Optimal Taxation of Intermediate Goods in the Presence of Externalities: A Survey Towards the Transport Sector

Joakim Ahlberg*

May 31, 2006

Abstract

The paper surveys the literature on optimal taxation with emphasis on intermediate goods, or, more specific, freight (road) transport. There are two models frequently used, first, the one emanated from Diamond & Mirrlees' (1971) paper, where the production efficiency lemma made it clear that intermediate goods was not to be taxed. And, second, the Ramsey-Boiteux model where a cost-of-service regulation imposes a budget constraint for the regulated firm. In the latter model, in contrast to the first, freight transports (intermediate goods) are to be taxed in the Ramsey tradition, and thus trades the production efficiency lemma against a budget restriction.

The paper also discusses welfare effects due to environmental tax reforms, with emphasis to what has become to known as the double dividend hypothesis. Finally, administrative costs in the context of optimal taxation is touched upon, a subject that is to a large degree repressed in optimal tax theory.

*The author is grateful to Gunnar Lindberg, Jan-Eric Nilsson and Lars Hultkrantz for comments and support.

Contents

1 Summary

The present paper surveys the literature on optimal taxation on intermediate goods, starting from Diamond & Mirrlees' (D & M) famous papers from 1971 (even though it includes Ramsey's 1927 paper for completeness).

The first of the two papers from D & M made it clear that there were no scope for intermediate good taxation (at least not in a competitive economy producing with constant returns to scale) since, in the absence of profits, taxation on intermediate goods must be reflected in changes in final good prices. Therefore, the revenue could have been collected by final good taxation, causing no greater change in final good prices and avoiding production inefficiency.

D & M did not include externalities in their analysis, nor did they deal with administrative matters of the tax. Sandmo's (1975) pioneering work integrates the theory of optimal taxation with the analysis of the use of indirect taxes (on final goods) to counteract negative external effects, i.e. Pigovian taxes. He concludes that the social damages, generated by the externality-creating commodity, enters the tax formula additively for that commodity and does not enters the tax formula for the clean commodities.

Bovenberg & Ploeg (1994), Bovenberg & Goulder (1996) and Mayeres & Proost (1997) (among others) then extend the theory in different ways. The first couple introduce the concept of *net social Pigovian tax* whereas the second couple include intermediate goods. They find that, in the presence of distortionary taxes, optimal environmental tax rates are generally below the rates suggested by the Pigovian principle, even when revenues from environmental taxes are used to cut distortionary taxes. Moreover, intermediate inputs are not to be taxed for revenue-raising issues, they are to be taxed for their environmental impact solely, this in agreement with Diamond & Mirrlees' desirability of aggregate production efficiency. While the last couple incorporate both externalities of congestion type and income distributions. They show that the results still stand; intermediate goods are not to be equipped with a Ramsey term (i.e. they are not to be taxed for revenue-raising issues) and the additively property indicated by Sandmo is still valid.

The conclusion from the above is that one should not levy any Ramsey tax on intermediate goods, at least if production exhibits constant return to scale. But, if the Ramsey-Boiteux model is employed, where a cost-of-service regulation imposes a somewhat ad hoc budget constraint for the regulated firm, one is confronted by a different problem and, by that, different solutions.

One of the implicit conclusions of Boiteux (1971) is that there are gains to be made by imposing a single budget constraint across as broad a range of public enterprise activities as possible, rather than treating them as separate compartments required to meet individual constraints. This is due to one important caveat with the Boiteux-Ramsey pricing: it is an application of optimal tax theory to only a subset of the economy.

Borger (1997) investigates pricing rules for a budget-constrained and externality-generating public enterprise which provides both final and intermediate goods. The pricing schemes extracted from this model prescribes a somewhat different rule than the previous ones. Here, the intermediate goods are also to be taxed in the Ramsey tradition, that is, the input goods are equipped with a revenue-generating term.

There are several concerns with this model as a paradigm for regulation. Or, as Laffont & Tirole (1993) pointed out: Under linear pricing the firm's fixed cost should not enter the charges to consumers so as not to distort consumption, and therefore it ought to be paid by the government. But the Ramsey-Boiteux model exogenously rules out transfers from the government to the firm, so prices in general exceed marginal costs (which basic economic principles have made clear is efficient). In the end, a tension is uncovered between the benevolent regulator, on the one hand, and that the regulator is not given free rein to to operate transfers to the firm and to obtain efficiency, on the other.

The paper also discusses the welfare effects that environmental tax reforms produce, with emphasis giving to what has become to known as the *double dividend* hypothesis. Which claims that a revenue-neutral *green* tax reform may not only improve the environment, it may also reduce the distortion of the existing tax system. The revenues from the first dividend (environmental taxes) make it practicable feasible to achieve the second dividend (a less distortionary tax system).

Concerning the welfare improvement, it has been shown theoretically and illustrated numerically that returning revenues via labour taxes rather that lump-sum, unambiguously reduces the marginal cost of the policy reform. This suggests that reducing freight taxes are more desirable if recycling is through labour taxes than via lump-sum taxes. This is the so called weak form of the double dividend. Regarding the strong form of the double dividend, which is defined as the effect an environmental tax reform has on the non-environmental welfare cost of the whole tax system, it is very much in doubt even though there may be scope for it if a green tax reform helps eliminate pre-existing inefficiencies in the non-environmental tax system.

One criticism to the models above is that they do not incorporate administrative costs in their framework. Slemrod (1990) distinguishes between the theory of *optimal taxation* and *optimal tax systems*. Optimal taxation is usually restricted to the optimal setting of a given set of tax rates, ignoring other social costs of taxation. When optimising tax systems one has to consider all the elements of the problem.

Slemrod & Yitzhaki (1996) distinguish five components of the cost of taxation:

> **Administrative costs**: The cost of establishing and/or maintaining a tax administration.
>
> **Compliance costs**: The costs imposed on the taxpayer to comply with the law.
>
> **Regular deadweight loss**: The inefficiency caused by the reallocation of activities by taxpayers who switch to non-taxed activities. (Optimal tax theory is focused on minimising the deadweight loss due to substitution between commodities.)
>
> **Excess burden of tax evasion**: The risk borne by taxpayers who are evading.
>
> **Avoidance costs**: The cost incurred by a taxpayer who searches for legal means to reduce tax liability.

The classification is sometimes arbitrary and may depend on the interpretation of the agents' intention, but, nevertheless, the classification could be important in avoiding double counting of social costs. However, even though the issue of tax evasion has been theoretically investigated, relatively little analytical work incorporating tax administration has been done, mainly because administrative issues are hard to analyse with continuous differentiable functions and, therefore, they require complex modelling.

The literature on the subject is few in number and also (frequently) in a positive manner. For use as policy recommendations the theory must come to term with such issues as the choice of tax instruments, the optimal design of enforcement policy, the tax treatment of financial strategies and more generally, must develop a descriptive and normative framework in which to evaluate the issues of tax arbitrage. In this more general framework of optimal tax system (once it is accomplished), optimal taxation could emerges as a special case in which the set of tax instrument is fixed and enforcement of any available instrument is cost-less.

2 Introduction

Marginal cost based pricing for the use of transport infrastructure has been a pillar of the Swedish transport policy for decades and is nowadays also included in the European Common Transport Policy. The development of new technology in the form of mobile communication and global positioning system creates the possibility to charge road users on a very detailed level. In the near future we will probably see a widespread use of different forms of road pricing systems from urban congestion charging, kilometre charges for heavy goods vehicles and pay-as-you-go insurance premiums. Already today we witness introduction of road pricing system based on less advanced technologies as the Swiss Heavy Duty Fee on trucks and the congestion charging schemes in London and the plan for Stockholm. In addition, the policy development in the European railway sector where horizontal disintegration has been imposed, following British and Swedish tracks, has created regulations advocating marginal cost based pricing for the use of tracks but with possibility for mark-ups in certain cases. Sweden has recently adjusted the legislation regulating track charges.

Empowered with these new charging and taxation instruments it would be naïve to believe that policy makers only will consider them for marginal cost based pricing. Other policy objectives could also be legitimate to pursue through these instruments. One such objective is to raise general tax revenues at the lowest cost and another is to promote internal efficiency in various agencies with cost-of-service regulation. Discussions on regional road funds, additional rail charges to recover investment expenditures and new organisational structures in the transport sector can be expected. In many of these cases mark-ups on the marginal cost based price will be necessary. This is already the case for the Maritime and Civil Aviation Administration in Sweden.

However, even with a very brief knowledge of the optimal taxation literature it can be recalled that intermediate goods should not be burdened with financing taxes according to work done in the 70s. If this is the case, no upward deviation from marginal cost based pricing should be acceptable on transport of intermediate goods and services. Mark-ups should not be levied on a large part of the freight transport sector. While this in turn will raise a number of interesting and complicating second-best issues about the pricing of passenger and freight transport the purpose of this paper is to review the optimal taxation literature, up to the most recent contributions, and to conclude on the case for taxation of intermediate goods.

3 Ramsey (1927)

Ramsey (1927) tackled the following question in his 1927 article:

> A given revenue is to be raised by linear taxes on some or all uses of income, the taxes on different uses being possibly at different rates; how should these be adjusted in order that the decrement of utility may be at a minimum?[1]

What he showed was that, under certain premises, in raising revenue by linear taxes on given commodities, the taxes should be such as to diminish in the same proportion the production of each commodity taxed. This leads, among other things, to the today famous inverse elasticity formula.

3.1 The Theory

In a n-commodity economy with *net* utility $u = F(x_1, \ldots, x_n)$ he starts with postulating an equilibrium without taxes and call these values for $\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n$ or collectively the point P. Then at P one have:

$$\begin{aligned} \frac{\partial u}{\partial x_r} &= 0, \quad r = 1, 2, \ldots, n \quad \text{and that} \\ d^2 u &= \sum \sum \frac{\partial^2 u}{\partial x_r \partial x_s} dx_r dx_s \quad \text{is a negative definitite form.} \end{aligned}$$

As can be seen, and will be seen, he focuses on differentials, that is, a revenue-neutral change that leaves u unchanged.

Suppose taxes at rates $\lambda_1, \ldots, \lambda_n$ per unit (whose marginal utility is unity) is levied on the commodities. Then the new equilibrium is determined by $\partial u / \partial x_r = \lambda_r$, for $r = 1, \ldots, n$.

The problem is then: given R (revenue) how should the λ's be chosen in order to maximise u, mathematically:

$$\max_{\lambda_1, \ldots, \lambda_n} u \; : \; \sum \lambda_r x_r = R. \tag{1}$$

This is equivalent to:

$$0 = du = \sum \lambda_r dx_r \quad \text{for any values of } dx_r$$

subject to:

$$0 = dR = \sum \lambda_r dx_r + \sum \sum x_s \frac{\partial \lambda_s}{\partial x_r} dx_r.$$

The solution to this problem is:

$$\frac{\lambda_1}{\sum x_s \frac{\partial \lambda_s}{\partial x_1}} = \cdots = \frac{\lambda_n}{\sum x_s \frac{\partial \lambda_s}{\partial x_n}} = \frac{R}{\sum \sum \frac{\partial \lambda_s}{\partial x_1} x_r x_s} = -\theta \; (say). \tag{2}$$

[1] This problem was in fact suggested to Ramsey by A. C. Pigou

Equation (2) determines values of the x's which are critical for u. Ramsey shows that if R is small enough they will determine an unique solution which tends to P as $R \to 0$ and that this solution makes u a true maximum. He also shows that $\theta > 0$.

Now, suppose that R and the λ's can be regarded as infinitesimal; then putting $\lambda_r = \sum_s (\partial\lambda_r/\partial x_s)dx_s$ equation (2) gives:

$$\frac{dx_1}{x_1} = \frac{dx_2}{x_2} = \cdots = \frac{dx_n}{x_n} = -\,\theta < 0. \tag{3}$$

That is to say, *the production of each commodity should be diminished in the same proportion.*

Ramsey then extends these results to the case of a given revenue to be raised by taxing certain commodities only.

