Tax avoidance and the optimal provision of public goods

Alessandro Balestrino[a] and Umberto Galmarini[b]

[a]Department of Economics, University of Pisa, Italy

[b]Institute of Economics and Finance, Catholic University of Milan, Italy

February 19, 2001

Abstract

Our aim in this paper is to investigate whether tax avoidance affects the optimal provision of public goods. We approach this question using a framework in which the government uses a general income tax and controls the resources devoted to tax enforcement. We first derive the income tax structure and the tax enforcement rule, and then a generalized Samuelson rule. We argue that, under tax avoidance, it is always desirable to distort public-good supply downwards, both in the sense that the sum of marginal rates of substitution between public and private consumption must exceed unity (rule-underprovision), and in the sense that the actual level of the public good is optimally lower with tax avoidance than without it (level-underprovision).

Keywords: Tax avoidance, Public goods, Income taxation, Tax enforcement

JEL: H21, H26, H41

Corresponding author: Alessandro Balestrino, Department of Economics, Political Sciences Faculty Bldg, University of Pisa, Via Serafini 3, I56126 Pisa, Italy (e-mail: sandro@specon.unipi.it, fax: +39050920450)

1 Introduction

In this paper we extend the self-selection approach to optimal income taxation proposed by Stiglitz (1982) including income tax *avoidance*[1] and public-good provision. Actually, both these extensions have already been explored separately in the literature: Boadway and Keen (1993), Edwards *et al.* (1994) and Nava *et al.* (1996) deal with optimal public-good provision, while Boadway *et al.* (1994) investigate the optimal direct-indirect tax mix in a model with tax avoidance. Tax *evasion* under general income taxation has been studied by Cremer and Gahvari (1995) and Schroyen (1997). To the best of our knowledge, there is only one paper which, like ours, deals with tax avoidance and public-good provision at the same time, namely Hindricks (1999). However, our emphasis is on somewhat different issues. In particular, while Hindricks derives the same rule for public-good provision as we do, he does not go on to identify the factors which differentiate this rule from the one that would hold in a standard model with full tax compliance; moreover, he discusses the question whether the public good is over- or underprovided in second-best only in terms of first-order conditions, whereas we focus also on the levels of public-good provision.

Tax avoidance is an activity which has a great relevance in the real world economies. Indeed, the common practice of assuming that the tax liability can be ascertained and collected at no cost is a very poor description of reality. For instance, the IRS reports that about 17% of the USA income tax liability is not paid and that about 40% of US households underpay their taxes (Andreoni *et al.*, 1998). Tax collection costs are also important, adding up to about 10% of tax revenue (Slemrod and Bakija, 1996). In the present paper, we build on the growing body of research on the economics of tax evasion and tax avoidance to address the following question: what are the implications of tax avoidance for public expenditure? That is, do tax avoidance activities suggest to

[1] It is customary to distinguish tax *avoidance* from tax *evasion* (Cowell, 1990, pp. 10–14). Tax avoidance is costly and riskless, whereas tax evasion is risky because taxpayers face the possibility of an audit and, in case of discovered tax evasion, are subject to the payment of a fine. As a consequence, tax enforcement in the case of avoidance consists of all activities that limit taxpayers' opportunities for risk-free tax dodging, whereas, in the case of tax evasion, is a measure of the probability of being audited by tax authorities.

deviate from the level of public expenditure which would arise if these activities were not present?[2] Interestingly enough, it turns out that tax avoidance implies a tendency towards a reduction of public expenditure. This result thus could be interpreted as offering an argument in favour of the downsizing of State intervention which is not based on a priori reasons, but on a theoretical analysis of the effects of some relevant features of real economies.

To answer the proposed question, we provide a generalized Samuelson rule which extends the modification proposed by Boadway and Keen (1993). They demonstrated that deviations from the first-best rule for public good provision are useful as long as they can be used to relax the self-selection constraint, thereby allowing a Pareto-improving change in the income tax schedule. We confirm this result and show how the pattern of the deviations is modified to account for tax avoidance. Specifically, we note a general tendency to induce a "downward" distortion, i.e. what is usually called an "underprovision" (relative to the first-best) of the public good. We argue, relying an some recent results by Gaube (1999), that this tendency translates itself into an effectively *lower* level of public-good provision, both relative to the first-best and, importantly, to the second-best with full tax compliance.

The paper is structured as follows. Section 2 introduces the model and characterizes consumer behaviour. Section 3 sketches the derivation of the optimal general income tax and the optimal level of tax enforcement in the presence of tax avoidance. Section 4 focuses on how the *rule* for public-good provision is affected by tax avoidance. Section 5 studies instead how tax avoidance affects the *level* of public-good provision. Finally, Section 6 sums up.

2 The model

The economy is inhabited by two types of individuals, distinguished by their ability. There are n^h individuals of type h, $h = 1, 2$, with $w^2 > w^1$, where w^h is the wage rate

[2] A similar question has been investigated in a model of tax evasion (rather than tax avoidance) and assuming linear income taxation by Falkinger (1991). Our setup and our results are however entirely different.

(taken, as usual, to represent ability).[3] The utility function is concave, identical across types and defined over a composite consumption good, x, labour supply, l, and a public good, G; it is written $u = u(x, l, G)$. Production is linear and uses labour as the only input; production prices are normalized to unity for both x and G.

The government has at its disposal a general income tax, but it does not observe gross income, wl. Rather, it has to rely on reported income, which may differ from wl. Misreporting income is lucrative because it reduces the tax liability without taking risks, but is done at a cost. Let $C(H, wl; A)$ denote the cost of avoidance,[4] where H, $0 \leq H \leq wl$, is the amount of income which is not reported to the tax authorities and A represents a measure of administrative costs devoted to the enforcement of the income tax. Note that we assume that negative reports ($H > wl$) are not feasible, since true income can never be negative. We also rule out overreporting ($H < 0$), which can be shown to never occur in equilibrium.

