further except by noting that "what measure of evasion is in fact practised is a matter of controversy" – a statement that remains true to this day. Thirty years later, Musgrave's famous treatise (Musgrave 1959) made no mention of tax evasion, and it was not until the 1970s that the subject became an important area of research for public economists. This was the start both of a long series of empirical studies and of a formalized theory, beginning with a paper on income tax evasion by Allingham and Sandmo (1972).¹ It is natural to take my point of departure from this model and start in Section 2 with a brief outline of its main assumptions and results. Section 3 discusses some interpretations and reformulations of the basic model such as its implications for the "extensive margin" of tax evasion, its robustness with respect to alternative assumptions about taxpayer preferences, and its extension to include variable labour supply and work in the hidden economy. Section 4 considers topics related to the evasion of taxes on capital and of indirect taxes, which also raises the problem of how to model tax evasion decisions by firms. In Section 5 we consider the tax compliance decision in a systems context by discussing general equilibrium effects, issues of optimal taxation and the tax administration perspective. Section 6 concludes with a summary of the main results and a brief discussion of the significance of tax evasion when judging the performance of the modern welfare state.

This paper should be read as a survey of the *issues* involved in theorizing about the hidden economy, not as a survey of the extensive literature. Recent survey articles that include further references to the literature include Andreoni, Erhard and Feinstein (1998), Slemrod and Yitzhaki (2002) and Sandmo (2005). An excellent early exposition of the theory is Cowell (1990).

2 Income tax evasion: The basic model

¹ A related article by Srinivasan (1973) addressed many of the same issues but assumed risk neutrality rather than risk aversion on the part of the taxpayer.

Taxpayers have incentives to escape from the private burden of taxation either by engaging in avoidance or evasion activities. The distinction between the two concepts is essentially that tax avoidance is legal while evasion is not. An implication of this is that in the case of avoidance activities, such as exploiting loopholes in the tax law, it is essential for the taxpayer to report his transactions openly and accurately in order to make sure that he receives the reductions in his tax liability that he claims.² By contrast, in the case of evasion the taxpayer is concerned with hiding his activities from the tax authorities. In the present paper our concern will be with tax evasion.

The model in Allingham and Sandmo (1972) can be seen as an adaptation of a simple model of portfolio choice with two assets, one safe and one risky. There is an amount of true income which is assumed to be exogenous and unknown to the tax collector. The amount reported corresponds to the safe asset in portfolio theory: The tax to be paid on this amount is known with certainty. The amount not reported corresponds to the risky asset: With some probability the attempt at evasion will be discovered, in which case the taxpayer has to pay tax on this amount at a penalty rate which is higher than the regular tax rate, but it is also possible that the underreporting will not be discovered, in which case no tax will be paid. The taxpayer has to decide how much of his income to report and how much not to report, taking account both of the expected tax rate and its degree of uncertainty. He is assumed to make this choice with the objective of maximizing expected utility. In the simplest version of the model, after-tax income is the only argument in the utility function.

I will not go into any analytical detail, but in order to fix ideas it may be useful to introduce a bit of mathematical notation. Let true income be represented by W , and let t be the statutory tax rate which is assumed to be constant. Further, let e be the amount not reported and θ the penalty rate. If the attempted evasion is not discovered, after-tax income is

$$Y = W - t(W - e) = (1 - t)W + te. \quad (1)$$

In the case where the attempted evasion is discovered by the tax authority, income after tax becomes

² However, he may attempt to bend the rules that regulate the opportunities for avoidance, thereby crossing the border between avoidance and evasion.

$$Z = (1-t)W + te - \theta e = (1-t)W - (\theta-t)e. \quad (2)$$

If U is the utility function of income and p is the probability of detection, expected utility can be written as

$$V = (1-p)U(Y) + pU(Z). \quad (3)$$

The taxpayer is assumed to be risk averse, so that the second-order derivative of U is negative. In addition, I shall also make the natural assumption that the measure of absolute risk aversion, $-U''(\cdot)/U'(\cdot)$, is decreasing in income; a standard assumption in the economics of uncertainty. The taxpayer chooses the level of e that maximizes expected utility, which is determined by the first order condition

$$(1-p)U'(Y)t - pU'(Z)(\theta-t) = 0. \quad (4)$$

Let us first assume that e is strictly positive: The taxpayer finds that expected utility maximization calls for some amount of tax evasion. The optimum conditions can then be used to derive hypotheses about the taxpayer's reactions to changes in the values of the exogenous parameters of the model, i.e. W , t , θ and p . With the assumption of decreasing absolute risk aversion, tax evasion will be an increasing function of income. The amount of evasion will be a decreasing function of the parameters θ and p ; deterrence can be achieved both through a high probability of detection and by a high penalty in the event of discovery. As regards the effect of the statutory tax rate, there is an ambiguity which can be expressed as a Slutsky decomposition into a positive substitution effect and a negative income effect: On the one hand, a higher tax rate creates a stronger incentive to evade taxes; on the other hand, higher taxation makes the taxpayer less wealthy and accordingly less prone to expose himself to risk. This ambiguity reflects the assumption that the tax is proportional which confounds the effects of the average and marginal tax rates. If instead one assumes that the tax is linear as in the case of the negative income tax, one could separate the two effects and conceptually increase the marginal rate of tax while holding the average rate constant. In that case the prediction would be that an increase of the marginal tax rate unambiguously increases tax evasion.

Needless to say, the model is built on a number of simplifying assumptions. In general, the theoretical perspective on the tax evasion decision is a very narrow one in regarding the tax compliance decision in complete isolation from all other economic choices made by the individual such as occupational choice, labour supply, saving and portfolio decisions. Extensions of the model in these directions will be further discussed below.

