Optimal Taxation, Imperfect Competition and Tax Enforcement Policies

Umberto Galmarini

Ph.D. thesis

University of Warwick
Department of Economics

July 1993

Contents

Acknowledgements

I wish to thank Norman Ireland and Gareth Myles for supervising my research at the University of Warwick. I am indebted to them for helpful comments and criticisms. I also benefited from extensive discussions with Flavia Ambrosanio, Massimo Bordignon and Piero Giarda at the Catholic University in Milan. I am grateful to the examiners, Michael Keen and Jonathan Thomas, for valuable comments. All errors are mine.

A previous version of chapter 1 was presented at the Young Economists' session at the Royal Economic Society Annual Conference 1991, University of Warwick, 8th–11th April.

Earlier versions of the following chapters circulated as working papers. Chapter 1: *Ad valorem and specific taxation in a model of vertically related oligopolies,* Warwick Economic Research Paper n. 383, University of Warwick, Coventry, and Quaderno n. 17, Dipartimento di Scienze Economiche, Università Cattolica di Milano. Chapter 2: *Tax reform in a model of vertically related oligopolies,* Quaderno n. 25, Dipartimento di Scienze Economiche, Università Cattolica di Milano.

Summary

This thesis contains four papers in the area of Public Economics.

Chapter 1 looks at producers' taxation in a model of vertically related oligopolies. Both ad valorem and specific taxes are considered and formulae expressing their effects on prices and profits are derived, showing how these depend on factors such as demand conditions, technology and market structure. Conditions for taxation to cause price overshifting and to raise profits are given. Also, tax instruments are compared in terms of the amount of revenue collected and the effect on the price for the final good.

Chapter 2 applies the results of the previous paper to the analysis of tax reforms. Vertically related oligopolies result in welfare loss for two reasons. Firstly, upstream oligopolists set the price of the intermediate good above marginal cost and this causes *aggregate* production inefficiency. Secondly, downstream oligopolists introduce an additional price-cost margin. The analysis focuses on tax reforms, where the government aims at reducing the welfare loss by levying taxes and subsidies on producers while raising no revenue.

Chapter 3 focuses on the design of income tax enforcement policies in a principal-agent framework. The existing literature assumes risk neutral taxpayers while this chapter considers the case of risk averse agents by assuming a kinked linear utility function. When individuals have the same attitude towards risk, it is shown that the optimal policy is such that income reports below a given threshold are audited at the probability level just sufficient to induce truthful reporting, whereas those above it are not audited. This makes the effective tax schedule to be quite regressive. Instead, if attitudes towards risk vary across taxpayers, the numerical results show that the optimal audit policy causes only a limited regressive bias, for income reports above the threshold meet a positive probability of audit.

Chapter 4 examines the Presumptive Income Coefficients (PIC) audit policy, a scheme recently introduced in the Italian tax code and aimed at reducing tax evasion in the non-corporate sector. The tax agency applies the PIC to observable production costs to get an estimate of taxpayer's income, or presumptive income. The probability of audit is then dependent on the gap between presumptive and reported income. This issue is examined in a setting where the *game* between the taxing authority and taxpayers is modelled in a principal-agent framework.

Introduction

This thesis contains four papers in the area of Public Economics. Chapters 1 and 2 deal with tax incidence and welfare improving tax reforms in a model of vertically related oligopolies. Chapters 3 and 4 focus on the optimal design of income tax enforcement policies in a principal-agent framework. Although each chapter can be read independently from the others, chapter 2 is the natural continuation of chapter 1. The order of presentation follows the chronological development of my research.

Tax incidence and tax reforms

Chapter 1 examines the impact of ad valorem and specific taxes levied on producers in a model of vertically related oligopolies, where a downstream industry produces a final good using the output of an upstream industry as an input. The scope of the analysis is twofold. The first objective is to characterize tax incidence on producers' net prices and profits and to determine the conditions under which taxes are overshifted and profits are increased. The second goal is to compare the various tax instruments in terms of revenue collected and the effect on the price of the consumption good.

Market structure is modelled following Katz and Rosen (1983), Seade (1985) and Stern (1987). These authors examine the effects of taxing producers in the homogeneous-product conjectural- variation oligopoly framework. With perfect competition price overshifting never occurs and therefore also the possibility of profits raising with taxation is ruled out. When the market is monopolized taxation may increase producer's net price but a positive effect on profits can never occur. Their main finding is that in oligopoly both the net price and profits may increase with taxation.

In accordance with these previous studies, which consider only the downstream stage of production, tax incidence in the downstream industry turns out to be governed by final demand conditions and downstream market structure. The novel results concern the effects of taxation on the intermediate good price and upstream profits. These are shown to be related to upstream and downstream market structure, final demand conditions and input substitution in downstream production. Also, specific taxation is more likely to cause price overshifting and to raise profits than ad valorem taxation.

Central to the analysis is the comparison between ad valorem and specific taxes. Suits and Musgrave (1953) compare ad valorem and specific taxation under monopoly, showing that if a specific and an ad valorem tax result in the same final good price, then the latter raises a higher revenue.[1] Chapter 1 generalizes this approach to the case of vertically related oligopolies. It is shown that the revenue from an ad valorem tax is always higher than the revenue from a specific tax levied in the same industry, when both taxes result in the same price for the consumption good. This result holds also when downstream producers price at marginal cost, provided that upstream producers price above it. The comparison between taxes of the same type levied in both industries is made assuming Leontief technology and isoelastic final demand. The revenue from downstream

[1]The recognition that they are not equivalent under monopoly dates back to Cournot (1838) and Wicksell (1896).

ad valorem taxation can be higher or lower than the yield from upstream ad valorem taxation, and which case applies depends on upstream and downstream market structure, production coefficients and the elasticity of final demand. Finally, it is shown that upstream and downstream specific taxation always raise the same revenue, when both taxes result in the same price for the consumption good.

Chapter 2 applies the framework of the previous chapter to the problem of tax reforms. The market allocation generates two kind of inefficiencies that result in welfare loss. Firstly, upstream oligopolists set the price of the intermediate good above marginal cost and provided that downstream technology allows inputs substitution this causes *aggregate* production inefficiency. Secondly, downstream oligopolists introduce an additional price-cost margin. The analysis focuses on tax reforms, where the government aims at reducing the welfare loss by means of distortionary commodity taxation while raising no revenue. Both ad valorem and specific taxes levied on producers in both industries are considered. The pattern of actions the government has to take in order to improve welfare are shown to be related to downstream and upstream market structure, final demand conditions and input substitution in downstream production. Also, it is shown that under some circumstances the tax reform improves overall market performance by introducing additional aggregate production inefficiency.

The contents of chapter 2 are closely related to those of a recent work by Myles (1989), who considers in a general equilibrium framework the two polar cases of downstream (upstream) monopoly and upstream (downstream) competitive sector. His model is general equilibrium because it takes account of the impact of monopoly profits on final demand. In the model presented here distributed profits are assumed not to enter final demand but a wider range of downstream and upstream market structures are considered.

The results are summarized as follows. When downstream oligopolists price above marginal cost and the upstream industry is competitive, then specific taxation cannot improve welfare. Instead, *ad valorem* taxes are effective: the welfare loss can be reduced by taxing the final good and subsidizing the intermediate good.

When both downstream and upstream oligopolists price above marginal cost, and under exponential final demand, specific taxation may improve welfare. The condition under which taxing downstream producers and subsidizing upstream firms raises welfare is shown to depend on the elasticity of final demand, the elasticity of input substitution and the degree of market power in the downstream industry.

When both downstream and upstream oligopolists price above marginal cost, and under isoelastic final demand, ad valorem taxation may improve welfare. The design of the tax reform is shown to depend on the values of the elasticity of final demand and the elasticity of input substitution. Welfare can be improved also when downstream technology is Leontief and downstream producers price at marginal cost. Moreover, when the elasticity of input substitution is low, the tax reform may improve welfare by raising the price of the intermediate good and by doing so generates additional *aggregate* production inefficiency.

The design of income tax enforcement policies

The theory of optimal income taxation in the presence of costly enforcement has been the focus of a seminal paper by Reinganum and Wilde (1985). Subsequent works include Border and Sobel (1987), Cremer, Marchand and Pestieu (1990), Sanchez and Sobel (1990) and Chander and Wilde (1992). These authors assume that taxpayers are risk neutral and are endowed with different pre-tax incomes which are exogenously given. Since the taxing authority cannot directly observe taxable incomes, taxpayers are required to report their own income by filling in an income declaration form. The audit policy consists of a function which gives a probability of investigation for each level of reported income. The objective of the tax agency is to maximize total tax revenue net of audit cost while taking taxes and penalties as given.[2] The standard result is that the optimal audit policy is a step function which divides taxpayers into two groups on the basis of reported income: reports below a given threshold are audited at the probability level just sufficient to induce a risk neutral taxpayer to behave honestly whereas income reports above the threshold are never audited.

Clearly, the optimal audit policy makes the effective tax payments to be quite regressive. On one hand, the relatively poor individuals (with income below the threshold) are forced to behave honestly so that they pay in full the legislated income tax. On the other hand, the relatively rich individuals (with income above the threshold) report an income just sufficient to escape the audit (the threshold) and safely evade the difference.

All papers referred to above assume that taxpayers are risk neutral. Chapter 3 considers the case of risk averse agents by assuming a kinked linear utility function. This specification implies that the marginal utility of the prospective gains from evasion is lower (in absolute value) than the marginal disutility of the corresponding prospective losses. The kinked utility function can be given two economic interpretations. One is in terms of non-pecuniary costs of evasion. The other is in terms of portfolio selection theory, where the measure of risk is the expected value of loss.

---

[2]When considering the general problem, the government controls taxes and penalties as well as the audit instruments.

When individuals have the same attitude towards risk, it is shown that the nature of the optimal audit policy under risk neutrality is preserved under risk aversion: income reports below a given threshold are audited at the probability level just sufficient to induce truthful reporting, whereas those above it are not audited. As illustrated above, the outcome is that the *effective* tax schedule is quite regressive.

Instead, if attitudes towards risk vary across taxpayers, the numerical results show that the optimal audit policy causes only a limited regressive bias, for income reports above the threshold meet a positive probability of audit.

A recent innovation of the Italian tax code gives rise to chapter 4. In 1989 the Italian government introduced a tax enforcement mechanism, called Presumptive Income Coefficients (PIC), with the objective of reducing tax evasion in the non-corporate sector.

Consider a homogeneous group of professionals or firms operating in the same sector of activity. The rationale of the mechanism is simple and is based on two considerations. The first is that sales revenues are easily concealed to the taxing authority whereas production costs are not. The second is that when examining various firms, one observes a similar relation between production costs and the corresponding revenues. Hence it is possible to estimate actual revenues and income by applying simple algorithms on observed production costs. The choice of the Italian tax administration has been that of determining the PIC by regressing revenues on production costs over a sample of taxpayers.

The PIC audit policy works as follows. For each group of taxpayers the taxing authority publishes the list of PIC, one for each type of expenditure. The application of the coefficients to production costs then determines the so called *presumptive income.* If reported income is at least as high as the presumptive income, the taxpayer is subjected to random audits. Otherwise the taxpayer is expected to pay the tax on presumptive income, unless he or she demonstrates that the actual income is lower than presumptive income.

The purpose of chapter 4 is to provide a theoretical analysis of the PIC tax enforcement policy. Following Reinganum and Wilde (1985), the *game* between the taxing authority and taxpayers is modelled in a principal-agent framework. The audit policy is restricted to the simple mechanism where income reports below and above a given mark-up on production costs are audited at different probabilities. When the mark-up coefficient is meant to give an *estimate* of taxable income, its interpretation is that of *presumptive income coefficient.*

When labour supply is fixed, it is shown that the net revenue maximizing audit policy is such that income reports below the the corresponding mark-up on production costs are audited at the probability level just sufficient to induce truthful reporting, whereas those above it are never audited. Also, when the mark-up coefficient is constrained to be a presumptive income coefficient, the best audit policy may collapse into random audits.

The numerical computations of the model with endogenous labour supply show that the properties of the optimal policy crucially depend on the size of taxpayers' average income relative to the size of unit audit costs. If average income is relatively high the best policy is random audits. If average income is relatively low, the two probability policy is effective: income reports below the corresponding mark-up on production costs meet the probability of audit just sufficient to induce honest behaviour, whereas those above it face a lower, but positive, level of enforcement. Also, the optimal mark-up coefficient is generally low and is lower than the presumptive income coefficients. This is an important result, especially when one considers that in taxpayers' minds a high value of the mark-up is likely to be perceived as *unfair.*

References

Border, Kim C. and Joel Sobel, 1987, Samurai accountant: a theory of auditing and plunder, Review of Economic Studies 54, 525–540.

Chander Parkash and Louis Wilde, 1992, A general characterization of optimal income taxation and enforcement, California Institute of Technology, Social Science Working Paper 791.

Cournot, A., 1838, Récherches sur les principes mathématiques de la théorie des richesses, L. Hachette, Paris.

Cremer, H., Marchand, M. and P. Pestieu, 1990, Evading, auditing and taxing: the equity-compliance tradeoff, Journal of Public Economics 43, 67–92.

Katz, Michael L. and Harvey S. Rosen, 1983, Tax analysis in an oligopoly model, NBER Working Paper, No. 1088.

Myles, Gareth D., 1989, Imperfect competition and the taxation of intermediate goods, Public Finance 44, 62-74.

Reinganum, Jennifer F. and Louis L. Wilde, 1985, Income tax compliance in a principal-agent framework, Journal of Public Economics 26, 1–18.

Sanchez, Isabel and Joel Sobel, 1990, Hierarchical design and enforcement of income tax policies, Southern European Economics Discussion Series, Universidad del Pais Vasco, D.P. 86.

Seade, Jesus, 1985, Profitable cost increases and the shifting of taxation: Equilibrium responses of markets in oligopoly, Warwick Economic Research Paper, No. 260.

Stern, Nicholas, 1987, The effects of taxation, price control and government contracts in oligopoly and monopolistic competition, Journal of Public Economics 32, 133–158.

Suits, D.B. and R.A. Musgrave, 1953, Ad valorem and unit taxes compared, Quarterly Journal of Economics 67, 598-604.

Wicksell, Knut, 1896, Finanztheoretische Untersuchungen und das Steuerwesen Schwedens, Fisher, Jena.

Chapter 1

Ad Valorem and Specific Taxation in a Model of Vertically Related Oligopolies

1.1 Introduction

This chapter examines the impact of ad valorem and specific taxes levied on producers in a model of vertically related oligopolies, where a downstream industry produces a final good using the output of an upstream industry as an input. The scope of the analysis is twofold. The first objective is to characterize tax incidence on producers' net prices and profits and to determine the conditions under which taxes are overshifted and profits are increased. The second goal is to compare the various tax instruments in terms of revenue collected and the effect on the price of the consumption good.

Katz and Rosen (1983), Seade (1985) and Stern (1987) examine the effects of levying *specific* taxes on producers in the homogeneous-product conjectural-variation oligopoly framework. The basic model assumes a fixed number of identical firms and constant returns to scale. With perfect competition price overshifting never occurs and therefore also the possibility of profits raising with taxation is ruled out. When the market is monopolized taxation may increase producer's net price but a positive effect on profits can never occur. Their main finding is that in oligopoly both the net price and profits may increase with taxation. Within the same framework, a recent paper by Delipalla and Keen (1992) focuses on the comparison between ad valorem and specific taxation, showing that the latter is more likely to cause price overshifting and to raise profits than the former.

In accordance with these previous studies, which consider only the downstream stage of production, tax incidence in the downstream industry turns out to be governed by final demand conditions and downstream market structure. The novel results concern the effects of taxation on the intermediate good price and upstream profits. These are shown to be related to upstream and downstream market structure, final demand conditions and input substitution in downstream production. Also, specific taxation is more likely to cause price overshifting and to raise profits than ad valorem taxation.

Stern (1987) condiders also a model with free entry where in equilibrium profits are zero and the focus is on the effects of taxation upon output per firm, price and the number of firms. The impact of entry for tax incidence is also analyzed by Besley (1989) and Delipalla and Keen (1992). They show that taxation may induce entry and that price overshifting is more likely in the free entry case than with firms fixed in number. Myles (1987) explores the effects of taxation in an oligopolistic industry with free entry when taxes are levied in a related competitive industry, the interaction between the two markets arising via consumers' tastes.

Tax incidence (only specific taxation) in vertically related industries has been analysed by Myles (1989), who considers in a general equilibrium framework[1] the two polar cases of downstream (upstream) monopoly and upstream (downstream) competitive sector. He then focuses on tax reforms, where the government aims at improving welfare by levying taxes (and subsides) on the final and intermediate goods while raising no revenue. Konishi (1990) examines a model in which a number of competitive industries supplies inputs to a free entry Cournot oligopoly producing a final good, showing that market performance can be improved by taxing the intermediate inputs. Tax reforms within the framework developed in this chapter are examined in chapter 2.

Central to the analysis is the comparison between ad valorem and specific taxes. Suits and Musgrave (1953) compare ad valorem and specific taxation under monopoly, showing that if a specific and an ad valorem tax result in the same final good price, then the latter raises a higher revenue.[2] This chapter generalizes this approach to the case of vertically related oligopolies. Propositions 1 and 2 show that the revenue from an ad valorem tax is always higher than the revenue from a specific tax levied in the same industry, when both taxes result in the same price for the consumption good. This result holds also when downstream producers price at marginal cost, provided that upstream producers price above it. The comparison between taxes of the same type levied in both industries is made assuming Leontief technology and isoelastic final demand. The revenue from downstream ad valorem taxation can be higher or lower than the yield from upstream ad valorem taxation, and which case applies depends on upstream and downstream market structure, production coefficients and the elasticity of final demand, see proposition 3. Finally, proposition 4 shows that upstream and

[1]Myles's model is general equilibrium because it takes account of the impact of monopoly profits on final demand.

[2]The recognition that ad valorem and specific taxes are not equivalent under monopoly dates back to Cournot (1838) and Wicksell (1896).

downstream specific taxation always raise the same revenue, when both taxes result in the same price for the consumption good.