We make several assumptions on C. Firstly, we posit $C_H > 0$ and $C_{HH} > 0$ for $H > 0$, that is the marginal cost of tax avoidance is positive and increasing in the level of concealed income; $C(0, wl; A) = C_H(0, wl; A) = 0$, that is full reporting is costless and the marginal cost of concealing the first unit of income is zero; and $C_H(wl, wl; A) > 1$, that is the marginal cost of concealing the last unit of income is greater than one. Secondly, we assume that $C_{(wl)} \leq 0$, which implies that, for given H, an increase in true income does not increase the cost of avoidance; this is what Slemrod (1998) refers to as the "avoidance-facilitating" value of true income. Thirdly, we set $C_A(0, wl; A) = 0$, and $C_A > 0$ for $H > 0$, that is tighter enforcement of the tax code makes tax avoidance more costly, provided that some fraction of true income is actually hidden from the tax agency. Finally, we assume that $C(.)$ is homogeneous of degree one in H and wl, i.e. $C(\delta H, \delta wl; A) = \delta C(H, wl; A)$ for all $\delta > 0$; in words, the tax avoidance technology exhibits constant returns to scale. This latter assumption allows us to redefine the cost

[3] Superscripts denote households, while subscripts denote partial derivatives; however, to simplify the notation, we will drop superscripts whenever possible.

[4] We follow the common practice of not describing the tax avoidance activity in a specific way. For an example in which tax avoidance is modelled explicitly as a form of arbitrage between taxed and untaxed assets, see Agell and Persson (2000).

of tax avoidance in a form which is very convenient to solve the model and interpret the results. Let $H = (1 - s)wl$, where 0 □ s □ 1 is the fraction of actual income which is reported to the tax authorities. Then define the per-unit-of-true-income tax-avoidance-cost as $c(s; A) \equiv C[(1 - s)wl, wl; A]/wl = C(1 - s, 1; A)$. The function $c(.)$ depends on how large is the fraction of reported income and on administrative costs; it is independent of wl. It is immediate to see that the assumptions made on $C(.)$ imply the following

Assumption 1 *$c_s < 0$, $c_{ss} > 0$, $c_A > 0$ for $s < 1$, $c\,(1; A) = c_s\,(1; A) = c_A\,(1; A) = 0$, $c_s\,(0; A) < -1$.*

The households choose x^h, l^h and s^h optimally, thereby positioning themselves on a given point of the tax schedule. Following Boadway *et al.* (1994), we can linearize the after-tax budget constraint at that point. Each linearized budget will be defined by a household-specific marginal tax rate, t^h □ 1, and lump-sum tax, T^h. We can then define the so-called *virtual* budget constraint as:

$$x^h = \left[1 - t^h s^h - c(s^h; A)\right] w^h l^h - T^h, \qquad h = 1, 2. \tag{1}$$

Each household maximizes its utility function subject to (1) by choice of x^h, l^h and s^h. The first-order conditions immediately lead to:

$$-\frac{u_l^h}{u_x^h} = w^h \left(1 - t^h s^h - c^h\right), \qquad t^h = -c_s^h, \qquad h = 1, 2, \tag{2}$$

where $1 - t^h s^h - c^h > 0$ for all t^h □ 1, since the second part of (2) implies that $t^h > t^h s^h + c^h$.[5] Notice that c_s^h has the opposite sign of the marginal tax rate; then, given our assumptions on the cost-of-avoidance function, we have that $t^h \in (0, 1]$ implies $s^h < 1$ and $t^h = 0$ implies $s^h = 1$. In other words, the household underreports its income only when income is taxed at the margin. From the maximisation problem, we can derive an indirect utility function in the usual way, denoted $v(t^h, T^h, G, A)$.

[5] To see this, consider the $-c_s$ curve, which is decreasing in the $[0, 1]$ interval. The optimal s is given by the intersection between that curve and a straight line representing the value of t. Now, we have that $ts + c \equiv \int_0^s tdz - \int_s^1 c_s(z; A)dz < t \equiv \int_0^1 tdz$, since $-c_s(z; A) < t$ for $z \in (s, 1)$ from the first order condition for s. Clearly, if $ts + c < t$, then $ts + c < 1$ for t □ 1.

In the self-selection models of income taxation, the social planner can observe gross income and thus also net income, which is gross income minus taxes. Thus, the self-selection constraint says that the tax structure must be such that the (*gross income, net income*) pair of the low-ability household is not attractive for the high-ability household (in the "normal" case in which the redistribution goes from the high- to the low-ability household —more on this later). This will prevent the type-2 households from reducing their labour supply to mimic the behaviour of type 1 households. Here, however, the government can only observe *reported* income, and thus mimicking takes the form of *reporting* the same income of the low-ability person, not of *earning* the same gross income. As a consequence, mimicking may be accomplished either working less or by concealing more income. Then, in the present setting, the self-selection constraint prevents the high-ability households from *reporting* the same gross income as the low-ability ones.

To misrepresent its type, a high-ability household should obey:

$$y^1 \equiv s^1 w^1 l^1 = \hat{s}^2 w^2 \hat{l}^2 \equiv \hat{y}^2, \tag{3}$$

where the "hat" denotes the variables pertaining to the mimicker. (3) tells us that, in order to have the same reported income as the low-ability household, the mimicker chooses either its labour supply or the fraction of income to report, as the two are not independent. Its utility is $\hat{u}^2 = u(\hat{x}^2, \hat{l}^2, G)$, and its budget constraint is:

$$\hat{x}^2 = \left[1 - t^1 \hat{s}^2 - c(\hat{s}^2; A)\right] w^2 \hat{l}^2 - T^1, \tag{4}$$

where $\hat{s}^2 = y^1/(w^2 \hat{l}^2)$ depends on $\hat{l}^2$ because of (3). The mimicker will choose $\hat{x}^2$ and $\hat{l}^2$ to maximize its utility subject to (4). This yields the first order condition:

$$-\frac{\hat{u}_l^2}{\hat{u}_x^2} = w^2 \left(1 + \hat{c}_s^2 \hat{s}^2 - \hat{c}^2\right), \tag{5}$$

where we have substituted for $\partial \hat{s}^2/\partial \hat{l}^2 = -\hat{s}^2/\hat{l}^2$. (5) reflects the interdependency between the choice of l and s for the mimicker: it represents the trade-off between the disutility of labour supply and the advantages (net of costs) of income concealment. The mimicker's indirect utility will be $\hat{v}^2\left(t^1, T^1, G, A\right)$.