It should be kept in mind that the theory is based on the assumption that discovery is equivalent to conviction. In real life, however, the discovery of evasion does not immediately translate itself into a conviction in a court of law since the court may not uphold the initial assessment of the amount of income evaded. Although the taxpayer may know the provisions in the tax law regarding the penalties attached to the amount of income that he does not report, the fact that he does not know for certain what the court's decision will be means that the penalty itself will be uncertain; this complication has been further discussed by Snow and Warren (2005).

3 Further aspects and extensions of the basic model

Conceptually, the amount of tax evasion undertaken can be related both to the number of taxpayers that engage in evasion (the extensive margin) and to the extent of non-compliance by the individual evader (the intensive margin). In the literature on tax evasion following the Allingham-Sandmo article considerable attention has been devoted to the model's implications for the extensive margin and to the question of whether the expected utility approach adequately captures the nature of the taxpayer's decisions. A related issue is whether the simple portfolio framework may be misleading in terms of understanding the interaction with black labour markets and the "unofficial" economy. These issues will be discussed below.

3.1 The extensive margin

Changes in the amount of tax evasion in the economy can occur both on the intensive and extensive margins. The individuals who engage in tax evasion may choose to report more or less of their actual income as their perception of the policy parameters changes, but the total amount of evasion may also change because the fraction of the population that evades income may increase or decrease. So far the analysis has been based on the assumption that the

individual taxpayer has decided not to report part of his income ($e > 0$); analytically, the model is assumed to have an interior solution. If the taxpayer is regarded as a representative agent, he represents an economy in which everyone evades some tax. But the decision to become a tax evader ought to be part of the predictions of the model and should be derived as an implication of it, not stated as an assumption. However, it is in fact easy to show that for an interior solution to be optimal it must be the case that

$$t > p\theta. \quad (5)$$

If the regular tax rate is higher than the expected penalty rate, the risk averse taxpayer will decide not to report some of his income. It is worth noting that this condition makes no reference to the taxpayer's degree of risk aversion or more generally to his preferences. The *extent* of his underreporting, on the other hand, will obviously depend on the shape of his utility function and his degree of risk aversion.

When interpreting this condition it should be kept in mind that the three parameters t , p and θ must be expected to vary across the population and to depend on the taxpayer's income, profession, family status etc. However, many observers have pointed out that even when these complications are taken into account the inequality is likely to be satisfied for the large majority of taxpayers, implying that almost everyone is a tax evader. This, they argue, is unrealistic, so there must be some important element of reality that is missing from the model.

In arguing the implausibility of condition (5), critics have mainly pointed out that it implies an unrealistic critical value for the probability of detection. To get a feeling for the magnitudes involved, suppose that θ is twice the value of t . Then the individual will become a tax evader if the probability of detection is less than one half. But realistic estimates of the frequency of audits indicate a number that is well below 0.10, so that the prediction that almost everyone will underreport their income seems pretty safe. Or does it? It has to be kept in mind that what matters for the taxpayer's decision is not the objective frequency of audit but his subjective perception of what the probability of discovery is. There are in fact empirical data indicating that this might be much higher than the objective probability, and this might go some way towards explaining the otherwise apparent inconsistency between theory and observation. But there are also other reasons why the number of tax evaders in the population is less than may be suggested by condition (5) and the numerical example that was used above.

One such reason is that the possibility of successful evasion depends crucially on the source of income. In most Western countries, labour income is reported directly to the tax collecting agency by the employer. In that case it is pointless for the employee to underreport his income because he is virtually certain to be discovered: $p=1$. Although some workers may have access to additional employment in the black market economy, many have not, and this could account for much of the discrepancy that is claimed to exist between a literal interpretation of the model and the alleged fact of the relative infrequency of tax evasion.

When it comes to capital income the situation is different because the opportunities for hiding true income from the tax collector are substantially higher than in the case of labour. However, at least in the countries that I know well, a good part of capital income is also reported from its source; this would be true of interest on bank deposits and mortgages and of stock dividends in publicly held companies. Even though the opportunities for successful evasion clearly are greater for capital than for labour income, it seems likely that for the large number of taxpayers whose financial assets and liabilities consist of bank deposits and housing mortgages they are fairly limited. My conclusion is therefore that a flexible interpretation of condition (4) does not necessarily justify the inference that the model predicts an unrealistically high share of tax evaders in the economy. (I leave aside the question of how many evaders there actually are; this is clearly an empirical question whose answer is likely to vary from one country to another.)

3.2 Broadening the view of taxpayer preferences

A more fundamental reason why both economists and others who are concerned with these issues might feel uneasy about the implications of the portfolio approach is that it gives a picture of the tax evasion decision which may seem too calculating and cynical. Now there may be good reasons for believing that tax evaders really *are* calculating and cynical, but the objection to the theory must be taken more seriously with regard to the extensive margin. Is it really only their possibly exaggerated estimate of the probability of detection that prevents the majority of the population from becoming tax evaders?

In order to provide a more convincing approach to this issue, there are two avenues that can be explored. One consists in broadening the perspective provided by the expected utility

approach through assuming more general preferences. Another is to discard the use of expected utility theory altogether and base the theory of tax evasion on alternative assumptions about taxpayer preferences.

As regards the first alternative, there are at least two simple extensions of the theory that can provide a more convincing explanation of the condition that must be satisfied for the taxpayer to engage in illegal underreporting. One of these was explored in Allingham and Sandmo (1972): There may be a social stigma attached to the discovery of tax fraud, so that the utility of a given income is lower in the state of discovery. This assumption can be captured by writing expected utility as

$$V = (1-p)U(Y) + pU(Z, s), \quad (6)$$

where s represents the social stigma so that U depends negatively on s . The implication of this generalization is that, given that the state of discovery occurs, the taxpayer derives less utility from his after tax income than he would have done had his fraud not been discovered. An alternative assumption (Sandmo 2005) is that the act of underreporting gives the taxpayer a bad conscience, so that the amount of attempted evasion enters directly into his utility function³. A simple way to represent this is to write expected utility as

$$V = (1-p)U(Y) + pU(Z) - B(e), \quad (7)$$

where $B(e)$ is the disutility or bad conscience attached to the act of underreporting. Comparing the formulations (6) and (7), we see that they represent different psychological assumptions about the preferences of the taxpayer. However, when it comes to the characterization of the extensive margin their implications are similar. Using (7), we can rewrite condition (5) as

$$t > p\theta + B'(0)/U'(W(1-t)). \quad (5')$$

One sees immediately that the assumption of disutility of evasion has the effect of making the condition for positive underreporting to be optimal more restrictive; there is a “conscience tax” which requires that the gap between the statutory tax rate and the expected penalty must

