



Separability and public finance

Stéphane Gauthier^a, Guy Laroque^{a,b,*}

^a CREST-INSEE, France

^b University College London and Institute for Fiscal Studies, United Kingdom

ARTICLE INFO

Article history:

Received 30 October 2008

Received in revised form 26 June 2009

Accepted 8 September 2009

Available online 16 September 2009

JEL classification:

H21

H11

Keywords:

Separability

Second best optimality

Indirect taxes

Samuelson rule

Pigovian taxation

Individual production

ABSTRACT

In a second best environment, the optimal policy choice sometimes follows first best rules. This paper presents a formal general argument which allows to unify much of the literature. It lays down the information structure and separability assumptions under which the results hold in a variety of setups, with extensions to preference heterogeneity and individual production sets.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

In a first best setting, all information is publicly available and the government redistributes income at will. Efficiency can be achieved independently of equity concerns through lump sum transfers without recourse to indirect taxes. Pigovian taxes are used to correct externalities, and the Samuelson rule applies to the provision of public goods. In a second best environment, the social planner faces additional constraints, e.g., incentive constraints when some individual characteristics are not publicly observable. These new constraints narrow the scope of possible redistribution. The optimal rules of taxation then typically differ from the first best ones, and efficiency can no longer be disconnected from equity considerations. Indirect taxes then have a redistributive function, the Pigovian formula typically is not satisfied in the presence of externalities, and the Samuelson rule does not apply anymore.

Still there are a number of circumstances in the public finance literature where second best rules have a first best flavor. The best known instance is the Atkinson and Stiglitz (1976) theorem, according to which indirect taxation is useless when tastes for consumption goods are identical across agents, separable from leisure, and nonlinear income

taxation is allowed. In the same vein, also under separability assumptions, the optimal provision of public goods follows the Samuelson rule in Christiansen (1981) and Kaplow (1996), and Pigovian taxation is the appropriate way to handle externalities in Cremer et al. (1998) and Kaplow (2006b). Kaplow (2004, 2008b) provide a synthesis of this literature. All these results hinge on some form of separability, coupled with specific informational assumptions. The purpose of this paper is to put forward their common underlying structure in a unified setting. It appears that in all these examples one can isolate in the second best program a part which has a first best shape: conditional on the values taken by some variables, the remaining ones are solutions of a first best program from which the incentive constraints are absent. Thus the variables that are determined in this part of the program satisfy standard properties of first best allocations.

Our argument is of a global nature. It avoids solving for the optimal allocation and using first order conditions, thus bypassing the difficulties associated with second order conditions and bunching. The technique allows to provide a concise synthesis of the existing results in the literature and to extend some of them. First keeping Atkinson and Stiglitz (1976) separability assumptions, indirect taxation remains superfluous when tastes for consumption goods are heterogeneous. Two conditions are required for this extension to be valid: the differences in the agents' preferences must be observable by the government (this may be the case for, e.g., a handicap or family size) and the income tax schedule must be allowed to depend on these differences. The result holds both when these differences are

* Corresponding author. INSEE-CREST, J360 15 boulevard Gabriel Peri, 92245 Malakoff, France.

E-mail address: laroque@ensae.fr (G. Laroque).

exogenously determined (handicap), as previously shown by [Kaplow \(2008a\)](#), or are the outcome of the agents endogenous choices (number of children).

Second, we look more closely at the production side of the economy. The technology in [Atkinson and Stiglitz \(1976\)](#) is linear with fixed transformation rates between labor and the commodities, but the result actually holds in economies with a general aggregate constant returns production set, provided that wages are independent of aggregate consumption. We also consider economies with individual production sets which depend on the private ability of the entrepreneur. Under a separability assumption which makes input/output trade-offs in commodities separable from both labor and hidden characteristics, the choice of the (separable) input/output combinations should not be distorted by differential taxation. Suppose for instance that the quantity of labor required to produce one final good depends on another input, say education services, and on a private ability parameter. Under the (demanding) separability assumption the return on education (the increase in output which follows a marginal increase in education, holding labor input constant) is independent of the ability parameter. Then education should be provided competitively, with no subsidies neither to the teachers nor to the students, as in [Bovenberg and Jacobs \(2005\)](#). All redistribution takes place through a nonlinear income tax, which should only depend on incomes, not on the education levels of the workers, even if these levels are observable.

The paper is organized as follows. The next section presents a regularity assumption that bears on the constrained allocations under study. Three different models are then examined in turn, dealing with indirect taxation, the provision of public goods and individual production.

2. A non-satiation property

We define a non satiation property which states that extra resources lead to a second best Pareto improvement. This property is a corner stone of the analysis and may hold in a variety of setups. We present the definition in a general framework, which encompasses the models used in the rest of the paper.