If the quantities of the commodities to be taxed is denoted by x and those not to be taxed by y ($\mu_r = \partial u/\partial y_r = 0$ is the tax per unit on the y's) the extension of equation 2 is:

$$\cdots = \frac{\lambda_r}{\sum_{s=1}^{n}\left(\frac{\partial\lambda_s}{\partial x_r} + \sum_{t=1}^{m}\chi_{tr}\frac{\partial\lambda_s}{\partial y_t}\right)} = \cdots \tag{4}$$

where χ_{tr} solves:

$$\frac{\partial\mu_t}{\partial x_r} + \sum_{u=1}^{m}\frac{\partial\mu_t}{\partial y_u}\chi_{ur} = 0 \qquad \begin{cases} r = 1, 2, \ldots, n \\ u = 1, 2, \ldots, m \end{cases}$$

These equations give a maximum of u with the same sort of limitations as equation (2) do.

As before, suppose that the λ's are infinitesimal, then by letting the λ's again be split in differentials of the taxed and the untaxed commodities one can again show that the solution to equation (4) are the same as equation 3, i.e. the taxes should be such as to reduce in the same proportion the production of each *taxed* commodity.

Ramsey then assumes that the utility can be described by a non-homogeneous quadratic function of the x's, or that the λ's are linear. However, it is not necessary to suppose the utility function to be quadratic for all values of the variables; one need only suppose it for a certain range of values round the point P, such that there is no question of imposing taxes large enough to move the production point, i.e. the x's, outside this range. If the commodities are independent, this is the same as the taxes are small enough to treat the supply and demand curves as straight lines.

Letting the utility be $u = Const. + \sum a_r x_r + \sum\sum \beta_{rs} x_r x_s$ and regard the x's as rectangular Cartesian coordinates, Ramsey then deduce, by an geometrical analysis, the same as above, that is to say; the taxes should be such as to diminish the production of all commodities in the same proportion.

This is now valid not merely for an infinitesimal revenue but for any revenue which it is possible to raise at all. Moreover, the maximum revenue will be obtained by diminishing the production of each commodity to one-half of its previous amount, i.e. to the production point $\frac{1}{2}\bar{x}_1, \frac{1}{2}\bar{x}_2, \dots, \frac{1}{2}\bar{x}_n$. It is also shown that the taxes at the optimum (were the revenue is maximised) would be $\lambda_r = \frac{1}{2}\, a_r$.

He then finish the theory by consider the more general problem:

> A given revenue is to be raised by means of fixed taxes $\mu_1, \dots, \mu_m$ on m commodities and by taxes to be chosen at discretion on the remainder. How should they be chosen in order that utility may be a maximum?

Again he gives a geometrical solution which says that: the desired production point satisfies:

$$\frac{x_{m+1}}{\bar{x}_{m+1}} = \frac{x_{m+2}}{\bar{x}_{m+2}} = \cdots = \frac{x_n}{\bar{x}_n},$$

i.e, the whole system of taxes must be such as to reduce in the same proportion the production of the commodities taxed at discretion.

3.2 Special Cases

In this section the results above is explained in certain special cases. First, suppose that all the commodities are independent and have their own supply and demand equations and the tax *ad valorem* (reckoned on the price got by the producer) on the rth commodity is μ_r, then $\lambda_r = \mu_r p_r$ where p is the producer price. If the elasticities of demand and supply are denoted by ϵ_r and ρ_r the following tax schedule can be derived (provided the revenue is small enough as discussed above):

$$\mu_r = \frac{\left(\frac{1}{\rho_r} + \frac{1}{\epsilon_r}\right)\theta}{1 - \frac{\theta}{\epsilon_r}}. \tag{5}$$

For infinitesimal taxes, θ is infinitesimal and

$$\frac{\mu_1}{\frac{1}{\rho_1} + \frac{1}{\epsilon_1}} = \frac{\mu_2}{\frac{1}{\rho_2} + \frac{1}{\epsilon_2}} = \cdots = \frac{\mu_n}{\frac{1}{\rho_n} + \frac{1}{\epsilon_n}}. \tag{6}$$

That is, the tax *ad valorem* on each commodity should be proportional to the sum of the reciprocals of its supply and demand elasticities.

Three things can be seen from equation (6). First, the same rule applies if the revenue is to be collected from certain commodities only, which have supply and demand schedules independent of each other and all other commodities, even when the other commodities are not independent of one another. Second, the rule does not justify any subsidies since, in a stable equilibrium, although ρ_r^{-1} may be negative, $\rho_r^{-1} + \epsilon_r^{-1}$ must be positive. And third, if any one commodity is absolutely inelastic, either for supply or demand, the whole revenue should be collected from it. If there is several such commodities the whole revenue should be collected from them and it does not matter in what proportions.

Next, the case in which all the commodities have independent demand schedules but are complete substitutes for supply is investigated. The process brings about equation (5) and (6) again but with all the ρ_r's changed to ρ. In this case one see that if the supply of labour is fixed, i.e. absolutely inelastic $\rho \to \infty$, the taxes should be at the same ad valorem rate on all commodities.

If some commodities only are to be taxed, as in the end of the first section, one gets (when working with infinitesimal revenue) as before, that between two commodities, the one with the least elasticity of demand is to be taxed the most but that if the supply of labour is absolutely inelastic all the commodities should be taxed equally.

In the appendix Ramsey also derive the same solutions, i.e. equations (5) and (6), to the more general problem in which the State wishes to raise revenue for two purposes; first, as before, a fixed money revenue which is transferred to rentiers or otherwise without effect on the demand schedules; and secondly, an additional revenue sufficient to purchase fixed quantities of each commodity.

The theory could be useful in the following cases; first, if a commodity is produced by several different methods or in several different places between which there is no mobility of resources, it is shown that it will be advantageous to discriminate between them and tax most the source of supply which is least elastic. Second, if several commodities which are independent for demand require precisely the same resources for their production, the tax should be highest where the elasticity of demand is the least. Third, in taxing commodities which are rivals for demand, the rule to be observed is that the taxes should be such as to leave unaltered the proportions in which they are consumed.

Ramsey also emphasises in conclusion that the results about infinitesimal taxes can only claim to be approximately true for small taxes, how small depending on the data which are not obtainable. It is perfectly possible that a tax of 500 % on whisky could for the present purpose be regarded as small. The unknown factors are the curvatures of the supply and demand curves; if these are zero the results will be true for any revenue whatever but the greater the curvature, the narrower the range of “small” taxes.

4 Diamond and Mirrlees (1971)

Diamond and Mirrlees' two articles from 1971 discuss optimal taxation in the absence of externalities. The first article states the desirability of aggregate production efficiency in many circumstances provided that taxes are set at the optimal level. The second examines the optimal tax structure at the optimal level. Their conclusion is that production efficiency is desirable even though a full Pareto optimum not can be achieved. In the optimum position, the presence of commodity taxes implies that marginal rate of substitution are not equal to marginal rate of transformation.

It is a second-best solution since lump-sum taxation is regard as not feasible, i.e. the income distribution will not be the best that can be achieved. Yet, the presence of optimal commodity taxes is shown to imply the desirability of aggregate production efficiency.

4.1 Part I (Production Efficiency)

In an economy without lump-sum transfers, but with linear taxes or subsidies on each commodity which can be adjusted independently, it is shown that any second-best optimum of a Paretian social welfare function entails efficient production. That is to say, the marginal rate of transformation in public production must equal the marginal rate of transformation in private production and thus aggregate production efficiency.

If the prices faced by private producers is denoted by p_i, the consumers' prices become $q_i = p_i + t_i$ where t_i represents the indirect tax (faced by consumers) on commodity x_i. Then, if $v = u(x(q))$ is the indirect utility function and $f(y)$ and $g(z)$ the private respective public production function and the conditions that all markets clear (Walras' law) are $x_i(q) = y_i + z_i$ (y_i is then private output and z_i public), the Lagrangian of the problem can be formulated as (after some manipulation):

$$L = v(q) - \lambda\left(x_1(q) - f(x_2 - z_2, \ldots, x_n - z_n) - g(z_2, \ldots, z_n)\right)$$

where $y_1 = f(y_2, \ldots, y_n)$ and $z_1 = g(z_2, \ldots, z_n)$. (As can be seen, the constraints have been reduced to $x_1(q) = y_1 + z_1$, this without loss of generality.) Differentiating L with respect to z_k one has:

$$\lambda(f_k - g_k) = 0, \qquad k = 2, \ldots, n.$$

That is: *aggregate production efficiency.*

The optimal tax structure which ensures the efficiency has the following appearance:

$$\frac{\partial v}{\partial q_k} = \lambda \sum_{i=1}^{n} p_i \frac{\partial x_i}{\partial q_k} = -\lambda \frac{\partial}{\partial t_k}\left(\sum_{i=1}^{n} t_i x_i\right), \quad k = 1, 2, \ldots, n \tag{7}$$

where λ reflects the change in welfare from allowing a government deficit financed from some outside source. It says that the impact of a price rise, of commodity k, on social welfare (first term) is proportional to the cost of meeting the change in demand induced by the price rise (second term). Alternatively (third term), since $q_i = p_i + t_i$, the impact of a tax increase on social welfare is proportional to the induced change in tax revenue (all calculated at fixed producer prices). There is also a long discussion why $\lambda \neq 0$. The above and below equation is calculated for an one-consumer economy but the analysis carries over to the many-consumer economy which will be seen in Part II below.

If the welfare function is individualistic the (above) first-order conditions become:

$$x_k = \frac{\lambda}{\alpha} \frac{\partial(\sum t_i x_i)}{\partial t_k}, \quad k = 1, 2, \ldots, n \tag{8}$$

where α is the marginal utility of income. These equations say that: for all commodities the ratio of marginal tax revenue from an increase in the tax on that commodity to the quantity of the commodity is a constant. (Here it is assumed that $\alpha \neq 0$.) The ratio λ/α then gives the marginal cost of raising revenue. This first-order condition shows the information needed to test whether a tax structure is optimal.

Introducing further taxes do not alter the efficiency argument, the optimal production must still be on the production frontier. They argue that whatever the class of possible tax systems, if all possible commodity taxes are available to the government, then, in general, and certainly if a poll subsidy is possible, optimal production is weakly efficient, i.e. that the production plan is on the production frontier. The conclusion is not to be expected valid if there were constraints on the possibilities of commodity taxation, or more generally, on the possible relationship between producer prices and consumer demand, e.g. the presence of pure profits.

They also prove rigorously the existence of an optimum and the efficiency of optimal production where they assume a finite set of consumers with continuous single-valued demand functions, e.g. strictly convex consumers' preferences, and continuous demand functions.[2]

The model leaves no scope for intermediate good taxation (in a competitive economy producing with constant returns to scale) since, in the absence of profits, taxation on intermediate goods must be reflected in changes in final good prices. Therefore, the revenue could have been collected by final good taxation, causing no greater change in final good prices and avoiding production inefficiency.

[2] Hammond (2000) have extended the analyses to a continuum of consumers with original assumptions greatly relaxed such as non-linear pricing for consumers and individual non-convexities.

However, when there is decreasing returns to scale Dasgupta & Stiglitz (1972) conclude that production efficiency is only desirable if the range of government instrument is sufficiently great, in effect, only if profits can be taxed at appropriate rates. Myles (1989) explores how the removal of the perfect competition assumption effect the production efficiency lemma. The major result is that, if there is imperfect competition, there is a strong case for including intermediate goods in the tax system. The only general exception for this rule appears to be the case of Leontief technology.