³ A version of this model was also examined by Gordon (1989).

be larger than in the original model; how much larger it must be depends on the taxpayer's preferences. Assumption (6) leads to a similar characterization although the exact form of the condition will be different. Thus, both of the two alternative assumptions (6) and (7) provide us with an explanation why taxpayers may refrain from exploiting opportunities for tax evasion that are apparently profitable on the basis of a straightforward calculation of expected gains. It is worth emphasizing, however, that neither alternative leads to any change in the comparative statics predictions of the model for the case of an interior solution; the difference that they make relates solely to the explanation of the extensive margin of tax compliance behaviour.

An alternative to enriching the expected utility approach is to discard this hypothesis altogether and adopt a different rationalization of choice under uncertainty. Thus, Bernasconi (1998) has explored the implications of the rank-dependent expected utility hypothesis for the explanation of tax evasion behaviour, and Dhami and al-Nowaihi (2007) have analyzed the implications of using prospect theory. Like the modifications of expected utility theory, this leaves the analysis of the intensive margin unaffected while the explanation of the extensive margin has to be modified in the same direction as above. Yet another assumption that makes only minimal assumptions about taxpayer rationality is the state preference approach that stems from the work of Arrow (1953) and Debreu (1959). Here one simply postulates that the taxpayer has a preference ordering over income in the two states of discovery and non-discovery which can be represented by the ordinal utility function

$$V = \Phi(Y, Z). \tag{8}$$

In this case, the explanation of the extensive margin allows for a number of interpretations and can be made consistent with a wide range of empirical observations. The comparative statics results for the case of an interior solution are remarkably similar to those of the expected utility model. Thus, the Slutsky decomposition of the effect of the statutory tax rate has exactly the same form in the two cases: There is a positive substitution effect and a negative income effect. The generality of this approach is also to some degree its weakness; in particular, it does not distinguish conceptually between "pure" preferences and probability beliefs. The marginal rate of substitution between Y and Z – the slope of the indifference curve – reflects in part the desire for income in the two states, in part the taxpayers' beliefs concerning the likelihood of the two states occurring. This weakness may make the approach

less appealing to the economist concerned with explaining the role played by the probability of detection, but the state preference approach is still of interest in providing a benchmark against which the expected utility hypothesis as well as other approaches may be judged.

My own conclusion regarding the merits of alternative explanations of tax compliance behaviour is that I will continue to use the expected utility approach but with the extensions represented by the social stigma or the conscience tax; this seems to me to be both analytically simple and with considerable explanatory power. However, the ongoing study of people's actual behaviour in risk-taking situations may well lead to further progress and new insights, so that this conclusion may have to be revised in the future.

3.3 Income tax evasion and labour supply

The Allingham-Sandmo portfolio model is so stylized that it is hard to decide what kind of economic agent the evader is and how the analysis can be related to the more standard issues in public finance, like the effects of taxation on labour supply and saving. It does not specify whether the taxpayer's income derives from labour or capital, and although the true tax base has been referred to as "income" it could equally well have been interpreted as wealth and the tax rate as wealth taxation. In particular, this implies that the analysis provides no direct connection to the study of labour supply in the hidden economy, which must clearly be counted as a serious shortcoming of the theory. When extending it to take account of black market labour, however, we should not jump to the conclusion that income from labour constitutes the most important part of tax evasion; this remains an empirical question⁴.

Sandmo (1981) extends the analysis of the portfolio model to cover the case of variable labour supply⁵. In this model there are two classes of taxpayers, evaders and non-evaders. The non-evaders are individuals who have no opportunity to evade taxes. The evaders are individuals of the same type that was described in the portfolio model; they may or may not evade taxes, depending on the values of the relevant tax parameters. Thus, even the evaders may choose not to evade taxes; the difference between them is that while the non-evaders are honest by

⁴ For recent surveys of studies of the extent of the hidden or shadow economy see the articles by Friedrich Schneider (2005) and Lars P. Feld and Friedrich Schneider (2010).

⁵ This issue has also been considered by Baldry (1979) and Pencavel (1979), among others. The related problem of tax evasion in the context of occupational choice has been studied by Pestieau and Possen (1991).

necessity (their probability of detection equals one) the evaders, if honest, are so by choice. In the following I focus on the modelling of the evaders' behaviour.

The standard textbook model of labour supply assumes that the consumer maximizes a utility function $U(C, L)$ with consumption and leisure as arguments, and this is a natural starting point for a model of black market work. Using the notation from the portfolio model, consumption in the cases of non-discovery and discovery will be denoted by Y and Z , respectively. Expected utility can then be written as

$$V = (1-p)U(Y, L) + pU(Z, L). \quad (9)$$

The time constraint in this model is that the sum of leisure time, hours of work in the regular economy and hours worked in the hidden economy must equal the consumer's time endowment, and this can be written as

$$L + H + h = T,$$

where H represents hours worked in the regular economy and h the hours supplied to the hidden sector.