Consider an economy with different consumers. The consumers are indexed by a possibly multidimensional characteristics n . They buy private goods c , supply labor ℓ and enjoy the public good g . The utility of consumer n is denoted $U(c_n, \ell_n, g, n)$. It is increasing in each component of private consumption and decreasing with labor supply. The distribution of characteristics in the population is described through the cumulative distribution function of n , denoted $F(\cdot)$. The economy may have a productive sector. The activity of firm j is described with a couple (z_j, ℓ_j) , where z_j is a vector of private goods whose positive components are outputs, and negative are inputs. The positive labor input is ℓ_j . Technology is represented through production sets satisfying standard assumptions. The public good g is produced from the inputs in private good and labor (c_g, ℓ_g) . The aggregate resource constraint on private goods (omitting labor) can be written as

$$\int_n c_n dF(n) + c_g \leq \int_j z_j d\tilde{F}(j),$$

where the inequalities hold componentwise.

In a first best setup, extra resources in private goods lead to a Pareto improvement. An allocation that satisfies this property is *non-satiated*. In a second best situation, the government has more limited power than in the first best, since it faces additional constraints, e.g., incentive compatibility. We refer to these additional constraints as *second best constraints*, and to an allocation which satisfies them as a *constrained allocation*. An optimal feasible constrained allocation is a second best allocation.

We shall consider constrained allocations which are non-satiated. Formally

Definition 1. A feasible constrained allocation is non-satiated when an increase dx , $dx \geq 0$, $dx \neq 0$, in the aggregate resources of private goods, leading to the aggregate resource constraint

$$\int_n c_n dF(n) + c_g \leq \int_j z_j d\tilde{F}(j) + dx,$$

allows a Pareto improvement while satisfying the second best constraints.

The production efficiency lemma of [Diamond and Mirrlees \(1971\)](#) states conditions under which the non-satiation property is satisfied at the optimum in an economy with linear taxation. This property also holds in the non linear taxation [Mirrlees \(1971\)](#) setup, under the single-crossing condition ([Hellwig \(2007\)](#)). The government then must take into account incentive constraints, in addition to the feasibility constraints. Under the usual single-crossing condition, the incentive constraints which make sure that high productivity individuals do not mimic the less-well-off are binding and prevent transfers from the rich to the poor. But it is always possible to give the extra resources to the richest agents without violating incentives. In the absence of single-crossing conditions with multiple consumption goods, there may exist circular no-envy conditions, such that every agent would like to imitate someone else. In this situation, the increase in aggregate resources must be used to simultaneously change the allocation of the concerned agents to induce a Pareto improvement while preserving incentive compatibility. Such a Pareto improvement is implementable in the (few) cases that we have studied. It seems plausible that constrained allocations often are non-satiated, but we have no proof of this property under general assumptions.

3. Indirect taxes

3.1. The Atkinson Stiglitz setup

In the setup of [Atkinson and Stiglitz \(1976\)](#), the preferences of agent n are represented by a utility function $U(V(c), \ell, n)$, where c is her consumption vector, ℓ her labor supply, and n is a non negative scalar which denotes labor productivity (there is no public good). This function is separable between consumption goods on the one hand, and labor supply on the other hand. The function $V(c)$, identical across agents, is assumed to be increasing and quasi-concave. When it is differentiable, the marginal rate of substitution between any two consumption goods does not depend on labor supply nor on productivity.

Technology is linear. The feasibility constraint takes the form

$$\int_n p c_n dF(n) \leq \int_n n \ell_n dF(n), \tag{1}$$

where p is a fixed vector of producer prices.

The government observes individual incomes $y = n\ell$, but not separately individual productivities nor labor supplies. It announces a non linear schedule $R(\cdot)$ which relates before tax income y to after tax income $R(y)$. It also can impose linear taxes $q-p$ on consumption goods, where q is the vector of consumer prices. Given q and $R(\cdot)$, an agent whose before tax income is y maximizes her utility $U(V(c), \ell, n)$ subject to the budget constraint $qc = R(y)$. From the separability assumption, the consumption bundle chosen by the agent depends only on her income y , not separately on her productivity n or her labor supply ℓ . This bundle indeed maximizes $V(c)$ on the budget set $qc = R(y)$. The solution to this problem is the demand function $\gamma(q, R(y))$.

The government chooses a vector of consumer prices q , an income profile (y_n) and the corresponding after tax income profile $R(y_n)$,

which maximize a weighted sum of utilities subject to the feasibility constraint (1) and the incentive compatibility constraints

$$U\left(V(\gamma(q, R(y_n))), \frac{y_n}{n}, n\right) \geq U\left(V(\gamma(q, R(y_m))), \frac{y_m}{n}, n\right)$$

for every (m, n) .

Lemma 1. 1. Consider a non-satiated second best allocation in which agent n has before tax income y_n^* and consumes c_n^* . Given $(y_n^*), (c_n^*)$ is a first best allocation of the economy where all the agents have the same quasi-concave and increasing utility function $V(\cdot)$ and the aggregate production set is

$$\int_n p c_n dF(n) \leq \int_n y_n^* dF(n). \tag{2}$$

2. Consider a non-satiated constrained allocation in which agent n has before tax income y_n^* . Given (y_n^*) , it is (weakly) Pareto dominated by another non-satiated constrained allocation in which consumption (c_n^*) is a first best allocation of the economy where all the agents have the same quasi-concave and increasing utility function $V(\cdot)$ and the aggregate production set is defined by (2).