Delipalla and Keen (1992) address four different issues regarding the comparison between ad valorem and specific taxation. First, they examine how changing the balance between the two forms of taxation, while raising the same revenue, affects prices and profits, the result being that a local shift towards ad valorem taxation reduces both the consumer price and profits. Second, taxation is used for correcting market inefficiency due to imperfect competition, and it is shown that both ad valorem and specific taxes can be used to attain marginal cost pricing, provided that the government is unrestricted in its ability to levy lump sum taxation. Third, they look at a Ramsey type problem, where the objective is to determine the balance between the two forms of taxes that minimizes the welfare loss of raising a given revenue. Finally, they search for the mix of taxes that maximizes government's revenue.

Kay and Keen (1983) look at the optimal balance between ad valorem and specific taxes in a model of perfect competition with endogenous quality and in a model of monopolistic competition with product variety. They show that prices are affected more by specific than ad valorem taxes, while ad valorem taxes are more powerful in controlling product quality and variety. Diericks, Matutes and Neven (1988) compare specific and ad valorem taxation in a model of Cournot oligopoly where firms differ in productive efficiency, showing that ad valorem taxes are more likely to improve aggregate productive efficiency by raising market shares of low cost firms.

The work is organized as follows. Section 1.2 describes the model and derives equilibrium prices and profits. Comparative statics begins with section 1.3, which illustrates the effects of taxation on the final and intermediate good prices. Section 1.3.1 gives the conditions for price overshifting. To illustrate the role of market structure, final demand conditions and downstream technology for tax incidence, section 1.3.2 examines three typical final demand curves, namely isoelastic, exponential and linear demand. The formulae expressing the impact of taxation on profits and the conditions for profitable tax increases are contained in section 1.4. The comparison of tax instruments in terms of revenue collected and the impact on the final good price is made in section 1.5. Conclusions are given in section 1.6.

1.2 Equilibrium prices and profits

The model consists of two vertically related industries, where a downstream industry produces a final good using two inputs: labour and an intermediate good produced by an upstream sector. Labour is the only input in upstream production. In both industries product is homogeneous, technology is constant returns to scale and no entry / exit occurs so that the number of firms is fixed.[3] Each oligopolist has a conjecture about the way the other firms in the same industry will change their output levels in response to changes in its own output. These conjectural variations capture the degree of competition among oligopolists belonging to the same industry. This approach has the merit of encompassing a wide range of market structures as particular cases, including the two polar cases of monopoly and competitive pricing. Symmetry is also assumed: oligopolists belonging to the same industry have identical cost functions and conjectures.

Competition between industries is modelled assuming that downstream oligopolists act as perfect competitors in the market for the intermediate good. This

[3] Quoting Katz and Rosen (1983, p. 10): "One may think of the model in two ways. First, it can be viewed as a short run analysis of a market in which capital stocks are fixed. Second, it can be viewed as a long run analysis of a market in which existing firms can adjust the levels of all productive inputs but sufficiently high barriers exist to preclude entry of new firms".

means that upstream oligopolists move first by setting the price for the intermediate good and downstream oligopolists move second by setting the price of the final good taking the price of input as given. This is what Waterson (1980a) refers to as *arms-length* pricing. The model is solved in three steps. First, downstream profits are maximized for given price of the intermediate good. Second, the derived demand facing the upstream industry is obtained and finally upstream profits are maximized.

The market inverse demand function for the final good is

$$q = q(mx) \tag{1.1}$$

where m is the number of downstream firms, x the output of each firm and q its price; mx is total output. It is assumed that $q(mx) \in C^3$, $q(0) > 0$ and $q'(mx) < 0$.

Each firm is assumed to maximize profits

$$\pi^d = \Big[(1 - t_v)q(mx) - t_s - C(p, w)\Big]x \tag{1.2}$$

where p is the price of the intermediate good, w the wage rate and $C(p, w) \in C^3$ is the average (and marginal) cost function, which is assumed to be nondecreasing, linearly homogeneous and concave in input prices. Two kind of taxes are levied on downstream producers: a specific tax t_s and an ad valorem tax at a tax inclusive rate $t_v < 1$. The cost function embodies the assumption that downstream producers are price takers in the input markets. The assumption of arms-length pricing implies that in the market for the intermediate good there is no market power on the demand side. The labour market is assumed to be competitive and the supply for labour to be perfectly elastic, so that the wage rate is given.

Each oligopolist conjectures that changes in its own output will cause the other firms to respond by changing outputs levels as well. Assuming linear responses the conjectural variation is constant and is defined as

$$\frac{\partial(m-1)x}{\partial x} = v^d, \qquad -1 \leq v^d \leq m-1 \tag{1.3}$$

$v^d = -1$ corresponds to Bertrand behaviour or marginal cost pricing; Cournot conjectures give $v^d = 0$; when oligopolists collude $v^d = m-1$; if the industry is monopolized then $v^d = 0$ and $m = 1$.

The first order condition for a maximum of (1.2) is that *perceived* marginal profits are zero

$$(1-t_v)\Big[q(mx) + xq'(mx)(1+v^d)\Big] - t_s - C(p,w) = 0 \tag{1.4}$$

This can be written as

$$q(mx)\left[1 - \frac{\gamma^d}{\varepsilon^d(mx)}\right] - \frac{C(p,w)+t_s}{1-t_v} = 0 \tag{1.5}$$

where $\gamma^d \equiv (1+v^d)/m$, thus $\gamma^d \in [0,1]$. $\varepsilon^d(mx) \equiv -q(mx)/[mx\, q'(mx)]$ is the market elasticity of demand, whereas each oligopolist perceives the elasticity to be $\varepsilon^d(mx)/\gamma^d$.

The second order condition is

$$E^d(mx) < \frac{2}{\gamma^d} \tag{1.6}$$

where $E^d(mx) \equiv -mx\, q''(mx)/q'(mx)$ is Seade's (1985) elasticity of the slope of inverse demand.

Eq. (1.5) implicitly defines the equilibrium output mx. The condition for this to be positive, Stern's (1987, eq. 5) existence condition, is

$$\begin{aligned} \varepsilon^d(mx) &> \gamma^d \\ t_v &< 1 \\ t_s &> -C(p,w) \end{aligned} \tag{1.7}$$

This requires the perceived elasticity to be greater than one; market elasticity can instead be less than one if $\gamma^d < 1$.

Following Seade (1985, eq. 14), a stability condition can be the following

$$E^d(mx) < \frac{1}{\gamma^d} + 1 \tag{1.8}$$

This ensures stability in response to common (symmetric) disturbances to the equilibrium. In other words, were all firms forced to expand (reduce) output away from the equilibrium by the same amount $\delta > 0$ ($\delta < 0$), then the stability condition (1.8) would ensure that perceived marginal profits become negative (positive), thus giving the oligopolists an incentive to reduce (increase) output. Notice that $1/\gamma^d + 1 \leq 2/\gamma^d$, thus the stability condition implies the second order condition.

The comparative statics below will make extensive use of the equivalent stability condition

$$F^d(q) > 1 - \frac{\varepsilon^d(q)}{\gamma^d}, \qquad \text{where} \quad F^d(q) = \frac{q\varepsilon^d_q(q)}{\varepsilon^d(q)} \tag{1.9}$$

see Stern (1987, eq. 6); $\varepsilon^d(q) \equiv -q\,\chi'(q)/\chi(q)$ is the elasticity of the direct demand $mx = \chi(q)$, ε^d_q is the derivative of ε^d with respect to q and $F^d(q)$ is the price elasticity of the elasticity of direct demand. The relation between (1.8) and (1.9) comes from $\varepsilon^d(q)E^d(mx) = 1 + \varepsilon^d(q) - F^d(q)$, see appendix A.1.

In terms of the final good price the first order condition (1.5) can be written as

$$q\left[1 - \frac{\gamma^d}{\varepsilon^d(q)}\right] - \frac{C(p,w) + t_s}{1 - t_v} \equiv f(q,p,w,t_v,t_s) = 0 \tag{1.10}$$

This is an implicit function of q. If the stability condition (1.9) is met, then the partial derivative of (1.10) with respect to q is positive (see (1.12) below), therefore the conditions of the implicit function theorem apply and the equilibrium price can be defined as

$$q = \phi(p,w,t_v,t_s) \tag{1.11}$$

Letting subscripts denote partial derivatives (so f_q is the partial derivative of $f(.)$ with respect to q)

$$\begin{aligned}
f_q &= 1 - \frac{\gamma^d}{\varepsilon^d(q)}[1 - F^d(q)] > 0 \\
\phi_p &= -\frac{f_p}{f_q} = \frac{C_p(p,w)}{1-t_v}\frac{1}{f_q} > 0 \\
\phi_{tv} &= -\frac{f_{tv}}{f_q} = \frac{C(p,w)+t_s}{(1-t_v)^2}\frac{1}{f_q} = \frac{q}{1-t_v}\left[1 - \frac{\gamma^d}{\varepsilon^d(q)}\right]\frac{1}{f_q} > 0 \\
\phi_{ts} &= -\frac{f_{ts}}{f_q} = \frac{1}{1-t_v}\frac{1}{f_q} > 0
\end{aligned} \tag{1.12}$$

where $f_q > 0$ from the stability condition (1.9).

The next step is to determine the equilibrium price for the intermediate good. Consider n identical upstream firms. Assuming that downstream oligopolists are the only customers of upstream firms, the market demand facing the upstream industry is given by the sum of downstream conditional input demands for the intermediate good. By Shephard's lemma the (direct) derived demand for the intermediate good is equal to $ny = C_p(p,w)\chi(q)$, where y is the output of each upstream firm. After substituting for (1.11) this becomes

$$ny = C_p(p,w)\chi\big(\phi(p,w,t_v,t_s)\big) \equiv \psi(p,w,t_v,t_s) \tag{1.13}$$

Derived demand is negatively sloped if

$$\psi_p(p,w,t_v,t_s) = C_{pp}\chi + C_p\chi'\phi_p < 0 \tag{1.14}$$

A sufficient condition for ψ_p to be negative and finite for each $p > 0$ is that inputs are not perfect substitutes in downstream production. Then $C_{pp} \leq 0$ (and finite), $C_p > 0$, $\phi_p > 0$; also $\chi' = 1/q' < 0$ from (1.1). If $\psi_p < 0$ then (1.13) has an inverse, the inverse derived demand

$$p = p(ny,w,t_v,t_s) \tag{1.15}$$

Let i_s be the specific tax and i_v the ad valorem tax levied on upstream producers. Profits of a representative oligopolist are

$$\pi^u = \Big[(1-i_v)p(ny,w,t_v,t_s) - i_s - wa_{Ly}\Big]y \tag{1.16}$$

where wa_{Ly} is the constant average (and marginal) cost of labour, which is assumed to be the only input.

The first order condition for a maximum of (1.16) can be written as

$$p\left[1-\frac{\gamma^u}{\varepsilon^u(p,w,t_v,t_s)}\right]-\frac{wa_{Ly}+i_s}{1-i_v} \equiv g(p,w,t_v,t_s,i_v,i_s)=0 \tag{1.17}$$

where

$$\begin{aligned}\varepsilon^u(p,w,t_v,t_s) &\equiv -\frac{p\psi_p}{\psi} = \\ &= \frac{wC_w(p,w)}{C(p,w)}\sigma(p,w)+\frac{pC_p(p,w)}{C(p,w)+t_s}\varepsilon^d(q)\omega(q)\end{aligned} \tag{1.18}$$

and

$$\omega(q) \equiv \frac{1-\frac{\gamma^d}{\varepsilon^d(q)}}{1-\frac{\gamma^d}{\varepsilon^d(q)}[1-F^d(q)]} > 0 \tag{1.19}$$

$\varepsilon^u(.)$ is the price elasticity of the direct derived demand (1.13);[4] $\gamma^u \equiv (1+v^u)/n$, the conjectural variation of upstream oligopolists being defined as $v^u = \partial[(n-1)y]/\partial y$, where $-1 \le v^u \le n-1$ and constant; thus $\gamma^u \in [0,1]$. $\sigma(p,w)$ is the elasticity of input substitution between labour and the intermediate good in

[4]The elasticity of derived demand under profit maximizing monopoly has been first presented by Yeung (1972); Waterson (1980b) extends it to the constant returns, homogeneous-product conjectural-variation oligopoly model. Assuming isoelastic final demand they show that the expression for the elasticity of derived demand is identical to the corresponding expression under perfect competition. This point is taken up by de Meza (1982), who shows that with a general final demand curve market structure does influence the elasticity of derived demand. To see this in terms of eq. (1.18) notice that downstream market structure, represented by γ^d, influences $\omega(q)$ and q, see eqs. (1.19) and (1.10). If final demand is of constant elasticity then ε^d is constant, $F^d = 0$ and $\omega(q) \equiv 1$, thus (1.18) is independent of γ^d.

downstream production, $q = \phi(p, w, t_v, t_s)$ from eq. (1.11) and $\omega(q)$ is positive by the existence and stability conditions (1.7) and (1.9). The elasticity (1.18) is derived in appendix A.2.

The existence and stability conditions are respectively

$$\varepsilon^u > \gamma^u, \qquad i_s > -wa_{Ly}, \qquad i_v < 1 \tag{1.20}$$

$$F^u > 1 - \frac{\varepsilon^u}{\gamma^u} \qquad \text{where} \qquad F^u \equiv \frac{p\varepsilon_p^u}{\varepsilon^u} \tag{1.21}$$

F^u is the price elasticity of the elasticity of derived demand (1.13); ε_p^u is the derivative of ε^u with respect to p.

When the stability condition (1.21) is met $g_p > 0$, thus the conditions of the implicit function theorem apply and eq. (1.17) defines the equilibrium price for the intermediate good

$$p = \xi(w, t_v, t_s, i_v, i_s) \tag{1.22}$$

The equilibrium price for the final good is determined by inserting (1.22) into (1.11). Substituting for the equilibrium prices into the respective demand functions gives the equilibrium quantities. Finally, plugging equilibrium prices and quantities into (1.2) and (1.16) gives the equilibrium profits of downstream and upstream oligopolists respectively.

This framework can now be employed to carry out the analysis of tax incidence. Eq. (1.17) determines the effect of taxes on the intermediate good price; the impact of taxation on the final good price is governed by eqs. (1.17) and (1.10). Katz and Rosen (1983), Seade (1985), Stern (1987) and Delipalla and Keen (1992) consider only one stage of production, which in the context of the present model corresponds to the downstream sector, and therefore their analysis is limited to the comparative statics of eq. (1.10) with p exogenous. The original feature of the model presented here is the introduction of the upstream oligopoly setting the price of the intermediate good which makes p endogenously determined by eq. (1.17).

1.3 The impact of taxation on prices

The effect of taxation on upstream and downstream prices is shown in table 1.1. The first three rows contain the determinants of tax incidence. All these elements have been defined in the previous section. The first row shows the stability condition for downstream firms and the elasticity of the elasticity of final demand. The second contains the elasticity of derived demand. The third row gives the stability condition for upstream firms and the elasticity of the elasticity of derived demand. The rest of the table shows in the first column the partial derivatives expressing the impact of taxation on q and p. The second and the third columns contain respectively the *indicator shifting coefficients* and the conditions for price overshifting that will be described in subsection 1.3.1.

The partial derivatives with respect to the taxes levied on upstream producers are all positive: i_v and i_s raise upstream costs, thus p raises as well and this in turn represents a cost shift for downstream oligopolists (provided that $C_p > 0$) who then increase q.

Next consider the effect of taxes levied on downstream producers on the price for the intermediate good. Obviously, if upstream producers price at marginal cost ($\gamma^u = 0$) then t_v and t_s do not affect p. When upstream oligopolists price above marginal cost ($\gamma^u \neq 0$) then $\partial p/\partial t_v \gtreqless 0$ iff $\varepsilon^u_{tv} \lesseqgtr 0$ and $\partial p/\partial t_s \gtreqless 0$ iff $\varepsilon^u_{ts} \lesseqgtr 0$. In words, a tax on downstream oligopolists raises p if and only if it lowers the elasticity of derived demand. The expressions for ε^u_{tv} and ε^u_{ts} are

$$\varepsilon^u_{tv} = \frac{pC_p}{C + t_s}(F^d\omega + q\omega_q)\frac{\varepsilon^d}{q}\phi_{tv} \tag{1.23}$$

Table 1.1: The impact of taxation on prices

$f_q = 1 - \frac{\gamma^d}{\varepsilon^d}(1 - F^d) > 0, \qquad F^d = \frac{q\varepsilon^d_q}{\varepsilon^d}$		
$\varepsilon^u = \frac{wC_w(p,w)}{C(p,w)}\sigma(p,w) + \frac{pC_p(p,w)}{C(p,w)+t_s}\varepsilon^d(q)\omega(q), \qquad \omega(q) = \frac{1 - \frac{\gamma^d}{\varepsilon^d(q)}}{1 - \frac{\gamma^d}{\varepsilon^d(q)}[1 - F^d(q)]}$		
$g_p = 1 - \frac{\gamma^u}{\varepsilon^u}(1 - F^u) > 0, \qquad F^u = \frac{p\varepsilon^u_p}{\varepsilon^u}$		
$\frac{\partial p}{\partial t_v} = -\gamma^u \frac{p\varepsilon^u_{tv}}{(\varepsilon^u)^2}\frac{1}{g_p}$		$\frac{\partial p}{\partial t_v} > 0$ iff $F^d < 0^a$
$\frac{\partial p}{\partial t_s} = -\gamma^u \frac{p\varepsilon^u_{ts}}{(\varepsilon^u)^2}\frac{1}{g_p}$		if $F^d \leq 0$ then $\frac{\partial p}{\partial t_s} > 0^a$
$\frac{\partial p}{\partial i_v} = \frac{1}{1-i_v}ps^u_{iv}$	$s^u_{iv} = \left(1 - \frac{\gamma^u}{\varepsilon^u}\right)\frac{1}{g_p}$	$s^u_{iv} > 1$ iff $F^u < 0$
$\frac{\partial p}{\partial i_s} = \frac{1}{1-i_v}s^u_{is}$	$s^u_{is} = \frac{1}{g_p}$	$s^u_{is} > 1$ iff $F^u < 1$
$\frac{\partial q}{\partial t_v} = \frac{1}{1-t_v}\left(q + C_p\frac{\partial p}{\partial t_v}\right)s^d_{tv}$	$s^d_{tv} = \frac{q\left(1 - \frac{\gamma^d}{\varepsilon^d}\right) + C_p\frac{\partial p}{\partial tv}}{q + C_p\frac{\partial p}{\partial tv}}\frac{1}{f_q}$	$s^d_{tv} > 1$ iff $F^d < 0^b$
$\frac{\partial q}{\partial t_s} = \frac{1}{1-t_v}\left(1 + C_p\frac{\partial p}{\partial t_s}\right)s^d_{ts}$	$s^d_{ts} = \frac{1}{f_q}$	$s^d_{ts} > 1$ iff $F^d < 1$
$\frac{\partial q}{\partial i_v} = \frac{1}{1-t_v}\frac{1}{1-i_v}pC_p s^u_{iv}s^d_{iv}$	$s^d_{iv} = \frac{1}{f_q}$	$s^d_{iv} > 1$ iff $F^d < 1$
$\frac{\partial q}{\partial i_s} = \frac{1}{1-t_v}\frac{1}{1-i_v}C_p s^u_{is}s^d_{is}$	$s^d_{is} = \frac{1}{f_q}$	$s^d_{is} > 1$ iff $F^d < 1$

(a) assuming that $\mid F^d\omega \mid > \mid q\omega_q \mid$
(b) see appendix A.3

$$\varepsilon_{ts}^{u} = -\frac{pC_p}{(C+t_s)^2}\varepsilon^d\omega + \frac{pC_p}{C+t_s}(F^d\omega + q\omega_q)\frac{\varepsilon^d}{q}\phi_{ts} \tag{1.24}$$

Consider first the tax t_v. Eq. (1.23) takes the sign of $F^d\omega + q\omega_q$: ω and q are positive but F^d and ω_q may take positive and negative values. Notice that when final demand is of constant elasticity then $F^d = 0$, which implies $\omega \equiv 1$ and $\omega_q = 0$, see eq. (1.19), thus $\varepsilon_{tv}^u = 0$ and t_v does not affect p. In general, a relation between the signs of F^d and ω_q cannot be found; also, there is no economic interpretation for ω_q, which contains the derivative of F^d and therefore the third derivative of the demand function $\chi(q)$. However, if one assumes that $\mid F^d\omega \mid > \mid q\omega_q \mid$, then $\varepsilon_{tv}^u \gtreqless 0$ iff $F^d \gtreqless 0$, which implies the condition $\partial p/\partial t_v > 0$ iff $F^d < 0$ shown in table 1.1. Section 1.3.2 below shows that this result occurs when some typical demand functions are considered.