4.2 Part II (Tax Rules)

In this part the structure of taxation is explored in more detail, starting out with an economy with one consumer with an individualistic welfare function. First, changes in demand due to a tax change are examined. This is done by assuming that both price derivatives of demand and production prices are constant, i.e. $x(q)$ is linear. The actual changes in demand for good k induced by a tax structure are:

$$\frac{\sum_i \frac{\partial x_k}{\partial q_i} t_i}{x_k} = \frac{\alpha}{\lambda} - 1 - \sum_i t_i \frac{\partial x_i}{\partial I} - \frac{\partial x_k}{\partial I} \frac{\sum_i t_i x_i}{x_k}, \tag{9}$$

where the income derivatives, $\partial x_j / \partial I$, $j \in \{i, k\}$, come from the Slutsky equation. The first three terms (on the right hand side) are independent of k which is the good investigated. So by looking at the fourth term one sees that the changes differ from proportionality with a larger than average percentage fall in demand for goods with a large income derivative.[3]

In the case of a three-good economy they then obtain an expression for the relative (linear) tax rates when one good is untaxed, e.g. labour. The conclusion is that; the tax rate is proportionally greater for the good with the smaller cross-elasticity of compensated demand with the price of labour, the untaxed good. The economic interpretation of this is that since labour (or leisure) is untaxed, one can tax it indirectly by taxing the commodities that are substitute for labour (or complementary with leisure).

If the ordinary demand elasticities, ϵ_{ik}, are used in the optimal tax formula, eq (7) above, it can be written as:

$$\frac{q_k}{p_k} = -\frac{\lambda}{\alpha} \sum_i \frac{p_i x_i}{p_k x_k} \epsilon_{ik}, \tag{10}$$

again assuming individualistic welfare functions. If there exists a good whose price does not affect other demands the equation simplifies to:

[3] Sandmo (1976) has an intuitive interpretation of this, namely: Tax increases have both income and substitution effects, and the income effects are analogous to the changes that would have resulted if the revenue had been raised by lump-sum taxes. Since the latter effects are non-distortionary, so are the pure income effects and one should therefore reduce the demand most for the commodities where these effects dominate.

$$\frac{q_k}{p_k} = \frac{\lambda}{\alpha}.$$

Thus, since $q_k p_k^{-1}$ equals one plus the percentage tax rate, the optimal tax rate on such a good gives the cost to society of raising the marginal dollar of tax.

To pursue the analysis further, and to incorporate many consumers, the individualistic welfare function now has each individual's utility function as argument, i.e. $V(q) = W(v^1(q), v^2(q), ..., v^H(q))$. Then the corresponding equation to eq. (7) and (8) (together), for the many-consumers case, is:

$$-\frac{\partial V}{\partial q_k} = \sum_h \beta^h x_k^h = \lambda \frac{\partial T}{\partial t_k}, \quad \text{where} \quad T = \sum t_i x_i.$$

Since α^h is the marginal utility of consumer h, $\beta^h = \frac{\partial W}{\partial u^h}\alpha^h$ becomes the increase in social welfare from a unit increase in the income of consumer h. The necessary condition for optimal taxation makes $\partial V / \partial q_k$ proportional to the marginal contribution to tax revenue from raising the tax on good k. The derivative is, still, evaluated at constant producer prices, i.e. on the basis of consumer excess demand function alone. It can also be written as:

$$\sum_h \beta^h x_k^h = -\lambda \sum_i p_i \frac{\partial x_i}{\partial q_k}.$$

In an example where each consumer has a Cobb-Douglas utility function and assuming an individualistic welfare function the optimal tax rate is determined. In this example, if the social marginal utilities, β^h, are independent of taxation, e.g. if $W = \sum_h v^h$, the optimal tax rates can be read off at once. It is noticed that, although each household's social marginal utility of income is unaffected by taxation, it is desirable to have taxation in general. Because, if household's with relative low social marginal utility of income predominate among purchasers of a commodity, that commodity should be relatively highly taxed. Although such taxation does nothing to bring social marginal utilities of income closer together, it does increase total welfare. (If, for example, the welfare function treats all individuals symmetrically and if there is diminishing social marginal utility with income, then there is greater taxation on goods purchased more heavily by the rich.)

The corresponding equation to eq. (10), with Cobb-Douglas utility functions, is:

$$\frac{q_k}{p_k} = \lambda \frac{\sum_h x_k^h}{\sum_h \beta^h x_k^h}, \quad k = 2, 3, \ldots, n.$$

From this equation one can identify two cases where optimal taxation is proportional. If the social marginal utility of income is the same for everyone, i.e. $\beta^h = \beta$ for all h, then it reduces to $q_k p_k^{-1} = \lambda/\beta$ as above. In this case there is no welfare gain to be achieved by redistributing income, and so no need to tax differently, on average, the expenditures of different individuals.

The second case leading to proportional taxation occurs when demand vectors are proportional for all individuals. When all individuals demanding goods in the same proportions, it is impossible to redistribute income by commodity taxation implying that the tax structure assumes the form it had in a one-consumer economy.

Analysis of the change in demand is also carried out, the equivalence to eq. (9) for the many-consumer case is:

$$\frac{\sum_h \sum_i t_i \frac{\partial x_i^h}{\partial q_i}}{\sum_h x_k^h} = \frac{1}{\lambda} \frac{\sum_h \beta^h x_k^h}{\sum_h x_k^h} - 1 + \frac{\sum_h \left(\sum_i t_i \frac{\partial x_i^h}{\partial I}\right) x_k^h}{\sum_h x_k^h} - \frac{\sum_h \left(\sum_i t_i x_i^h\right) \frac{\partial x_k^h}{\partial I}}{\sum_h x_k^h}.$$

With constant producer prices the equation gives the change in demand as a result of taxation for a good with constant price-derivatives of the demand function, i.e. for small taxes. Considering two such goods, one can see that the percentage decrease in demand, LHS in the equation, is greater for the good the demand for which is concentrated among:

> Term 1 in the RHS above: Individuals with low social marginal utility of income, λ.
>
> Term 3: Individuals with small decreases in taxes paid with a decrease in income.
>
> Term 4: Individuals for whom the product of the income derivative of demand for good k and taxes paid are large.

Then they include income taxation in the model and conclude; at the optimum, for any two different kinds of change in income tax structure, the social-marginal-utility changes in taxation (consumer behaviour held constant) are proportional to the changes in total tax revenue (both income and commodity tax revenue, calculated at fixed producer prices, with consumer behaviour responding to the price change).

Following a discussion of public consumption, the Optimal Taxation Theorem is presented formally. The section provides a rigorous analysis of conditions under which the tax formula, eq. (7) (for the many consumer case), are indeed necessary conditions for an optimum and also provides economically meaningful assumptions that ensure the Lagrange multipliers validity. This, under the assumptions that the welfare and the demand functions are continuously differentiable; and that the production set is convex and has a non-empty interior. They also discuss some extensions when the production set is not convex and some uniqueness problems that may arise.

The article can be viewed as a major generalisation and extension of the Ramsey formulation. The text gives great insight into policy problems, even though it omits administration costs, as well as tax evasion, for the tax-structure derived. The standard constant-return-to-scale, price-taking and profit-maximising behaviour are also assumed in private production. Pure profits (or losses) associated with the violation of these assumption imply that private production decisions directly influence social welfare by affecting household incomes. In such a case, it would presumably be desirable to add profit tax to the set of policy instruments. Nevertheless, aggregate production efficiency would no longer be desirable in general; although it may possible to get close to the optimum with efficient production if pure profits are small.

5 Externalities and Intermediate Goods

D & M did not include externalities in their analysis, nor did they deal with administrative matters of the tax. Sandmo's (1975) pioneering work integrates the theory of optimal taxation with the analysis of the use of indirect taxes to counteract negative external effects, i.e. Pigovian taxes. He also considers the problem of distributional impact of taxation in the special case of individuals with identical preferences and a utilitarian social welfare function.

Bovenberg & Ploeg (1994) extend the above analysis in some ways, e.g. they consider the impact of environmental externalities not only on the optimal tax structure but also on the optimal level and composition of public spending. In doing so they integrate environmental externalities and the optimal provision of the public good of the natural environment.

Bovenberg & Goulder (1996) then extend the analysis by considering pollution taxes imposed on intermediate inputs. They also investigate second-best optimal environmental taxes numerically and with the help of this numerical approach they also examine optimal environmental tax policies in the presence of (realistic) policy constraints.

Mayeres & Proost (1997) examine externalities whose level are determined by the total use of some commodity and for which the externality level itself affects the private use of certain commodities. i.e. there are feedback effects. Intermediate goods are now represented by road (freight) transport.

5.1 Sandmo (1975)

Sandmo uses a simple model in which there are n-consumers and $m+1$ consumer goods and where consumption of good m, x_m, creates a negative externality which is a function of the total consumption of that good, $X_m = \sum x_m = nx_m$. This externality enters the utility function as its $(m+1)$th argument. The first-best solution gives the familiar result that the producer and consumer prices should be equated for the first $m-1$ goods and for the externality-creating good, good x_m, the optimal tax rate should reflects its marginal social damage which, in this analysis, is the sum of the marginal rates of substitution between good m as a private good (x_m) and as a public good (X_m) or, mathematically:

$$\theta_m = \frac{t_m}{q_m} = -n\frac{u_{m+1}}{u_m},$$

where u_i is the marginal utility of good i and n in the formula represents the n consumers.

But since the government needs other, distortionary taxes, in order to satisfy its revenue requirements, the second-best solution to the same problem becomes a bit more complex. The main result is that the Pigovian principle holds in a modified form in this case as well.

The optimisation problem can be formulated as maximisation of the sum of the indirect utility functions with respect to consumer prices subject to the budget constraint $\sum t_i x_i = n\sum(q_i - p_i)x_i = T$. The Lagrangian becomes:

$$L = nv(q) - \beta\left[n\sum(q_i - p_i)x_i - T\right].$$

Sandmo concludes that the optimal tax structure is now characterised by what might be called an additivity property; the marginal social damage of commodity m enters the tax formula for that commodity additively, and does not enter the tax formulas for the other commodities, regardless of the pattern of complementarity and substitutability. Thus, the fact that a commodity involves a negative externality is not in itself an argument for taxing other commodities which are complementary with it, nor for substitutes. The structure has the following mathematical form:

$$\begin{aligned} \theta_k &= (1-\mu)\left[-\frac{1}{q_k}\frac{\sum_{i=1}^m x_i J_{ik}}{J}\right], \quad \text{for } k \neq m \\ \theta_m &= (1-\mu)\left[-\frac{1}{q_m}\frac{\sum_{i=1}^m x_i J_{im}}{J}\right] + \mu\left[-n\frac{u_{m+1}}{u_m}\right] \end{aligned} \tag{11}$$

where J_{ik} is the cofactor of the Jacobian matrix of the demand functions for the taxed goods and J the determinant of that Jacobian. The μ can be interpreted as the marginal rate of substitution between private and public income;[4] the higher μ is, the higher the marginal value of private income compared with public income, and the lower the tax requirements, given that this is itself derived from an underlying optimisation criterion. What

[4] μ is the inverse of what is often referred to as the marginal cost of public funds, which will be discussed further below.

can also be seen is that with increasing μ, the proportionality factor of the efficiency terms in the formula decrease, and the marginal social damages comes to dominate the tax on good m. If $\mu > 1$, the efficiency terms become formulas for optimal subsidies instead, and if $\mu = 1$, one is back at the first-best solution. This is the fortunate case where the Pigovian tax alone happens to satisfy the tax requirement exactly.

Consider the case of independent demands, that is $\partial x_j / \partial q_k = 0$ for $j \neq k$, then the formula above reduces to:

$$\begin{aligned} \theta_k &= (1-\mu)\left[-\frac{1}{\epsilon_k}\right], \quad \text{for } k \neq m \\ \theta_m &= (1-\mu)\left[-\frac{1}{\epsilon_m}\right] + \mu\left[-n\frac{u_{m+1}}{u_m}\right]. \end{aligned} \tag{12}$$

The top equation is the familiar inverse elasticity formula, originally derived by Ramsey (eq (6) on page 9), which says that the highest tax should be levied on commodities where the elasticity of demand is the lowest. The bottom equation shows that the optimal tax rate for the externality-creating commodity is a weighted average of the inverse elasticity and the marginal social damage.