Proof. The same argument applies to both parts of the Lemma. In the interest of space, we focus on the first part: the reference allocation $((y_n^*, c_n^*))$ is a second best allocation. Let $V_n^* = V(c_n^*)$ be the sub-utility derived from consumption. The incentive constraints take the form

$$U\left(V_n^*, \frac{y_n^*}{n}, n\right) \geq U\left(V_m^*, \frac{y_m^*}{n}, n\right) \text{ for every } (m, n).$$

Any allocation $((c_n, y_n^*))$ that yields the profile (V_n^*, y_n^*) satisfies the incentive constraints.

Consider a fictitious economy where agents have preferences $V(\cdot)$ and where the feasible set is defined by (2). From the second welfare theorem (Mas-Colell et al. (1995), Proposition 16.D.1), a first best optimum of this economy can be decentralized as a quasi-equilibrium with an appropriate choice of (q, R_n) . Note that R_n can be written as a function $R(y_n)$. Indeed, consider two agents m and n with the same before tax income y . Then, by incentive compatibility they must have the same sub-utility V , and therefore the same after tax income R .

Now, we proceed by contradiction. Suppose that the reference allocation is not a first best optimum of the fictitious economy. Then one can achieve the same utility profile (V_n^*) with aggregate resources less than $\int_n y_n^* dF(n)$, and the government has a positive dy at its disposal. The non-satiation property gives the desired contradiction. \square

Indirect taxation is therefore superfluous. Unlike Atkinson and Stiglitz (1976), the argument used here does not depend on the optimality of the initial allocation, as already noted by Kaplow (2006a) and Laroque (2005). The proof of Kaplow is based on a marginal change, using first order conditions. The proof here is in the spirit of Laroque (2005), with more emphasis on non satiation and a more abstract general equilibrium type of argument (it appeals to the second welfare theorem, rather than properties of the expenditure functions). Indeed, consider an incentive compatible allocation with distorting indirect taxes. Separability allows to keep constant the utilities derived from consumption while suppressing indirect taxes, by adapting the non linear after tax income schedule. When the indirect taxes are distortive, this transformation yields a surplus (part 2 of the lemma). As a consequence, any non satiated second best allocation (as in part 1 of the lemma) has non distortive commodity taxes. In the sequel of the paper, we focus our attention on the properties of second best allocations, but similar arguments would show that applying first best rules would yield a Pareto improvement on constrained (inefficient) allocations.

Remark 1. Lemma 1 holds in economies with a general production set Y , where the primary input is the aggregate supply of labor, measured in efficiency units, provided profits if any are taxed away by the government. One just needs to replace (2) with

$$\int_n (-y_n, c_n) \in Y.$$

The production set Y can be derived from a more detailed structure, involving intermediary inputs that are produced and used in the production process. Then Lemma 1 implies no price distortions in the allocation of these inputs.¹

Remark 2. The previous argument extends to any situation where labor supply can be written as a function of income y and of the unobservable parameter n . Instead of the Atkinson and Stiglitz (1976) productivity model, suppose that n is a multidimensional characteristic. Suppose also that in order to get income y , agent n must supply a quantity of labor $\ell = L(y, n)$, where L is a fixed known function. Then the reader can easily check that Lemma 1 carries over to this more general setup. However with a general production set, such as that in the previous remark, it is natural to have wage, or labor productivity, function of aggregate variables, e.g., aggregate capital or production, so that the function L could depend on aggregate consumption in the economy. As discussed in Stiglitz (1985) and Naito (1999), Lemma 1 then typically does not carry through.

3.2. An example with unobservable heterogeneous preferences

There are two agents and two consumption goods, whose quantities are designated respectively with lower case c and upper case C . Agent i has productivity n_i , and we assume $n_2 > n_1$ so that 2 is the richest. We consider a situation where the assumption of Atkinson and Stiglitz (1976) does not hold: the tastes for goods of the two agents depend on the hidden productivity parameter. Agent 1 likes consuming commodity c , but does not care for good C , while this is the converse for agent 2. Their utility functions respectively are

$$\ln c_1 - \frac{y_1}{n_1},$$

$$\ln C_2 - \frac{y_2}{n_2}.$$

With an appropriate choice of units the feasibility constraint takes the form

$$c_1 + C_2 \leq y_1 + y_2.$$

The planner's objective is to maximize the sum of the agents' utilities. It is easy to check that at the first best optimum the more productive agent does all the work, $y_1 = 0, c_1 = C_2 = n_2$ and $y_2 = 2n_2$.