As for the specific tax t_s, notice that if final demand is isoelastic ($F^d = 0$, $\omega_q = 0$) then $\varepsilon_{ts}^u < 0$ and t_s raises p. Assuming $\mid F^d\omega \mid > \mid q\omega_q \mid$ implies that if $F^d < 0$ then $\varepsilon_{ts}^u < 0$ and $\partial p/\partial t_s > 0$, which is only a sufficient condition because the first term in (1.24) is negative.

Finally, consider the impact of t_v and t_s on the price for the final good

$$\frac{\partial q}{\partial t_v} = \phi_{tv} + \phi_p\frac{\partial p}{\partial t_v} = \frac{1}{1-t_v}\left[q\left(1-\frac{\gamma^d}{\varepsilon^d}\right) + C_p\frac{\partial p}{\partial t_v}\right]\frac{1}{f_q} \tag{1.25}$$

$$\frac{\partial q}{\partial t_s} = \phi_{ts} + \phi_p\frac{\partial p}{\partial t_s} = \frac{1}{1-t_v}\left(1 + C_p\frac{\partial p}{\partial t_s}\right)s_{ts}^d \tag{1.26}$$

In both derivatives, the first term is positive and expresses the effect of increased tax liabilities, or the *direct* impact of taxation. The second term represents the *indirect* effect that goes through the change in the price of the intermediate good. As illustrated above, if $\gamma^u = 0$ then this term is zero, thus (1.25) and (1.26) are positive. In general, since $\partial p/\partial t_v$ and $\partial p/\partial t_s$ may take negative values, it cannot be excluded that (1.25) and (1.26) are negative as well. However, this seems a remote possibility and one may reasonably assume that the direct effect dominates the indirect effect so that $q(1-\gamma^d/\varepsilon^d)+C_p(\partial p/\partial t_v)>0$ and $1+C_p(\partial p/\partial t_s)>0$.

For the sake of comparison, notice that if $\partial p/\partial t_s = 0$, the expression for $\partial q/\partial t_s$ equals the corresponding expression in Seade (1985), Stern (1987) and Delipalla and Keen (1992). Also, if $\partial p/\partial t_v = 0$, the derivative $\partial q/\partial t_v$ equals the corresponding expression in Delipalla and Keen (1992).

1.3.1 The indicator shifting coefficients

The derivatives in the first column of table 1.1 express the impact of taxation on producers' *gross* prices q and p. From the point of view of producers what really matters is the difference between the effect on the gross price and the tax induced cost shifts; in other words, the effect on the *net* price.

Consider first the upstream industry. For each unit of intermediate good produced and sold the manufacturer receives p from his or her customer, pays $i_s + i_v p$ to the government and pays wa_{Ly} to the workers. The price p is the *gross* price, or revenue per unit of output. Let $G^u(p, w, i_s, i_v) \equiv i_s + i_v p + wa_{Ly}$ be total outlays per unit of output, where $p = \xi(w, t_v, t_s, i_v, i_s)$, see eq. (1.22). The producers' net price is defined as the difference between the gross price p and total outlays G^u, or profits per unit of output. A tax is said to cause *price overshifting* if it increases producers' net price.

Following Seade (1985) define the *shifting coefficient* of a tax $\tau \in T = \{i_v, i_s\}$ as the ratio between the change in the gross price p and the change in total outlays G^u

$$S^u_\tau \equiv \frac{dp}{dG^u} = \frac{\partial p/\partial \tau}{i_v \frac{\partial p}{\partial \tau} + \frac{\partial G^u}{\partial \tau}}, \qquad \tau \in T = \{i_v, i_s\} \tag{1.27}$$

Price overshifting occurs if and only if $S^u_\tau > 1$. Applying the following monotonically increasing transformation to the shifting coefficients (1.27) one obtains the *indicator shifting coefficients* which are reported in table 1.1

$$s_\tau^u \equiv \frac{(1-i_v)S_\tau^u}{1-i_v S_\tau^u} \equiv (1-i_v)\frac{\partial p/\partial \tau}{\partial G^u/\partial \tau}, \qquad \tau \in T = \{i_v, i_s\} \tag{1.28}$$

Since $i_v < 1$, the property of these coefficients is that $s_\tau^u \gtreqless 1$ if and only if $S_\tau^u \gtreqless 1$; also, if $i_v = 0$ then $s_\tau^u = S_\tau^u$. The indicator shifting coefficients are convenient because they fit into the derivatives in the first column of table 1.1 and because they give the conditions for price overshifting shown in the third column.

If upstream oligopolists price at marginal cost ($\gamma^u = 0$), then taxes do not affect the net price and $s_{iv}^u = s_{is}^u = 1$. If $0 < \gamma^u \leq 1$ then overshifting may occur; note that the stability condition (1.21) does not rule out this possibility. Specific taxation i_s causes overshifting if and only if $F^u < 1$, whereas the corresponding condition for ad valorem taxation i_v is $F^u < 0$. Hence specific taxation is more likely to cause overshifting than ad valorem taxation.

Next consider downstream oligopolists. Let $G^d(q, p, w, t_s, t_v) \equiv t_s + t_v q + C(p, w)$ be total outlays per unit of output of each downstream producer, where $p = \xi(w, t_v, t_s, i_v, i_s)$ and $q = \phi(p, w, t_v, t_s)$, see eqs. (1.22) and (1.11). Again, price overshifting occurs when a tax causes the gross price q to increase more than total outlays. For any tax $\tau \in T = \{t_v, t_s, i_v, i_s\}$ the corresponding shifting coefficient is defined as

$$S_\tau^d \equiv \frac{dq}{dG^d} = \frac{\partial q/\partial \tau}{t_v \frac{\partial q}{\partial \tau} + C_p \frac{\partial p}{\partial \tau} + \frac{\partial G^d}{\partial \tau}}, \qquad \tau \in T = \{t_v, t_s, i_v, i_s\} \tag{1.29}$$

so that price overshifting occurs if and only if $S_\tau^d > 1$. The indicator shifting coefficient is

$$s_\tau^d \equiv \frac{(1-t_v)S_\tau^d}{1-t_v S_\tau^d} \equiv (1-t_v)\frac{\partial q/\partial \tau}{C_p \frac{\partial p}{\partial \tau} + \frac{\partial G^d}{\partial \tau}}, \qquad \tau \in T \tag{1.30}$$

Again, $t_v < 1$ implies that $s_\tau^d \gtreqless 1$ if and only if $S_\tau^d \gtreqless 1$; also, when $t_v = 0$ then $s_\tau^u = S_\tau^u$.

If downstream oligopolists price at marginal cost ($\gamma^d = 0$) then $s^d_\tau = 1$ for all $\tau \in T$. If $0 < \gamma^d \leq 1$ the conditions for price overshifting in the downstream industry depend entirely on the characteristics of final demand, as expressed by the price elasticity of the elasticity F^d. Table 1.1 shows that the ad valorem tax t_v causes overshifting if and only if $F^d < 0$. This result is not immediate, the proof and the assumptions introduced to obtain it are given in appendix A.3. The taxes t_s, i_v and i_s raise downstream unit profits if and only if $F^d < 1$. Therefore, specific downstream taxation t_s is more likely to cause price overshifting than ad valorem downstream taxation t_v, a result also noticed by Delipalla and Keen (1992, p. 357).

1.3.2 The determinants of tax incidence

It has been shown that tax incidence is governed by the price elasticities of final and derived demand. These factors are now examined in more detail.

Consider first the final good price. The conditions for overshifting shown in table 1.1 are $s^d_{tv} > 1$ iff $F^d < 0$ and $s^d_\tau > 1$ iff $F^d < 1$, $\tau = t_s, i_v, i_s$, where

$$F^d(q) = 1 + \varepsilon^d(q) - E^d(q), \qquad F^d \equiv \frac{q\varepsilon^d_q}{\varepsilon^d}, \quad E^d \equiv -\frac{q\chi''}{\chi'} \tag{1.31}$$

E^d is the elasticity of the slope of (direct) demand, see appendix A.1. If final demand is concave or linear ($\chi'' \leq 0$) then overshifting never occurs because $E^d \leq 0$ and $F^d > 1$. Overshifting may occur only with convex demand functions ($\chi'' > 0$). To illustrate, tables 1.2 and 1.3 show the results for three typical final demand curves: isoelastic (ISO), exponential (EXP) and linear (LIN) demand. The first two are strongly convex. ISO has the property that $F^d = 0$, thus $s^d_{tv} = 1$ and $s^d_\tau > 1$, $\tau = t_s, i_v, i_s$; EXP has the property that $F^d = 1$, hence $s^d_{tv} < 1$ and $s^d_\tau = 1$, $\tau = t_s, i_v, i_s$; finally, LIN has the property that $F^d > 1$ so that all taxes are undershifted.

Table 1.2: Isoelastic (ISO), Exponential (EXP) and Linear (LIN) final demand

	$mx = \chi(q)$	$\varepsilon^d(q)$	$F^d(q)$	elasticity of derived demand ε^u
ISO	$mx = (bq)^{-\bar{\varepsilon}}$	$\bar{\varepsilon}$	0	$\varepsilon^u(p,w,t_s) = \frac{wC_w}{C}\sigma + \frac{pC_p}{C+t_s}\bar{\varepsilon}$
EXP	$mx = A\exp^{-bq}$	bq	1	$\varepsilon^u(p,w,t_v) = \frac{wC_w}{C}\sigma + \frac{pC_p}{C+t_s}(\varepsilon^d - \gamma^d) \quad = \frac{wC_w}{C}\sigma + pC_p\frac{b}{1-t_v}$
LIN	$mx = A - bq$	$\frac{q}{A/b - q}$	$\frac{A/b}{A/b-q} > 1$	$\varepsilon^u(p,w,t_v,t_s) = \frac{wC_w}{C}\sigma + \frac{pC_p}{C+t_s}\frac{\varepsilon^d - \gamma^d}{1+\gamma^d} \quad = \frac{wC_w}{C}\sigma + \frac{pC_p}{(1-t_v)A/b - C - t_s}$

	ε^u_p	ε^u_{tv}	ε^u_{ts}
ISO	$(1-\sigma)(\bar{\varepsilon}-\sigma)\frac{wC_wC_p}{C^2} \qquad \forall t_v,\ \ t_s = 0$	0	–
EXP	$(1-\sigma)(\varepsilon^d - \gamma^d - \sigma)\frac{wC_wC_p}{C^2} + \frac{pC_p^2}{C}\frac{b}{1-t_v}$	+	0
LIN	$(1-\sigma)\left(\frac{\varepsilon^d-\gamma^d}{1+\gamma^d} - \sigma\right)\frac{wC_wC_p}{C^2} + \frac{pC_p^2}{C}\frac{\varepsilon^d_q}{(1-t_v)(1+\gamma^d)^2}$	+	+

Table 1.3: Isoelastic (ISO), Exponential (EXP) and Linear (LIN) final demand

	$\frac{\partial p}{\partial t_v}$	$\frac{\partial p}{\partial t_s}$	$s^d_{tv} > 1$ iff $F^d < 0$	$s^d_\tau > 1$ iff $F^d < 1$ $\tau = t_s, i_v, i_s$	$s^u_{iv} > 1$ iff $F^u < 0$	$s^u_{is} > 1$ iff $F^u < 1$
ISO	0	+	$s^d_{tv} = 1$	$s^d_\tau > 1$	$s^u_{iv} > 1$ iff $1 < \sigma < \bar{\varepsilon}$ or $\bar{\varepsilon} < \sigma < 1$	if $\bar{\varepsilon} \geq \sigma - 2$ then $s^u_{is} > 1$
EXP	–	0	$s^d_{tv} < 1$	$s^d_\tau = 1$	if $\sigma \leq 1$ and $\varepsilon^d \geq \sigma + \gamma^d$ or $\sigma \geq 1$ and $\varepsilon^d \leq \sigma + \gamma^d$ then $s^u_{iv} < 1$	if $\varepsilon^d \geq \sigma - 2 - \gamma^d$ then $s^u_{is} > 1$
LIN	–	–	$s^d_{tv} < 1$	$s^d_\tau < 1$	if $\sigma \leq 1$ and $\varepsilon^d \geq (1+\gamma^d)\sigma + \gamma^d$ or $\sigma \geq 1$ and $\varepsilon^d \leq (1+\gamma^d)\sigma + \gamma^d$ then $s^u_{iv} < 1$	

Next consider the intermediate good price. Tax incidence is determined by the partial derivatives of the elasticity of derived demand (1.18). The derivative of ε^u with respect to p is

$$\varepsilon^u_p = (1-\sigma)(\varepsilon^d\omega - \sigma)\frac{wC_wC_p}{C^2} + \frac{pC_p}{C}(F^d\omega + q\omega_q)\frac{\varepsilon^d}{q}\phi_p \tag{1.32}$$

This expression has been computed assuming that σ, the elasticity of substitution between labour and the intermediate good in downstream production, is constant. Hence $C(p, w)$ is a CES. When $\sigma = 0$ technology is Leontief, $\sigma = 1$ gives the Cobb-Douglas case, if $\sigma \to \infty$ inputs are perfect substitutes. The derivatives ε^u_{tv} and ε^u_{ts} are shown in (1.23) and (1.24) respectively.

Tax incidence on p is governed by (i) final demand conditions, expressed by ε^d and F^d, (ii) downstream taxes t_v and t_s, (iii) input substitution in downstream production σ, (iv) downstream market structure γ^d and (v) upstream market structure γ^u. With a general final demand curve though, it is difficult to identify the contribution of each factor, thus the three special cases will be examined.

A common property of the three demand functions, see table 1.2, is that ε^u is independent of γ^d. This means that downstream market structure does not affect the equilibrium price p as well as tax incidence.[5]

ISO implies that ε^u and p are independent of t_v whereas t_s raises p. EXP implies that t_s does not affect p, whereas t_v lowers p. With LIN both t_v and t_s decrease p. These results show that the *ad hoc* assumption introduced on p. 22, namely that $| F^d\omega | > | q\omega_q |$, is reasonable.

As for the ad valorem tax i_v, table 1.1 shows that $s^u_{iv} > 1$ iff $F^u < 0$. From the definition of F^u an equivalent condition is $s^u_{iv} > 1$ iff $\varepsilon^u_p < 0$. With ISO demand $\varepsilon^u_p < 0$ iff $(1-\sigma)(\bar{\varepsilon} - \sigma) < 0$, thus overshifting occurs iff $\bar{\varepsilon} > 1$ and $1 < \sigma < \bar{\varepsilon}$ or $\bar{\varepsilon} < 1$ and $\bar{\varepsilon} < \sigma < 1$. With EXP and LIN demands ε^u_p contains a term

[5] The result that with linear final demand the elasticity of derived demand is independent of downstream market structure is also noticed by de Meza (1982).

(the second) which is positive, thus only sufficient conditions for undershifting ($s^u_{iv} < 1$) can be found, see table 1.3.

As for the specific tax i_s, the overshifting condition is $s^u_{is} > 1$ iff $F^u < 1$, which can be written $p\varepsilon^u_p < \varepsilon^u$. After some algebra one finds that

$$\text{ISO} \quad s^u_{is} > 1 \quad \text{iff} \quad \sigma(\sigma - \bar{\varepsilon} - 2) < \frac{wC_w}{pC_p}\sigma + \frac{pC_p}{wC_w}\bar{\varepsilon}$$

$$\text{EXP} \quad s^u_{is} > 1 \quad \text{iff} \quad \sigma - \varepsilon^d - \gamma^d - 2 < \frac{wC_w}{pC_p}$$

Both terms on the right hand side are positive, thus a sufficient condition for overshifting is that the term on the left hand side is negative or zero. Corresponding conditions for LIN cannot be found.

1.4 The impact of taxation on profits

Taxation causes gross prices p and q to increase and this lowers equilibrium outputs, for final and derived demands are downward sloping. If producers' net price (profit per unit of output) falls then also profits must fall. But if taxation causes price overshifting (an increase in profits per unit of output) then the possibility of a positive impact on profits arises.