We can now characterize the behaviour of the mimicker relative to that of the low-ability household, establishing a result which will be useful in what follows.[6]

Lemma 1 *If x is a normal good, then $\hat{s}^2 < s^1$, $\hat{x}^2 > x^1$ and $w^2\hat{l}^2 > w^1l^1$.*

Proof. Consider what happens when $\hat{s}^2 = s^1$. We have that $\hat{l}^2 = \hat{l}_0^2 < l^1$ such that $w^2\hat{l}_0^2 = w^1l^1$. Hence, $\hat{c}^2 = c^1$ and $\hat{x}^2 = x^1$. If x is a normal good, then $\hat{x}^2 = x^1$ and $\hat{l}_0^2 < l^1$ imply that the indifference curve of the mimicker is flatter than that of type 1 in the (x, l)-plane (this follows because normality of x requires $\partial MRS_{xl}/\partial l > 0$). Also, the slope of the mimicker's budget constraint is (at least locally) steeper than the budget constraint of type 1, since $w^2 > w^1$, $\hat{s}^2 = s^1$, $w^2\hat{l}_0^2 = w^1l^1$ and $\hat{c}^2 = c^1$. Formally:

$$-\frac{u_l(\hat{x}^2, \hat{l}_0^2, G)}{u_x(\hat{x}^2, \hat{l}_0^2, G)} < -\frac{u_l\left(x^1, l^1, G\right)}{u_x\left(x^1, l^1, G\right)} = w^1(1 + s^1c_s^1 - c^1) < w^2(1 + \hat{s}^2\hat{c}_s^2 - \hat{c}^2).$$

The equality comes from the first order conditions for type 1. This implies that the mimicker chooses $\hat{l}^2 > \hat{l}_0^2$ and $\hat{s}^2 < s^1$, and thus also $\hat{x}^2 > x^1$ and $w^2\hat{l}^2 > w^1l^1$. ■

Notice that nothing can be said as to whether $\hat{l}^2 > l^1$ or *viceversa*; this fact bears important implications for the results that follow.

3 Optimal taxes and tax enforcement

We will study Pareto-optimal arrangements, whereby the policy variables are chosen so as to maximize the utility of one household type, subject to the other type's utility being constant, plus the self-selection and revenue constraints. As common, we focus on the "normal" case, in which the social planner wishes to redistribute from the high-ability to the low-ability households. Furthermore, we impose the standard single-crossing condition, that in our context means that the high-ability consumer indifference curves in the (*pre-tax reported income, post-tax reported income*)-space are flatter.[7] The combined effect of these assumptions is that the only binding incentive

[6] Our Lemma 1 is an extension of Lemma 1 in Boadway *et al.* (1994).

[7] Adapting the proof provided by Boadway *et al.* (1994), we can show that the single-crossing condition holds when the private good and leisure are non-inferior.

constraint will be the one ruling out the possibility that the high-ability type mimics the low-ability one by reporting the same income.

The optimal policy problem is:

$$\begin{aligned} \max_{\{t^h,T^h\},G,A} \quad & v^1\left(t^1,T^1,G,A\right) \\ \text{s.t.} \quad & v^2\left(t^2,T^2,G,A\right) \geq \bar{v}^2, & [\lambda] \\ & v^2\left(t^2,T^2,G,A\right) \geq \hat{v}^2\left(t^1,T^1,G,A\right), & [\sigma] \\ & \sum_{h=1}^{2} n^h\left(t^h s^h w^h l^h + T^h\right) = G + A. & [\mu] \end{aligned} \tag{6}$$

Notice that the (unit) cost of tax enforcement has been normalized to one and that the Lagrange multipliers on the three constraints are in brackets.

The subsequent analysis of public-good provision presupposes that an optimal tax system is implemented. We will therefore briefly note a few results concerning the tax rates and the tax enforcement rule. Following the procedure outlined by Boadway *et al.* (1994),[8] it is possible to show that the marginal tax rates are as follows:

$$t^1 = \frac{\sigma}{\mu n^1}\hat{\alpha}^2\left(c_s^1 - \hat{c}_s^2\right) > 0; \qquad t^2 = 0. \tag{7}$$

(7) states that the top-earners should be undistorted and therefore that they do not misreport their income – see (2); and that, since σ, $\hat{\alpha}^2$, μ, and n^1 are all greater than zero, the marginal tax rate on the low-earners has the same sign as the term $\left(c_s^1 - \hat{c}_s^2\right)$. This term represents the difference between the marginal cost of (mis)reporting for the low-ability households and for the mimicker, and is strictly positive, for $\hat{s}^2 < s^1$ (see Lemma 1) and $c_{ss} > 0$ imply that $\hat{c}_s^2 < c_s^1 < 0$; thus the low-ability households will be taxed at the margin ($t^1 > 0$). An intuitive understanding of this latter result is easily gained by noting that it is a simple consequence of the fact that marginal avoidance costs are higher for the mimicker. Indeed, it is desirable to set $t^1 > 0$ and to induce tax avoidance —see (2)— just because this will hurt the mimicker more than the true low-earner, and therefore the self-selection constraint will be relaxed. By contrast, setting

[8] A detailed derivation of all the policy rules is contained in an Appendix available from us upon request.

$t^2 > 0$ would be useless, as there is nobody interested in mimicking the high-ability type.