At first glance, the first best allocation is unlikely to be attainable in a second best environment. Agent 2 would rather not work, and get the same income as agent 1, which however she would spend on good C rather than good c . This intuition is incorrect when the government can implement linear indirect taxes, on top of a non linear income tax. To see this, let q and Q respectively be the consumer prices of goods c

¹ The lemma also extends to a multi-period model. Suppose that there are a finite number of dates $t = 1, \dots, T$. At date t , the typical agent consumes c^t , supplies a quantity of labor ℓ^t and has productivity n^t , with $y^t = n^t \ell^t$. Her utility under the separability assumption takes the form

$$U[V(c^1, \dots, c^T), \ell^1, \dots, \ell^T, n^1, \dots, n^T].$$

Then a non-satiated allocation can be decentralized with intertemporal prices which apply both to production and consumption, provided after tax income can be made a function $R(y^1, \dots, y^T)$ of the before tax income profile. Intermediary products, capital in particular, need not be taxed.

and C. Define after tax incomes $R(0) = qn_2$ and $R(2n_2) = Qn_2$. Then the incentive constraints at the first best allocation can be written as follows. Agent 1 would buy Qn_2/q units of good c if he worked as much as agent 2, so that he will not imitate her when

$$\ln n_2 \geq \ln \frac{Qn_2}{q} - \frac{2n_2}{n_1}.$$

Similarly agent 2 will stay at work if

$$\ln n_2 - 2 \geq \ln \frac{qn_2}{Q}.$$

This last constraint is satisfied if Q/q is larger than $\exp(2)$, when the relative price of the good preferred by agent 2 is large enough. The incentive constraint of agent 1 does not bind provided the relative price is not too large, Q/q smaller than $\exp(2n_2/n_1)$. Differences in tastes allow the government to reach the first best allocation through indirect taxation. Note that in the absence of indirect taxes ($q = Q$), the intuition mentioned above is valid, agent 2 incentive constraint binds, and the first best allocation cannot be reached. When tastes for consumption goods depend on the parameter n , indirect taxes are typically part of the second best fiscal policy: when hidden heterogeneity affects preferences for consumption goods, Lemma 1 fails.

3.3. Observed heterogeneous preferences

However, as noted by Kaplow (2008a), Lemma 1 holds when preferences for consumption goods depend on observable and verifiable individual characteristics, provided that income tax can be made conditional on these characteristics. In order to make this statement precise, let v be the multidimensional type of the agent, where one component of v is productivity n . Let a be a parameter that describes her tastes for consumption goods. Under the separability assumption, the agent utility function takes the form $U(V(c, a), \ell, a, v)$. We assume a to be observable, think of a as, e.g., family size. We allow in turn a to be a fixed exogenous characteristics as in Kaplow (2008a) (i.e. $v = (a, n)$), or to be endogenously chosen from some given set, say A . Production is described as before in Section 3.1, and the feasibility constraint (1) stays unchanged.

Lemma 2. Suppose that the after tax income schedule can be made dependent on the observable characteristics a . Consider a non-satiated second best allocation² in which agent v has characteristics a_v , before tax income y_v and consumes c_v . Given (a_v, y_v) , (c_v) is a first best allocation of the economy where, for all v , agent v has preferences for goods given by the quasi-concave and increasing utility function $V(\cdot, a_v)$ and the aggregate production set is

$$\int_v p c_v dF(v) \leq \int_v y_v^* dF(v). \tag{3}$$

Proof. Let $\gamma(q, R, a)$ be the demand for goods of an agent of characteristics a , facing the budget constraint $qc = R$. We have $c_v = \gamma(q, R(y_v, a_v), a_v)$. When a is exogenous, the incentive compatibility constraints are

$$U\left(V(c_v, a), \frac{y_v}{n(v)}, a, v\right) \geq U\left(V(c_\mu, a), \frac{y_\mu}{n(v)}, a, \mu\right),$$

for all a, μ and v such that $a(\mu) = a(v) = a$. When a is endogenous, agent v can mimic any other agent μ , provided she aligns her choice of a with that of μ . There are more incentive compatibility constraints

$$U\left(V(c_v, a_v), \frac{y_v}{n(v)}, a_v, v\right) \geq U\left(V(c_\mu, a_\mu), \frac{y_\mu}{n(v)}, a_\mu, v\right).$$

Let V_v^* be the sub-utility obtained by the agent at the optimum. In both cases, any allocation that keeps the profile (V_v, y_v, a_v) unchanged satisfies the incentive constraints.

The same argument as in the proof of Lemma 1 then applies. The consumption profile (c_v^*) is a first best allocation of the economy with utilities $(V(\cdot, a_v^*))$ and aggregate endowment $\int_v y_v^* dF(v)$. Appealing to the second theorem of welfare economics, (c_v^*) can be decentralized through a fiscal policy which does not involve indirect taxes ($q^* = p$).

It may be worthwhile to remark that the argument works because the variables that come out of the second welfare theorem (i.e. consumptions c , or prices and after tax incomes) do not bear on the choice of a , so that the original profile a_v^* is compatible with the Pareto improvement. Here this comes from the assumption that a is chosen from a given set A . As pointed out by Hellwig (2008a), the result extends to the case where A depends on y , e.g. $a = y$. □

Indirect taxation is thus superfluous in situations where the separable preferences for goods are heterogeneous, provided that this heterogeneity is both publicly observable and can be made a determinant of the income tax schedule. In some cases, this last condition is not a stringent requirement. For instance, transfers such as family benefits depend on family composition, and there are specific social assistance schemes based on age or disability. Still, there are cases where this kind of requirement may be considered discriminatory, say when based on gender or race.