Considering only one stage of production, Seade (1985) and Stern (1987) have shown that taxation has always a negative impact on profits in the two polar cases of monopoly and perfect competition.[6] Profitable tax increases arise as a distinct possibility when the market is oligopolistic. Assuming constant returns to scale, Seade (1985) finds that specific taxation increases profits when the elasticity of the slope of inverse final demand is greater than two. The same condition is

[6]Taxation decreases profits in perfect competition provided that equilibrium profits are positive, for which decreasing returns to scale are required. If constant returns to scale are assumed, then equilibrium profits are zero and taxes have no effect on profits.

expressed by Stern (1987) as $F < 1 - \varepsilon$, where ε is the elasticity and F is the price elasticity of the elasticity of final demand.

1.4.1 Downstream profits

Table 1.4 shows the derivatives expressing the impact of taxation on profits and the corresponding conditions for profitable tax increases, see appendix A.4 for the computations.

When producers price at marginal cost ($\gamma^d = 0$) then all derivatives are zero, because profits are zero in equilibrium.[7] If oligopolists collude, or if the industry is monopolized ($\gamma^d = 1$), then the derivatives are negative and taxation reduces profits.

When the downstream market is oligopolistic ($0 < \gamma^d < 1$), the taxes t_s, i_v and i_s raise profits if and only if $F^d < 1 - \varepsilon^d$. As for ad valorem taxation t_v, only a sufficient condition can be found: if $F^d < \gamma^d - \varepsilon^d < 0$ then profits increase. With EXP and LIN demands $F^d \geq 0$, thus all taxes reduce profits. With ISO demand $F^d = 0$, hence t_v lowers profits, whereas the other taxes are profitable if and only if $\bar{\varepsilon} < 1$.

1.4.2 Upstream profits

If $\gamma^u = 0$ then $\partial p/\partial t_v = 0$, $\partial p/\partial t_s = 0$, $s^u_{iv} = 1$ and $s^u_{is} = 1$ so that all derivatives are zero. If $\gamma^u = 1$ taxation reduces profits.

When the upstream market is oligopolistic ($0 < \gamma^u < 1$), the taxes i_v and i_s increase profits if and only if $F^u < \gamma^u - \varepsilon^u$ and $F^u < 1 - \varepsilon^u$ respectively. The effect of downstream taxes t_v and t_s is difficult to assess. The derivatives contain a negative term, the second, which represents the fall in profits that arises from

[7] In terms of the formulae of table 1.4, this simply follows from $s^d_\tau = 1$, $\tau = t_v, t_s, i_v, i_s$.

Table 1.4: The impact of taxation on profits

Downstream profits	$0 < \gamma^d < 1$
$\dfrac{\partial(m\pi^d)}{\partial t_v} = [(1-\gamma^d)s^d_{tv} - 1]\left(q + C_p \dfrac{\partial p}{\partial t_v}\right) mx$	if $F^d < \gamma^d - \varepsilon^d$ then $\dfrac{\partial(m\pi^d)}{\partial t_v} > 0$
$\dfrac{\partial(m\pi^d)}{\partial t_s} = [(1-\gamma^d)s^d_{ts} - 1]\left(1 + C_p \dfrac{\partial p}{\partial t_s}\right) mx$	$\dfrac{\partial(m\pi^d)}{\partial t_s} > 0$ iff $F^d < 1 - \varepsilon^d$ assuming $1 + C_p(\partial p/\partial t_s) > 0$
$\dfrac{\partial(m\pi^d)}{\partial i_v} = [(1-\gamma^d)s^d_{iv} - 1]\dfrac{1}{1-i_v} s^u_{iv} p C_p mx$	$\dfrac{\partial(m\pi^d)}{\partial i_v} > 0$ iff $F^d < 1 - \varepsilon^d$
$\dfrac{\partial(m\pi^d)}{\partial i_s} = [(1-\gamma^d)s^d_{is} - 1]\dfrac{1}{1-i_v} s^u_{is} C_p mx$	$\dfrac{\partial(m\pi^d)}{\partial i_s} > 0$ iff $F^d < 1 - \varepsilon^d$
Upstream profits	$0 < \gamma^u < 1$
$\dfrac{\partial(n\pi^u)}{\partial t_v} = (1-i_v)\left[(1-\gamma^u)\dfrac{\partial p}{\partial t_v} + -\dfrac{p}{1-t_v}\dfrac{\gamma^u}{\varepsilon^u}(\varepsilon^d - \gamma^d)\dfrac{1}{f_q}\right] C_p mx$	if $\dfrac{\partial(n\pi^u)}{\partial t_v} > 0$ then $\dfrac{\partial p}{\partial t_v} > 0$ if $\dfrac{\partial p}{\partial t_v} \le 0$ then $\dfrac{\partial(n\pi^u)}{\partial t_v} < 0$
$\dfrac{\partial(n\pi^u)}{\partial t_s} = (1-i_v)\left[(1-\gamma^u)\dfrac{\partial p}{\partial t_s} + -\dfrac{p}{1-t_v}\dfrac{\gamma^u}{\varepsilon^u}\dfrac{\varepsilon^d}{q} s^d_{ts}\right] C_p mx$	if $\dfrac{\partial(n\pi^u)}{\partial t_s} > 0$ then $\dfrac{\partial p}{\partial t_s} > 0$ if $\dfrac{\partial p}{\partial t_s} \le 0$ then $\dfrac{\partial(n\pi^u)}{\partial t_s} < 0$
$\dfrac{\partial(n\pi^u)}{\partial i_v} = [(1-\gamma^u)s^u_{iv} - 1]pC_p mx$	$\dfrac{\partial(n\pi^u)}{\partial i_v} > 0$ iff $F^u < \gamma^u - \varepsilon^u$
$\dfrac{\partial(n\pi^u)}{\partial i_s} = [(1-\gamma^u)s^u_{is} - 1]C_p mx$	$\dfrac{\partial(n\pi^u)}{\partial i_s} > 0$ iff $F^u < 1 - \varepsilon^u$

The stability conditions require $F^d > 1 - \varepsilon^d/\gamma^d$ and $F^u > 1 - \varepsilon^u/\gamma^u$.
The existence conditions require $\varepsilon^d > \gamma^d$ and $\varepsilon^u > \gamma^u$.

output falling: downstream taxation increases q and reduces final output mx, and this in turn causes a fall in upstream production. The first term can instead take positive and negative values, depending respectively on whether taxation raises or decreases p. Therefore the net impact is undetermined. With EXP and LIN demands $\partial p/\partial t_v < 0$ and $\partial p/\partial t_s \leq 0$, thus upstream profits fall; with ISO demand t_v reduces profits (because $\partial p/\partial t_v = 0$) whereas t_s may be profitable (because $\partial p/\partial t_s > 0$).

1.5 Taxation and government's revenue

The two preceding sections have examined the effects of taxation on producers' net prices and total profits. This section takes the view of the government. When comparing a pair of tax instruments, the taxing authority considers the amount of revenue collected and the effect on the price for the final good. When both taxes result in the same price, the criteria is to regard as a better tax that raising more revenue. This is the approach adopted by Suits and Musgrave (1953). They compare ad valorem and specific taxation in monopoly, showing that the revenue from a specific tax is always smaller than the revenue from an ad valorem tax, when both taxes result in the same final good price (proposition 1, p. 598).[8]

[8]They also show that this result bears two implications. Firstly, the maximum revenue that can be obtained with a specific tax is smaller than the maximum revenue obtainable with an ad valorem tax (proposition 2, p. 599). Secondly, when the same revenue is raised using a specific and an ad valorem tax, the former always brings about a higher final price (proposition 3, p. 599). Delipalla and Keen (1992) examine a local version, or tax reform version, of the latter approach, termed "P-shift", consisting of a tax change "which has the feature of tilting the balance towards ad valorem taxation whilst leaving total tax payments at the initial equilibrium price unchanged" (Delipalla and Keen, 1992, pp. 357–58). They show, see proposition 3, that a P-shift from specific to ad valorem taxation leads to a strict reduction in the consumer price.

Chapter 2 will focus on a different but related approach, where tax instruments are compared in terms of the revenue collected and the effect on social welfare.

The first two propositions deal with the comparison between ad valorem and specific taxes levied in the same industry.

Proposition 1 *Consider a tax* i_v^* *and a tax* i_s^* *resulting in the same final good price* q^*. *Let* $R(i_v^*)$ *and* $R(i_s^*)$ *be the corresponding revenues. Then* $R(i_v^*) > R(i_s^*)$ *iff* $\gamma^u \in (0,1]$ *and* $R(i_v^*) = R(i_s^*)$ *iff* $\gamma^u = 0$.

Proof. The equilibrium price of the consumption good is defined by eq. (1.11), $q = \phi(p, w, t_v, t_s)$. If i_v^* and i_s^* are to result in the same price q^*, they must result in the same price for the intermediate good, say p^*, so that $q^* = \phi(p^*, w, t_v, t_s)$. From eq. (1.17) defining the equilibrium price of the intermediate good

$$\begin{aligned} p^* \left[1 - \frac{\gamma^u}{\varepsilon^u(p^*, w, t_v, t_s)}\right] &= \frac{wa_{Ly}}{1 - i_v^*} \\ p^* \left[1 - \frac{\gamma^u}{\varepsilon^u(p^*, w, t_v, t_s)}\right] &= wa_{Ly} + i_s^* \end{aligned}$$

Therefore the condition for i_v^* and i_s^* to result in q^* is

$$i_v^* = \frac{i_s^*}{p^*[1 - \gamma^u/\varepsilon^u(p^*, w, t_v, t_s)]} \tag{1.33}$$

Tax revenues are

$$\begin{aligned} R(i_v^*) &= i_v^* p^* C_p(p^*, w)\chi(q^*) \\ R(i_s^*) &= i_s^* C_p(p^*, w)\chi(q^*) \end{aligned}$$

Using (1.33)

$$\begin{aligned} R(i_v^*) - R(i_s^*) &= (i_v^* p^* - i_s^*) C_p(p^*, w)\chi(q^*) = \\ &= i_s^* \frac{\gamma^u}{\varepsilon^u(p^*, w, t_v, t_s) - \gamma^u} C_p(p^*, w)\chi(q^*) \geq 0 \end{aligned}$$

and this completes the proof. *Q.E.D.*

Obviously, if upstream producers price at marginal cost $\gamma^u = 0$ and the two taxes raise the same revenue. When upstream oligopolists price above marginal cost, the yield from i_v is always higher than the yield from i_s, whatever market structure in the downstream industry.

Proposition 2 *Consider a tax t_v^* and a tax t_s^* resulting in the same final good price q^*. Let $R(t_v^*)$ and $R(t_s^*)$ be the corresponding revenues. Let $p^{t_v^*}$ and $p^{t_s^*}$ be the price of the intermediate good under t_v^* and t_s^* respectively. Then $p^{t_v^*} \leq p^{t_s^*}$. Also, $R(t_v^*) = R(t_s^*)$ iff $\gamma^d = 0$ and $\gamma^u = 0$. In all other cases, i.e. if in at least one market producers price above marginal cost, $R(t_v^*) > R(t_s^*)$.*

Proof. From eq. (1.10) defining the equilibrium price of the final good

$$\begin{aligned} q^*\left[1 - \frac{\gamma^d}{\varepsilon^d(q^*)}\right] &= \frac{C(p^{t_v^*}, w)}{1 - t_v^*} \\ q^*\left[1 - \frac{\gamma^d}{\varepsilon^d(q^*)}\right] &= C(p^{t_s^*}, w) + t_s^* \end{aligned}$$

The condition for t_v^* and t_s^* to result in q^* is

$$t_v^* = \frac{C(p^{t_s^*}, w) - C(p^{t_v^*}, w) + t_s^*}{q^*[1 - \gamma^d/\varepsilon^d(q^*)]} \tag{1.34}$$

Tax revenues are

$$\begin{aligned} R(t_v^*) &= t_v^* q^* \chi(q^*) \\ R(t_s^*) &= t_s^* \chi(q^*) \end{aligned}$$

Using (1.34)

$$\begin{aligned} R(t_v^*) - R(t_s^*) &= (t_v^* q^* - t_s^*)\chi(q^*) = \\ &= t_s^* \frac{\gamma^d}{\varepsilon^d(q^*) - \gamma^d}\chi(q^*) + \frac{C(p^{t_s^*}, w) - C(p^{t_v^*}, w)}{1 - \gamma^d/\varepsilon^d(q^*)} \end{aligned} \tag{1.35}$$

The first term is greater than or equal to zero. The second term is positive or zero if and only if $p^{t_s^*} \geq p^{t_v^*}$. The equilibrium price for the intermediate good is determined by

$$p\left[1 - \frac{\gamma^u}{\varepsilon^u(p, w, t_v, t_s)}\right] = \frac{w a_{Ly} + i_s}{1 - i_v} \tag{1.36}$$

where

$$\varepsilon^u(p,w,t_v,t_s) = \frac{wC_w(p,w)}{C(p,w)}\sigma(p,w) + \frac{pC_p(p,w)}{C(p,w)+t_s}\varepsilon^d(q^*)\omega(q^*) \tag{1.37}$$

If $\gamma^u = 0$ then $p^{t_v^*} = p^{t_s^*}$. When $\gamma^u \in (0,1]$, from eq. (1.37) it follows that, for all $p > 0$, $\varepsilon^u(p,w,t_v^*,0) > \varepsilon^u(p,w,0,t_s^*)$. Therefore from eq. (1.36) $p^{t_v^*} < p^{t_s^*}$. This shows that eq. (1.35) is positive or zero. *Q.E.D.*

Both taxes result in the same final good price q^* but $p^{t_v^*} < p^{t_s^*}$ (provided that $\gamma^u \neq 0$) because the elasticity of derived demand is lower under t_s^* than under t_v^*. This result can be explained as follows. A variation, say an increase, of p represents a cost shift for downstream producers. From the first order condition for profit maximization (1.4), under ad valorem taxation a marginal cost shift $C_p dp$ must be matched by a corresponding increase of the perceived after-tax marginal revenue $\mathrm{MR}^{t_v} = (1-t_v)[q + xq'(1+v^d)]$ which, from the second order condition, is obtained by a reduction of output such that $-d\mathrm{MR}^{t_v} = C_p dp$. On the other hand, under specific taxation the perceived after-tax marginal revenue is $\mathrm{MR}^{t_s} = q + xq'(1+v^d) - t_s$, so that for any marginal reduction of output $-d\mathrm{MR}^{t_v} < -d\mathrm{MR}^{t_s}$, the reason being that with ad valorem taxation part of the increase in marginal revenue is absorbed by the government. Therefore, a given marginal cost shift causes a larger output reduction under ad valorem than under specific taxation, i.e. derived demand is more elastic under ad valorem than under specific taxation.

In terms of revenue collected, the two taxes are equivalent when *both* upstream *and* downstream producers price at marginal cost. Pricing above marginal cost in at least one market is sufficient for t_v^* to raise more revenue than t_s^*. Notice that $R(t_v^*) > R(t_s^*)$ also when downstream producers price at marginal cost, provided that upstream firms price above it. Recall that when only the downstream market is considered, with exogenous input prices, as in Suits and Musgrave's (1953) paper, ad valorem and specific taxes are always equivalent if producers price at marginal cost.

The next two propositions deal with the comparison between downstream and upstream taxes of the same type. To proceed, a pair of restrictions are imposed. First, assume that downstream technology is Leontief. Let $C_p(p,w) = a_{yx}$ and $C_w(p,w) = a_{Lx}$ be respectively the units of commodity y and the units of labour services employed to produce one unit of the consumption good x. Second, assume that final demand is of constant elasticity $\bar{\varepsilon}$. Under these assumptions, equilibrium prices are

$$q = \frac{\bar{\varepsilon}}{\bar{\varepsilon}-\gamma^d}\frac{pa_{yx}+wa_{Lx}+t_s}{1-t_v} \tag{1.38}$$

$$p = \frac{\bar{\varepsilon}}{\bar{\varepsilon}-\gamma^u}\left(\frac{wa_{Lx}+t_s}{a_{yx}\bar{\varepsilon}}\gamma^u + \frac{wa_{Ly}+i_s}{1-i_v}\right) \tag{1.39}$$

First consider the comparison between upstream and downstream ad valorem taxation.

Proposition 3 *Assume that downstream technology is Leontief and that final demand is of constant elasticity. Consider a tax t_v^* and a tax i_v^* resulting in the same final good price q^*. Let $R(t_v^*)$ and $R(i_v^*)$ be the corresponding revenues. Then*

$$R(t_v^*) \gtrless R(i_v^*) \quad \textit{iff} \quad (1-i_v^*)\frac{a_{Lx}}{a_{yx}a_{Ly}} \lesseqgtr \frac{\bar{\varepsilon}}{\bar{\varepsilon}-\gamma^d}\frac{\gamma^d}{\gamma^u} \tag{1.40}$$

Proof. Equilibrium prices are

$$p^{t_v^*} = \frac{\bar{\varepsilon}}{\bar{\varepsilon}-\gamma^u}\left(\frac{wa_{Lx}}{a_{yx}}\frac{\gamma^u}{\bar{\varepsilon}} + wa_{Ly}\right) \tag{1.41}$$

$$q^* = \frac{\bar{\varepsilon}}{\bar{\varepsilon}-\gamma^d}\frac{p^{t_v^*}a_{yx}+wa_{Lx}}{1-t_v^*} \tag{1.42}$$

$$p^{i_v^*} = \frac{\bar{\varepsilon}}{\bar{\varepsilon}-\gamma^u}\left(\frac{wa_{Lx}}{a_{yx}}\frac{\gamma^u}{\bar{\varepsilon}} + \frac{wa_{Ly}}{1-i_v^*}\right) \tag{1.43}$$

$$q^* = \frac{\bar{\varepsilon}}{\bar{\varepsilon}-\gamma^d}(p^{i_v^*}a_{yx}+wa_{Lx}) \tag{1.44}$$

Tax revenues are

$$\begin{aligned} R(t_v^*) &= t_v^* q^* \chi(q^*) \\ R(i_v^*) &= i_v^* p^{i_v^*} a_{yx} \chi(q^*) \end{aligned}$$

From eqs. (1.42) and (1.44) one gets

$$t_v^* q^* = \frac{\bar{\varepsilon}}{\bar{\varepsilon} - \gamma^d} a_{yx} (p^{i_v^*} - p^{t_v^*})$$

Therefore

$$\begin{aligned} R(t_v^*) - R(i_v^*) &= (t_v^* q^* - i_v^* p^{i_v^*} a_{yx}) \chi(q^*) = \\ &= \left(\frac{\bar{\varepsilon}}{\bar{\varepsilon} - \gamma^d} (p^{i_v^*} - p^{t_v^*}) - i_v^* p^{i_v^*} \right) a_{yx} \chi(q^*) \end{aligned} \tag{1.45}$$

Substituting (1.41) and (1.43) into (1.45) one obtains

$$R(t_v^*) - R(i_v^*) = \left(\frac{a_{Ly}}{1 - i_v^*} \frac{\gamma^d}{\bar{\varepsilon} - \gamma^d} - \frac{a_{Lx}}{a_{yx}} \frac{\gamma^u}{\bar{\varepsilon}} \right) \frac{\bar{\varepsilon}}{\bar{\varepsilon} - \gamma^u} w a_{yx} i_v^* \chi(q^*)$$

and this implies (1.40). *Q.E.D.*

Condition (1.40) is a function of (i) upstream and downstream market structure, (ii) elasticity of final demand and (iii) technological coefficients. If $\gamma^d = 0$ and $\gamma^u = 0$ then the two taxes are equivalent. If $\gamma^d \neq 0$ and $\gamma^u = 0$ then $R(t_v^*) > R(i_v^*)$. If $\gamma^d = 0$ and $\gamma^u \neq 0$ then $R(t_v^*) < R(i_v^*)$. The tax t_v^* is more likely to raise more revenue than i_v^* the lower is the elasticity of final demand. As for technological coefficients, notice that a_{Lx} and $a_{yx}a_{Ly}$ are respectively the units of labour *directly* and *indirectly* employed to produce one unit of the final good. The higher is a_{Lx} relative to $a_{yx}a_{Ly}$, the more likely is i_v^* to raise a greater revenue than t_v^*.