There could be no distributional problem in the above analyse since every individual was alike. If, instead, one let them have unequal productivities, but the same preferences, the result becomes a bit different, but the striking factor is that the additive property carries over to this more general case. It is still true that the marginal social damages is only an argument in the tax formula for good m. If μ_j is defined as the marginal rate of substitution between private and public income for consumer j, the corresponding equations for eq. 11 now becomes a sum over all μ_j.

In the case with independent demands, the analog to eq. 12 now becomes:

$$\begin{aligned} \theta_k &= \frac{\sum_j (1-\mu_j)x_{kj}}{\sum_j x_{kj}}\left[-\frac{1}{\epsilon_k}\right], \quad \text{for } k \neq m \\ \theta_m &= \frac{\sum_j (1-\mu_j)x_{kj}}{\sum_j x_{kj}}\left[-\frac{1}{\epsilon_m}\right] + \sum_j \mu_j\left[-n\frac{u_{m+1}}{u_m}\right]. \end{aligned}$$

The proportionality factor, or the *distributional characteristic* of good k, has now become a weighted average across individuals of the factor $(1-\mu_j)$, the weights being in each case the amount of the commodity in question consumed by individual j. $(1-\mu_j)$ varies positively with the level of income, being low for low-income individuals and high for high-income individuals. Thus, this proportional factor takes a low value if the consumption of commodity k is concentrated among low-income individuals and a high value if it is mainly consumed by high-income individuals. Which, by itself, comes from the fact that a utilitarian social welfare function has been used.

Distributional factors also enters in the social damages term, for the externality-generating commodity, since each individual's marginal rate of substitution is weighted by the factor μ_j, which varies negatively with income. Thus, the social damages term will be high if those who suffer the most from the externality tend to have low incomes and low if they are concentrated among the high-income groups.

The paper is not designed as a practical guide to the use of Pigovian taxes, the models are too stylised for that. The purpose is more to show that the Pigovian taxation principle can be validated as part of a more comprehensive system of indirect taxation, and the author demonstrated that it holds in a modified form even when distributional considerations enter as correctives to the efficiency principles of taxation. The fact that the social damages, generated by the externality creating commodity, enters the tax formula (additively) for that commodity does not mean that it enters the private commodities, i.e. the clean commodities.

This result has obvious relevance for economic policy and is not evident from the viewpoint of the second-best theory. Dixit (1985) has referred to Sandmo's result as an instance of the more general *principle of targeting.* The idea is that one should best counter a distortion by the tax instrument that acts on it directly (i.e. at the relevant margin).

5.2 Bovenberg & Ploeg (1994)

In Bovenberg & Ploeg's (1994) article, the representative consumer derives utility from consumption of clean and dirty private goods, leisure, clean and dirty public goods and the quality of the environment, i.e. $U = u(C, D, V, X, Y, E)$. After deriving optimal taxation in a first-best world and a second-best world without externalities (i.e. Ramsey tax schemes), the authors derive optimal labour and dirt taxes when environmental externalities are present in consumption, i.e. $E = e(nD, Y, A)$ where n is the number of private agents and A stands for the governments abatement activities.[5] Labour and dirt taxes are employed not only to internalise environmental externalities but also to finance public spending.

B & P derive a similar result as Sandmo with the use of compensated demand elasticities. The optimal tax becomes the sum of the Ramsey and externality-correcting (Pigovian) terms, in accordance with Sandmo. If ϵ_{ik} is defined as the compensated elasticity of demand for commodity i with respect to the price of commodity k, μ the marginal disutility of financing public spending and λ' the marginal *social* [6] utility of private income, the

[5]Since a labour tax is equivalent to a uniform tax on clean and dirty private production and a dirt tax is the natural candidate for inducing private agents to pollute less, one can assume that the clean good is untaxed.

[6]λ' may exceed the marginal *private* utility of income λ as it takes account of the

optimal dirt (and labour) tax becomes:

$$\begin{aligned}\theta_D &= \frac{t_D}{1-t_D} = \left(\frac{\epsilon_{CL}-\epsilon_{DL}}{\epsilon_{CD}-\epsilon_{DD}}\right)\theta_L + \theta_{DP}, \quad \theta_{DP} = \frac{t_{DP}}{1-t_D} \\ \theta_L &= \frac{t_L}{1-t_L} = \left(\frac{\epsilon_{LD}-\epsilon_{DD}}{\epsilon_{LD}\epsilon_{DL}-\epsilon_{DD}\epsilon_{LL}}\right)\left(\frac{\mu-\lambda'}{\mu}\right), \end{aligned} \tag{13}$$

where t_{DP} stands for the externality correcting tax. It can be written as:

$$t_{DP} = \left(\frac{-ne_{ND}u'_E}{u_C}\right)\left(\frac{1}{\eta}\right) \tag{14}$$

where u'_E stands for the marginal *social* [7] utility of the environment, e_{ND} the environmental damage per unit of dirty private consumption ($e_{ND} < 0$), u_C is the marginal utility of clean private products and $\eta = \mu/\lambda$ the marginal costs of funds (i.e. η is equal to Sandmo's μ^{-1} as was said in footnote on page 17).

What can be seen from the above equations is that the government should levy a dirt tax (on top of the labour tax) the sign of which (the Ramsey term) depends on the cross-elasticities with leisure. The Ramsey tax is positive if clean goods are better substitutes for leisure than dirty goods are ($\epsilon_{CL} > \epsilon_{DL}$). In that case, dirty goods are the relative complement to leisure. Accordingly, it is optimal to levy an additional tax on the product that is most complementary to leisure.

Even if the compensated elasticities of the demand for clean and dirty goods with respect to the price of leisure are identical, i.e. $\epsilon_{CL} = \epsilon_{DL}$, a zero dirt tax is not optimal due to the fact of a separate non-distortionary (or externality-correcting) term which corrects for the environmental externality. If $\eta = 1$ and $u'_E = u_E$ the non-distortionary component of the dirt tax, t_{DP}, coincide with the Pigovian tax in the first-best world, i.e. where lump-sum transfers is an option.

If the marginal cost of funds exceeds unity ($\eta > 1$), the optimal non-distortionary component falls below the Pigovian tax (i.e. the marginal social damage of pollution as measured by the sum of the marginal rates of substitutions between environmental quality and clean private consumption, eq 14). The reason is that the optimal non-distortionary tax measures the social costs of pollution in terms of public rather than private income. In particular, the optimal environmental tax equates the social costs of pollution to the social benefit of the public goods that can be financed by the additional revenue generated by the pollution tax. This implies that each

increased tax revenues resulting from additional private expenditures.

[7] u'_E accounts not only for the direct impact of the environment on utility ($u_E > 0$), but also for the indirect effects of an improved environment on the tax base. The two measures coincides if environmental quality is weakly separable from the other arguments in social utility.

unit of pollution does not have to yield as much public revenue to offset the environmental damage if this revenue becomes more valuable as measured by a higher marginal cost of public funds. Intuitively, the government employs the tax system to simultaneously accomplish two objectives; first, to raise public revenues to finance public goods (other than the environment), and, second, to internalise pollution externalities, thereby protecting the public good of the natural environment. If public revenues become scarcer, as indicated by a higher marginal cost of public funds (η), the optimal tax system focuses more on generating revenues and less on internalising pollution externalities.

In contrast to e.g. Sandmo, this definition of the non-distortionary dirt tax incorporates a second factor that may cause the Pigovian tax to deviate from the sum of the marginal rates of substitution. In particular, the environmental quality may directly impact the consumption of taxed commodities. For example, if labour supply is taxed and an improved environment induces people to enjoy more leisure and work less, the social value of environmental protection is reduced and the optimal environmental tax falls. In principle it is possible, albeit unlikely, that the Pigovian component of the dirt tax is negative, namely if tax rates are high and if a better environmental quality substantially reduces the demands for taxed goods.[8]

The authors also deal with, among other things, the subject of choice between public and private goods. They recognise that the marginal rate of transformation between private and public goods no longer corresponds to the sum of the marginal rates of substitution between private and public goods. One of the reasons is that, if public goods are complementary to taxed commodities, rising public spending alleviates the excess burden of distortionary taxation by boosting the consumption of taxed commodities. For example, the construction of public highways between suburbs and cities may induce some agents to work more and, therefore, pay more tax on their labour income. Moreover, they may buy more heavily taxed commodities, such as petrol and cars. Public libraries work the other way around, and eroding the tax base, since they encourage agents to enjoy more leisure.

However, it is only the dirt tax net of the distortionary tax term ($t_D - t_{DP}$) that enters the formula. This since, if public highways are complementary to the consumption of taxed gasoline, the construction of highways boosts gasoline consumption. Whereas the additional consumption of gasoline boosts tax collections, it also pollutes the environment. The social cost

[8]Ng (1980) explores the sign of the optimal pollution tax. He finds that, in the presence of environmental externalities, the pollution tax is typically positive. However, if the revenue requirement is small and falls short of the revenues from the Pigovian tax, the optimal pollution tax may actually be negative. In this counterintuitive case, a lower consumption wage must be very effective in reducing dirty consumption, compared to a higher consumption price for dirty consumption. Hence, the combination of a wage tax and a subsidy on dirty consumption reduces pollution.

of the environmental damage is measured by the additional revenue collected from the non-distortionary (i.e. externality-correcting) component of the gasoline tax. Hence, only to the extent that the revenues from the Ramsey (distortionary) component of the gasoline tax rise, does the widening of the tax base yield a net social benefit, thereby reducing the social cost of financing highways.

The focus of the article is not what the present paper focuses on, nonetheless, they have extended both Sandmo's (1975) and Ng's (1980) articles and provided more insight on the optimal tax theory in areas such as; how the well-known Ramsey formula for optimal taxes is altered when one incorporates consumption commodities that generates externalities. They also emphasise Sandmo's finding, the principle of targeting as Dixit (1985) named it, that is; the presence of the externality does not change the structure of second-best taxes on private (non externality creating) goods or income. That is to say, the model shows, with linear taxes, that the formula for the labour income tax must remain unaffected by the tax on the externality-creating good as well as it stresses Sandmo's additively property of the dirt tax.

5.3 Bovenberg & Goulder (1996)

Bovenberg & Ploeg's (1994) model is now extended by incorporation of intermediate inputs. Output derives from a constant-return-to-scale production $F(L, x_C, x_D)$ with inputs of labour, L, and clean and dirty products, x_C respective x_D. Output can be devoted to public consumption, G, to a clean or dirty consumption good, C_C and C_D. Hence, the commodity market equilibrium is given by: $F(L, x_C, x_D) = G + x_C + x_D + C_C + C_D$ (units are normalised so that the constant rates of transformation between the produced commodities are unity).

The representative household maximises utility $U(C_C, C_D, l, G, Q) = u(N(H(C_C, C_D), l), G, Q)$. Private utility $N(\cdot)$ is homothetic, while commodity consumption $H(\cdot)$ is weakly separable from leisure, l.[9] In addition, private utility is weakly separable from public consumption, G, and environmental quality, Q.[10] The environmental quality is directly related to quantity used of dirty intermediate and dirty consumption goods; thus, $Q = q(x_D, C_D)$, with negative derivatives. The household faces the budget constraint $C_C + (1 + t_D^C)C_D = (1 + t_L)wL$, where the t:s are taxes and the government budget constraint is $G = t_C^x x_C + t_D^x x_D + t_D^C C_D + t_L wL$.

[9]This assumption, the weakly separable one, is not an intuitive or harmless one. But the authors of the article have made it since it matches the numerical model which they use.

[10]This imply that the compensated elasticities ϵ_{CL} and ϵ_{DL} are identical, see e.g. equation (13) on page 20, which, in turn, makes the assumption debatable.