As in the portfolio model, there is a budget constraint for each of the two cases of discovery and non-discovery. If the wage rates of the two sectors are denoted by w and w^* , respectively, we can write the consumption levels for the two outcomes as

$$Y = (1-t)wH + w^*h, \quad (10)$$

$$Z = (1-t)wH + (1-\theta)w^*h. \quad (11)$$

I shall skip the details of the conditions for utility maximization and proceed directly to a discussion of the comparative statics results for the case of an interior solution (the intensive margin). Since the present model is more general than the pure portfolio model, having more arguments in the utility function and more choice variables, it is to be expected that the results will be more complex, and this turns out to be true. Nevertheless, the results concerning the effects of the penalty rate and the probability of detection are quite similar to the previous

model. Of special interest is the prediction that the model yields concerning the effect of an increase in the marginal tax rate: Does it encourage work in the hidden economy? We have seen that the answer is ambiguous even in the portfolio model because of conflicting income and substitution effects. But even if we disregard income effects and focus on the substitution effects there emerges another type of ambiguity in the present model: The primary effect of an increase in the tax rate is to discourage work in the regular economy, but without further restrictive assumptions the model is not able to predict whether this is reflected in more work in the hidden economy, more leisure, or both.

As regards the extensive margin, it turns out that condition (5) in the portfolio model must be extended to take account of the difference in wages between the regular and the hidden economy. The condition for some work in the hidden economy to be optimal from the taxpayer's point of view is

$$w(1-t) < w^*(1-p\theta). \quad (12)$$

If the after-tax wage rate in the regular labour market is less than the expected after-tax wage in the hidden economy, the worker will supply some work to the shadow economy. One sees immediately that this condition is identical to (5) for the special case where the wage rates in the two sectors are identical, i.e. $w = w^*$. In a broader analytical framework, it would be interesting to study in more detail the general equilibrium determinants of the wage gap between the two sectors, including the analysis of the incidence of the various policy parameters.

It may be realistic to assume that in the majority of cases labour's marginal productivity is lower in the hidden economy, so that under competitive conditions $w > w^*$. Production is likely to be more capital intensive in the regular economy and this indicates that the marginal productivity of labour is higher there than in the hidden economy. Under this assumption it is interesting to note from (12) that the opportunity for tax evasion may cause the worker to switch some of his labour effort from the high productivity to the low productivity sector, involving a loss of production efficiency for the economy as a whole.

Although the simultaneous analysis of labour supply and tax evasion is an interesting one and clearly an exercise with policy relevance, one may wonder whether it adequately captures the

sequential nature of taxpayer decisions. The individual who in tax year t , inspired by his mental vision of inequality (12), engages in work in the hidden economy may have second thoughts when he comes to fill in his income tax return in year $t+1$. At that stage, his income from both types of labour in the past year is given and there is usually nothing that prevents him from honest reporting, even if his intention last year was to have been dishonest. In this situation the realistic representation of his choice process may after all be the portfolio model, so that the modelling of tax compliance and the hidden economy may be more complex than either of the two models is able to capture, calling for a dynamic sequential approach.

This discussion has assumed that the decision to work in the hidden economy is one that is taken by a worker who acts on his own. However, for a better understanding of the structure of the hidden economy it may be necessary to study the case where workers and employers act in collusion: Workers get their wages free of tax, while employers get the benefit of lower wages, both in the form of evasion of the payroll tax and of lower net wage payments. How the gains from such a transaction get divided between the two parties to the transaction is an interesting problem in the theory of tax incidence which has so far received little attention. The incidence of the tax policy instruments will presumably depend not only on the elasticities of demand and supply, as in the conventional case, but also on the legal allocation of responsibility for tax payment between worker and employer.

An interesting aspect of this problem has been studied in a recent unpublished paper by Tone Ognedal, who argues that some firms will benefit from the hidden economy by offering jobs that can easily be combined with shadow work, exploiting the workers' need for some regular income in order to escape detection. She shows that this mechanism may imply misallocation of labour in the regular economy by leading too many workers to be employed in low-productivity firms. Thus, tax evasion may involve a loss of productive efficiency both because it encourages low-productivity work in the hidden economy as well as inside the regular economy.

4 Capital and indirect tax evasion

The extensions of the basic model discussed in the previous section all focus on the decisions of the individual taxpayer and at least by implication on the evasion of income from labour. In

the present section of the paper we broaden the view to take account of the evasion of income from capital and of the non-compliance with indirect taxes.

4.1 Wealth and capital income tax evasion

In the theoretical literature, the evasion of taxes on labour income has received considerably more attention than the evasion of taxes on capital. It is not obvious why this should be so; as already noted, it is difficult to argue that capital income evasion is of less empirical importance. One reason may be that the dividing line between evasion and avoidance is less clear as regards capital income than in the case of income from labour. Another reason might be that the analysis is seen as too easy. The study of taxation and risk taking has, ever since the classic study by Domar and Musgrave (1944), been set in the context of a model of portfolio choice, similar in structure to the one later developed by Arrow (1974) and others and used for the analysis of tax evasion by Allingham and Sandmo (1972). On this background, the study of evasion of capital taxes may seem too obvious to need further elaboration. However, the general model may be made more specific in a number of ways, so that some further discussion may be in order.

The simplest case to analyze in the context of the model of Section 2 is a wealth tax. If W is defined as real wealth and e as the amount of wealth not reported the analysis may proceed exactly as when the model is interpreted in terms of labour income. The evasion of capital income tax evasion is only slightly more complex, at least as long as the rate of return on wealth is assumed to be certain. Assume that utility depends on final wealth, that W is the amount of initial wealth and that r is the rate of return. Final wealth in the two outcomes of non-discovery and discovery can then be written as

$$Y = W(1+r) - t(W-e)r = W[1+r(1-t)] + ter, \quad (13)$$

$$Z = W(1+r) - t(W-e)r - \theta er = W[1+r(1-t)] - (\theta-t)er. \quad (14)$$

Expected utility maximization yields the first order condition

$$(1-p)U'(Y)tr - pU'(Z)(\theta-t)r = 0. \quad (15)$$

After factoring out r , we see that this condition has exactly the same form as (4), so that the comparative statics properties of the two models are essentially the same.