The last proposition is about specific taxation and shows that levying upstream or downstream producers brings about the same revenue.

Proposition 4 *Assume that downstream technology is Leontief and that final demand is of constant elasticity. Consider a tax t_s^* and a tax i_s^* resulting in the same final good price q^*. Let $R(t_s^*)$ and $R(i_s^*)$ be the corresponding revenues. Then $R(t_s^*) = R(i_s^*)$.*

Proof. Equilibrium prices are, see eqs. (1.38)–(1.39),

$$
\begin{aligned}
p^{t_s^*} &= \frac{\bar{\varepsilon}}{\bar{\varepsilon}-\gamma^u}\left(\frac{wa_{Lx}+t_s^*}{a_{yx}\bar{\varepsilon}}\gamma^u + wa_{Ly}\right) &(1.46)\\
q^* &= \frac{\bar{\varepsilon}}{\bar{\varepsilon}-\gamma^d}(p^{t_s^*}a_{yx} + wa_{Lx} + t_s^*) &(1.47)\\
p^{i_v^*} &= \frac{\bar{\varepsilon}}{\bar{\varepsilon}-\gamma^u}\left(\frac{wa_{Lx}}{a_{yx}}\frac{\gamma^u}{\bar{\varepsilon}} + wa_{Ly} + i_s^*\right) &(1.48)\\
q^* &= \frac{\bar{\varepsilon}}{\bar{\varepsilon}-\gamma^d}(p^{i_s^*}a_{yx} + wa_{Lx}) &(1.49)
\end{aligned}
$$

Tax revenues are

$$
\begin{aligned}
R(t_s^*) &= t_s^*\chi(q^*)\\
R(i_s^*) &= i_s^* a_{yx}\chi(q^*)
\end{aligned}
$$

Substitute (1.46) into (1.47) and (1.48) into (1.49). Equating the resulting equations, one finds that t_s^* and i_s^* bring about the same final good price q^* if and only if

$$t_s^* = a_{yx} i_s^*$$

Therefore $R(t_s^*) - R(i_s^*) = 0$. *Q.E.D.*

Of course, one must bear in mind that this equivalence result is obtained under the restrictive assumptions of Leontief technology and isoelastic final demand.

1.6 Conclusions

The impact of ad valorem and specific taxes in a model of vertically related oligopolies has been examined. Following Seade (1985) the analysis of tax incidence has focused on finding the conditions for taxes to increase producers' net price and profits.

It has been shown that tax incidence on downstream price and profits is governed by the price elasticity of the elasticity of final demand and downstream market structure. Specific downstream taxation is more likely to cause price overshifting and to raise profits than ad valorem downstream taxation. Also, ad valorem and specific taxes on upstream producers are equivalent, in terms of incidence, to the specific tax on downstream production.

Tax incidence on upstream price and profits has been shown to depend on the price elasticity of the elasticity of derived demand, which in turn depends on upstream and downstream market structure, final demand conditions and input substitution in downstream production. Taxes on downstream producers may raise or lower the price of the intermediate good, but the ad valorem tax is more likely to reduce it. Specific taxes on upstream producers are more likely to raise upstream unit and total profits than upstream ad valorem taxation.

Following Suits and Musgrave's (1953) approach, tax instruments have been compared in terms of revenue collected and the effect on the price of the consumption good. The revenue from an ad valorem tax is always higher than the revenue from a specific tax levied in the same industry, when both taxes result in the same price for the consumption good. The revenue from downstream ad valorem taxation can be higher or lower than the yield from upstream ad valorem taxation. Upstream and downstream specific taxation always raise the same revenue, when both taxes result in the same price for the consumption good.

The next chapter will apply the framework developed here to the analysis of tax reforms in vertically related industries.

A Appendix

A.1 The relation between $E(mx)$ and $F(q)$

Let $mx = \chi(q) \in C^2$ be direct demand, with $\chi'(q) < 0$ for each $q > 0$. The inverse demand is defined as $q = q(mx)$. To ease the notation let $m = 1$. Then $x_0 \equiv \chi[q(x_0)]$ and $q_0 \equiv q[\chi(q_0)]$. Two properties concerning the derivatives of direct and inverse demands which will be used below are the following. For each (q_0, x_0) such that $x_0 = \chi(q_0)$

$$q'(x_0)\chi'(q_0) = 1 \tag{1.50}$$

$$q''(x_0) = -\frac{\chi''(q_0)}{[\chi'(q_0)]^3} \tag{1.51}$$

The elasticity of demand can be expressed either as a function of quantity or as a function of price

$$\varepsilon(q_0) \equiv -\frac{\chi'(q_0)q_0}{\chi(q_0)} = -\frac{q(x_0)}{x_0 q'(x_0)} \equiv \varepsilon(x_0) \tag{1.52}$$

The elasticity of the slope of inverse demand is defined as

$$E(x_0) \equiv -q''(x_0)\frac{x_0}{q'(x_0)}$$

and the elasticity of the slope of direct demand is

$$E(q_0) \equiv -\chi''(q_0)\frac{q_0}{\chi'(q_0)}$$

Using (1.50), (1.51) and (1.52) the relation between $E(x_0)$ and $E(q_0)$ turns out to be

$$E(x_0) = -q''\frac{x_0}{q'} = \frac{\chi''}{(\chi')^3}\chi\chi' = \frac{\chi'' q_0}{\chi'}\frac{\chi}{\chi' q_0} = \frac{E(q_0)}{\varepsilon(q_0)} \tag{1.53}$$

The price elasticity of the elasticity of direct demand is

$$F(q_0) \equiv \frac{d}{dq}\varepsilon(q_0)\frac{q_0}{\varepsilon(q_0)} = \left(\frac{\varepsilon\chi''}{\chi'} + \frac{\varepsilon}{q_0} - \frac{\varepsilon\chi'}{\chi}\right)\frac{q_0}{\varepsilon} = -E(q_0) + 1 + \varepsilon(q_0)$$

Finally, using (1.53)

$$F(q_0) = -\varepsilon(q_0)E(x_0) + 1 + \varepsilon(q_0)$$

which is the result given on p. 16.

A.2 The elasticity of derived demand

The elasticity of the derived demand (1.13) is defined by

$$\varepsilon^u(p, w, t_v, t_s) \equiv -\frac{p\psi_p}{\psi} = -\frac{pC_{pp}}{C_p} - p\frac{\chi'}{\chi}\phi_p \tag{1.54}$$

where ψ_p is given in (1.14).

The elasticity of substitution between labour and the intermediate good in downstream production is defined as

$$\sigma = -\frac{d\log(y/L^d)}{d\log(p/w)} = -\frac{pC_{pp}C}{wC_wC_p}$$

where L^d and y are respectively the conditional input demands for labour and the intermediate good. Thus the first term in (1.54) may be written

$$-\frac{pC_{pp}}{C_p} = \frac{wC_w}{C}\sigma$$

Substituting for ϕ_p from (1.12) the second term in (1.54) becomes

$$-p\frac{\chi'}{\chi}\phi_p = -\frac{\chi' q}{\chi}\frac{pC_p}{q}\frac{1}{1-t_v}\frac{1}{f_q}$$

Substituting for q from (1.10)

$$\frac{pC_p}{C+t_s}\varepsilon^d(q)\left[1 - \frac{\gamma^d}{\varepsilon^d(q)}\right]\frac{1}{f_q}$$

Finally, to obtain the elasticity of derived demand (1.18) define

$$\left[1 - \frac{\gamma^d}{\varepsilon^d(q)}\right]\frac{1}{f_q} \equiv \omega(q)$$

A.3 Ad valorem downstream taxation

The tax t_v causes q to overshift if and only if the corresponding indicator shifting coefficient is greater than one

$$s_{tv}^d \equiv \frac{q\left(1-\frac{\gamma^d}{\varepsilon^d}\right)+C_p\frac{\partial p}{\partial t_v}}{q+C_p\frac{\partial p}{\partial t_v}}\frac{1}{f_q} > 1$$

Assuming that $q(1-\gamma^d/\varepsilon^d)+C_p(\partial p/\partial t_v) > 0$, so that $\partial q/\partial t_v > 0$, see eq. (1.25), this condition can be written

$$s_{tv}^d > 1 \quad \text{iff} \quad C_p\frac{\partial p}{\partial t_v} > \left(q+C_p\frac{\partial p}{\partial t_v}\right)F^d$$

The term in brakets is positive. Assuming that $\mid F^d\omega \mid > \mid q\omega_q \mid$, see on p. 22, then $\partial p/\partial t_v \gtrless 0$ iff $F^d \lesseqgtr 0$, which gives $s_{tv}^d > 1$ iff $F^d < 0$.

A.4 Proofs of the results in table 1.4

Downstream profits. The profit function is

$$m\pi^d = \Big\{(1-t_v)q(w,T) - t_s - C\big(p(w,T),w\big)\Big\}\chi\big(q(w,T)\big) \tag{1.55}$$

where $T = (t_v, t_s, i_v, i_s)$ is the vector of tax instruments, $p(w,T) = \xi(w, t_v, t_s, i_v, i_s)$ is eq. (1.22) defining the equilibrium price of the intermediate good and $q(w,T) = \phi(p(w,T), w, t_v, t_s)$ defines the equilibrium price for the final good, see eq. (1.11).

A preliminary result is the following

$$\begin{aligned}
& [(1-t_v)q - t_s - C(p,w)]\chi' + (1-t_v)\chi = \\
= \;& (1-t_v)\chi'\left[q - \frac{C(p,w)+t_s}{1-t_v} + \frac{\chi}{\chi'}\right] = \\
= \;& (1-t_v)\chi'\left[q\left(1-\frac{1}{\varepsilon^d}+\frac{\gamma^d}{\varepsilon^d}-\frac{\gamma^d}{\varepsilon^d}\right) - \frac{C(p,w)+t_s}{1-t_v}\right] = \\
= \;& (1-t_v)\chi' q\frac{\gamma^d-1}{\varepsilon^d} = && (1.56)\\
= \;& (1-t_v)(1-\gamma^d)\chi && (1.57)
\end{aligned}$$

where (1.56) makes use of the first order condition (1.10) and (1.57) uses $\chi' q = -\varepsilon^d \chi$.

Consider the tax t_s. Differentiating (1.55)

$$\begin{aligned}\frac{\partial(m\pi^d)}{\partial t_s} &= \left\{\left[(1-t_v)q - t_s - C(p,w)\right]\chi' + (1-t_v)\chi\right\}\frac{\partial q}{\partial t_s} + \\ &\quad - \chi - C_p\frac{\partial p}{\partial t_s}\chi \\ &= (1-t_v)(1-\gamma^d)\chi\frac{\partial q}{\partial t_s} - \left(1 + C_p\frac{\partial p}{\partial t_s}\right)\chi\end{aligned}$$

Let $\chi = mx$. Substituting for $\partial q/\partial t_s$ from table 1.1

$$\frac{\partial(m\pi^d)}{\partial t_s} = \left[(1-\gamma^d)s^d_{ts} - 1\right]\left(1 + C_p\frac{\partial p}{\partial t_s}\right)mx$$

which is the derivative contained in table 1.4. The condition for this expression to be positive follows from $(1-\gamma^d)s^d_{ts} > 1$ and from the definition of s^d_{ts}.

The computations of the derivatives with respect to t_v, i_v and i_s are similar and are omitted.

The sufficient condition for t_v to raise $m\pi^d$ is obtained as follows. Consider

$$\frac{\partial(m\pi^d)}{\partial t_v} = \left[(1-\gamma^d)s^d_{tv} - 1\right]\left(q + C_p\frac{\partial p}{\partial t_v}\right)mx$$

This is positive if and only if $s^d_{tv}(1-\gamma^d) > 1$. Substituting for s^d_{tv} from table 1.1 this condition can be written

$$\left(q - \frac{\gamma^d}{\varepsilon^d}q + C_p\frac{\partial p}{\partial t_v}\right)(1-\gamma^d) > \left(q + C_p\frac{\partial p}{\partial t_v}\right)\left(1 - \frac{\gamma^d}{\varepsilon^d} + \frac{\gamma^d}{\varepsilon^d}F^d\right)$$

and after some passages one obtains

$$q(F^d - \gamma^d + \varepsilon^d) + C_p\frac{\partial p}{\partial t_v}(F^d - 1 + \varepsilon^d) < 0 \qquad (1.58)$$

Price overshifting is a necessary condition for profits to increase: from appendix A.3 $s^d_{tv} > 1$ iff $F^d < 0$, assuming $\partial p/\partial t_v > 0$. Notice that if $F^d - \gamma^d + \varepsilon^d < 0$ then $F^d - 1 + \varepsilon^d < 0$. Thus $F^d < \gamma^d - \varepsilon^d$ is a sufficient condition for the unequality (1.58) to hold.

Upstream profits. The profit function is

$$n\pi^u = \Big[(1-i_v)p(w,T) - i_s - wa_{Ly}\Big]\psi\Big(p(w,T), w, t_v, t_s\Big) \tag{1.59}$$

where ψ is defined in eq. (1.13).

A preliminary result is the following

$$\begin{aligned}
& [(1-i_v)p - i_s - wa_{Ly}]\psi_p + (1-i_v)\psi = \\
= \; & (1-i_v)\psi_p\left(p - \frac{wa_{Ly}+i_s}{1-i_v} + \frac{\psi}{\psi_p}\right) = \\
= \; & (1-i_v)\psi_p\left[p\left(1 - \frac{1}{\varepsilon^u} + \frac{\gamma^u}{\varepsilon^u} - \frac{\gamma^u}{\varepsilon^u}\right) - \frac{wa_{Ly}+i_s}{1-i_v}\right] = \\
= \; & (1-i_v)\psi_p p\frac{\gamma^u - 1}{\varepsilon^u} = \qquad (1.60) \\
= \; & (1-i_v)(1-\gamma^u)\psi \qquad (1.61)
\end{aligned}$$

where (1.60) makes use of the first order condition (1.17) and (1.61) uses $\psi_p p = -\varepsilon^u\psi$.

Consider the tax i_s. Differentiating (1.59)

$$\begin{aligned}
\frac{\partial(n\pi^u)}{\partial i_s} &= \left\{\Big[(1-i_v)p - i_s - wa_{Ly}\Big]\psi_p + (1-i_v)\psi\right\}\frac{\partial p}{\partial i_s} - \psi \\
&= (1-i_v)(1-\gamma^u)\psi\frac{\partial p}{\partial i_s} - \psi
\end{aligned}$$

Let $\psi = C_p mx$. Substituting for $\partial p/\partial i_s$ one gets the expression in table 1.4

$$\frac{\partial(n\pi^u)}{\partial i_s} = \Big[(1-\gamma^u)s^u_{is} - 1\Big]C_p mx$$

The condition for this derivative to be positive follows from $(1-\gamma^u)s^u_{is} > 1$ and from the definition of s^u_{is}. The computation of the derivative with respect to i_v is similar and is omitted.

Consider the tax t_s. Differentiating (1.59)

$$\begin{aligned}
\frac{\partial(n\pi^u)}{\partial t_s} &= \left\{\Big[(1-i_v)p - i_s - wa_{Ly}\Big]\psi_p + (1-i_v)\psi\right\}\frac{\partial p}{\partial t_s} + \\
&\quad + \Big[(1-i_v)p - i_s - wa_{Ly}\Big]\psi_{ts}
\end{aligned}$$

Using (1.61), the first term is equal to

$$(1 - i_v)(1 - \gamma^u)\frac{\partial p}{\partial t_s} C_p m x$$

As for the second term, using the first order condition (1.17), it is equal to

$$(1 - i_v) p \frac{\gamma^u}{\varepsilon^u} \psi_{ts}$$

The expression for ψ_{ts} is obtained differentiating (1.13)

$$\begin{aligned} \psi_{ts} &= C_p \chi' \frac{1}{1 - t_v} \frac{1}{f_q} = \\ &= -\varepsilon^d C_p m x \frac{1}{1 - t_v} \frac{1}{f_q} \frac{1}{q} \end{aligned}$$

Finally

$$\frac{\partial(n\pi^u)}{\partial t_s} = (1 - i_v)\left[(1 - \gamma^u)\frac{\partial p}{\partial t_s} - \frac{p}{1 - t_v}\frac{\gamma^u}{\varepsilon^u}\frac{\varepsilon^d}{q} s_{ts}^d\right] C_p m x$$

which is the expression in table 1.4. The computation of the derivative with respect to t_v is similar and is omitted.

References

Besley, Timothy, 1989, Commodity taxation and imperfect competition. A note on the effects of entry, Journal of Public Economics 40, 359-367.

Cournot, A., 1838, Récherches sur les principes mathématiques de la théorie des richesses, L. Hachette, Paris.