The government chooses values of its four tax instruments t_L, t_D^C, t_C^x and t_D^x to maximise:

$$u[V(w_N, t_D^C), G.q(x_D, C_D)] + \mu[t_C^x x_C + t_D^x x_D + t_D^C C_D + t_L wL - G]$$

where V represents indirect private utility and μ denotes the marginal utility associated with the public goods consumption made possible by one additional unit of public revenue.

When deriving optimal tax rates, the analysis reveals that the clean intermediate input should not be taxed, i.e. $t_C^x = 0$. This is in accordance with the well-known result of Diamond & Mirrlees. It demonstrates that, if production exhibits constant returns to scale, an optimal tax system should not distort production.[11]

The optimal tax on the dirty intermediate input is:

$$t_D^x = \left[\frac{\frac{\partial U}{\partial Q}\left(-\frac{\partial q}{\partial x_D}\right)}{\frac{\partial U}{\partial C_C}} \right] \frac{1}{\eta}, \tag{15}$$

where $\eta = \mu / \frac{\partial U}{\partial C_C}$ is referred to as the marginal cost of public funds. The term between the square brackets is the marginal environmental damage from this input. Analogously, the optimal tax on the dirty consumption good is the marginal environmental damage from the use of this good divided by the marginal cost of public funds:

$$t_D^C = \left[\frac{\frac{\partial U}{\partial Q}\left(-\frac{\partial q}{\partial C_D}\right)}{\frac{\partial U}{\partial C_C}} \right] \frac{1}{\eta}. \tag{16}$$

The assumption that private goods are weakly separable from public consumption and environmental quality, so that environmental quality and public consumption do not directly affect private demand, imply that the tax on dirty consumption goods, i.e. eq (16), has no Ramsey term as e.g. equation (13), it is only endowed with a Pigovian term. This is not carried over to the intermediate input, which is only to be taxed at marginal external cost, i.e. eq (15). This implies that, consistent with Diamond & Mirrlees, production efficiency is maintained and that there is no additional revenue-generating role for taxes on intermediate inputs.[12]

Equations (15) and (16) indicate how the presence of distortionary taxation affects the optimal environmental tax rate. In general, an optimal pollution tax induces the level of emissions at which the marginal benefit from emissions reductions equals the marginal welfare cost of achieving

[11] Under decreasing returns to scale, production efficiency continues to be optimal so long as a 100 % tax on pure profits is available

[12] This is also true when environmental quality enters the households utility in a non-separable fashion.

such reduction. In the special case of a first-best world without distortionary taxes, a one-unit reduction in emissions involves a welfare cost corresponding to the loss of tax revenue due to the erosion of the base of the pollution tax; thus, the pollution tax rate represents the marginal welfare cost of emissions reductions. Hence, in a first-best setting, optimality requires that the pollution tax be set equal to the marginal benefit from pollution reduction, which is given by the term in the square brackets above. This is the Pigovian tax rate.

The second term, η^{-1}, reveals how the presence of distortionary taxes requires a modification of the Pigovian principle. In particular, it shows that the Pigovian rate is optimal if and only if η is unity. A unitary η means that public funds are no more costly than private funds. The higher the η, the greater the cost of public consumption goods, including the public good of environmental quality. When these goods are more costly, the government finds it optimal to cut down on public consumption of the environment by reducing the pollution tax.

In a second-best world with distortionary taxes, the marginal cost of public funds is given by:

$$\eta = [1 - t_L \epsilon_{LL}]^{-1}, \tag{17}$$

where ϵ_{LL} is the uncompensated wage elasticity of labour supply. η exceeds unity if *i)* ϵ_{LL} is positive *ii)* the distortionary tax on labour, t_L, is positive (which is required if Pigovian taxes are not sufficient to finance public consumption). Combining equation (17) with equation (15) or (16), one finds that the presence of distortionary labour taxation reduces the optimal pollution tax below its Pigovian level if and only if ϵ_{LL} is positive. In a second-best setting, environmental taxes are more costly because they exacerbate the distortions imposed by the labour tax. In particular, by reducing the real after-tax wage, they decrease labour supply if the uncompensated wage elasticity of labour supply is positive. In the presence of a distortionary labour tax, the decline in labour supply produces a first-order loss in welfare by eroding the base of the labour tax. This additional welfare loss raises the overall welfare cost associated with a marginal reduction in emissions. As a result, and in contrast with the first-best case, the marginal welfare cost of a unit of emissions reduction exceeds the pollution tax rate. Thus, to equate marginal welfare costs and marginal social benefits from emissions reduction, the optimal environmental tax must be set below the marginal social benefit, that is, below the Pigovian rate.

The analytical results is then tested by a numerical model of the U.S economy to examine further the issues of second-best environmental taxation. The focus is on the policy of a carbon tax, which is a tax on fossil fuels in proportion to their carbon content and, since carbon dioxide emissions are proportional to the carbon content of these fuels, a tax based on carbon content is effectively a tax on carbon dioxide emissions.

The model that has been used in the article indicates that in the presence of distortionary taxes, optimal environmental tax rates are generally below the rates suggested by the Pigovian principle, even when revenues from environmental taxes are used to cut distortionary taxes. The numerical simulations supports this analytical result. Under certain values for parameters, optimal carbon tax rates from the numerical model are between six and twelve percent below the marginal environmental damages. In addition, the numerical model shows that in the presence of (realistic) policy constraints, optimal carbon taxes are far below the marginal environmental damages and may even be negative.

Moreover, intermediate inputs are not to be taxed for revenue-raising issues, they are to be taxed for their environmental impact solely, this in agreement with Diamond & Mirrlees' desirability of aggregate production efficiency. Also, Sandmo's weights, here represented as the marginal cost of public funds, are still valid and attached to the relevant terms and make the tax rates differ from the Pigovian tax.

5.4 Mayeres & Proost (1997)

Mayeres & Proost (1997) construct a similar theoretical model as Bovenberg & Ploeg (1994) above but with income distribution aspects when they addresses the problem of what they call externalities of congestion types. It considers an externality that affects consumers and producers simultaneously and that has a feedback effect on their decisions.

Precisely as Sandmo (1975) there is a commodity M that generate externality level Z, but now some firms use good M as input ($y_M < 0$) and thereby generate congestion, while the other firms produce it as an output ($y_M > 0$) and do not generate congestion. All firms are characterised by constant return to scale for given externality parameters. One can think of the good M as trucks for the producers and cars for the n consumers. The level of Z (i.e. congestion) can be reduced by public investment R (say, road investments), thus $Z = Z(\sum x^i_M, \sum y^j_M, R)$ where the sum is over all n consumers (good x^i_M) and all m firms (good y^j_M).

If an individualistic social welfare function $W(V^1(q,T,Z),\ldots,V^n(q,T,Z))$ (where T stands for a uniform poll transfer) is maximised with respect to taxes and public investments (and under the usual material balance constraints) the following optimal tax rules emerges (where S is the matrix of aggregated compensated price effects defined for constant Z, $\underline{S}$ is the determinant of this matrix and $\underline{S}_{jk}$ the cofactor of the element in the mth row,

kth column of $\underline{S}$):

$$t_k = \sum_{j=2}^{M}(X_j\phi_j)\frac{\underline{S}_{jk}}{\underline{S}}, \quad k=2,3,\ldots,M-1$$

$$t_M = \sum_{j=2}^{M}(X_j\phi_j)\frac{\underline{S}_{jk}}{\underline{S}} + Y. \tag{18}$$

Y is the net social Pigovian tax and ϕ is defined as the normalised covariance between the consumption of good k and the net social marginal utility of income. It equals the distributional characteristic of good k minus 1. The distributional characteristic measures the extent to which good k is consumed by people with a high net social marginal utility of income. (Good 1 represents leisure and is taken to be the numéraire and to be untaxed.)

They notice that the optimal tax structure has Sandmo's additivity property; the net social Pigovian tax only enters the expression for the indirect tax of the externality generating good.[13] The tax formula for the other goods remains unchanged. Note that in spite of this, the introduction of a tax on externalities necessitates a change in all taxes to satisfy the budget constraint and to re-optimise the welfare distribution, i.e. distributional considerations play a role in both components.

The first (Ramsey) term is related to revenue raising. It makes a trade-off between efficiency and equity considerations. This is best illustrated when the cross-price effects of demands are zero. Assume that the government want to reduce inequality and gives a higher weight to lower income groups. In that case the component of the tax will be lower the more sensitive transport demand is to price changes (efficiency) and if the transport good is consumed proportionally more by lower income groups (equity).

In the more general case, when the cross-price elasticities are non-zero, efficiency requires that the tax is higher for those goods which are more complementary with leisure. This is important in the transport pricing debate. People travel for different purposes. In general a distinction can be made between leisure and commuting trips. If it is possible to tax these trips differently, theory suggests that one should tax leisure trips, which are complementary with leisure, more than commuting trips, which are complementary to labour.

[13] Cremer et al. (1999) show that this property depends on the information in the economy and the assumptions made about the feasible tax instruments. They explore how externalities affect the optimal structure of the nonlinear income tax. It is showed that externalities may change the formula for the optimal marginal income tax rates if commodity transactions are anonymous and the government therefore cannot levy nonlinear commodity taxes. Intuitively, in the absence of sufficiently rich instruments to control the externality directly through a tax on pollution, it is second-best optimal to address the externality indirectly through the nonlinear income tax.

The second term of equation (18), i.e. the net social Pigovian tax, differs from the first-best Pigovian tax in several aspects. It consists of three terms (which not are displayed here):

> i) A weighted average of the costs that congestion, environmental effects and safety effects cause to households, corrected by the marginal cost of funds.
>
> ii) The marginal external costs for the firms, related to congestion, air pollution and accidents.
>
> iii) The effect of the transport externalities on net government tax revenue, it represents the productivity loss associated with a marginal increase in congestion.

In the model, the government can influence the level of congestion by infrastructure investment. The conclusion is that the government should provide additional road infrastructure up to the point where the cost of an additional unit of road infrastructure equals its benefit. Or, the cost of increasing road capacity should be equated to the benefit of a reduction in congestion, where the latter equals the net social Pigovian tax expressed per unit of congestion.

Finally they show that intermediate inputs which do not cause congestion are not to be subject to any tax; cf. Diamond and Mirrlees (1971). However, decentralisation on production decisions requires an excise on the input of the externality-generating good M. The excise is similar to the net social Pigovian tax defined for consumer purchases of good M.

Even though the optimal tax model now is extended to incorporate both externalities of congestion type and income distributions, the results still stand; intermediate goods are not to be equipped with a Ramsey term and the additively property as well as the weighting property indicated by Sandmo are still valid. The concept of net social Pigovian introduced by Bovenberg and Ploeg (1994), see page 19 in the present paper, is also extended.

6 Cost-of-Service Regulations

The conclusion from the above papers is that one should not levy any Ramsey tax on intermediate goods, at least if production exhibits constant return to scale. This is in accordance with the production efficiency lemma which says that production is not to be distorted, that is, the tax system should not directly alter the relative prices of intermediate inputs. But, if the Ramsey-Boiteux model is employed, where a cost-of-service regulation imposes a somewhat ad hoc budget constraint for the regulated firm, one is confronted by a different problem and, by that, different solutions.

One of the implicit conclusions of Boiteux (1971) is that there are gains to be made by imposing a single budget constraint across as broad a range of public enterprise activities as possible, rather than treating them as separate compartments required to meet individual constraints. The social loss from pricing above marginal costs is minimised if users are charged according to willingness to pay for the service as a whole, whether or not there are alternative suppliers seeking their business. This is due to one important caveat with the Boiteux-Ramsey pricing: it is an application of optimal tax theory to only a subset of the economy.

6.1 Borger (1997)

Borger (1997) investigates pricing rules for a budget-constrained and externality-generating public enterprise, or a regulated sector, which provides both final and intermediate goods. It is a general equilibrium model where the externalities affect both consumers and producers, and they include congestion-type externalities which generate feedback effects into final goods demand of consumers as well as factor demand of producers. The model does not impose constant returns to scale on private production, and allows for distributional effects of publicly determined prices and it also handles private sector profits.