It is also straightforward to show, using the first order conditions w.r.t. t^1, t^2 and A, and the properties of the cost-of-avoidance function, that the optimal level of tax enforcement satisfies:

$$1 + n^1 w^1 l^1 c_A^1 = \frac{\sigma \hat{\alpha}^2}{\mu} \left(w^2 \hat{l}^2 \hat{c}_A^2 - w^1 l^1 c_A^1 \right). \tag{8}$$

The l.h.s. of this expression represents the marginal administrative costs of tax enforcement (which are constant and normalized to one) plus the marginal effect of A on type-1 households tax avoidance costs (given the equilibrium level of income reporting, s^1). The marginal benefits of tax enforcement are shown on the r.h.s. of (8). These are positive if tighter enforcement increases the mimicker's tax avoidance costs by more than it increases type-1 household's tax avoidance costs, i.e. if $w^2 \hat{l}^2 \hat{c}_A^2 > w^1 l^1 c_A^1$. Since by Lemma 1 $w^2 \hat{l}^2 > w^1 l^1$, a sufficient condition for this is that $\hat{c}_A^2 \geq c_A^1$. Given our assumptions on $c(.)$, this latter condition does not hold true in general. Sufficient conditions are however rather weak: for instance, if $c_{sA} > 0$ for all (s, A), then $\hat{s}^2 < s^1$ (see Lemma 1) implies $\hat{c}_A^2 \geq c_A^1$. Intuitively, a marginal increase in tax enforcement is worthwhile as long as it allows to relax the self-selection constraint, and this occurs whenever an increase in A harms the mimicker more than type-1 households.

4 A modified Samuelson rule for public-good provision

The rule for optimal public-good provision can be obtained from problem (6) by manipulating the first order conditions in a way similar to Boadway and Keen (1993)[9] — see also Hindricks (1999):

$$n^1 MRS_{Gx}^1 + n^2 MRS_{Gx}^2 = 1 + \frac{\sigma \hat{\alpha}^2}{\mu} \left(\widehat{MRS}_{Gx}^2 - MRS_{Gx}^1 \right). \tag{9}$$

[9] Actually, Boadway and Keen (1993) derive a Samuelson rule also for the case in which labour, as opposed to consumption, is used as the *numeraire*, and show that the choice of the *numeraire* is not without consequences for the analysis. An explicit consideration of these issues would however take us too far afield; a comprehensive treatment can be found in Gaube (1999).

(9) looks misleadingly the same as the Boadway-Keen modification of the Samuelson rule. One could discuss, following Hindricks (1999), the decentralisation issue on the basis of (9), but it seems more instructive to identify the factors which differentiate the public-good provision rule under tax avoidance from the one which would arise in the standard model with full tax compliance. The crucial difference is that the *MRS* for the mimicker and the low-ability household are not evaluated at the same after-tax income (i.e. consumption), as was the case in the Boadway and Keen (1993) model. Let $\widehat{MRS}^2_{Gx}|_{x^1}$ denote the mimicker's marginal valuation of G evaluated at the low-ability type consumption level. Since $\hat{x}^2 \neq x^1$ and $\hat{l}^2 \neq l^1$ by Lemma 1, the second term on the r.h.s. can be decomposed by adding and subtracting $\widehat{MRS}^2_{Gx}|_{x^1}$ so as to obtain:

$$\begin{aligned} & n^1 MRS^1_{Gx} + n^2 MRS^2_{Gx} = \\ & = 1 + \frac{\sigma\hat{\alpha}^2}{\mu}\left(\widehat{MRS}^2_{Gx}|_{x^1} - MRS^1_{Gx}\right) + \frac{\sigma\hat{\alpha}^2}{\mu}\left(\widehat{MRS}^2_{Gx} - \widehat{MRS}^2_{Gx}|_{x^1}\right). \end{aligned} \tag{10}$$

The second term on the r.h.s. of (10) (call it the "BK-term") is similar to that in Boadway and Keen (1993), as labour supply is different while consumption (net income) is the same for type-1 and the mimicker. The third term ("TA-term", from "tax avoidance") is new, as labour supply is the same but consumption is different. Note that, by Lemma 1, $\hat{x}^2 > x^1$; instead, $\hat{l}^2$ can be lower or higher than l^1.

Before discussing and interpreting this condition, we need two pieces of terminology:

- we say that G is *overprovided* (*underprovided*) with respect to the first best rule if the sum of the marginal rates of substitution — the l.h.s. of (10) — falls short of (exceeds) the marginal rate of transformation — unity in our case, the first term on the r.h.s of (10);[10]

- we say that G is an *L-substitute* (*L-complement*) if MRS_{Gx} decreases (increases) as leisure, denoted by L, increases.

[10] Given that the standard concavity properties are not generally satisfied in this sort of second best policy problems, overprovision in this sense does not imply that the provision level is necessarily higher than in first best. Recently, Gaube (1999) has however identified the conditions which determine under- and overprovision in level terms. We will return to this issue in Section 5.

We can now decompose our provision rule in three parts. If we consider only the first term on the r.h.s, we just have the standard Samuelson condition. Adding the second term, we have the Boadway-Keen modification, which states that deviations from the first best rule are justified inasmuch as the mimicker's marginal valuation of the public good differs from that of the low-ability household, due to the difference in their labour supply levels ($\hat{l}^2 \neq l^1$); if the BK-term is negative (positive), G should be overprovided (underprovided). The third term is specific to our model and represents the effect of tax avoidance, namely the fact that the mimicker's marginal valuation of the public good is not evaluated at the same consumption level of the low-ability household ($\hat{x}^2 \neq x^1$): overprovision (underprovision), then, requires that the algebraic sum of the two last terms is negative (positive).

What can we say about the sign of the BK- and the TA-term? Taking the former first, note that: *i*) G can be an L-substitute or an L-complement, and *ii*) the labour supply of the mimicker can be lower or higher than type-1's labour supply. If the mimicker's labour supply is lower than that of the low-ability household, then, for equal consumption, the BK-term is positive if G is an L-complement (thus tending to favour underprovision of G); on the other hand, this term is negative if G is an L-substitute (thus tending to favour overprovision of G). This corresponds exactly to what happens in the standard case, that is optimal income taxation without tax avoidance. However, the opposite signs for the BK-term, and thus the opposite implications for the provision of G, are obtained when the mimicker's labour supply is higher than type-1's labour supply.