Delipalla, Sofia and Michael Keen, 1992, The comparison between ad valorem and specific taxation under imperfect competition, Journal of Public Economics 49, 351–367.

Dierickx, I., Matutes, C. and D. Neven, 1988, Indirect taxation and Cournot equilibrium, International Journal of Industrial Organization 6, 385-399.

Katz, Michael L. and Harvey S. Rosen, 1983, Tax analysis in an oligopoly model, NBER Working Paper, No. 1088.

Kay, J.A. and M.J. Keen, 1983, How should commodities be taxed ?, European Economic Review 23, 339-358.

Konishi, Hideki, 1990, Final and intermediate goods taxation in an oligopolistic economy with free entry, Journal of Public Economics 42, 371-386.

de Meza, David, 1982, Generalized oligopoly derived demand with an application to tax induced entry, Bulletin of Economic Research 34, 1-16.

Myles, Gareth D., 1987, Tax design in the presence of imperfect competition: An example, Journal of Public Economics 34, 367-378.

———, 1989, Imperfect competition and the taxation of intermediate goods, Public Finance 44, 62-74.

Seade, Jesus, 1985, Profitable cost increases and the shifting of taxation: Equilibrium responses of markets in oligopoly, Warwick Economic Research Paper, No. 260.

Stern, Nicholas, 1987, The effects of taxation, price control and government contracts in oligopoly and monopolistic competition, Journal of Public Economics 32, 133–158.

Suits, D.B. and R.A. Musgrave, 1953, Ad valorem and unit taxes compared, Quarterly Journal of Economics 67, 598-604.

Waterson, Michael, 1980a, Price-cost margins and successive market power, Quarterly Journal of Economics 94, 135-150.

———, 1980b, Oligopoly and derived demand, Economics Letters 5, 115–118.

Wicksell, Knut, 1896, Finanztheoretische Untersuchungen und das Steuerwesen Schwedens, Fisher, Jena.

Yeung, Patrick, 1972, A note on the rules of derived demand, Quarterly Journal of Economics 86, 511-517.

Chapter 2

Tax Reforms in a Model of Vertically Related Oligopolies

2.1 Introduction

This chapter examines tax reforms in a model of vertically related oligopolies, where a downstream industry produces a final good using the output of an upstream industry as an input. The basic framework is taken from chapter 1 and the analysis will make extensive use of the results on tax incidence contained there.

The market allocation generates two kind of inefficiencies that result in welfare loss. Firstly, upstream oligopolists set the price of the intermediate good above marginal cost and provided that downstream technology allows inputs substitution this causes *aggregate* production inefficiency.[1] Secondly, downstream oligopolists introduce an additional price-cost margin. The analysis focuses on tax reforms, where the government aims at reducing the welfare loss by means of distortionary commodity taxation while raising no revenue. Both ad valorem and

[1]The meaning of *aggregate* production inefficiency is illustrated in detail in section 2.4.

specific taxes levied on producers in both industries are considered. The pattern of actions the government has to take in order to improve welfare are shown to be related to downstream and upstream market structure, final demand conditions and input substitution in downstream production. Also, it is shown that under some circumstances the tax reform improves overall market performance by introducing additional aggregate production inefficiency.

Tax reforms (only specific taxation) in vertically related industries has been recently analysed by Myles (1989), who considers in a general equilibrium framework the two polar cases of downstream (upstream) monopoly and upstream (downstream) competitive sector. His model is general equilibrium because it takes account of the impact of monopoly profits on final demand. In the model presented here distributed profits are assumed not to enter final demand but a wider range of downstream and upstream market structures are considered. Competition among oligopolists belonging to the same industry is modelled using the conjectural variations framework, which encompasses monopoly, Bertrand oligopoly (marginal cost pricing) and Cournot oligopoly as particular cases.

With downstream monopoly and upstream perfect competition, Myles (1989, proposition 2) shows that *specific* taxation cannot improve welfare if the monopolist produces with constant returns to scale.[2] The same result is contained here in proposition 6, extended to the case of downstream oligopoly. However, a new result is given in proposition 1, namely that with upstream marginal cost pricing and downstream oligopoly, tax reforms with *ad valorem* taxes are effective. The welfare loss can be reduced by taxing the final good and subsidizing the intermediate good.

With downstream perfect competition and upstream monopoly, Myles (1989)

[2]With increasing (decreasing) returns welfare can be improved by taxing (subsidizing) the intermediate good and subsidizing (taxing) the final good.

shows that a) if downstream technology is Leontief then *specific* taxation cannot raise welfare (proposition 3) and b) if input substitution occurs then welfare rises by subsidizing the intermediate good and taxing the final good (proposition 4). Assuming exponential final demand, the work presented here generalizes this result (proposition 7) to the case where both downstream and upstream producers price above marginal cost. The condition under which taxing downstream producers and subsidizing upstream firms raises welfare is given a more precise characterization, being related to the elasticity of final demand, the elasticity of input substitution and the degree of market power in the downstream industry.

Assuming isoelastic final demand, propositions 2– 5 in this paper deal with *ad valorem* taxation when both downstream and upstream oligopolists price above marginal cost. The design of the tax reform is shown to be contingent on the values of the elasticity of final demand and the elasticity of input substitution. Proposition 3 shows that welfare can be improved also when downstream technology is Leontief and downstream producers price at marginal cost. Moreover, proposition 5 shows that when the elasticity of input substitution is low, the tax reform may improve welfare by raising the price of the intermediate good and by doing so generates additional *aggregate* production inefficiency.

The chapter is organized as follows. Section 2.2 describes the model and summarizes the results on tax incidence, contained in chapter 1, that are relevant for the analysis of tax reforms. Section 2.3 derives the formulae expressing the effect of taxes on consumer's welfare and government's revenue. The rationale of tax reforms is illustrated in section 2.4, where particular attention is dedicated to the concept of *aggregate* production inefficiency. Tax reforms with ad valorem taxes are examined in section 2.5, where the particular case of isoelastic final demand is also considered. Section 2.6 examines tax reforms with specific taxes and look at the special case of exponential final demand. Numerical computations of optimal ad valorem taxes are presented in section 2.7. Finally, conclusions are given in section 2.8.

2.2 The model

The basic framework is taken from chapter 1. This section briefly describes the model and summarizes the results on tax incidence that will be used in the analysis of tax reforms.

The economy consists of one consumer, two vertically related oligopolies and the government. The downstream industry produces a final good using two inputs: labour and an intermediate good produced by the upstream sector. Labour is the only input in upstream production. Profits are distributed to the consumer, who demands the final good and supplies labour. The labour market is assumed to be competitive and the wage rate is taken as the *numeraire*. When in at least one market the producers price above marginal cost, the resulting allocation is not efficient. The objective of the government is to improve market performance by levying ad valorem and specific taxes on producers while raising zero revenue.

In both industries product is homogeneous, technology is constant returns to scale and no entry / exit occurs so that the number of firms is fixed. Each oligopolist has a conjecture about the way the other firms in the same industry will change their output levels in response to changes in its own output. The conjectural variations capture the degree of competition among oligopolists belonging to the same industry, encompassing a wide range of market structures as particular cases, including the two polar cases of monopoly and competitive pricing. Symmetry is assumed: oligopolists belonging to the same industry have identical cost functions and conjectures.

Competition between industries is modelled assuming that downstream oligopolists act as perfect competitors in the market for the intermediate good. This means that upstream oligopolists move first by setting the price for the intermediate good and downstream oligopolists move second by setting the price of the final good taking the price of input as given. The model is solved in three steps. First, downstream profits are maximized for given price of the intermediate good. Second, the derived demand facing the upstream industry is obtained and finally upstream profits are maximized.

Let x be the output of each downstream firm, L the supply of labour, q the price of the final good, w the wage rate, m, π^d and n, π^u the number and profits of downstream and upstream firms respectively. The consumer solves the following problem:

$$\begin{aligned} \max_{mx,L} \quad & U(mx) - \beta L \\ \text{s.t.} \quad & qmx = wL + m\pi^d + n\pi^u \end{aligned} \tag{2.1}$$

It is assumed that $U(mx) \in C^4$, $U'(0) > 0$ and $U''(mx) < 0$. The parameter $\beta > 0$ is the (constant) marginal disutility of labour.

From the first order conditions of (2.1) the inverse demand for the final good is

$$q = \frac{w}{\beta} U'(mx) \equiv q(mx) \tag{2.2}$$

β/w is the (constant) marginal utility of lump sum income, in this case profits.

Direct demand is denoted by

$$mx = \chi(q) \tag{2.3}$$

and the supply of labour is

$$L = \frac{q}{w}\chi(q) - \frac{m\pi^d + n\pi^u}{w} \tag{2.4}$$

Separability and linearity in labour supply of the utility function imply that distributed profits affect labour supply but not final demand.[3]

Profits of each downstream oligopolist are

$$\pi^d = \Big[(1-t_v)q(mx) - t_s - C(p,w)\Big]x \tag{2.5}$$

where p is the price of the intermediate good and $C(p,w)$ is the average (and marginal) cost function. The production function is assumed to be a CES, the elasticity of substitution between the labour input L^d and the intermediate good input y^d being $\sigma \equiv -d\log(y^d/L^d)/d\log(p/w)$. When $\sigma = 0$ technology is Leontief; the Cobb-Douglas case corresponds to $\sigma = 1$; inputs are perfect substitutes when $\sigma \to \infty$. Two kind of taxes are levied on downstream producers: a specific tax t_s and an ad valorem tax at a tax inclusive rate $t_v < 1$. Both t_s and t_v may be negative, in which case the government is subsidizing producers.

In terms of q the first order condition for a maximum of (2.5) is

$$q\left[1 - \frac{\gamma^d}{\varepsilon^d(q)}\right] - \frac{C(p,w)+t_s}{1-t_v} \equiv f(q,p,w,t_v,t_s) = 0 \tag{2.6}$$

where $\varepsilon^d \equiv -q\chi'/\chi$ is the price elasticity of direct final demand and $\gamma^d = (1+v^d)/m$ is a term expressing the degree of market power, $v^d = \partial[(m-1)x]/\partial x$ being the conjectural variation. Thus $\gamma^d = 0$ ($v^d = -1$) corresponds to Bertrand conjectures or marginal cost pricing; Cournot behaviour gives $\gamma^d = 1/m$ ($v^d = 0$); when oligopolists collude or the industry is a monopoly $\gamma^d = 1$ ($v^d = m-1$ and $v^d = 0, m = 1$ respectively). In general $\gamma^d \in [0,1]$, a higher value representing a higher degree of market power. Eq. (2.6) defines the equilibrium price for the final good $q = \phi(p,w,t_v,t_s)$.

[3]Since final demand depends only on own price, profit maximization by downstream oligopolists is a simple problem. Suppose instead that distributed profits affect final demand. Then the profit function of downstream oligopolists have profits among its arguments, which makes profit maximization a more complex problem.

Let y be the output of each upstream firm. By Shephard's lemma the (direct) derived demand for the intermediate good is

$$ny = C_p(p, w)\chi\big(\phi(.)\big) \equiv \psi(p, w, t_v, t_s) \tag{2.7}$$

Let wa_{Ly} be the constant average (and marginal) cost of labour in upstream production, i_s the specific and i_v the ad valorem tax levied on upstream producers. Profits of each upstream oligopolist are

$$\pi^u = \Big[(1 - i_v)p(ny, w, t_v, t_s) - i_s - wa_{Ly}\Big]y \tag{2.8}$$

where $p(ny, w, t_v, t_s)$ is the inverse derived demand.

In terms of p the first order condition for a maximum of (2.8) is

$$p\left[1 - \frac{\gamma^u}{\varepsilon^u(p, w, t_v, t_s)}\right] - \frac{wa_{Ly} + i_s}{1 - i_v} \equiv g(p, w, t_v, t_s, i_v, i_s) = 0 \tag{2.9}$$

where $\varepsilon^u \equiv -p\psi_p/\psi$ is the price elasticity of the direct derived demand and $\gamma^u \in [0, 1]$ is the term expressing the degree of market imperfection in the upstream industry.

Eqs. (2.6) and (2.9) define the equilibrium prices.[4] Substituting for the equilibrium prices into the respective demand functions gives the equilibrium quantities. Finally, plugging equilibrium prices and quantities into (2.5) and (2.8) gives the equilibrium profits of downstream and upstream oligopolists respectively.

2.2.1 The impact of taxation on prices

The comparative statics results about the impact of taxation on q and p are summarized in table 2.1. Rows 1–3 contain the determinants of tax incidence.

[4] The existence and stability conditions for the downstream industry are respectively $\varepsilon^d > \gamma^d$ and $F^d > 1 - \varepsilon^d/\gamma^d$, where $F^d = q\varepsilon^d_q/\varepsilon^d$. The corresponding conditions for the upstream sector are $\varepsilon^u > \gamma^u$ and $F^u > 1 - \varepsilon^u/\gamma^u$, where $F^u = p\varepsilon^u_p/\varepsilon^u$. The stability condition is sufficient for a maximum, for it implies the second order condition.

Table 2.1: The impact of taxation on prices

$f_q = 1 - \frac{\gamma^d}{\varepsilon^d}(1 - F^d) > 0, \qquad F^d = \frac{q\varepsilon^d_q}{\varepsilon^d}$		
$\varepsilon^u = \frac{wC_w(p,w)}{C(p,w)}\sigma + \frac{pC_p(p,w)}{C(p,w)+t_s}\varepsilon^d(q)\omega(q), \qquad \omega(q) = \frac{1 - \frac{\gamma^d}{\varepsilon^d(q)}}{1 - \frac{\gamma^d}{\varepsilon^d(q)}[1 - F^d(q)]}$		
$g_p = 1 - \frac{\gamma^u}{\varepsilon^u}(1 - F^u) > 0, \qquad F^u = \frac{p\varepsilon^u_p}{\varepsilon^u}$		
$\frac{\partial p}{\partial t_v} = -\gamma^u \frac{p\varepsilon^u_{tv}}{(\varepsilon^u)^2}\frac{1}{g_p}$		$\frac{\partial p}{\partial t_v} > 0$ iff $F^d < 0^a$
$\frac{\partial p}{\partial t_s} = -\gamma^u \frac{p\varepsilon^u_{ts}}{(\varepsilon^u)^2}\frac{1}{g_p}$		if $F^d \leq 0$ then $\frac{\partial p}{\partial t_s} > 0^a$
$\frac{\partial p}{\partial i_v} = \frac{1}{1-i_v} p s^u_{iv}$	$s^u_{iv} = \left(1 - \frac{\gamma^u}{\varepsilon^u}\right)\frac{1}{g_p}$	$s^u_{iv} > 1$ iff $F^u < 0$
$\frac{\partial p}{\partial i_s} = \frac{1}{1-i_v} s^u_{is}$	$s^u_{is} = \frac{1}{g_p}$	$s^u_{is} > 1$ iff $F^u < 1$
$\frac{\partial q}{\partial t_v} = \frac{1}{1-t_v}\left(q + C_p\frac{\partial p}{\partial t_v}\right) s^d_{tv}$	$s^d_{tv} = \frac{q\left(1 - \frac{\gamma^d}{\varepsilon^d}\right) + C_p\frac{\partial p}{\partial tv}}{q + C_p\frac{\partial p}{\partial tv}}\frac{1}{f_q}$	$s^d_{tv} > 1$ iff $F^d < 0^b$
$\frac{\partial q}{\partial t_s} = \frac{1}{1-t_v}\left(1 + C_p\frac{\partial p}{\partial t_s}\right) s^d_{ts}$	$s^d_{ts} = \frac{1}{f_q}$	$s^d_{ts} > 1$ iff $F^d < 1$
$\frac{\partial q}{\partial i_v} = \frac{1}{1-t_v}\frac{1}{1-i_v} pC_p s^u_{iv} s^d_{iv}$	$s^d_{iv} = \frac{1}{f_q}$	$s^d_{iv} > 1$ iff $F^d < 1$
$\frac{\partial q}{\partial i_s} = \frac{1}{1-t_v}\frac{1}{1-i_v} C_p s^u_{is} s^d_{is}$	$s^d_{is} = \frac{1}{f_q}$	$s^d_{is} > 1$ iff $F^d < 1$

(a) assuming that $\mid F^d\omega \mid > \mid q\omega_q \mid$

(b) see appendix A.3 in chapter 1

The first shows the stability condition for downstream firms and defines the elasticity of the elasticity of final demand F^d. The second contains the elasticity of derived demand ε^u. The third gives the stability condition for upstream firms and defines the elasticity of the elasticity of derived demand F^u.

The second column of rows 6–11 contains the *indicator shifting coefficients.* For the sake of illustration, take s^u_{iv}: the way this coefficient is constructed is such that $s^u_{iv} > 1$ if and only if a marginal increase of the tax i_v causes upstream producers' net price, or profits per unit of output, to increase, in which case i_v is said to cause *price overshifting*. Producers' net price is unaffected if and only if $s^u_{iv} = 1$ and it goes down if and only if $s^u_{iv} < 1$. The third column of the table gives the conditions for price overshifting and shows that these are governed by F^d and F^u. The indicator shifting coefficients and the conditions for overshifting will play a central role in the analysis of tax reforms.

2.3 Consumer's welfare and tax revenue

Consumer's welfare is measured by the indirect utility function

$$\mathcal{V}(q, m\pi^d, n\pi^u) = U\big(\chi(q)\big) - (\beta/w)q\chi(q) + (\beta/w)(m\pi^d + n\pi^u) \qquad (2.10)$$

Let $T = \{t_v, t_s, i_v, i_s\}$ be the set of tax instruments. By the usual properties of indirect utility $\partial\mathcal{V}/\partial q = -(\beta/w)mx$ and $\partial\mathcal{V}/\partial(m\pi^d) = \partial\mathcal{V}/\partial(n\pi^u) = \beta/w$. Thus the marginal impact on consumer's welfare of a tax $\tau \in T$ is

$$\frac{\partial\mathcal{V}}{\partial\tau} = -(\beta/w)mx\frac{\partial q}{\partial\tau} + (\beta/w)\left(\frac{\partial(m\pi^d)}{\partial\tau} + \frac{\partial(n\pi^u)}{\partial\tau}\right), \qquad \tau \in T \qquad (2.11)$$

A marginal increase of a tax instrument raises the price of the consumption good and this lowers welfare. However, as the analysis of chapter 1 has shown, taxation may increase profits and thus raise welfare. The partial derivatives expressing the impact of taxes on profits are given in table 1.4 of chapter 1.