4. Public good provision

There is an extensive literature which studies whether one can dissociate equity from efficiency considerations when providing public good, including Boadway and Keen (1993), Guesnerie (1995), Kaplow (1996), Nava et al. (1996) and Slemrod and Yitzhaki (2001). Intuition, as discussed by Sandmo (1998), suggests that the marginal cost of public funds should be larger, the more distortionary the taxes used to finance the good. In fact, under appropriate separability assumptions, this intuition is not valid and a version of the first best Samuelson rule applies. Of course this does not mean that the quantities of public good are equal at the first and second best: willingness to pay differ in the two worlds.

The economy has a single good which can be used either for private or public consumption and labor. Labor can be transformed into good at constant returns to scale, one unit of labor giving one unit of good. Everyone in the economy has the same utility function $U(V(c, g), \ell)$, where c denotes private consumption and g public good, but productivities differ. As in Section 3, agent n must work $\ell = y/n$ hours to get before tax income y . The aggregate feasibility constraint is

$$\int_n c_n dF(n) + g = \int_n y_n dF(n).$$

The government observes individual incomes y , but neither individual productivity n nor labor ℓ separately. It chooses a quantity of public good g and an after tax income schedule $R(y)$.

Given g , person n chooses an income level y that maximizes $U(V(R(y), g), y/n)$.

Lemma 3. Consider a non-satiated second best allocation $((y_n^*, c_n^*), g^*)$. Then, given (y_n^*) , $((c_n^*), g^*)$ is a first best allocation of the economy with utility functions $V(c_n, g)$ and production set

$$\int_n c_n dF(n) + g = \int_n y_n^* dF(n).$$

As a consequence, when V is differentiable, the provision of public good satisfies the Samuelson rule

$$\int_n \frac{\partial V / \partial g}{\partial V / \partial c} dF(n) = 1.$$

² As in Lemma 1, the result also applies to a constrained (not optimal) allocation.

The sum of the willingness to pay for public good of all consumers is equal to the marginal cost of production. The allocation rule is unaffected by the redistributive concerns of the government.

Proof. It is similar to the proof of Lemma 1. Let $V_n^* = V(R^*(y_n^*), g)$ be the sub-utility profile associated with the second best allocation. Given (y_n^*) , any consumption schedule $((c_n), g)$ which keeps (V_n^*) unchanged satisfies the incentive compatibility constraints. It is therefore attainable by the second best government provided it satisfies the feasibility constraint

$$\int_n c_n dF(n) + g = \int_n y_n^* dF(n).$$

This implies that the allocation $((c_n^*), g^*)$ is first best for the economy with utility function V and aggregate resources $\int_n y_n^* dF(n)$. Otherwise, the same utility profile could be achieved with less resources, a contradiction with the non-satiated hypothesis. \square

Remark 3. It is possible to have more heterogeneity in tastes for public good, as in Hellwig (2008b). Indeed, when preferences take the form $u(\phi(c), \ell) + \psi(g, \zeta)$, the incentive constraints are independent of the heterogeneity ζ . The problem separates into pieces. Given the before tax income profile (y_n^*) , any $((c_n), g)$ such that $\phi(c_n) = \phi(c_n^*)$ satisfies the incentive compatibility constraints. This implies, under non-satiation, that the allocation $((c_n^*), g^*)$ is first best for the economy in which agents have utilities $u(\phi(c), y_n^*/n) + \psi(g, \zeta)$ and the feasibility constraint is

$$\int_n p c_n dF(n) + g = \int_n y_n^* dF(n).$$

When the government knows the joint distribution of (n, ζ) , the optimal provision of public goods is governed by the Samuelson rule.

Remark 4. A similar argument applies in the presence of externalities as in Cremer et al. (1998) and Kaplow (2006b). Suppose the agents' utility functions are of the type

$$U(V(c, e), \ell)$$

where c is the individual consumption vector of private goods $i = 1, \dots, k$, and \mathbf{c} denotes the collection of consumptions in the economy, which may influence individual welfare but is treated as an externality. As in Section 3, one unit of good i is produced from p_i units of labor according to a linear technology. The government can non linearly tax income through an after tax income schedule $R(\cdot)$, and linearly tax goods by quoting a vector of consumer prices q different from p . Then a non-satiated second best allocation can be supported with a first best Pigovian tax designed for the economy with utility functions V and feasibility constraint $\int_n p c_n dF(n) \leq \int_n y_n^* dF(n)$.

To fix ideas, consider the simple case where all rates of transformation are equal to 1, and the only externality, a nuisance, comes from the aggregate consumption C_1 of good 1. Then the sub-utility function simplifies into $V(c, C_1)$, and all supporting consumer prices can be taken equal to 1, except for that of the first good which supports a Pigovian tax

$$q_1 = 1 + \int_n \frac{\partial V / \partial C_1}{\partial V / \partial c_1} dF(n).$$

This tax is anonymous, although the poor agents may rely on commodity 1 more than the rich ones. Anonymity would not hold, with possibly a nonlinear tax depending on income, in the absence of separability. Note that the Pigovian formula implicitly depends on the full second best allocation, so that the government redistributive motive influences the aggregate consumption of good 1.