Moreover, if taxes can be avoided, net incomes (consumption levels) differ for the mimicker and the low-ability household. This is reflected in the TA-term, whose sign depends on whether the mimicker values G more at its net income level or at the low-ability household's net income level. Using Lemma 1, we can prove the following:

Lemma 2 *If G is a normal good, then the TA-term is always positive.*

Proof. If G is a normal good, then:

$$\frac{\partial MRS_{Gx}}{\partial x} = \frac{u_x u_{Gx} - u_G u_{xx}}{(u_x)^2} > 0.$$

Therefore, since $x^1 < \hat{x}^2$ by Lemma 1, we have that:

$$\widehat{MRS}^2_{Gx} \equiv \frac{u_G(\hat{x}^2, \hat{l}^2, G)}{u_x(\hat{x}^2, \hat{l}^2, G)} > \frac{u_G(x^1, \hat{l}^2, G)}{u_x(x^1, \hat{l}^2, G)} \equiv \widehat{MRS}^2_{Gx}|_{x^1}.$$

■

In words, the mimicker's marginal willingness to pay for the public good (in terms of the consumption good) is higher at its own income level. This means that tax avoidance always tends to favour the underprovision of G.

To interpret this discussion of (10), we can adapt the intuition provided by Boadway and Keen (1993). Suppose that we start from a situation in which the public good is provided according to the Samuelson condition. We now reduce G infinitesimally, and simultaneously reduce the income tax liabilities of both types by their MRS. This way, neither the welfare levels nor the government budget will be affected. Suppose, however, that the mimicker, *if he or she had the same net income as the low-ability household*, would value G more, i.e. would have a larger MRS; then, the BK-term is positive, that is the reduction in the mimicker's tax liability (MRS^1) does not compensate its loss from the reduced availability of G. Moreover, we know that the mimicker's actual MRS, that is the one evaluated at its income level, is larger than the one evaluated at the low-ability household income level, so that the mimicker's total loss is underestimated by the BK-term: adding the TA-term gives us the correct measure. Therefore, the overall effect of this policy change is that of making the mimicker worse-off: the reduction in public good provision relaxes the self-selection constraint, enabling the realization of a Pareto-improving change in the income tax schedule. A similar intuition applies to the case in which the policy change consists of an increase in the provision of G.

In the Boadway and Keen (1993) setup, when preferences are identical, differences in the valuation of the public good arise only because the mimicker enjoys more leisure than the low-ability household (they have the same income, but the mimicker has a higher wage). Therefore, the mimicker values G less than the low-ability household if the public good is an L-substitute, in which case G will be overprovided (by a symmetric argument, it will be underprovided if it is an L-complement).[11] In our case, the relations between the various forces at work is less straightforward, for two reasons.

[11] Boadway and Keen (1993) develop their model without imposing identical preferences, and then

Table 1: Public-good provision

		BK term	*TA term*	*Provision of G*
$l^1 > \hat{l}^2$	L-compl.	$+$	$+$	Underprovision
	L-subst.	$-$	$+$	Ambiguous
	L-indep.	0	$+$	Underprovision
$l^1 < \hat{l}^2$	L-compl.	$-$	$+$	Ambiguous
	L-subst.	$+$	$+$	Underprovision
	L-indep.	0	$+$	Underprovision
$l^1 = \hat{l}^2$		0	$+$	Underprovision

First, we do not know whether the mimicker has a larger or a smaller labour supply than the low-ability household; second, they have different consumption (net income) levels. The first effect implies that the sign of the BK-term may be reversed compared to the standard case; the second effect, instead, offers a rationale for distorting the provision of the public good (in the direction of underprovision) which, remarkably, does not depend on L-complementarity or L-substitutability. Indeed, with tax avoidance, underprovision occurs even if G is L-independent (i.e. MRS_{Gx} does not depend on L).[12] It is easy to see that, in that case, the BK-term is zero, while the TA-term is, as always, positive.

As a guide to the reader, we provide an overview in Table 1. The BK-term and the TA-term may have opposite signs (when $l^1 > \hat{l}^2$ and G is an L-substitute, or when $l^1 < \hat{l}^2$ and G is an L-complement), so that the net effect on the provision of G is ambiguous. In these cases tax avoidance tends to weaken the distortion required by self-selection, reducing the extent to which the public good is overprovided (and

derive the correspondence between L-substitutability (L-complementarity) and overprovision (underprovision) for the case in which households have the same utility function in their Corollary 2.

[12] The condition of L-independence corresponds to the case in which the utility function is weakly separable between leisure and all the other goods; in a model without tax avoidance that type of utility function ensures that the Samuelson condition holds at the second best optimum (*cf.* Corollary 1 in Boadway and Keen, 1993; the result is originally due to Christiansen, 1981).

possibly reverting the direction of the distortion if the TA-term dominates the BK-term). In the other cases, the BK-term and the TA-term work in the same direction, leading to underprovision of the public good.

5 The level of public-good provision

The analysis so far has given us some insights into the direction in which public good provision is optimally distorted. Still, as we noted in fn. 9, we cannot be sure that over- or underprovision in our sense corresponds to actual over- or underprovision of the *level* of the public good. The so-called "Pigou's conjecture" states that the level of public-good provision is lower in second- than in first-best, due to the distortionary nature of second-best taxes. There are many contributions on the subject, some employing specific functional forms, like Atkinson and Stern (1974) and Wilson (1991a, 1991b), others focusing on linear tax systems, like Chang (2000) and Gaube (2000), and most of them show, in different ways, that this conjecture is not necessarily verified. The first to discuss Pigou's conjecture in a non-linear tax context (and with a general utility function) is however Gaube (1999). While we refer the reader to Gaube's own work for the formal analysis, we will employ his intuitive arguments to investigate the question whether tax avoidance induces level- as well as rule-underprovision.