The model is phrased in terms of a state-owned, or regulated, transport sector offering both passenger transportation (a final consumer good) and freight transport (an intermediate good). Two modes are available, in which mode 1 (e.g. road transport) implies a congestion-type of externality. A benevolent government chooses all state-controlled transport prices to maximise a social welfare function of Bergson-Samuelson type, subject to two constraints; a market equilibrium restriction and a budget restriction imposed on the transport sector.

If q is the output price for the *aggregated* production good y of the private sector and p_1^F and p_2^F is the respective input price for the two modes of freight transportation. The consumers face prices p_1^H and p_2^H for passenger transportation and have $I = \delta^h \pi$ as non-labour income, that is, δ^h represents h's share in private sector profit. And, finally, if $E = \rho z_1^H + \beta z_1^F$ represents the external effect imposed by passenger transport (z_1^H) and freight transport (z_1^F), where ρ and β are constants, the government solves:

$$\max_{q,p_1^H,p_2^H,p_1^F,p_2^F} \quad W[v^1(q,p_1^H,p_2^H,I,E),\cdots,v^n(q,p_1^H,p_2^H,I,E)] \ :$$

$$\pi^*(q,p_1^H,p_2^H,p_1^F,p_2^F,E) = 0 \quad (\lambda)$$

$$\sum_h x^h(q,p_1^H,p_2^H,I,E) = y(q,p_1^F,p_2^F.E) \quad (\mu)$$

where the transport sector's profit, π^*, is given by

$$\pi^* = p_1^H z_1^H + p_1^F z_1^F + p_2^H z_2^H + p_2^F z_2^F - G(z_1^H, z_1^F, z_2^H, z_2^F)$$

and where G is a joint cost function for all transport modes.[14]

If the above expression is optimised, a comparative observation reveals that the presence of a production externality implies that general equilibrium effects are non-zero even under constant returns to scale. An increase in public sector prices affects externality levels, and these in turn imply changes in marginal production costs and thus equilibrium prices on the private market. This implies that imposing the assumption of constant return to scale in the model does not reduce the results with respect to final good prices to those obtained by Sandmo (1975) and Bovenberg & Ploeg (1994). One has to assume that the external effects were a pure consumption externality, i.e. no feed-back effects into the production line, then their results would follow if constant returns to scale were imposed.

In order to display the pricing rules transparently, the private sector prices are assumed to be parametrically given. If distributional issues are ignored, and if the assumption of zero cross-price elasticities of the demand for final good and zero cross-price elasticities between private and intermediate goods are added[15] and where ϵ_{ii}^j is the own price elasticities, m the marginal utility of income, MPC_i^j the marginal private production cost associated with z_i^j and MEC_1^j reflects the marginal social damage due to an increase in the externality-generating good z_1^j, one gets:

$$\begin{aligned}
&\frac{p_1^H - (MPC_1^H + (m/\lambda)MEC_1^H))}{p_1^H} = \frac{m-\lambda}{\lambda}\frac{1}{\epsilon_{11}^H} \\
&\frac{p_2^H - MPC_2^H}{p_2^H} = \frac{m-\lambda}{\lambda}\frac{1}{\epsilon_{22}^H} \\
&\frac{p_1^F - (MPC_1^F + (m/\lambda)MEC_1^F))}{p_1^F} = \frac{m-\lambda}{\lambda}\frac{1}{\epsilon_{11}^F} \\
&\frac{p_2^F - MPC_2^F}{p_2^F} = \frac{m-\lambda}{\lambda}\frac{1}{\epsilon_{22}^F}.
\end{aligned} \tag{19}$$

To interpret these results consider the case where the welfare cost of the budget constraint exceeds the marginal utility of income (i.e. $\lambda > m$). This is probably the most relevant case since it reflects the distortionary effects of taxation. The welfare-optimal pricing rules then resemble Ramsey pricing, in which price is a markup over marginal private cost plus a fraction of marginal external cost. This is analogues to e.g. Oum and Tretheway (1988). (Whose results indicate that the fraction of externality costs varies with the

[14]The introduction of separate cost functions for the different modes does not affect the interpretation of the optimal pricing result.

[15]The latter assumption is somewhat artificial. As intermediate good prices affect private sector profits and consumer incomes, one can show that the assumption is equivalent to assuming zero income elasticities for the z_i^H. Without this assumption the more general results for intermediate public goods cannot be simplified in a straightforward manner.

size of the deficit that would result under private marginal cost pricing.) The finding that the markup should capture just a fraction of the marginal external cost, and not the full marginal external cost, can be explained by the fact that the budget restriction is specified in terms of private, not social, costs, whereas all externality costs are captured in the objection function. A markup of price over marginal private cost, necessary to satisfy the budget constraint, itself reduces the externality level. This implies that it suffices to introduce a markup which only captures part of the marginal cost, as the budget restriction itself contributes to externality reduction.

The pricing schemes above prescribes a somewhat different rule than the previous ones. Here, the intermediate goods are also to be taxed in the Ramsey tradition, that is, the input goods are equipped with a revenue-generating term. This is partly due to the balanced budget constrain imposed on the transport sector. The production efficiency lemma is traded against a cost-of-service regulation.

The insights of the Ramsey-Boiteux model are important. Yet there are several concerns with this model as a paradigm for regulation. Or, as Laffont & Tirole (1993) pointed out: Under linear pricing the firm's fixed cost should not enter the charges to consumers so as not to distort consumption, and therefore it ought to be paid by the government. But the Ramsey-Boiteux model exogenously rules out transfers from the government to the firm, so prices in general exceed marginal costs (which basic economic principles have made clear is efficient). In the end, a tension is uncovered between the benevolent regulator, on the one hand, and that the regulator is not given free rein to to operate transfers to the firm and to obtain efficiency, on the other.

Laffont & Tirole give emphasis to incentives. They actually rule out marginal cost pricing in favour for average cost pricing, since, first, a balanced-budget rule may commit the government not to automatically subsidise an inefficient firm. Second, passing the fixed costs along to taxpayers would not guarantee the cost would be closely monitored, while, on the other hand, charging the fixed cost to consumers gives them incentives to act as watch-dogs. Due to that they propose an incentive pricing model similar to the model by Ramsey-Boiteux with an incentive term entering the formula additively which reflects asymmetric information.

6.2 Borger, Coucelle & Swysen (2003)

Borger et al. (2003) studies welfare-optimal pricing of passenger and freight transport services in a federation. They use a model similar to that of Borger (1997), see page 28, to study external cost spillovers to other regions and incorporate the possibility to tax exporting and tax competition between countries. It is a two-region model in which each region can tax all passenger and freight transport flows occurring within its jurisdiction as well as decide on the level of externality-reducing investment expenditures. The model does not incorporate the labour market explicitly and is thus not a full-blown general equilibrium model.[16]

Freight flows in a given region come from both domestic and foreign firms but, for ease, passenger transport flows in a given region are assumed to be due to its own inhabitants only. Both freight and passenger transport flows are assumed to cause the externality but, on the other hand, transport demand is assumed to be independent of externality levels, i.e. there is no feedback effects in the model. In order to focus on tax exporting and tax competition in isolation they also assume zero cross-price effects in final demand, zero spillovers and equal marginal cost of funds in the two regions.

There are three different optima studied, namely; a local optimum with and without discrimination possibilities between domestic and international (transit) flows, and a global optimum. Tax competition can be illustrated by considering domestic freight transport and comparing globally and locally (the case with discrimination is used) optimal tax rules. If F_i^i and F_j^i denote freight transport services in region i and j associated with local production in region i, respectively, and the corresponding input prices are q_i^i and q_j^i. Then, if c_i^i and c_i^j are the marginal resource costs of domestic and international freight transport in region i, the tax rules become:

$$\begin{aligned} \left[(q_i^i - c_i^i) - \frac{MEC_i^i}{1+\lambda}\right]\left(\frac{dF_i^i}{dq_i^i}\right) &= -\frac{\lambda}{1+\lambda}F_i^i - \left[(q_j^i - c_j^i) - \frac{MEC_j^i}{1+\lambda}\right]\left(\frac{dF_j^i}{dq_i^i}\right) \\ \left[(q_i^i - c_i^i) - \frac{MEC_i^i}{1+\lambda}\right]\left(\frac{dF_i^i}{dq_i^i}\right) &= -\frac{\lambda}{1+\lambda}F_i^i. \end{aligned} \tag{20}$$

The global optimum from the federal viewpoint, i.e. the first row, takes account of both the impact of the tax change in i on tax revenue in region j and its effect on external costs there. Both these effects are ignored at the local level, i.e. the second row .

If it is assumed that an increase in the domestic freight tax has a positive effect on tax revenues abroad and that the tax abroad exceeds the marginal external costs corrected for the shadow cost of funds, the effect will be to lower the domestic freight tax rate. One reason could be that firms in i can to some extent substitute domestic freight transport and transport abroad. Alternatively, in the case of fuel taxes, for example, higher domestic tax induces domestic trucks to buy fuel abroad. In both cases foreign tax revenue rise. This positive fiscal externality is ignored in the local optimum in region i and implies a lower tax as compared to the global optimum. However, if

[16]This is not innocuous, because it is well known that optimal taxes on final and intermediate goods strongly depend on labour market distortions induced by suboptimal labour taxes (see Bovenberg & Goulder (1996) on page 22).

domestic freight and transport abroad are complements, or if foreign taxes are below external cost, then the individual country will set a higher tax rate than is desirable from a global viewpoint.

To see the role of tax exporting the tax rules for domestic and international (or transit) freight flows are compared. The tax structure for domestic freight flow, eq. (20), can be rearranged to yield:

$$\frac{q_i^i - (c_i^i + \frac{1}{1+\lambda} MEC_i^i)}{q_i^i} = -\frac{\lambda}{1+\lambda} \frac{1}{\epsilon_i^i}, \tag{21}$$

if ϵ_i^i is defined as the own-price elasticity of domestic freight demand with respect to the domestic prices. (cf. also with eq. (19) on side 29.) However, the locally optimal tax rule for international (or transit) flows is given by:

$$\frac{q_i^j - (c_i^j + \frac{1}{1+\lambda} MEC_i^j)}{q_i^j} = -\frac{1}{1+\lambda} \frac{1}{\epsilon_i^j}, \tag{22}$$

where the ϵ_i^j is the cross-price elasticity of freight demand abroad with respect to the domestic price. The tax on transit flows implies monopoly pricing corrected for externalities. The optimal tax simply maximises the differences between the tax revenues and the relevant social costs. Further, even in the special case of lump-sum taxation, tax exporting persists. In that case $\lambda = 0$ and the domestic tax rule, eq. (21), simply reduces to local marginal cost pricing. However, transit would be taxed at more than marginal cost, as shown by eq. (22). The conclusion is clear: tax exporting induces countries to set higher taxes on transit than on domestic flows.

If no discrimination is feasible (for technical or legal reasons) the common tax rule for transit and domestic freight flows yield:

$$\frac{q_i^i - (c_i^i + \frac{1}{1+\lambda} MEC_i^i)}{q_i^i} = -\frac{\frac{\lambda}{1+\lambda} + \frac{F_i^j}{F_i^i}}{\epsilon_i^i + \epsilon_i^i \frac{F_i^j}{F_i^i}}.$$

To interpret these results, note that the optimal tax rate should be higher the more inelastic is the demand for both domestic and international freight transport, and the larger the proportion of the international flows. The local planner now faces a trade-off between increased tax revenue obtained from transit flows on her territory and increases in excess burden associated with domestic freight flows facing prices that do not reflect true social costs.

To summarise, it is *a priori* unclear whether tax exporting and tax competition leads countries to impose taxes that are too high or too low from the viewpoint of the federation as a whole. Tax exporting typically raises taxes in countries where international transport constitutes a substantial share of total flows. The tax competition was shown to be *a priori* undetermined.