Allowing for the rate of return on saving being uncertain also preserves the structure and predictions of the original portfolio model. This is hardly surprising, since this generalization does not introduce any new margin of choice for the taxpayer. When one allows the taxpayer to choose the composition of his savings portfolio, however, matters get more complicated, since the taxpayer can now choose the probability distribution of his future wealth both via his portfolio choice and through the amount of capital income evaded. Within this framework, there are several ways in which one can model the structure of the decision to evade taxes, but these alternative specifications will not be pursued any further in the present context.

In recent years the international dimension of evasion of taxes on capital has become of increasing importance. A series of empirical studies have documented the importance of tax havens, but their emergence also offers interesting challenges to the economic theorist. While the positive economics of tax havens can to a large extent be analyzed by means of existing models of tax evasion, the normative problems of optimal policy choice raise some new issues. The core problem is the design of policies in a world of multiple jurisdictions when the different countries have partly conflicting interests, and this leads into the study of international policy coordination and international tax treaties⁶.

4.2 Indirect tax evasion

The modelling of income tax evasion relies heavily on the theory of individual decision making in the presence of risk aversion. The evasion of indirect taxes, on the other hand, is most naturally regarded as a decision made by firms, and this raises the question of how to model the objectives of the firm. Should they be assumed to be risk averse? Risk aversion on the part of firms is an assumption that does not fit well into the framework of general equilibrium theory, but one could still make a case for it in a situation of incomplete and imperfect risk markets. Both risk aversion (Marrelli 1984) and risk neutrality (Sandmo 2002) have been explored in the literature.

⁶ An interesting survey of the basic facts of tax havens and of the main policy issues can be found in the report of a committee appointed by the Norwegian Ministry of Foreign Affairs and chaired by my colleague Guttorm Schjelderup; see Norwegian Ministry of Foreign Affairs (2009).

As in the case of income tax evasion, the scope for individual evasion is limited by the extent of third party reporting. As regards indirect taxes, this is particularly important when considering alternative systems for commodity taxation. As regards sales taxes, no third party reporting is involved. By contrast, the VAT is less vulnerable to evasion because in inter-firm transactions buyers and sellers have opposite interests regarding the reporting of the transactions for tax purposes⁷.

A particularly case of commodity taxation that is becoming of increasing importance for policy purposes is that of environmental or Pigouvian taxes. If these taxes are levied on the emission of pollutants, there are no market transactions that can be used to identify the magnitude of the firm's tax liability; instead, the tax base of the individual firm must be based on its own reporting or on direct government controls. In both cases there is scope for underreporting and evasion of taxes. This problem was first studied by Harford (1978) and further analyzed by Sandmo (2002) and others. One concern that one may have is that the existence of opportunities for tax evasion may weaken the efficiency case for environmental taxes. In a number of important cases, however, this turns out not to be the prediction of the theoretical models. The profit maximizing firm – even when risk aversion leads it to maximize the expected utility of profits – will set the marginal cost of reducing emissions equal to the Pigouvian tax rate, while the amount of evasion will be determined by a tax arbitrage condition by which the firm will equalize the regular tax rate to the expected penalty rate⁸. Pollution control and the control of evasion can therefore be regarded as separate policy objectives where the optimal policy design for each objective can be determined without regard for the spillover effects on the other.

Another insight that emerges from this type of model concerns the equivalence of taxes and quotas. As taxes can be evaded, quotas can be violated, and the violations can be studied by the theoretical tools used in the analysis of tax evasion. In this perspective, the marginal fine for quota violations can be shown to have many of the properties of a tax on emissions, so that the difference between the two types of policy may be less than one is otherwise likely to infer from the literature, particularly as regards the properties of non-tradable quotas.

⁷ A theoretical analysis of the relative merits of VAT and sales taxes is Boadway and Sato (2009).

⁸ This assumes that the penalty is non-linear in the amount of evasion with an increasing marginal rate.

4.3 Smuggling

Smuggling can be regarded as a special case of indirect tax evasion in being an attempt to evade import duties and quota regulations. In fact, import duties were once much more important for government revenue than regular indirect taxes, primarily because the administrative costs were lower and the opportunities for evasion were less. Perhaps because of its longer history and because smuggling is a more colourful activity than tax evasion, the empirical work on the subject is much larger and covers a longer historical period. The theoretical literature, on the other hand, is of more modest proportions. Its origin goes back to the early 1970s which saw the publication of several related papers by Bhagwati and Hansen (1973), Bhagwati and Srinivasan (1973) and others, the most important of which have been collected in Bhagwati (1974). These contributions are strikingly different from the tax evasion literature that started to develop at about the same time. There is little attempt to analyze the behaviour of the individual smuggler, and virtually no attempt is paid to the analysis of smuggling as involving choice under uncertainty. On the other hand, there is a strong focus on the general equilibrium effects of smuggling on relative prices and the structure of industry. Although some attempts (Martin and Panagariya 1984) have later been made to strengthen the microeconomic foundations of the theory of smuggling, the disparity between the two approaches to what are actually closely related problems still remains.

The difference may be characteristic in a more general sense of the two fields of public economics and international trade. Historically, the field of international economics has had a much stronger focus on general equilibrium effects of trade policy, while public economics has paid closer attention to the effects of tax policy on individual incentives and allocative efficiency. As regards the particular sub-fields of tax evasion and smuggling, part of the difference in approach probably also stems from the fact that smuggling concerns specific goods, while tax evasion theory has been mostly concerned with the evasion of direct taxation where the links to industry structure and relative prices are less obvious.

5 Systems perspectives: General equilibrium, optimal taxation and tax administration

The discussion of smuggling serves as a reminder that while the main part of the theory of tax evasion is concerned with the individual decision to evade taxes there is an obvious need to extend the theory by relating individual decisions to the study of the effects of tax evasion on

the economy as a whole. In the present section we consider some of the broader issues that arise in this context.