To focus the analysis, we concentrate on the case of L-independence, i.e. the one in which leisure enters separately in the utility function. Then, the Samuelson rule holds at the second-best optimum with full tax compliance (the Boadway-Keen model). What does this imply in level terms? Gaube (1999) suggests, as a first step, to think of the sum of the MRS^i_{Gx} as an "implicit relative price" of the public good in terms of the *numeraire*; since weak separability ensures that this "price" has not changed in the passage from the first- to the second-best, it is clear that the household's (compensated) "demand" of G is not changed either. The second step is to note that the transition from the first- to the second-best also affects the relative price of leisure in terms of private consumption, because low-ability households are now confronted with a positive marginal rate of income tax that reduces their compensated labour supply; since households enjoy more leisure in second-best, their "demand" for G is *reduced* because weak separability implies that leisure and the public good are Hicksian substitutes. In addition to this substitution effect, there is also a negative income effect that further reduces the demand for the public good (assuming normality). Hence, with a weakly separable utility function, the level of public-good provision in second-best is necessarily lower than in first-best — Pigou's conjecture is verified.

What happens if we now move from the second-best with full tax compliance to the present model with tax avoidance? We know from the previous analysis (see the last row of Table 1) that in this case the sum of the MRS^i_{Gx} exceeds unity; that is, the "price" of G has gone up, and consequently, the (compensated) "demand" for G has gone down. If we could show that the marginal tax rate of income tax for the low-wagers has gone up too, then we could complete the argument and show that the transition to the tax avoidance model implies a further reduction in the level of G. One could indeed presume that, being mimicking easier under tax avoidance (the mimicker can either work less or conceal more income), a higher marginal income tax rate is needed to deter it, but the general analysis in Section 3 above does not permit to reach this sort of conclusion for certain. To gain further insights into the matter we resort therefore to a numerical simulation.

We use a "modified" Cobb-Douglas utility function of the form:

$$u = \beta \lg x + (1-\beta)(1+G)^{\vartheta} \lg(1-l) + \gamma \lg G, \tag{11}$$

where $0 < \beta < 1$, $\gamma > 0$. The parameter ϑ captures the degree of L-substitutability (when $\vartheta < 0$) and L-complementarity (when $\vartheta > 0$); in fact, it is immediate to verify that MRS_{Gx} is decreasing (increasing) in $L = 1 - l$ when $\vartheta < 0$ ($\vartheta > 0$).[13] For $\vartheta = 0$, (11) reverts to a standard Cobb-Douglas form, which satisfies the L-independence, or weak separability, condition. Importantly, for ϑ □ 0, G and L are Hicksian substitutes, while for $\vartheta > 0$ they are Hicksian complements. The unit cost of avoidance is assumed to be quadratic in s, $c = (1+A)(1-s)^2/2$, so that marginal cost, $c_s = -(1+A)(1-s)$, is linear in s. We choose an economy characterized by the following structural param-

[13] When $\vartheta > 1$, u_{GG} can be positive for large values of G; a sufficient condition for $u_{GG} < 0$ for all G is $\vartheta < 1$.

Table 2: A comparison between the Boadway-Keen and the tax avoidance models

	ϑ	G	t^1	A	$\bar{v}^2$	*LHS*	*BK-term*	*TA-term*
TA	-0.10	16.02	0.323	1.137	2.223	1.0834	-0.0124	0.0958
TA	-0.05	14.39	0.329	1.071	2.155	1.0964	-0.0069	0.1033
TA	0.00	12.80	0.334	1.001	2.086	1.1118	0.0000	0.1118
TA	0.05	11.29	0.337	0.930	2.017	1.1294	0.0081	0.1213
TA	0.10	9.89	0.339	0.861	1.950	1.1492	0.0174	0.1318
BK	-0.10	18.07	0.207		2.232	0.9887	-0.0113	
BK	-0.05	16.26	0.230		2.164	0.9931	-0.0069	
BK	0.00	14.49	0.245		2.095	1.0000	0.0000	
BK	0.05	12.79	0.254		2.026	1.0092	0.0092	
BK	0.10	11.22	0.258		1.959	1.0203	0.0203	

eters:

$$\beta = 0.5,\ \gamma = 0.3,\ w^1 = 50,\ w^2 = 100,\ n^1 = 2/3,\ n^2 = 1/3.$$

Five values of ϑ are considered: –0.1 and –0.05 (L-substitutability), 0 (L-independence), 0.05 and 0.1 (L-complementarity). The results are reported in Table 2.[14]

Interestingly, we have that, for all ϑ, $G_{TA} < G_{BK}$. The transition from the second-best setting with tax compliance (BK model) to the one with tax avoidance (TA model) always induces a downward distortion in the level of public-good provision. To see why this is so, simply note that:

1. level-underprovision is always matched by rule-underprovision, as $n^1 MRS^1_{Gx} + n^1 MRS^1_{Gx}$ (*LHS* in the table) in the TA model is larger than one, and is also

[14] To solve the model we need to set a value for $\bar{v}^2$, i.e. the welfare level of type 2 households. We do this by first computing, for any given value of G, the welfare levels of types 1 and 2 households under poll tax financing ($T^1 = T^2$, $t^1 = t^2 = 0$). Let these welfare levels be $\tilde{v}^1(G)$ and $\tilde{v}^2(G)$, respectively. We then set $\bar{v}^2(G) = \tilde{v}^2 - \pi(\tilde{v}^2 - \tilde{v}^1)$ where $\pi \in (0,1)$ is a parameter capturing the degree of redistribution from type-2 to type-1 households ($\pi = 0.3$ in Table 2). A complete description of the algorithm employed to solve numerically the model is available from the authors upon request.

larger than its corresponding value in the BK model;

2. the marginal tax rate on low-wage households is, in accordance with the intuition given above, higher with tax avoidance than without it;[15]

Then, for ϑ □ 0, both these factors work in favour of a reduction of the "demand" for G; for $\vartheta > 0$, the presence of Hicksian complementarity between L and G means that the second factor works against it, but the effect is clearly not large enough to outweigh the downward distortion implied by the first factor.

6 Concluding remarks

In the real world, people do often their best to evade or avoid taxes, and most governments fight a constant battle against these activities. While several early and current contributions on optimal taxation have focused on the case in which the tax agencies ascertain the tax liability at no cost, there is by now a well-established body of knowledge on the economics of tax evasion and avoidance. We have built on this stream of work in order to investigate the question whether tax avoidance affects the optimal provision of a public good. In a period in which the legitimacy of government intervention is strongly questioned, and a thorough reform of the Welfare state is called for, it would seem interesting to know whether the explicit consideration of tax avoidance gives us some theoretical reason for advocating the expansion or the reduction of the scope of public action.