5. Individual production sets

At a second best optimum which satisfies the non-satiation property, any production process implemented through an aggregate production function independent of individual characteristics is efficient. When production sets are individual and depend on the abilities of their owners, one does not expect this property to typically hold anymore. For instance, if individual production depends on individual education, the government may want to subsidize education to improve the welfare of the more deserving part of the population. In developing countries, fertilizers are sometimes massively subsidized, the subsidy possibly being a function of the size of the harvest or the farmland. In India, these subsidies go to both farmers and producers of fertilizers.

To discuss the usefulness of taxes and subsidies in this context, we retain most of the framework of the model of the previous sections. There is a finite number of goods and labor, and the preferences of agent n are represented by the utility function $U(c, \ell, n)$. For now, the preferences for goods c do not necessarily separate from (ℓ, n) . We depart from the previous sections by assuming that agent n has access to an individual production function. The same physical good can be produced or used as an input, and we want to allow for different buying and selling prices. The outputs of the production process are reported in the vector z_+ and the inputs in the vector z_- , both of the dimension of the number of goods. The net amount produced is $z_+ - z_-$. Production is described by the vector $z = (z_+, -z_-)$. Formally the feasible production vectors are represented by a cost function, which gives the quantity of labor required from individual n to produce the vector z :

$$\ell = \tilde{G}(z, n).$$

The government is assumed to observe the vector z_n , an assumption which will be discussed later in Remark 5. The before tax income of agent n is $p(z_n)z_n = p_+(z_n)z_{n+} - p_-(z_n)z_{n-}$, where $p(z)$ denotes the vector of possibly nonlinear production prices chosen by the government for redistribution purposes. For instance, one good can be general education e . When acquired, it shows up as an input of education services, a component of z_- . Education services are produced by teachers (then a component of their z_+), themselves formed in the education system. The distinction between inputs and outputs, contrary to the usual convention to only show the difference $z_+ - z_-$, allows for different tax treatments. The payment $p_{e+}(z)$ to the teachers providers of education services ($e_+ > 0$) may be different from the (possibly zero) cost charged to the students, i.e. $p_{e-}(z) = 0$, or $p_{e-}(z) < p_{e+}(z)$. These two prices of education services a priori may depend on the whole vector of inputs and outputs, and not only on education.

In this setup, the feasibility constraint becomes

$$\int_n c_n dF(n) \leq \int_n [z_{n+} - z_{n-}] dF(n). \tag{4}$$

The government can impose anonymous linear taxes on consumption, leading to a consumption price vector q . The cost of consumption bundle c is qc . The government neither observes the characteristics n nor labor supply ℓ_n . It announces a non linear schedule $R(p(z)z)$. When agent n chooses the productive activities z_n , agent n gets an after tax income $R(p(z_n)z_n)$, to be spent on consumption goods, $q c_n = R(p(z_n)z_n)$.

In the sequel, we consider separable cost functions, so that

$$\tilde{G}(z, n) = G(H(z), n), \tag{5}$$

where G is increasing in its first argument, and H is increasing and quasi-convex. This is a strong assumption: at a given z , the marginal returns on inputs are independent of hidden characteristics n . When z comprises an education input and an associated output, separability implies that the marginal return on education is independent of ability n .

Lemma 4. *At a non-satiated second best allocation, the production profile (z_n^*) is efficient: there does not exist (z_n) that requires less labor input*

$$H(z_n) \leq H(z_n^*) \text{ for all } n$$

while

$$\int_n [z_{n+} - z_{n-}] dF(n) \geq \int_n [z_{n+}^* - z_{n-}^*] dF(n),$$

with some strict inequality. Therefore z_n^* is profit maximizing for some linear price vector p^* , with $p_+^* = p_-^*$, subject to the constraint $H(z) \leq H(z_n^*)$.

Proof. We first prove the preliminary property that, at any incentive compatible allocation, the after tax income schedule $R(p(z)z)$ is an increasing function of $H(z)$, say $\rho(H(z))$. Suppose that $H(z) < H(z')$ for two different production vectors z and z' . Recall that agent n chooses a bundle (c, ℓ, z) which maximizes her utility $U(c, \ell, n)$ subject to $\ell = G(H(z), n)$ and $qc \leq R(p(z)z)$. If $R(p(z)z) \geq R(p(z')z')$, then z is always preferred to z' (agent n works more with z' and has a lower after tax income). This shows that $R(p(z)z) < R(p(z')z')$, and thus proves the desired property.

We now proceed with the proof. Let $(q^*, (R_n^*))$ stand for the consumer prices and the after tax income profile at the second best optimum and $H_n^* = H(z_n^*)$. Denote $(c_{nm}, \ell_{nm}, z_{nm})$ the quantities that maximize consumer n utility $U(c, \ell, n)$ subject to the constraints $\ell = G(H(z), n)$, $q^*c \leq R(p(z)z)$ and $H(z) = H_n^*$, $R(p(z)z) = R_n^*$. Given $(q^*, (R_n^*))$, any production profile (z_n) such that $H(z_n) = H_n^*$ is incentive compatible, i.e., it satisfies

$$U(c_{nm}, \ell_{nm}, n) \geq U(c_{nm}, \ell_{nm}, n)$$

for all m and n . Let $\Delta_n^* = \{z_+ - z_- | H(z_+, z_-) \leq H_n^*\}$ be the convex set of net production vectors that can be realized without increasing the labor supply of agent n at the second best optimum. Suppose, by contradiction, that there exists a profile $(z_{n+} - z_{n-})$ in Δ_n^* which yields a larger aggregate production than $(z_{n+}^* - z_{n-}^*)$:

$$\int_n (z_{n+} - z_{n-}) dF(n) \geq \int_n (z_{n+}^* - z_{n-}^*) dF(n),$$

with at least one strict inequality.