The approach adopted herein is to express consumer's welfare in a form which is more convenient for the analysis of tax reform. Substituting $m\pi^d$ and $n\pi^u$ from (2.5) and (2.8) into (2.10) one finds that

$$\begin{aligned} V(q,p,T) &= U\big(\chi(q)\big) - \frac{\beta}{w}\Big[wC_w(p,w) + wa_{Ly}C_p(p,w)\Big]\chi(q) + \\ &\quad - \frac{\beta}{w}\Big[t_v q + t_s + (i_v p + i_s)C_p(p,w)\Big]\chi(q) \qquad (2.12) \end{aligned}$$

Under this alternative specification, each tax instrument affects welfare both directly and indirectly, the indirect effect going through the final and intermediate good prices.

Notice that, excluding β, the second and the third term of (2.12) must be equal to labour supply. In fact

$$\Big[C_w(p,w) + a_{Ly}C_p(p,w)\Big]mx + \frac{t_v q + t_s + (i_v p + i_s)C_p(p,w)}{w}mx \qquad (2.13)$$

is labour demand, which in equilibrium matches the corresponding supply. The first term of (2.13) is total demand for labour from the private sector: $C_w mx$ and $a_{Ly}C_p mx$ are labour inputs in the downstream and upstream industry respectively. The second term is government revenue evaluated in terms of the numeraire and this must correspond to the demand for labour from the public sector. This result is implied by Walras law and follows from the implicit assumption that the government does not to enter the markets for the final and intermediate goods. In other words, when the public sector raises a positive revenue, it demands labour services. Obviously, the government cannot run a budget deficit, for it cannot supply labour services. Notice also that a positive revenue causes a welfare loss, for the consumer supplies some labour services to the public sector which are transformed neither into tradeable goods nor into public goods.

Government's revenue is

$$\begin{aligned} R(q,p,T) &\equiv r(q,p,T)\chi(q) \\ r(q,p,T) &\equiv t_v q + t_s + (i_v p + i_s)C_p(p,w) \qquad (2.14) \end{aligned}$$

Table 2.2 shows the partial derivatives expressing the impact of taxation on $V(q,p,T)$, $R(q,p,T)$ and total profits $\Pi(q,p,T) = m\pi^d + n\pi^u$. Proofs are given in appendix A.1.

To illustrate, consider consumer's welfare. The formula shows that tax incidence is determined by two factors. One is $h(\tau)$, which is defined in the last row of the table. The other is a weighted sum of $\partial q/\partial\tau$ and $\partial p/\partial\tau$, the weights being θ_q and $\theta_p C_p$ respectively. Notice that θ_q depends on (i) the level of t_v, (ii) the elasticity of final demand and (iii) the share of profits in national income. The term θ_p is a function of (i) the tax rate i_v, (ii) the elasticity of substitution and the share of labour costs in downstream production and (iii) the share of profits in upstream revenues. A similar interpretation applies to the formulae expressing tax incidence on government's revenue and total profits.

2.4 Tax Reforms: the framework

Vertically related oligopolies result in welfare loss for two simple reasons. Firstly, upstream oligopolists set the price of the intermediate good above marginal cost and unless downstream technology is Leontief this causes *aggregate* production inefficiency. Downstream producers are still on the frontier of their production sets but the input mix is not efficient for society because the quantity of primary factor (labour) directly and indirectly employed to produce the final output is not minimized. Secondly, downstream oligopolists introduce an additional price-cost margin.

These points are illustrated in figure 2.1. The graph has final output (consumption) and labour inputs on the horizontal and vertical axis respectively. The line CBC is the Consumer's Budget Constraint

$$L = -\frac{m\pi^d + n\pi^u}{w} + \frac{q}{w}mx \tag{2.15}$$

Table 2.2:
The impact of taxation on consumer's welfare,
government revenue and total profits

$\frac{\partial V}{\partial q} = -\frac{\beta}{w}\theta_q mx$ $\frac{\partial V}{\partial p} = -\frac{\beta}{w}\theta_p C_p mx$	$\theta_q \equiv \frac{m\pi^d + n\pi^u}{qmx}\varepsilon^d + t_v$ $\theta_p \equiv \frac{n\pi^u}{pny}\frac{wC_w}{C}\sigma + i_v$
$\frac{\partial V}{\partial \tau} = -\frac{\beta}{w}\left(\theta_q \frac{\partial q}{\partial \tau} + \theta_p C_p \frac{\partial p}{\partial \tau} + h(\tau)\right) mx, \qquad \tau \in T$	
$\frac{\partial R}{\partial q} = \rho_q mx$ $\frac{\partial R}{\partial p} = \rho_p C_p mx$	$\rho_q \equiv t_v - \frac{r}{q}\varepsilon^d$ $\rho_p \equiv i_v - \frac{i_v p + i_s}{p}\frac{wC_w}{C}\sigma$
$\frac{\partial R}{\partial \tau} = \left(\rho_q \frac{\partial q}{\partial \tau} + \rho_p C_p \frac{\partial p}{\partial \tau} + h(\tau)\right) mx, \qquad \tau \in T$	
$\frac{\partial \Pi}{\partial q} = (1-\theta_q)mx$	$\frac{\partial \Pi}{\partial p} = -\theta_p C_p mx$
$\frac{\partial \Pi}{\partial \tau} = \left((1-\theta_q)\frac{\partial q}{\partial \tau} - \theta_p C_p \frac{\partial p}{\partial \tau} - h(\tau)\right) mx, \qquad \tau \in T$	
$h(t_v) = q \quad h(i_v) = pC_p \quad h(t_s) = 1 \quad h(i_s) = C_p$	

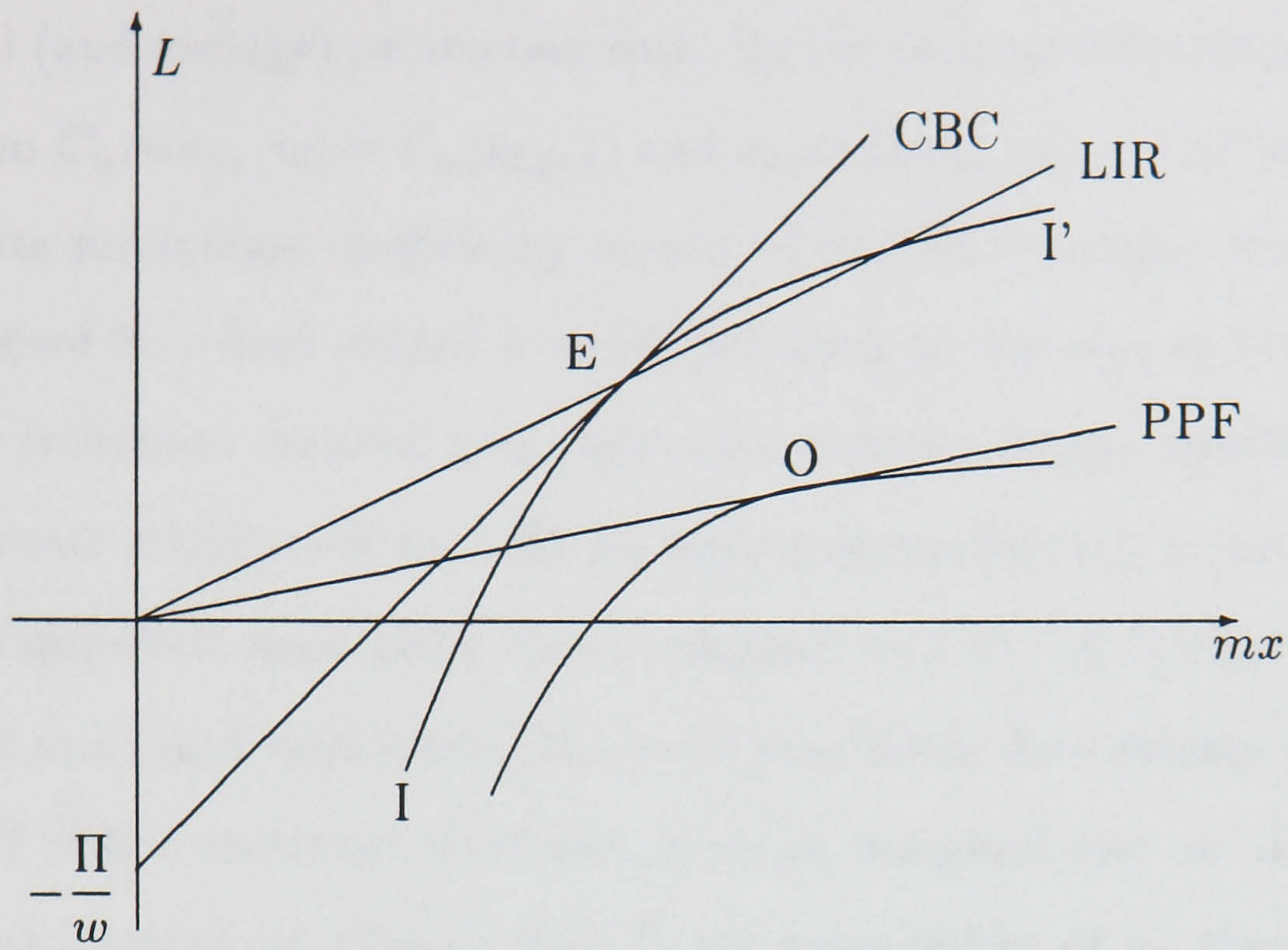

Figure 2.1:

The consumer's optimum is at E, the corresponding indifference curve being II'.

The line LIR (Labour Input Requirement) is the equation of labour demand (2.13), which has slope

$$\frac{L}{mx} = C_w(p,w) + a_{Ly}C_p(p,w) + \frac{r(q,p,T)}{w} \tag{2.16}$$

When $r = 0$ this gives the quantity of labour directly and indirectly demanded by the private sector to produce one unit of the final good at the current price p for the intermediate good. In equilibrium the labour market clears, thus LIR passes through the consumer's optimum E.

Finally, PPF is the (aggregate) Production Possibility Frontier, which has slope

$$\frac{L}{mx} = C_w(a_{Ly},1) + a_{Ly}C_p(a_{Ly},1) \tag{2.17}$$

This gives the minimum quantity of labour services that are directly and indirectly required to produce one unit of the final good, given the state of the technology. To obtain PPF, set $p = wa_{Ly}$, the price of the intermediate good at its marginal (and average) production cost. By the homogeneity property of the cost function $C_w(wa_{Ly}, w) = C_w(a_{Ly}, 1)$ and $a_{Ly}C_p(wa_{Ly}, w) = a_{Ly}C_p(a_{Ly}, 1)$.

Aggregate production inefficiency occurs when LIR is steeper than PPF, as shown in figure 2.1: final output is produced using an amount of labour which exceeds the minimum required given the state of technological knowledge. Two conditions must simultaneously hold for aggregate production to be inefficient: upstream oligopolists must price above marginal cost ($\gamma^u \neq 0$, thus $p > a_{Ly}w$ and $\pi^u > 0$) *and* input substitution has to be possible in downstream production ($\sigma \neq 0$). If either upstream producers price at marginal cost *or* downstream technology is Leontief (so that C_w and C_p are independent of p), then LIR is on PPF and aggregate production is efficient.

In summary, downstream and upstream market imperfections cause a welfare loss which in terms of figure 2.1 is represented by (a) LIR being steeper than PPF and (b) the CBC being steeper than LIR and having a negative intercept.[5] If both industries were perfectly competitive then CBC and LIR would be on PPF and the resulting Pareto efficient allocation would be at point O on PPF.

Consider the equilibrium with no taxation. Now suppose that the government selects two tax instruments $\tau_1, \tau_2 \in T$ and seeks to increase consumer's welfare by introducing one tax and one subsidy while raising no revenue.

For *small* taxes $\tau_1 = d\tau_1$ and $\tau_2 = d\tau_2$ the impact on welfare is

$$dV^\circ = \frac{\partial V^\circ}{\partial \tau_1} d\tau_1 + \frac{\partial V^\circ}{\partial \tau_2} d\tau_2$$

where V° denotes that the partial derivatives are evaluated at $t_v = i_v = t_s = i_s = 0$. If the tax reform has to raise no revenue then $d\tau_1$ and $d\tau_2$ are constrained by

$$dR^\circ = \frac{\partial R^\circ}{\partial \tau_1} d\tau_1 + \frac{\partial R^\circ}{\partial \tau_2} d\tau_2 = 0$$

[5]The absolute value of the intercept measures distributed profits in terms of the numeraire, see eq. (2.15).

Solving for $d\tau_1$ and substituting into the differential of the indirect utility function

$$dV^\circ = \left(\frac{\partial V^\circ}{\partial \tau_1} - \frac{\partial R^\circ/\partial \tau_1}{\partial R^\circ/\partial \tau_2}\frac{\partial V^\circ}{\partial \tau_2}\right) d\tau_1 \tag{2.18}$$

$$d\tau_2 = -\frac{\partial R^\circ/\partial \tau_1}{\partial R^\circ/\partial \tau_2} d\tau_1 \tag{2.19}$$

The element in brackets determines the direction of the tax reform. If this is positive then the government can increase consumer's welfare by introducing a tax τ_1 and a subsidy τ_2, while raising no revenue. A tax τ_2 and a subsidy τ_1 are needed if the term is negative. No welfare improving tax reform is feasible if it is equal to zero.

The effect on prices and total profits is

$$dq^\circ = \left(\frac{\partial q^\circ}{\partial \tau_1} - \frac{\partial R^\circ/\partial \tau_1}{\partial R^\circ/\partial \tau_2}\frac{\partial q^\circ}{\partial \tau_2}\right) d\tau_1 \tag{2.20}$$

$$dp^\circ = \left(\frac{\partial p^\circ}{\partial \tau_1} - \frac{\partial R^\circ/\partial \tau_1}{\partial R^\circ/\partial \tau_2}\frac{\partial p^\circ}{\partial \tau_2}\right) d\tau_1 \tag{2.21}$$

$$d\Pi^\circ = \left(\frac{\partial \Pi^\circ}{\partial \tau_1} - \frac{\partial R^\circ/\partial \tau_1}{\partial R^\circ/\partial \tau_2}\frac{\partial \Pi^\circ}{\partial \tau_2}\right) d\tau_1 \tag{2.22}$$

In terms of figure 2.1, the government is seeking to shift the CBC and the LIR lines by means of distortionary taxes in order to move the market equilibrium E to a higher indifference curve.

2.5 Tax Reform: ad valorem taxes

The effect on consumer's welfare, prices and total profits of a tax reform that makes use of ad valorem taxes and is constrained to raise no revenue is

$$dV^\circ = \frac{\beta}{w} mx \left(\theta_q s^d_{iv} Q + \theta_p P\right) dt_v$$

$$Q = qs^u_{iv} - q(1 - \gamma^d/\varepsilon^d) - C_p \frac{\partial p}{\partial t_v}$$

$$P = qs^u_{iv} - C_p \frac{\partial p}{\partial t_v}$$

$$dq^\circ = -s^d_{iv} Q\, dt_v, \quad dp^\circ = -\frac{1}{C_p} P\, dt_v, \quad d\,\Pi^\circ = mx[(\theta_q - 1)s^d_{iv} Q + \theta_p P]\, dt_v$$

$$di_v = -\frac{q}{pC_p} dt_v \tag{2.23}$$

The algebraic proofs are contained in appendix A.2. Recall that s^d_{iv}, s^u_{iv} and $\partial p/\partial t_v$ are defined in table 2.1 and that θ_q and θ_p are defined in table 2.2.

The sign of the term $\theta_q s^d_{iv} Q + \theta_p P$ determines the direction of the tax reform: if positive then $dV^\circ > 0$ with $dt_v > 0$ and $di_v < 0$. At the equilibrium without taxes $\theta_q \geq 0$ (it is equal to zero only if both downstream and upstream profits are zero), $s^d_{iv} > 0$ and $\theta_p \geq 0$ (it is equal to zero when upstream profits are zero or when downstream technology is Leontief). The terms Q and P can take positive and negative values and determine respectively the effect on the final and intermediate good prices: if $dt_v > 0$ then $dq \gtreqless 0$ iff $Q \lesseqgtr 0$ and $dp \gtreqless 0$ iff $P \lesseqgtr 0$. Notice that $Q = P - q(1 - \gamma^d/\varepsilon^d)$ and that $q(1 - \gamma^d/\varepsilon^d) > 0$.[6] Therefore there are four cases of tax reform to consider, which are summarized in table 2.3.

TR 1. $Q \geq 0$ is sufficient for $P > 0$. Welfare increases by taxing the final good and subsidizing the intermediate good. Both the final and the intermediate good prices fall;[7] the impact on total profits is undetermined.

When $0 < P < q(1 - \gamma^d/\varepsilon^d)$ and $Q < 0$ the two following cases occur.

TR 2. If $\theta_q s^d_{iv} Q + \theta_p P > 0$ then consumer's welfare can be increased by taxing the downstream industry and subsidizing the upstream sector. The price of the

[6] The condition $\varepsilon^d > \gamma^d$ must hold for the first order condition (2.6) to define a positive q.

[7] The price of the final good does not change if $Q = 0$.

Table 2.3: Tax Reform: ad valorem taxes

		dt_v	di_v	dq	dp	$d\Pi$	
$Q \geq 0 \Rightarrow P > 0$	$\theta_q s_{iv}^d Q + \theta_p P > 0$	$+$	$-$	$-$	$-$	?	*TR 1*
$0 < P < q(1 - \gamma^d/\varepsilon^d)$ $Q < 0$	$\theta_q s_{iv}^d Q + \theta_p P > 0$	$+$	$-$	$+$	$-$	$+$	*TR 2*
	$\theta_q s_{iv}^d Q + \theta_p P < 0$	$-$	$+$	$-$	$+$	?	*TR 3*
$P \leq 0 \Rightarrow Q < 0$	$\theta_q s_{iv}^d Q + \theta_p P < 0$	$-$	$+$	$-$	$-$	?	*TR 4*

intermediate good falls and the price of the final good rises. Therefore profits must increase, otherwise welfare would not improve.

TR 3. If $\theta_q s_{iv}^d Q + \theta_p P < 0$ then the government subsidizes downstream producers and taxes upstream production. The price of the final good falls whereas the price of the intermediate good increases; the impact on total profits is undetermined.