Inspection of the expressions describing the global and local optimum reveals that the overall effect will depend, among others, on the price sensitivity of the tax revenue and on the shadow cost of public funds.

7 Welfare Effects with Environmental Tax Reforms

This section discusses the welfare effects that environmental tax reforms produce, with emphasis giving to what has become to known as the *double dividend* hypothesis, which claims that a revenue-neutral *green* tax reform may not only improve the environment, it may also reduce the distortion of the existing tax system. The revenues from the the first dividend (environmental taxes) make it practicable feasible to achieve the second dividend (a less distortionary tax system).

Goulder (1995) distinguish between the strong and the weak form of the double dividend. The weak dividend occurs if tax revenues from a green tax reform is recycled in the form of lower distortionary taxes compared to the case when the revenues are recycled through lump-sum taxes. A corollary of this hypothesis, as Bovenberg (1999) pointed out, is that environmental taxes are more efficient instruments for environmental protection than environmental policy instruments that do not yield any revenues. This since the environmental taxes can be employed to cut distortionary taxes.

The strong form of the double dividend is defined as the effect an environmental tax reform has on the non-environmental welfare cost of the whole tax system. The tax reform enhances not only environmental quality but also non-environmental welfare. Whereas the weak double dividend thus compares two policy changes, the strong form compares the equilibrium after a single policy change with the status quo.

Regarding the weak form, the theory is not only easily derived, it has also strong support from numerical simulations.[17] But, concerning the strong form, many economists have showed that although the double dividend is-possible, it is unlikely to arise except under fairly unusual circumstances. This is due to the fact that revenue-neutral environmental tax policies lead to a reduction in employment. By harming the employment, pollution taxes narrow, rather than widen, the tax base. This means that the non-environmental component of welfare falls. Thus, the double dividend fails.

However, the absence of the double dividend does not mean that overall welfare falls, it could rise depending on the size of the environmental taxes. Hence, the (possible) failure of the double dividend claim does not imply that green tax reforms are inefficient, it simply states that environmental

[17] Mayeres and Proost (2001) show that distributional considerations can destroy this property. With a high inequality aversion and if rich people consume proportionally more of non-externality-generating commodities, a weak double dividend is less likely.

improvements comes at a (gross) cost.[18]

If the initial tax system is inefficient from a non-environmental point of view, an environmental tax reform may be able to reduce the overall burden of taxation and achieve the double dividend after all. The key requirement is that the revenue-neutral reform end up alleviating these prior inefficiencies and move the rest of the tax system closer to its non-environmental optimum. As a related result, Bovenberg & de Mooij (1994) have shown that the optimal environmental tax typically lie below the Pigovian tax. Intuitively, cutting pollution taxes below their Pigovian levels raise efficiency by alleviating non-environmental distortions. (See also section 5.2 and 5.3 in the present paper.)

Tax interactions are also crucial to the choice between environmental taxes and other, non-tax instruments for environmental protection. Non-auctioned pollution quotas produce the same costly tax-interaction effect that environmental taxes do, but, in contrast to environmental taxes whose revenues are used to finance cuts in distortionary taxes, such quotas fail to enjoy the beneficial revenue-raising effect. Moreover, the presence of uncertainty about abatement costs and benefits,and the associated costs of monitoring and enforcement, complicate the problem of instrument choice. Once these issues are taken into account, the efficiency ranking of taxes, quotas and other instruments (such as performance standards and mandated technologies) becomes less clear.

Regarding the transport sector within the EU, the emphasis of taxation has shifted towards taxation of (road) freight transport rather than passenger transport. This is to some extent due to the fact that charging passengers for the external costs they create seems infeasible at a European level for both political and technical reasons. An additional explanation is that international traffic flows throughout Europe to a large extent consist of freight. This raises a number of policy questions, for instance; given that the passenger transport is sub-optimality taxed, which seems to be the case in many European countries (Borger & Proost (2001)), under what conditions does it make sense to raise the tax on freight transport?

To see the role of the distortion on the passenger market more clearly, Calthrop, Borger & Proost (2003) studied the welfare implications of an externality-generating intermediate input (freight transport) in an general equilibrium model which accounts for congestion. They concluded that

[18] Two welfare effects underlie this conclusion. By driving up the price of polluting goods relative to leisure, environmental taxes tend to compound the factor-market distortions created by pre-existing taxes, thereby producing a negative welfare impact termed (by Parry (1995)) the *tax-interaction effect*. At the same time, environmental taxes whose revenues are recycled through cuts in marginal tax rates reduce the distortions caused by the pre-existing taxes, which contributes to a positive welfare impact. Typically, this *revenue-recycling effect* (Parry, again) only partly offsets the tax-interaction effect, implying that the overall impact of pre-existing taxes is to raise costs.

(when started from an initial situation where the tax on freight is equal to marginal external cost (MEC)) if a price increase of freight transport increases the demand for passenger transport, welfare increases when the tax on freight is lowered below MEC. Intuitively, untaxed passenger transport implies a distortion due to excessive congestion, reducing the tax on freight reduces this distortion. Likewise, if raising the price of freight reduces demand for passenger transport, welfare would increase if the tax on freight were raised above MEC.

However, if the tax (on freight transport) is above or below marginal external cost, the effects of changing the freight tax affects both the distortion on the passenger market and on the freight market. These effects must be traded off against one another. For example, even if increasing freight taxes raised passenger transport and hence congestion, it might still be welfare-improving to increase freight taxes if freight was also strongly under-priced in the initial equilibrium. In that case, although raising the freight tax increase congestion by passengers and hence the distortion on this market, it reduces the distortion on the freight transport market.

They conclude this discussion from an optimal taxation viewpoint. First, more congestion and hence a larger MEC does not necessarily imply a higher optimal freight transport tax. The reason is that, although the larger distortion on the freight market induces higher freight taxes, the higher MEC also raises the distortion on the passenger market. This may necessitate lower freight transport taxes if this helps to reduce distortions. Second, higher taxes on passenger transport may for the same reason both increase or decrease the optimal tax on freight transport. On the one hand, higher passenger transport taxes reduce congestion, which reduces the optimal freight transport tax. On the other hand, however, increasing the passenger transport tax reduces the distortion on the passenger market, which may induce higher freight transport taxes. In sum, even on the transport market alone it may be necessary to take into account important general equilibrium effects.

The above conclusions were derived under the assumption that recycling operated via the lump-sum transfer. They also consider the case when the revenues are recycled through cuts in the labour tax. Then, when freight is taxed at MEC, reducing the tax is welfare-improving if it increases labour supply. Whilst returning revenues through lump-sum taxes is likely to reduce labour supply (because labour is a normal good), using freight tax receipts to reduce labour taxes is much more likely to increase labour supply, assuming positive wage elasticities of labour supply. This suggests that reducing freight taxes are more desirable if recycling is through labour taxes than via lump-sum taxes. Alternatively, it reflects the facts that using freight revenues to reduce labour taxes weakens the cost of raising the freight tax relative to lump-sum instrument. Increasing the tax on freight result in higher welfare cost when revenues are returned via lump-sum transfers than via the labour tax.

8 Administrative Costs

Considerable effort has been devoted to characterising the manner in which the optimal tax rates differ across the taxable commodities. But the literature has rarely addressed the issue of which commodities should be included in the tax base. In fact, the standard optimal tax model cannot provide useful answer to this question, because it does not contain any of the resource costs normally associated with the collection of taxes.[19]

In adopting a set of taxes, governments are influenced by the relative costs of administering and enforcing each kind of tax. Similarly, costs of compliance and transactions for households and firms depend on the mix of taxes used. Slemrod (1990) recognise that administrative problems are often at the heart of why optimal commodity and profits taxation are not implemented, and thus opening the way for taxes which interfere with production efficiency. The cost of administering any tax system increases with the number of different tax rates that are imposed, so that only a small number of tax rates may be desirable. Thus, problems that arise in administering real tax systems may often make some forms of production tax appropriate, even if such a tax works against production efficiency.

Heller and Shell (1974) offer a first attempt at incorporating these important costs into a formal model of optimal taxation. Their work is an attempt to extend the Diamond and Mirrlees' result on the desirability of aggregate production efficiency, but their work is on a very abstract level. Their conclusion is: When transactions are costly, pure production efficiency is optimal only in very special and unlikely circumstances.

Slemrod also distinguishes between the theory of *optimal taxation* and *optimal tax systems*. Optimal taxation is usually restricted to the optimal setting of a given set of tax rates, ignoring other social costs of taxation. When optimising tax systems one has to consider all the elements of the problem. Clearly, any general solution that ignores components of the problem is doomed to fail if the omitted part of the problem is an important issue. For example, one of the major conclusions in optimal tax theory is that lump-sum taxation is the first-best solution. However, by introducing administrative costs and lump-sum taxes, lump-sum taxes ceases to be the optimal solution and that, in turn, implies that optimal taxation ceases to

[19]Among the writers who have explicitly modelled aspects of resource (or administrative) costs, Yitzhaki (1989) and Wilson (1989) examine the optimal number of taxed commodities when taxing each commodity involves fixed administrative costs, while Kaplow (1989) examines modifications to the Ramsey rule in the presence of administrative costs and Mayshar (1991) analyses both administrative activity by tax authorities and tax-resisting activity by taxpayers.

be a problem of second-best. The main reason for that is the cost of the lump-sum administration.[20]

What Slemrod wants to achieve is a shift for the normative theory of taxation from the structure of consumer preferences to the technology of collecting taxes and those aspects of the economy which affect tax collection, and from optimal tax rates structure to optimal tax systems. This because preferences are relative stable over time, but technology is not.

Slemrod & Yitzhaki (1996) distinguishes five components of the cost of taxation:

Administrative costs: The cost of establishing and/or maintaining a tax administration.

Compliance costs: The costs imposed on the taxpayer to comply with the law.

Regular deadweight loss: The inefficiency caused by the reallocation of activities by taxpayers who switch to non-taxed activities. (Optimal tax theory is focused on minimising the deadweight loss due to substitution between commodities.)

Excess burden of tax evasion: The risk borne by taxpayers who are evading.

Avoidance costs: The cost incurred by a taxpayer who searches for legal means to reduce tax liability.

The classification is sometimes arbitrary and may depend on the interpretation of the agents' intention, but, nevertheless, the classification could be important in avoiding double counting of social costs.

In reference to the first component above, administrative costs, it is correlated with information gathering. The cost of gathering information depends on how accessible the information is, and whether it can be easily hidden. The costs is also an increasing function of the complexity and lack of clarity of the tax law. Furthermore, they tend to be discontinuous and to have decreasing average costs with respect to the tax rate.

Regarding the compliance costs, the potential efficiency of involving taxpayers in the administrative process must be tempered with a practical consideration. Administrative costs must pass through a budgeting process, while compliance costs are hidden. Hence, there may be a tendency to view a decrease in administrative cost accompanied by an equal (or greater) increase in compliance costs as a decrease in social cost, because it results in

[20]See Hahn (1973) or Yitzhaki (1989) for more on that topic, or, in the context of freight transport and welfare effects of green tax reforms, see Calthrop, Borger & Proost (2003). The latter is summarized on page 34 in the present paper.

a decrease in government expenditure.[21] Mayshar (1991) addresses compliance cost with a formal normative model.

Any tax that creates a wedge between the relative prices that any two taxpayers face entails an efficiency loss. The deadweight loss created is an increasing and continuous function of the tax rates, but it is also a function of the combinations of taxes employed. If they are not given, the deadweight loss can not be assumed to be a continuous function.

Tax evasion has been investigated thoroughly starting with Allingham & Sandmo (1972), who concluded that tax evasion occurs only if the taxpayer expects to increase her expected income by evading taxes, including the expected fines that she would have to pay if she were detected. Or, as Vakneen & Yitzhaki (1987) concluded, assuming a risk-neutral government, the excess burden of tax evasion is equal to the risk premium that the taxpayer would be ready to pay in order to eliminate the exposure to risk.