5.1 General equilibrium effects

To some extent, the incidence of tax evasion and the hidden economy can be studied by means of the theory of tax incidence. Opportunities for tax evasion affect after tax prices in both commodity and factor markets, and this will have implications for output and factor use in different sectors of the economy and thereby for the distribution of income. This raises a whole set of issues that have so far received little attention in the theoretical literature. To illustrate, in Section 3.3 above we discussed the allocation of labour time between the regular and the hidden economy, taking the wage rates w and w^* as exogenously given. But in a broader analysis one would wish to determine the wage gap between the two segments of the labour market as a function of the policy parameters. This is of obvious importance both for understanding how the market mechanism works and for the design of policy. It is also important to realize that the wage gap between the regular and the hidden economy is likely to vary between industries, so that tax evasion will be of importance for the intersectoral allocation of labour and for relative commodity prices.

Another set of general equilibrium effects requires that we move beyond the conventional theory of tax incidence and study the effects of social interaction. The portfolio model and the various extensions of it that have been discussed so far consider the individual tax evader in isolation from the rest of the community. This is particularly striking in the case of the modelling of the notions of social stigma and bad conscience discussed in Section 3.2. Whether the tax evader feels stigmatized by being detected is likely to depend on how many others have been detected for similar violations of the tax law. Similarly, whether the taxpayer suffers from bad conscience in evading taxes is also likely to depend on his beliefs regarding the occurrence of evasion in society as a whole or in his own social reference group. It is easy to see that this might lead to an economy with multiple equilibria in terms of the extent of tax evasion and black market activities. An unstable equilibrium would be one in which an exogenous perturbation with regard to the number of evaders releases a process that leads either to a low-evasion equilibrium in which most people comply with the tax law because that is what most others do, or to a state where most people engage in tax evasion because that is what most people are perceived to do.

This analysis is very much in the spirit of Schelling (1978) who applies a similar line of reasoning to a wide range of social problems. In an area closely related to that of tax evasion and the hidden economy, models of this type have been applied to the study of corruption in an article by Andvig and Moene (1990). Their central insight is that the individual cost of being honest in a corrupt society is higher than in a society where most individuals are honest. This is why corruption may corrupt and why a large hidden economy is likely to be self-perpetuating.

5.2 Issues of optimal taxation

The theory of optimal taxation is essentially about the rational tradeoff between efficiency and distributive justice (as reflected in the choice of a social welfare function). Most of the literature in the field ignores tax evasion, and it is natural to ask what happens to the results of optimal tax theory once the presence of tax evasion is taken seriously. The incorporation of tax evasion in models of optimal taxation can be carried out in a number of different ways, and the following discussion is by necessity quite sketchy.

Intuitively, one might perhaps think that tax evasion in many cases would be good for efficiency. In the case of the income tax, in the absence of evasion the marginal tax rate serves to block some transactions between buyers and sellers of labour services that would otherwise have been Pareto improving. Allowing taxes to be evaded in areas like home repairs and day care for children in reality lowers the effective tax rate on these transactions and make these markets as well as related markets (such as the regular labour market) function more efficiently. However, such questions cannot be analyzed in a partial equilibrium context. If the government aims to raise a given amount of revenue in real terms, the loss of revenue from tax evasion must be compensated by higher taxes in other markets that normally involve new sources of price distortion and efficiency loss. How this compensation should optimally be carried out, ought to depend on the elasticity of the tax bases both in the markets that are especially prone to tax evasion and in other markets. Alternatively, the government might react to tax evasion by lowering its ambitions with regard to the raising of revenue. But this would involve less provision of public goods with other types of efficiency losses for the economy as a whole. The efficiency aspects of tax evasion and their implications for optimal taxation are therefore more complex than one's immediate intuition might lead one to believe.

As regards distributive justice economic intuition offers less in the way of suggestions regarding the design of optimal taxes. If black market labour is mostly carried out by low income workers, e.g. the unemployed, one might argue that a low effective tax rate on their income is an acceptable way in which to improve the situation of the least well off in society. At the other end of the scale, however, if one believes that evasion is especially important for capital and business income, this implies that there is in fact a low tax rate on the highest incomes, which would have less appeal to one with egalitarian social values.

A further complication arises from the fact that it is obviously artificial to analyze issues of optimal taxation in isolation from the choice of policies to ensure tax compliance. The issues discussed in the theory of optimal taxation should ideally be considered together with the analysis of optimal compliance policy, and this raises a challenging set of issues for the public finance economist. Should lower taxes on transactions where evasion is important go together with higher penalties and more use of resources on detection? Questions like this call for a more formalized approach, and efforts in this direction have been made by a number of authors.

The theory of optimal commodity taxation characterizes optimal taxes by means of the Ramsey rule of equal proportional reductions in demand or, as a special case, the inverse elasticity rule. The validity of these rules in the presence of tax evasion has been examined by Cremer and Gahvari (1993) who explicitly model indirect tax evasion as an activity that is undertaken by profit maximizing firms. The government chooses statutory tax rates and probabilities of detection which both affect the expected tax rates⁹ that are crucial for evaluating the welfare effects of tax rates. For the special case of independent demands, Cremer and Gahvari show how tax evasion leads to a modification of the Ramsey inverse elasticity rule: Optimal expected tax rates are lower in markets for which evasion is important.¹⁰ But it should be kept in mind that the importance of evasion is endogenous in the model, since it depends on the choice of probability of detection. This confirms the point made above: Optimal taxation and compliance policies should be considered in conjunction

⁹ The expected tax rate is the weighted average of the regular tax rate and the penalty rate, using the probabilities of non-detection and detection as weights.

¹⁰ This result is related to the argument more recently advanced by Feldstein (1999), Slemrod and Yitzhaki (2002) and Chetty (2009) that the deadweight loss from a tax can be measured by the elasticity of the tax base with respect to the tax, although their analysis does not use an explicit optimal taxation framework.

which each other; this leads to deeper insights but also to results that are not always easy to interpret. The Cremer-Gahvari analysis is valid for the case of a representative consumer economy; contributions that also incorporate distributional concerns include Chander and Wilde (1998) and Boadway and Sato (2000).