From standard results on production aggregation (Mas-Colell et al. (1995), Sections 5.E and 5.F), the production vector z_n can be chosen so that there exists a production price vector p such that net production $(z_{n+} - z_{n-})$ maximizes profit on Δ_n^* .

At the new production plan (z_n) , the after tax income can be taken equal to the reference profile (R_n^*) . Indeed, by construction, $H(z_n) = H(z_n^*)$ for all n is a function of the profit pz_n . Using the remark at the beginning of the proof, $R_n^* = \rho^*(H(z_n))$ thus is a function of pz_n . The allocation (c_{nm}, ℓ_{nm}, z_n) associated with the production profile (z_n) and after tax incomes (R_n^*) , given q^* , thus satisfies all the incentive and measurability constraints. By the non-satiation hypothesis, the extra resources

$$dx = \int_n (z_{n+} - z_{n-}) dF(n) - \int_n (z_{n+}^* - z_{n-}^*) dF(n)$$

would allow a Pareto improvement.

Therefore, $\int_n (z_{n+}^* - z_{n-}^*) dF(n)$ belongs to the boundary of the set $\int_n \Delta_n^* dF(n)$. A standard separation theorem leads to the desired conclusion. \square

In this setup, to achieve a given production combination z , the agents must provide different labor supplies depending on their individual characteristics n . These characteristics are not publicly observable,

however. This informational assumption prevents the government to make transfers conditional on n . One could have expected the government to exploit the fact that the choice of z typically reveals some information about n , by letting production prices p be a function of z . Under the separability assumption embedded in Eq. (5), this is not the case: Lemma 4 shows that production prices should be independent of z . Therefore the marginal rate of transformation between any two components of z is identical for all producers: production is efficient and profit maximizing given linear prices. Education is provided in a competitive way, with students paying the full cost of their education, and neither producers of fertilizers nor farmers are subsidized. The production prices should not depend on income or other components of the vector z . This property is reminiscent of the production efficiency lemma of Diamond and Mirrlees (1971). But their framework, with linear taxes in general equilibrium, is different.

So far we have been silent about the optimal consumer prices. Under the separability assumption of preferences of Atkinson Stiglitz, with separable production functions, one obtains the following result as a corollary of Lemma 1 and 4:

Corollary 1. *Assume separable labor cost functions (5). Assume furthermore that the preferences of the agent of characteristics n are represented by $U(V(c), \ell, n)$, where V is a smooth increasing quasi-concave function. Then, given $(z_n^*), (c_n^*)$ is a first best allocation of the economy in which agents have preferences given by the utility function V and the feasibility constraint is*

$$\int_n c_n dF(n) \leq \int_n [z_{n+}^* - z_{n-}^*] dF(n).$$

At the optimum, the producer prices p_+^* and p_-^* and consumer prices q^* can be taken to be equal.

Proof. This is a simple extension of Lemma 1. Let V_n^* be the sub-utility derived from consumption at the optimum, i.e., $V_n^* = V(\gamma(q^*, R^*(p^*z_n^*)))$. Any allocation that yields the same (V_n^*, H_n^*) profile is incentive compatible. Consider an economy where all agents have the same utility function $V(\cdot)$ and there is a family of production sets (Δ_n^*) . Then (c_n^*, z_n^*) is a first best allocation of this economy. Otherwise the government could achieve the same utility profile with less resources, a contradiction with the non-satiation hypothesis. From the second welfare theorem, the optimum can be decentralized with a price vector that applies both to consumer choices and producer choices. \square

With the linear production set of Atkinson and Stiglitz (1976), the relative producer prices are exogenously given. With a non linear production set, they become endogenous. The previous result shows that Lemma 1 extends to this situation under separability assumptions on technology. Then, indirect taxes are useless and consumer prices should be proportional to production prices.

An interesting case occurs when one assumes that production and consumption can only be taxed linearly, while the government can impose non linear taxes on income. Then, when utility is not separable, consumption prices will typically differ from production prices: $q^* \neq p_+^* = p_-^*$, and one gets something that looks like a VAT regime with rates differing across goods. Note that this assumes that the government can separate transactions for production purposes, z , from transactions for consumption purposes, c .

Remark 5. To implement a nonlinear price scheme $p(z)$, the government must observe the vector z . A less demanding informational setup is one where the government only observes the signs of the transactions, i.e. can only impose different buy and sell prices. Under production separability, Lemma 4 shows that at the optimum there is no wedge between the buy and sell prices, so that the government only has to observe the before tax income from the production activities to implement the second best.