We have approached this issue using a rather general framework in which the main instrument for redistribution is the non-linear income tax. Our analysis suggests that it is optimal to reduce public-good provision below its first-best level and also below the level which would arise in a second-best world without tax avoidance. The main reason is that, under a simple assumption of normality of private and public consumption, underprovision of the public good relaxes the self-selection constraint, thereby enabling

[15] Actually, one can show that also the *effective* marginal tax rate, that is the one actually faced by the type-1 households $(t^1 s^1 + c^1)$, albeit lower than the statutory tax rate, is still higher than the marginal tax rate without tax avoidance.

a Pareto-improving change in the income tax schedule. This follows because the mimicker, that is the high-ability household which modifies its level of reported gross income in order to misrepresent its type, ends up having a larger disposable income than the true low-ability type, and therefore has a stronger preference for the public good. This way, underprovision will hurt the mimicker more than the true low-ability household.

References

Agell, J., Persson, M., 2000. Tax arbitrage and labour supply. Journal of Public Economics, 78, 3-24.

Andreoni, J., Erard, B., Feinstein, J., 1998. Tax compliance. Journal of Economic Literature 36, 818–860.

Atkinson, A., Stern, N., 1974. Pigou, taxation and public goods. Review of Economic Studies 41, 119–128

Boadway, R., Keen, M., 1993. Public goods, self-selection and optimal income taxation. International Economic Review 34, 463–478.

Boadway, R., Marchand, M., 1990. The use of public expenditure for distributive purposes. Core Discussion Paper no. 9066, Université Catholique de Louvain.

Boadway, R., Marchand, M., Pestieau, P., 1994. Towards a theory of the direct-indirect tax mix. Journal of Public Economics 55, 71–88.

Chang, M., 2000, Rules and levels in the provision of public goods: the role of complementarities between the public goods and taxed commodities. International Tax and Public Finance 7, 83–91.

Christiansen, V., 1981. Evaluation of public projects under optimal taxation, Review of Economic Studies, 48, 447–457.

Cowell, F.A., 1990. Cheating the government. The MIT Press.

Cremer, H., Gahvari, F., 1995. Tax evasion and the optimum general income tax. Journal of Public Economics 60, 235–249.

Edwards, J., Keen, M., Tuomala, M., 1994. Income tax, commodity taxes and public good provision: a brief guide. Finanz Archiv 51, 472–487.

Falkinger, J., 1991. On optimal public good provision with tax evasion. Journal of Public Economics 45, 127–133.

Gaube, T., 1999, Efficient publig good provision with nonlinear income taxation. SFB Discussion Paper no. A-595, University of Bonn.

Gaube, T., 2000, When do distortionary taxes reduce the optimal supply of public goods? Journal of Public Economics 76, 151-180.

Hindricks, J., 1999, Optimal tax mix and public good supply under multilateral tax evasion, Discussion Paper in Economics no. 99/20, University of Exeter.

Nava, M., Schroyen, F., Marchand, M., 1996. Optimal fiscal and public expenditure policy in a two-class economy. Journal of Public Economics 61, 119–137.

Schroyen, F., 1997. Pareto-efficient taxation under costly monitoring, Journal of Public Economics, 65, 343-366.

Slemrod, J., 1994. Fixing the leak in Okun's bucket. Optimal tax progressivity when avoidance can be controlled. Journal of Public Economics 55, 41–51.

Slemrod, J., 1998. A general model of the behavioral response to taxation. NBER Working Paper no. 6582.

Slemrod, J., Bakija, J., 1996. Taxing ourselves. The MIT Press.

Stiglitz, J., 1982. Self-selection and Pareto-efficient taxation. Journal of Public Economics 17, 213–240.

Wilson, J., 1991a. Optimal public good provision in the Ramsey tax model. Economic Letters 35, 57–61.

Wilson, J., 1991b. Optimal public good provision with limited lump-sum taxation. American Economic Review 81, 153–166.

Appendix (not included in the paper)

The appendix gives a more detailed derivation of the optimal policy rules. We will use some properties of the indirect utility function $v(t, T, G, A)$; by the envelope theorem:

$$v_t = -\alpha swl; \quad v_T = -\alpha; \quad v_G = u_G; \quad v_A = -\alpha wlc_A. \tag{12}$$

The mimicker's indirect utility function can be written $\hat{v}^2 = v\left(t^1, T^1, G, A\right)$. Its derivatives, using (3), (4), and (5) are:

$$\frac{\partial \hat{v}^2}{\partial t^1} = -\hat{\alpha}^2 y^1 + m\frac{\partial y^1}{\partial t^1}, \tag{13}$$

$$\frac{\partial \hat{v}^2}{\partial T^1} = -\hat{\alpha}^2 + m\frac{\partial y^1}{\partial T^1}, \tag{14}$$

$$\frac{\partial \hat{v}^2}{\partial G} = \hat{u}_G^2 + m\frac{\partial y^1}{\partial G}, \tag{15}$$

$$\frac{\partial \hat{v}^2}{\partial A} = -\hat{\alpha}^2 w^2 \hat{l}^2 \hat{c}_A^2 + m\frac{\partial y^1}{\partial A}, \tag{16}$$

where

$$\frac{\partial \hat{v}^2}{\partial y^1} \equiv m = -\hat{\alpha}^2\left(t^1 + \hat{c}_s^2\right) = \hat{\alpha}^2\left(c_s^1 - \hat{c}_s^2\right) > 0. \tag{17}$$

The Lagrangian of the optimal policy problem (6) is

$$= v^1 + \lambda\left(v^2 - \bar{v}^2\right) + \sigma\left(v^2 - \hat{v}^2\right) + \mu\left(\sum_{h=1}^{2} n^h\left(t^h s^h w^h l^h + T^h\right) - G - A\right). \tag{18}$$