Several authors have examined the theory of optimal income taxation in the presence of tax evasion. Sandmo (1981) showed how evasion leads to modifications of the characterization of the marginal tax rate in the case of a linear income tax; however, the analysis did not lead to any firm conclusion as to whether the marginal tax rate should be higher or lower in the presence of evasion. This agnostic conclusion has also been confirmed in the later and more general study of Cremer and Gahvari (1996). In the light of the general discussion above, this ambiguity was to be expected; the joint analysis of tax design and compliance policy is too complex to result in simple and intuitive characterizations.

It should be emphasized that once one takes account of the distributive aspects of the hidden economy, there may be moral objections to tax evasion that are not adequately captured by the “welfarist” approach that treats individual utility of honest and dishonest taxpayers on an equal footing. Judgements about just or fair distribution of income and resources may be based not only on the *outcomes* for individuals in society but also the *process* by which the outcomes have been brought about.

5.3 The tax administration perspective

In the public finance literature, the concept of the costs of taxation usually refer to the costs of distortion of the set of competitive prices, while the more concrete costs of administering the price system are typically neglected, at least in the context of formal economic models. As emphasized e.g. by Slemrod and Yitzhaki (2002) the cost of administration is a crucial component in judging the feasibility and efficiency of alternative tax systems. Such judgements are likely to vary substantially over time and between countries; as always, what theory can contribute is a general framework for discussion and analysis.

A substantial part of the administrative costs of the tax system is the cost of ensuring compliance. The theory of tax evasion has laid the framework for thinking about ensuring compliance via the cost of detection, the use of penalties and the design of the regular tax

system. Different tax systems offer different opportunities for evasion, so that the choice of an optimal tax system should take account of the evasion activities that each system is likely to encourage. As shown in the discussion of the principles of optimal taxation above, if optimal tax rules are to be interpreted in terms of the elasticities of the different tax bases, one must take account of the fact that governments can not only choose tax rates but to some extent also the elasticities; a point emphasized by Slemrod and Kopczuk (2002). This follows from the insight that these reflect the opportunities for evasion which can be affected by government policy with regard to the magnitude of penalties and the resources spent on increasing the probabilities of detection.

The theory of tax evasion does not capture every aspect of the administrative costs of the tax system. E.g., a highly differentiated set of commodity taxes is likely to involve higher accounting costs than a system of uniform rates and the cost is likely to depend on the *number* of different rates rather than on their magnitudes. This may give rise to non-convex costs of administration which are difficult to analyze in formal optimization models. On the other hand, a highly differentiated set of tax rates may itself offer greater scope for tax evasion by the opportunities it offers for misrepresenting the true tax base.

A highly complex tax system, involving a differentiated system of tax rates and tax deductions, may also cause the taxpayer to pay more in taxes than required by the tax law. In other words, individual tax evasion may be negative. This case has received little attention in the literature, presumably because it is of minor empirical importance. Nevertheless, excess compliance may be a real issue in judging the distributive impact of complex tax systems.

6 Concluding remarks: The welfare state

The literature on tax evasion has undoubtedly added some important perspectives to the theoretical literature on public economics, and the present paper is an attempt to provide a survey of the main issues that the literature has emphasized. A focal point of the theory is the modelling of the tax evasion (or tax compliance) decision as an example of choice under uncertainty. We have seen how expected utility theory can be used to predict the responses of the taxpayer to changes in tax rates and penalties and to illuminate the crucial role of penalties and the probability of detection in determining the choice of whether or not to become a tax evader. We have also seen how the latter decision can be sensitive to the exact assumptions

made about the taxpayer's preferences, in particular to the role of moral concerns. It has further been demonstrated how the theory can be extended to take account of labour supply decisions, the evasion of taxes on wealth and income from capital as well as indirect tax evasion, including smuggling. Individual decisions about tax compliance affect the performance of the whole economic system, and we have seen how general equilibrium effects may appear both as conventional incidence effects and as the effects of social interaction. Tax compliance and evasion also lead to modifications of the characterization of optimal income and commodity taxes. Finally, it has been pointed out that the cost of ensuring tax compliance is an important component of the cost of tax administration, a concept that the theory of public finance at least until recently has found it difficult to come to grips with.

This set of modifications to the theory of public finance has important consequences for the analysis of public policy in the modern welfare state. Notions of the welfare state differ, but most economists would agree that the welfare state combines a concern for efficiency and economic growth with egalitarian values related to distributive justice (however defined) and economic security. Of the policies used to implement the aims of the welfare state, the tax system occupies a place of primary importance. A large public sector requires a high overall level of taxes, so that widespread evasion contributes to the erosion of the financial basis for the welfare state. Moreover, the income tax in particular is designed to ensure that the cost of welfare state policy gets distributed among individuals and socioeconomic groups in a way that is socially acceptable. Tax evasion implies that the real structure of the tax system differs from its design. When the tax base shrinks because of evasion tax rates must be set higher than otherwise in order to finance welfare state expenditure. And if the income tax base is distorted, e.g. so as to let the receivers of capital income escape from taxation, the tax becomes less progressive and less redistributive than legislators had intended.

These issues go beyond the purely economic considerations of targeting policies to social objectives. To be a sustainable social and economic system, the welfare state requires social legitimacy among the general public and the electorate. If many taxpayers perceive that they bear an unreasonably high share of the tax burden compared with individuals who obviously earn a much higher income than they do themselves, the political support for the welfare state is likely to diminish, particularly if it is believed that politicians fail to act against better knowledge. Admittedly, there are those that are sceptical to the welfare state, e.g. in its Scandinavian version. But even they would presumably agree that if the welfare state is to be

rolled back, it would be preferable if it happened through explicit political decisions rather than as side effects of growth of the hidden economy.