Last, avoidance costs, which result from voluntary action carried out by taxpayers whose intention is to reduce tax payment. Clearly, any activity to reduce the tax is a pure loss from the social points of view, and therefore creates deadweight loss. In practice, however, it is hard to distinguish between avoidance and compliance costs because the distinction between the two depends on the intention of the taxpayer.

Slemrod & Yitzhaki (1996) offer a methodology that evaluate marginal changes in tax systems that accounts for the above social costs and which they call the marginal efficiency cost of funds (MECF). The methodology is based on the concept of the marginal cost of public funds.

They argue that the MECF concept is useful for analysing minor tax reforms. For any change in the tax system that is considered, one has to evaluate the expected tax revenue and the expected leaked tax revenue. (In the simplest case, the sum of the two, divided by the former, is the MECF.) If the expected revenue is forecasted it can be used to calculate the relevant marginal excess burdens. Then, knowing these marginal excess burdens, one can identify a revenue neutral marginal change that will improve the tax system by lowering its total social cost.

The literature on the subject is few in number (compared to the literature on optimal taxation) and also (frequently) in a positive manner. For use as policy recommendations, in a broader context, the theory must come to term with such issues as the choice of tax instruments, the optimal design

[21] Consider the following problem (proposed by Slemrod & Yitzhaki (1996)): When is it optimal to delegate to employers the authority to collect taxes and convey information about employees, thus requiring the administration to audit both the taxpayer agent and the taxpayer herself, and when is it optimal to deal only with the employee? Clearly, given that the employer already has the necessary information, it would save administrative costs to require her to to pass it along to the tax administrator. This might also reduce total social costs if the cost of gathering the information by the administration is higher than the increase in cost caused by imposing a two-stage gathering system.

of enforcement policy, the tax treatment of financial strategies and more generally, must develop a descriptive and normative framework in which to evaluate the issues of tax arbitrage. In this more general framework of optimal tax system, optimal taxation could emerges as a special case in which the set of tax instrument is fixed and enforcement of any available instrument is cost-less.

9 Conclusions

Regarding intermediate goods and if the optimal tax literature is to be followed, freight transports are not to be taxed by revenue or investments reasons. But, if the (freight) transports create diseconomies, that is, have negative external effects, the tax formula is to be equipped with a Pigovian tax in order to reduce the marginal social damage of the external effects. Moreover, due to differences in cost for public and private funds (i.e. the presence of distortionary taxes), the Pigovian term is to be multiplied with the inverse of the marginal cost of public funds. One implication of this is that, if public revenues become scarcer, as indicated by a higher marginal cost of public funds, the optimal tax system focuses more on generating revenues and less on internalising pollution externalities.

Then there is the Ramsey-Boiteux model, which hinges a budget constraint on the sector, or regulated firm in question. The pricing scheme then becomes somewhat different. Pricing the intermediate good is still closely related to marginal cost, but it also depend on the elasticity of demand (the Ramsey term). Since the firm's deficit is socially costly, pricing between marginal cost and the monopoly price obtains. In particular, even though price exceeds marginal cost, pricing is unrelated to average cost pricing, since price is independent of the fixed cost.

The two model's share a common drawback since they create no incentives for cost reduction, which average cost pricing would, partly because closely monitoring by consumers would make it more difficult for the firm to inflate fixed costs, and partly since a balanced-budget constraint forces the firm to pay more attention to its fixed costs. Indeed, the Ramsey-Boiteux model has a budget constraint but the cost function is presumed to be exogenous and the firm's managers and employees have thereby no effect on cost.

One rationale, in light of the the latter paragraph, for Ramsey-Boiteux pricing is that, although it could be justified to induce the proper incentives, a formal budget restriction on the public sector may be inappropriate from a policy viewpoint. The reason is that the implied welfare cost of public funds may be substantially different from that obtained when funds come from other sources. While one rationale against it, is due to one of its caveat, that is, it is an application of optimal tax theory to only a subset of the economy.

Concerning the welfare improvement, it has been shown theoretically and illustrated numerically that returning revenues via labour taxes rather that lump-sum, unambiguously reduces the marginal cost of the policy reform. This suggests that reducing freight taxes are more desirable if recycling is through labour taxes than via lump-sum taxes. This is the so called weak form of the double dividend. Regarding the strong form of the double dividend (which is defined as the effect an environmental tax reform has on the non-environmental welfare cost of the whole tax system), it is very much in doubt even though there may be scope for it if a green tax reform helps eliminate pre-existing inefficiencies in the non-environmental tax system.

One criticism to the models above is that they do not incorporate administrative costs in their framework. Evasion and avoidance activities as well as administrative and compliance costs cause deadweight losses which not optimal tax theory considers, it only focuses on substitution costs between commodities. However, even though the issue of tax evasion has been theoretically investigated, relatively little analytical work incorporating tax administration has been done, mainly because administrative issues are hard to analyse with continuous differentiable functions and, therefore, they require complex modelling.

References

[1] Allington, Michael G. & Sandmo, Agnar (1972): Income Tax Evasion: A Theoretical Analysis; *Journal of Public Economics*, Vol. 1, Nov., pp. 323 - 338.

[2] Boiteux, M. (1971): On the Management of Public Monopolies Subject to Budgetary Constraints; *Journal of Economic Theory*, Vol. 3, pp. 219 - 240.

[3] Borger, Bruno de (1997): Public Pricing of Final and Intermediate Goods in the Presence of Externalities; *European Journal of Political Economy*, Vol. 13, pp. 765 - 781.

[4] Borger, Bruno de, Courcelle, Christophe & Swysen, Dider (2003): Optimal Pricing of Transport Externalities in a Federal System; *Journal of Transport Economics and Policy*, Vol. 37, Part 1, pp. 69 - 94.

[5] Borger, Bruno de & Proost, Stef (2001): Reforming Transport Pricing in the European Union: A Modelling Approach. *Edward Elgar.*

[6] Bovenberg, A. Lans (1999): Green Tax Reforms and the Double Dividend: An Updated Reader's Guide; *International Tax and Public Finance*, Vol. 6, pp. 421 - 443.

[7] Bovenberg, A. Lans & Goulder, Lawrence H. (1996): Optimal Environmental Taxation in the Presence of Other Taxes: General-Equilibrium Analyses; *The American Economic Review*, Vol. 86, No. 4, pp. 985 - 1000.

[8] Bovenberg, A. Lans & Mooij, Ruud A. de (1994): Environmental Levies and Distortionary Taxation; *American Economic Review*, Vol. 84, No. 4, pp. 1085 - 1089.

[9] Bovenberg, A. Lans & Ploeg, F. van der (1994): Environmental policy, public finance and the labour market in a second best world; *Journal of Public Economics*, Vol. 55, pp. 349 - 390.

[10] Calthrop, Edward, de Borger, Bruno & Proost, Stef (2003): Tax Reform for Dirty Intermediate Goods: Theory and an Application to the Taxation of Freight Transport; *ETE Working Paper Series, Centre of Economic Studies, K U Leuven*, No. 2003-02.

[11] Cremer, H., Gahvari, F. & Ladoux, N. (1999): Externalities and Optimal Taxation; *Journal of Public Economics*, Vol. 70, No. 3, pp. 343 - 364.

[12] Dasgupta, Partha & Stiglitz, Joseph (1972): On Optimal Taxation and Public Production; *The Review of Economic Studies*, Vol. 39, No. 1, pp. 87 - 103.

[13] Diamond, Peter A. & Mirrlees, James A. (1971): Optimal Taxation and Public Production I: Production Efficiency & II: Tax Rules; *American Economic Review*, Vol. 61 a & c, pp. 8 - 27 & pp 261 - 278.

[14] Dixit, A. (1985): Tax policy in open economies; In: Auerbach, A. & Feldstein, M. (Ed.) *Handbook of public economics*, Vol. 1, North-Holland, Amsterdam, pp. 313 - 374.

[15] Goulder, Lawrence H. (1995): Environmental Taxation and the "Double Dividend": A Reader's Guide; *International Tax and Public Finance*, Vol. 2, No. 2, pp. 157 - 183.

[16] Hammond, Peter J. (2000): Reassessing the Diamond/Mirrlees Efficiency Theorem; *Dept of Economics, Stanford University*; March 2000.

[17] Hahn, F. H. (1973): On Optimal Taxation; *Journal of Economic Theory*, Vol. 6 (Feb.), pp. 96 - 106.

[18] Heller, W. P. & Shell, K. (1974): On optimal Taxation with Costly Administration; *The American Economic Review*, Vol. 64, No. 2, pp. 338 - 345.

[19] Kaplow, Louis (1989): Optimal Taxation with Costly Enforcement and Evasion; *NBER Working Paper Series*, No. 2996

[20] Laffont, Jean.Jacques & Tirole, Jean (1993): A Theory of Incentives and Procurement and Regulation; *MIT Press, MA*.

[21] Mayeres, I. (2002): Taxes and Transport Externalities; *ETE Working Paper Series, Centre of Economic Studies, K U Leuven*, No. 2003-02.

[22] Mayeres, I. & Proost, S. (1997): Optimal Tax and Public Investment Rules for Congestion Type of Externalities; *Scandinavian Journal of Economics*, Vol. 99, No. 2, pp. 261 - 279.

[23] Mayeres, I. & Proost, S. (2001): Marginal tax reform, externalities and income distribution; *Journal of Public Economics*, Vol. 79, No. 2, pp. 343 - 363.

[24] Mayshar, Joram (1991): Taxation with Costly Administration; *Scandinavian Journal of Economics*, Vol. 93, No. 1, pp. 75 - 88.

[25] Myles, Gareth D. (1989): Imperfect Competition and the Taxation of Intermediate Goods; *Public Finance*, Vol. 44, pp. 62 - 74.

[26] Ng, Yew-Kwang (1980): Optimal Corrective Taxes or Subsidies when Revenue Raising Imposes an Excess Burden; *American Economic Review*, Vol. 70, No. 4, pp. 744 - 751.

[27] Oum, Tae Hoon & Tretheway, Michael W. (1988): Ramsey Pricing in the Presence of Externality Costs; *Journal of Transport Economics and Policy*, Vol. 22, pp. 307 - 316.

[28] Parry, Ian (1995); Pollution Taxes and Revenue Recycling; *Journal of Environmental Economics and Management*, Vol. 29, pp. 64 - 77.

[29] Parry, Ian & Bento, Antonio (2001); Revenue Recycling and the Welfare Effects of Road Pricing; *Scandinavian Journal of Economics*, Vol. 103, No. 4, pp. 645 - 671.

[30] Ramsey, F. P. (1927): A Contribution to the Theory of Taxation *The Economic Journal*, Vol. 37, No. 145 pp. 47- 61.

[31] Sandmo, Agnar (1976): Optimal Taxation: An introduction to the literature; *Journal of Public Economics*, Vol. 6, pp. 37 - 54.

[32] Sandmo, Agnar (1975): Optimal Taxation In the Presence of Externalities; *Swedish Journal of Economics*, Vol. 77, pp. 86 - 98.

[33] Slemrod, Joel (1990): Optimal Taxation and Optimal Tax System; *The Journal of Economic Perspective*, Vol. 4, No. 1, pp. 157 - 178.

[34] Slemrod, Joel & Yitzhaki, Shlomo (1996): The Costs of Taxation and the Marginal Efficiency Cost of Funds; *IMF Staff Papers*, Vol. 43, No. 1 (Mars).

[35] Vakneen, Yitzhak & Yitzhaki, Shlomo (1989): On the shadow price of a tax inspector; *Public Finance*, Vol. 44, No. 3, pp. 492 - 505.

[36] Wilson, J. Douglas (1989): On the Optimal Tax Base for Commodity Taxation; *The American Economic Review*, Vol. 79, No. 5, pp. 1196 - 1206.

[37] Yitzhaki, Shlomo (1979): A Note on optimal Taxation and Administrative costs; *The American Economic Review*, Vol. 69, No. 3, pp. 475 - 480.