Example 6. Bovenberg and Jacobs (2005) study whether one should subsidize education in a framework in which the agent of productivity n produces y units of a single good by using labor ℓ and an educational input e . Her individual production function is $y = \phi(\psi(\ell, n), e)$. Both labor and productivity are private information, while individual production and education are publicly observed by the government. Provided that individual production is increasing with efficient labor $\psi(\ell, n)$, the government can relate the hidden efficient labor to variables that are all publicly known, say $\psi(\ell, n) = H(y, e)$. Thus, when efficient labor increases with productivity, ℓ can be expressed as a function of $H(y, e)$ and n , which fits the functional form given in Lemma 4, with $z = (y, e)$. As a result, in this setting, education should be neither taxed nor subsidized.

Remark 7. Lemma 4 extends to the case where the H function depends on observable heterogeneous characteristics, provided that the after tax income can be made conditional on these characteristics. An example might be related to health expenditures. Suppose that individual health has two components (a, n) , where a is observable, while n is private knowledge. Suppose that the production function is separable

$$\ell = G(H(z, a), a, n).$$

Drugs and hospital care bought by the individuals are part of z . If the income tax can be made conditional on a , there should be no subsidies on health care. All redistribution should proceed through income taxation.

Acknowledgments

We thank Rafael Aigner, Martin Hellwig, Etienne Lehmann, Nicola Pavoni and seminar participants in Bonn and Paris for their thoughtful comments on the paper. We have greatly benefited from the remarks and suggestions of an anonymous referee and of the editor, Pierre Pestieau.

References

- Atkinson, Anthony, Stiglitz, Joseph, 1976. The design of tax structure: direct versus indirect taxation. *Journal of Public Economics* 6, 55–75.
- Boadway, Robin, Keen, Michael, 1993. Public goods, self selection and optimal income taxation. *International Economic Review* 34, 463–478.

- Bovenberg, A. Lans, Jacobs, Bas, 2005. Redistribution and education subsidies are Siamese twins. *Journal of Public Economics* 89, 2005–2035.
- Christiansen, Vidar, 1981. Evaluation of public projects under optimal taxation. *Review of Economic Studies* 48 (3), 447–457.
- Cremer, Helmuth, Gahvari, Firouz, Ladoux, Norbert, 1998. Externalities and optimal taxation. *Journal of Public Economics* 70, 343–364.
- Diamond, Peter, Mirrlees, James, 1971. Optimal taxation and public production, Part I: production efficiency. *American Economic Review* 61, 8–27.
- Guesnerie, Roger, 1995. A contribution to the pure theory of taxation. Cambridge University Press.
- Hellwig, Martin, 2007. A contribution to the theory of optimal utilitarian income taxation. *Journal of Public Economics* 91, 1449–1471.
- Hellwig, Martin, 2008a. A generalization of the Atkinson-Stiglitz (1976) theorem on the undesirability of nonuniform excise taxation. Technical report, Max Planck Institute for Research on Collective Goods.
- Hellwig, Martin, 2008b. Optimal income taxation, public goods provision and public sector pricing: a contribution to the foundations of public economics. Technical report, Max Planck Institute for Research on Collective Goods.
- Kaplow, Louis, 1996. The optimal supply of public goods and the distortionary cost of taxation. *National Tax Journal* 49, 513–533.
- Kaplow, Louis, 2004. On the (ir)relevance of distribution and labor supply distortion to government policy. *Journal of Economic Perspectives* 18 (4), 159–175.
- Kaplow, Louis, 2006a. On the undesirability of commodity taxation even when income taxation is not optimal. *Journal of Public Economics* 90 (6–7), 1235–1250 August.
- Kaplow, Louis, 2006b. Optimal control of externalities in the presence of income taxation. Technical Report 12339. NBER.
- Kaplow, Louis, 2008a. Optimal policy with heterogeneous preferences. *Journal of Economic Analysis and Policy: Advances* 1, 1–28.
- Kaplow, Louis, 2008b. The theory of taxation and public economics. Princeton University Press.
- Laroque, Guy, 2005. Indirect taxation is superfluous under separability and taste homogeneity: a simple proof. *Economics Letters* 87, 141–144.
- Mas-Colell, Andreu, Whinston, Michael D., Green, Jerry R., 1995. *Microeconomic Theory*. Oxford University Press.
- Mirrlees, James, 1971. An exploration in the theory of optimum income taxation. *Review of Economic Studies* 38, 175–208.
- Naito, Hisahiro, 1999. Re-examination of uniform commodity taxes under a non-linear income tax system and its implication for production efficiency. *Journal of Public Economics* 71, 165–188.
- Nava, Mario, Schroyen, Fred, Marchand, Maurice, 1996. Optimal fiscal and public expenditure policy in a two-class economy. *Journal of Public Economics* 61, 119–137.
- Sandmo, Agnar, 1998. Redistribution and the marginal cost of public funds. *Journal of Public Economics* 70, 365–382.
- Slemrod, Joel, Yitzhaki, Shlomo, 2001. Integrating expenditure and tax decisions: the marginal cost of funds and the marginal benefit of projects. Technical Report 8196. NBER.
- Stiglitz, Joseph, 1985. Inequality and capital taxation. Technical Report 457. IMSSS.