Using (12) and (13)–(16), the first order conditions can be written as follows:[16]

$$\frac{\partial}{\partial t^1} = -\alpha^1 y^1 + \sigma\left(\hat{\alpha}^2 y^1 - m\frac{\partial y^1}{\partial t^1}\right) + \mu n^1\left(y^1 + t^1\frac{\partial y^1}{\partial t^1}\right) = 0, \tag{19}$$

$$\frac{\partial}{\partial T^1} = -\alpha^1 + \sigma\left(\hat{\alpha}^2 - m\frac{\partial y^1}{\partial T^1}\right) + \mu n^1\left(1 + t^1\frac{\partial y^1}{\partial T^1}\right) = 0, \tag{20}$$

[16] As all second-best policy problems, ours is not necessarily well-behaved; following virtually all the literature on the subject, we will however assume that a solution exists where all the constraints bind.

$$\frac{\partial}{\partial t^2} = -\alpha^2 y^2 (\lambda + \sigma) + \mu n^2 \left(y^2 + t^2 \frac{\partial y^2}{\partial t^2} \right) = 0, \tag{21}$$

$$\frac{\partial}{\partial T^2} = -\alpha^2 (\lambda + \sigma) + \mu n^2 \left(1 + t^2 \frac{\partial y^2}{\partial T^2} \right) = 0, \tag{22}$$

$$\frac{\partial}{\partial G} = u_G^1 + u_G^2 (\lambda + \sigma) - \sigma \left(\hat{u}_G^2 + m \frac{\partial y^1}{\partial G} \right) + \mu \left(\sum_{h=1}^{2} n^h t^h \frac{\partial y^h}{\partial G} - 1 \right) = 0, \tag{23}$$

$$\begin{aligned} \frac{\partial}{\partial A} &= -\alpha^1 w^1 l^1 c_A^1 - \alpha^2 w^2 l^2 c_A^2 (\lambda + \sigma) + \\ &\quad + \sigma \left(\hat{\alpha}^2 w^2 \hat{l}^2 \hat{c}_A^2 - m \frac{\partial y^1}{\partial A} \right) + \mu \left(\sum_{h=1}^{2} n^h t^h \frac{\partial y^h}{\partial A} - 1 \right) = 0. \end{aligned} \tag{24}$$

Multiplying (22) by y^2 and subtracting from (21), and multiplying (20) by y^1 and subtracting from (19), one obtains, respectively:[17]

$$t^1 = \frac{\sigma}{\mu n^1} \hat{\alpha}^2 \left(c_s^1 - \hat{c}_s^2 \right) > 0 \tag{25}$$

$$t^2 = 0, \tag{26}$$

which are given as expression (7) in the main text.

Next, substituting t^1 from (25), t^2 from (26), and $c_A^2 = 0$ (see Assumption 1) into the first order condition for A, (24), we obtain:

$$-\alpha^1 w^1 l^1 c_A^1 + \sigma \hat{\alpha}^2 w^2 \hat{l}^2 \hat{c}_A^2 - \mu = 0. \tag{27}$$

Substituting (25) into (20) we get $\alpha^1 = \sigma \hat{\alpha}^2 + \mu n^1$. Plunging this latter expression into (27) we finally obtain (8).

Finally, we substitute (25) and (26) into (23) to arrive at

$$u_G^1 + u_G^2 (\lambda + \sigma) - \sigma \hat{u}_G^2 - \mu = 0 \tag{28}$$

Substituting (26) into (22) and (25) into (20) we obtain, respectively, $\alpha^2 (\lambda + \sigma) = \mu n^2$ and $\alpha^1 - \sigma \hat{\alpha}^2 = \mu n^1$. Multiplying and dividing the first term of (28) by α^1, the second

[17] To arrive at (26)–(25), we assume that $\frac{\partial y^h}{\partial t^h} - y^h \frac{\partial y^h}{\partial T^h} \neq 0$, $h = 1, 2$, i.e. the compensated change in reported income is different from zero.

term by α^2 and the third by $\hat{\alpha}^2$; adding and subtracting $\sigma\hat{\alpha}^2 u_G^1/\alpha^1$; using the above expressions obtained from (22) and (20), after some manipulations, we finally obtain (9). Note that (9) can be fully written as:

$$n^1 \frac{u_G\left(x^1, l^1, G\right)}{u_x\left(x^1, l^1, G\right)} + n^2 \frac{u_G\left(x^2, l^2, G\right)}{u_x\left(x^2, l^2, G\right)} = \\ = 1 + \frac{\sigma\hat{\alpha}^2}{\mu}\left(\frac{u_G(\hat{x}^2, \hat{l}^2, G)}{u_x(\hat{x}^2, \hat{l}^2, G)} - \frac{u_G\left(x^1, l^1, G\right)}{u_x\left(x^1, l^1, G\right)}\right).$$

Since $\hat{x}^2 \neq x^1$ and $\hat{l}^2 \neq l^1$ by Lemma 1, the second term on the r.h.s. can be decomposed so as to obtain

$$n^1 \frac{u_G\left(x^1, l^1, G\right)}{u_x\left(x^1, l^1, G\right)} + n^2 \frac{u_G\left(x^2, l^2, G\right)}{u_x\left(x^2, l^2, G\right)} = \\ = 1 + \frac{\sigma\hat{\alpha}^2}{\mu}\left(\frac{u_G(x^1, \hat{l}^2, G)}{u_x(x^1, \hat{l}^2, G)} - \frac{u_G\left(x^1, l^1, G\right)}{u_x\left(x^1, l^1, G\right)}\right) + \\ + \frac{\sigma\hat{\alpha}^2}{\mu}\left(\frac{u_G(\hat{x}^2, \hat{l}^2, G)}{u_x(\hat{x}^2, \hat{l}^2, G)} - \frac{u_G(x^1, \hat{l}^2, G)}{u_x(x^1, \hat{l}^2, G)}\right)$$

Let $\widehat{MRS}^2_{Gx}|_{x^1}$ denote the mimicker's marginal valuation of G evaluated at the low-ability type consumption level. Then the expression above can be written in compact form as (10).