References

- Allingham, M. G. & Sandmo, A. (1972). Income tax evasion: A theoretical analysis. *Journal of Public Economics* 1, 323-338.
- Andreoni, J., Erard, B. & Feinstein, J. (1998). Tax compliance. *Journal of Economic Literature* 36, 818-860.
- Andvig, J. C. & Moene, K. O. (1990). How corruption may corrupt. *Journal of Economic Behavior and Organization* 13, 63-76.
- Arrow, K. J. (1953). Le role des valeurs boursières pour la répartition la meilleure des risques. *Econometrie*, Paris. Reprinted as Chapter 4 in Arrow (1974).
- Arrow, K. J. (1974). *Essays in the Theory of Risk-Bearing*. Amsterdam: North-Holland.
- Baldry, J. C. (1979). Tax evasion and labour supply. *Economics Letters* 3, 53-56.
- Bernasconi, M. (1998). Tax evasion and orders of risk aversion. *Journal of Public Economics* 67, 123-134.
- Bhagwati, J. N. (Ed.) (1974). *Illegal Transactions in International Trade: Theory and Measurement*. Amsterdam: North-Holland.
- Bhagwati, J. N. & Hansen, B. (1973). A theoretical analysis of smuggling. *Quarterly Journal of Economics* 87, 172-187.
- Bhagwati, J. N. & Srinivasan, T. N. (1973). Smuggling and trade policy. *Journal of Public Economics* 2, 378-389.
- Boadway, R. & Sato, M. (2000). The optimality of punishing only the innocent. *International Tax and Public Finance* 7, 641-664.
- Boadway, R. & Sato, M. (2009). Optimal tax design and enforcement with an informal sector. *American Economic Journal: Economic Policy* 1 (1), 1-27.
- Chander, P. & Wilde, L. L. (1998). A general characterization of optimal income tax enforcement. *Review of Economic Studies* 65, 165-183.
- Chetty, R. (2009). Is the taxable income elasticity sufficient to calculate deadweight loss? The implications of evasion and avoidance. *American Economic Journal: Economic Policy* 1 (2), 31-52.
- Cowell, F. A. (1990). *Cheating the Government*. Cambridge, Mass.: MIT Press.
- Cremer, H. & Gahvari, F. (1993). Tax evasion and optimal commodity taxation. *Journal of Public Economics* 50, 261-275.

- Cremer, H. & Gahvari, F. (1996). Tax evasion and the optimum general income tax. *Journal of Public Economics* 60, 235-249.
- Debreu, G. (1959). *Theory of Value*. New York: Wiley.
- Dhami, S. & al-Nowaihi, A. (2007). Why do people pay taxes? Prospect theory versus expected utility theory. *Journal of Economic Behavior & Organization* 64, 171-192.
- Domar, E. D. & Musgrave, R. A. (1944). Proportional income taxation and risk-taking. *Quarterly Journal of Economics* 58, 388-422.
- Feld, L. P. & Schneider, F. (2010). Survey on the shadow economy and undeclared earnings in OECD countries. *German Economic Review* 11, 109-149.
- Feldstein, M. (1999). Tax avoidance and the deadweight loss of the income tax. *Review of Economics and Statistics* 81, 674-680.
- Gordon, J. P. F. (1989). Individual morality and reputation costs as deterrence to tax evasion. *European Economic Review* 33, 797-805.
- Harford, J. D. (1978). Firm behaviour under imperfectly enforceable pollution standards and taxes. *Journal of Environmental Economics and Management* 5, 26-43.
- Marrelli, M. (1984). On indirect tax evasion. *Journal of Public Economics* 25, 181-196.
- Martin, L. & Panagariya, A. (1984). Smuggling, trade and price disparity: A crime-theoretic approach. *Journal of International Economics* 17, 201-217.
- Musgrave, R. A. (1959). *The Theory of Public Finance*. New York: McGraw-Hill.
- Norwegian Ministry of Foreign Affairs (2009). *Tax Havens and Development*. NOU 2009: 19. Available at <http://www.regjeringen.no/nb/dep/ud/dok/nou-er/2009/nou-2009-19-2>.
- Pencavel, J. H. (1979). A note on income tax evasion, labor supply, and nonlinear tax schedules. *Journal of Public Economics* 12, 115-124.
- Pestieau, P. & Possen, U. M. (1991). Tax evasion and occupational choice. *Journal of Public Economics* 45, 107-125.
- Pigou, A. C. (1928). *A Study in Public Finance*. 3rd Edition, 1947. London: Macmillan.
- Sandmo, A. (1981). Income tax evasion, labour supply, and the efficiency-equity tradeoff. *Journal of Public Economics* 16, 265-288.
- Sandmo, A. (2002). Efficient environmental policy with imperfect compliance. *Environmental and Resource Economics* 23, 85-103.
- Sandmo, A. (2005). The theory of tax evasion: A retrospective view. *National Tax Journal* 58, 643-663.

- Schelling, T. (1978). *Micromotives and Macrobehavior*. New York: Norton.
- Schneider, F. (2005). Shadow economies around the world: what do we really know? *European Journal of Political Economy* 21, 598-642.
- Slemrod, J. & Kopczuk, W. (2002) The optimal elasticity of taxable income. *Journal of Public Economics* 84, 91-112.
- Slemrod, J. & Yitzhaki, S. (2002). Tax avoidance, evasion and administration. *Handbook of Public Economics, Volume 3*, Auerbach, A. J. & Feldstein, M. (Eds.), 1423-1470. Amsterdam: Elsevier Science.
- Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. London: Strahan and Cadell. Glasgow Bicentenary Edition, Oxford: Oxford University Press, 1976.
- Snow, A. & Warren, R. S. Jr. (2005). Tax evasion under random audits with uncertain detection. *Economics Letters* 88, 97-100.
- Srinivasan, T. N. (1973). Tax evasion: A model. *Journal of Public Economics* 2, 339-346.