TR 4. $P \leq 0$ is sufficient for $Q < 0$. The government subsidizes the downstream industry and taxes the upstream sector. Both prices fall;[8] total profits may rise or fall. Necessary (but not sufficient) condition for $P < 0$ is that $\partial p/\partial t_v > 0$.

One point of section 2.4 was to show that if upstream oligopolists price above marginal cost and if downstream technology allows input substitution, then the market allocation brings about aggregate production inefficiency. Hence *TR 3* looks rather surprising, for the tax reform raises the price of the intermediate good

[8]The price of the intermediate good does not change if $P = 0$.

and by doing so causes additional inefficiencies (in terms of figure 2.1 the LIR line becomes steeper). Proposition 5 below will show that under isoelastic final demand *TR 3* may arise when the degree of input substitution in downstream production is relatively low.

The goal of the analysis is to establish a relation between the four types of welfare improving tax reforms and factors such as market structure in the two industries, the characteristics of final demand and the degree of input substitution in downstream production.

Proposition 1 *(Ad valorem taxation) Assume that upstream firms price at marginal cost and downstream oligopolists price above marginal cost, i.e.* $\gamma^u = 0$ *and* $\gamma^d \in (0, 1]$. *Then TR 1 applies: welfare can be increased by taxing the final good and subsidizing the intermediate good. Both prices fall. If* $0 < \gamma^d < 1$ *then profits decrease, if* $\gamma^d = 1$ *then profits do not change.*

Proof. $\gamma^u = 0$ implies $\theta_p = 0$ (because $\pi^u = 0$), $\partial p/\partial t_v = 0$ and $s^u_{iv} = 1$. Thus $Q = q(\gamma^d/\varepsilon^d) > 0$ and $P = q > 0$, which correspond to *TR 1*. The impact on profits depends on the sign of $\theta_q - 1$; from the definition of θ_q and after some manipulations one finds that $\gamma^u = 0$ implies $\theta_q - 1 = \gamma^d - 1$. *Q.E.D.*

Proposition 1 is the only result that can be given assuming a general demand function. The welfare loss is caused by downstream producers pricing above marginal cost. Aggregate production is efficient, for upstream producers price at marginal cost. The tax on final output tends to increase q. However, the subsidy on upstream firms reduces p and in turn downstream production costs, and this tends to lower q. The tax reform works in improving welfare because the second effect more than compensates the first.

2.5.1 Isoelastic final demand

Assume that the utility function and the corresponding final demand are

$$U(mx) = \alpha \frac{\bar{\varepsilon}}{\bar{\varepsilon} - 1} (mx)^{(\bar{\varepsilon}-1)/\bar{\varepsilon}} - \beta L \tag{2.24}$$

$$mx = (bq)^{-\bar{\varepsilon}} \tag{2.25}$$

where $\alpha > 0$, $\beta > 0$, $b = \beta/(w\alpha)$ and $\bar{\varepsilon} > 1$. Isoelastic final demand is convenient because it simplifies the comparative statics derivatives of table 2.1: $F^d = 0$ and the elasticity of derived demand is

$$\varepsilon^u(p, w, t_s) = \frac{wC_w}{C}\sigma + \frac{pC_p}{C + t_s}\bar{\varepsilon} \tag{2.26}$$

Since ε^u is independent of t_v, $\partial p/\partial t_v$ vanishes in (2.23) and

$$Q = q(s^u_{iv} - 1 + \gamma^d/\bar{\varepsilon}) \tag{2.27}$$

$$P = qs^u_{iv} > 0 \tag{2.28}$$

Referring to the classification of table 2.3, *TR 4* is ruled out, for $P > 0$. The crucial element that determines the direction of the tax reform is the indicator shifting coefficient s^u_{iv}, expressing the impact of the ad valorem tax i_v on upstream producers' net price. If $s^u_{iv} \geq 1 - \gamma^d/\bar{\varepsilon}$ then *TR 1* occurs; if $s^u_{iv} < 1 - \gamma^d/\bar{\varepsilon}$ then either *TR 2* or *TR 3* occurs.

The goal is to derive necessary and sufficient conditions for the direction of the tax reform which are related to the following structural parameters: final demand elasticity $\bar{\varepsilon}$, elasticity of input substitution in downstream production σ, upstream market structure γ^u and downstream market structure γ^d.

The factor determining the direction of the tax reform is a function of the structural parameters and can be defined as

$$\Psi(\sigma, \bar{\varepsilon}, \gamma^d, \gamma^u) = \Psi_1(\sigma, \bar{\varepsilon}, \gamma^d, \gamma^u) + \Psi_2(\sigma, \bar{\varepsilon}, \gamma^d, \gamma^u) \tag{2.29}$$

where

$$\Psi_1(\sigma, \bar{\varepsilon}, \gamma^d, \gamma^u) = \theta_q s^d_{iv} Q \tag{2.30}$$

$$\Psi_2(\sigma, \bar{\varepsilon}, \gamma^d, \gamma^u) = \theta_p P \tag{2.31}$$

From (2.28) and from the definition of θ_p in table 2.2, $\Psi_2(0, \bar{\varepsilon}, \gamma^d, \gamma^u) = 0$ and $\Psi_2(.) > 0$ for all $\sigma > 0$, $\bar{\varepsilon} > 1$, $\gamma^d \in [0,1]$, $\gamma^u \in (0,1]$.

Lemma 1 $\Psi(\sigma, \bar{\varepsilon}, \gamma^d, \gamma^u)$ *is continuous in all its arguments.*

Proof. θ_q, s^d_{iv}, Q, θ_p and P are continuous functions of σ, $\bar{\varepsilon}$, γ^d, γ^u and of equilibrium prices and profits. The cost function is a CES and is continuous in σ. Final demand is continuous in $\bar{\varepsilon}$. Thus eqs. (2.6) and (2.9), determining equilibrium prices and profits, are continuous in σ, $\bar{\varepsilon}$, γ^d and γ^u. *Q.E.D.*

When $\gamma^u = 0$ then $s^u_{iv} = 1$, $Q \geq 0$, $\Psi_1(.) \geq 0$ and the result of proposition 1 applies. This subsection focuses on upstream producers pricing above marginal cost.

When $\gamma^u \neq 0$, table 2.1 shows that

$$s^u_{iv} > 1 \quad \text{iff} \quad F^u \equiv \frac{p\varepsilon^u_p}{\varepsilon^u} < 0 \tag{2.32}$$

Since $p, \varepsilon^u > 0$, the sign of F^u is determined by the sign of ε^u_p, the derivative of (2.26) with respect to p. When $t_s = 0$, the expression for ε^u_p is

$$\varepsilon^u_p = (1-\sigma)(\bar{\varepsilon} - \sigma)\frac{wC_wC_p}{C^2} \tag{2.33}$$

From (2.32)-(2.33) it follows that, for any $\bar{\varepsilon} > 1$, $\gamma^d \in [0,1]$, $\gamma^u \in (0,1]$,

$$\begin{aligned} s^u_{iv} &< 1 \quad \text{iff} \quad \sigma \in [0,1) \quad \text{or} \quad \sigma \in (\bar{\varepsilon}, \infty) \\ s^u_{iv} &= 1 \quad \text{iff} \quad \sigma = 1 \quad \text{or} \quad \sigma = \bar{\varepsilon} \\ s^u_{iv} &> 1 \quad \text{iff} \quad \sigma \in (1, \bar{\varepsilon}) \end{aligned} \tag{2.34}$$

The next proposition follows directly from these price overshifting conditions.

Proposition 2 *(Ad valorem taxation) Assume that final demand is of constant elasticity $\bar{\varepsilon} > 1$. Assume that $\gamma^d \in [0,1]$ and $\gamma^u \in (0,1]$. Assume that $\sigma \in [1, \bar{\varepsilon}]$. Then the welfare improving tax reform is TR 1.*

Proof. From (2.34), $s^u_{iv} \geq 1$ iff $\sigma \in [1, \bar{\varepsilon}]$. Since $\gamma^d \in [0,1]$, this is a sufficient condition for $Q = q(s^u_{iv} - 1 + \gamma^d/\bar{\varepsilon}) \geq 0$, $\Psi_1(.) \geq 0$. *Q.E.D.*

Putting together propositions 1 and 2, when $\sigma \in [1, \bar{\varepsilon}]$, the welfare improving tax reform is *TR 1*, whatever market structure in the two industries. The government levies a tax on downstream producers and subsidizes the upstream industry. Both prices fall.[9]

Next consider the case of downstream Leontief technology. Let $C_p(p, w) = a_{yx}$ and $C_w(p, w) = a_{Lx}$ be respectively the units of commodity y and the units of labour services employed to produce one unit of the consumption good x. The input mix is independent of input prices.

Proposition 3 *(Ad valorem taxation) Assume that final demand is of constant elasticity $\bar{\varepsilon} > 1$. Assume that $\gamma^d \in [0,1]$ and $\gamma^u \in (0,1]$. Assume that $\sigma = 0$. Then*

a. TR 3 applies iff $\dfrac{a_{Lx}}{a_{yx}a_{Ly}} > \dfrac{\bar{\varepsilon}}{\bar{\varepsilon} - \gamma^d}\dfrac{\gamma^d}{\gamma^u}$;

b. there is no welfare improving tax reform iff $\dfrac{a_{Lx}}{a_{yx}a_{Ly}} = \dfrac{\bar{\varepsilon}}{\bar{\varepsilon} - \gamma^d}\dfrac{\gamma^d}{\gamma^u}$;

c. TR 1 applies iff $\dfrac{a_{Lx}}{a_{yx}a_{Ly}} < \dfrac{\bar{\varepsilon}}{\bar{\varepsilon} - \gamma^d}\dfrac{\gamma^d}{\gamma^u}$.

Proof. Leontief technology ($\sigma = 0$) implies

$$\varepsilon^u = \frac{pC_p}{C}\bar{\varepsilon}, \quad \varepsilon^u_p = \frac{wC_wC_p}{C^2}\bar{\varepsilon}, \quad F^u = \frac{wC_w}{C}\bar{\varepsilon}$$

[9]The price of the final good does not change if $\gamma^d = 0$ and $\sigma = 1$ or $\gamma^d = 0$ and $\sigma = \bar{\varepsilon}$, for in both cases $Q = 0$.

so that

$$g_p \equiv 1 - \frac{\gamma^u}{\varepsilon^u}(1 - F^u) = 1 - \frac{\gamma^u}{\bar{\varepsilon}}$$

$$s^u_{iv} \equiv \left(1 - \frac{\gamma^u}{\varepsilon^u}\right)\frac{1}{g_p} = \frac{\frac{pC_p}{C}\bar{\varepsilon} - \gamma^u}{\frac{pC_p}{C}(\bar{\varepsilon} - \gamma^u)}$$

and

$$Q = s^u_{iv} - 1 + \frac{\gamma^d}{\bar{\varepsilon}} = -\frac{wC_w}{pC_p}\frac{\gamma^u}{\bar{\varepsilon} - \gamma^u} + \frac{\gamma^d}{\bar{\varepsilon}}$$

Let $C_p = a_{yx}$ and $C_w = a_{Lx}$. The equilibrium price of the intermediate good is

$$p = \frac{\bar{\varepsilon}}{\bar{\varepsilon} - \gamma^u}\left(\frac{a_{Lx}}{a_{yx}}\frac{\gamma^u}{\bar{\varepsilon}} + a_{Ly}\right) w$$

Therefore

$$Q = -\frac{\gamma^u a_{Lx}}{\gamma^u a_{Lx} + \bar{\varepsilon} a_{yx} a_{Ly}} + \frac{\gamma^d}{\bar{\varepsilon}} \tag{2.35}$$

so that $Q \gtreqless 0$, $\Psi_1 \gtreqless 0$ iff

$$\frac{a_{Lx}}{a_{yx} a_{Ly}} \lesseqgtr \frac{\bar{\varepsilon}}{\bar{\varepsilon} - \gamma^d}\frac{\gamma^d}{\gamma^u} \tag{2.36}$$

and this proves the proposition. *Q.E.D.*

Under Leontief technology, the welfare improving tax reform is either *TR 1* or *TR 3*. With the former, the government taxes downstream producers and subsidizes upstream firms, whereas with the latter the downstream industry is subsidized and the upstream sector is taxed.

The choice between *TR 1* and *TR 3* is contingent on eq. (2.36). Proposition 3 in chapter 1 gives an economic interpretation of this result. Consider a tax t^*_v and a tax i^*_v resulting in the same final good price q^*. Let $R(t^*_v)$ and $R(i^*_v)$ be the corresponding revenues. Then

$$R(t^*_v) \gtreqless R(i^*_v) \quad \text{iff} \quad (1 - i^*_v)\frac{a_{Lx}}{a_{yx} a_{Ly}} \lesseqgtr \frac{\bar{\varepsilon}}{\bar{\varepsilon} - \gamma^d}\frac{\gamma^d}{\gamma^u} \tag{2.37}$$

The relation between eqs. (2.36) and (2.37) is evident, for the former has been computed assuming $i_v = 0$. Hence, if t_v is more effective than i_v in raising revenue, then the government chooses *TR 1*, i.e. a tax $dt_v > 0$ and a subsidy $di_v < 0$.

Turning to the economic interpretation of eq. (2.36), notice that the left hand side depends on technological coefficients. The coefficient a_{Lx} gives the quantity of labour services *directly* employed to produce one unit of the final good. The term $a_{yx}a_{Ly}$ gives the quantity of labour services which are *indirectly* employed to produce one unit of the final good: each unit of output requires a_{yx} units of intermediate good, that in turn requires a_{Ly} units of labour. The right hand side depends on the elasticity of final demand and on market structure in the two industries. Notice that $\gamma^d = 0$ (downstream producers pricing at marginal cost) is sufficient for *TR 3* to occur. The tax reform *TR 3* ($dt_v < 0$, $di_v > 0$) is more likely to apply (i) the more labour intensive is the downstream sector relatively to the upstream sector, (ii) the higher is γ^u, i.e. market power in the upstream industry, (iii) the lower is γ^d, i.e. the more competitive is the downstream sector and (iv) the higher is the elasticity of final demand.

Having determined tax reforms for $\sigma = 0$ (Leontief) and $\sigma = 1$ (Cobb-Douglas), the next natural step is to consider the interval $\sigma \in (0,1)$. So far, the following results have been established.

Lemma 2 *For any* $\bar{\varepsilon} > 1$, $\gamma^d \in [0,1]$, $\gamma^u \in (0,1]$,

a. $\Psi_2(0,\bar{\varepsilon},\gamma^d,\gamma^u) = 0$ *and* $\Psi_2(1,\bar{\varepsilon},\gamma^d,\gamma^u) > 0$*;*

b. $Q(0,\bar{\varepsilon},\gamma^d,\gamma^u) < Q(1,\bar{\varepsilon},\gamma^d,\gamma^u)$*;*

c. $\Psi_1(0,\bar{\varepsilon},\gamma^d,\gamma^u) \gtreqless 0$ *iff* $\dfrac{a_{Lx}}{a_{yx}a_{Ly}} \lesseqgtr \dfrac{\bar{\varepsilon}}{\bar{\varepsilon}-\gamma^d}\dfrac{\gamma^d}{\gamma^u}$.

Proof. Part (a): if $\sigma = 0$, then $\theta_p = 0$ and if $\sigma = 1$, then $n\pi^u > 0$, $\theta_p > 0$. Part (b): if $\sigma = 1$, then $s^u_{iv} = 1$, thus $Q = \gamma^d/\bar{\varepsilon}$, whereas if $\sigma = 0$, then $Q < \gamma^d/\bar{\varepsilon}$, see eq. (2.35). Part (c): see proposition 3. *Q.E.D.*

Lemma 2 describes the behaviour of Ψ_1, Ψ_2 and Q at the endpoints of the interval $(0,1)$. The behaviour of these functions at interior points cannot be studied analytically. The following propositions make the *regularity* assumption that Ψ_1 and Ψ_2 are strictly increasing.[10] The justification for assuming $\partial\Psi_2/\partial\sigma > 0$ comes from point (a). More critical is to assume $\partial\Psi_1/\partial\sigma > 0$ from point (b), for $\Psi_1 = \theta_q s^d_{iv} Q$ and (b) just states that $Q(0,.) < Q(1,.)$. However, notice that $s^d_{iv} = \bar{\varepsilon}/(\bar{\varepsilon} - \gamma^d)$ independent of σ. As for θ_q, this term may be decreasing as well as increasing in σ.

Proposition 4 *(Ad valorem taxation) Assume that final demand is of constant elasticity $\bar{\varepsilon} > 1$. Assume that $\gamma^d \in (0,1]$ and $\gamma^u \in (0,1]$. Assume that Ψ_1 and Ψ_2 are strictly increasing in $\sigma \in (0,1)$ and that*

$$\frac{a_{Lx}}{a_{yx}a_{Ly}} \leq \frac{\bar{\varepsilon}}{\bar{\varepsilon} - \gamma^d}\frac{\gamma^d}{\gamma^u} \tag{2.38}$$

Then for any $\sigma \in (0,1)$ the welfare improving tax reform is TR 1.

Proof. Lemma 2–(a) and $\partial\Psi_2/\partial\sigma > 0$ imply that $\Psi_2 > 0$ for all $\sigma \in (0,1)$. From lemma 2–(c), condition (2.38) implies that $\Psi_1(0, \bar{\varepsilon}, \gamma^d, \gamma^u) > 0$; then $\partial\Psi_1/\partial\sigma > 0$ implies that $\Psi_1 > 0$ for all $\sigma \in (0,1)$. *Q.E.D.*

Proposition 5 *(Ad valorem taxation) Assume that final demand is of constant elasticity $\bar{\varepsilon} > 1$. Assume that $\gamma^d \in [0,1]$ and $\gamma^u \in (0,1]$. Assume that Ψ_1 and Ψ_2 are strictly increasing in $\sigma \in (0,1)$ and that*

$$\frac{a_{Lx}}{a_{yx}a_{Ly}} > \frac{\bar{\varepsilon}}{\bar{\varepsilon} - \gamma^d}\frac{\gamma^d}{\gamma^u} \tag{2.39}$$

[10] The numerical computations of section 2.7 will show that this assumption is plausible.

Then there exist a unique σ^* and a unique σ^{**}, both depending on $\bar{\varepsilon}$, γ^d and γ^u such that $0 < \sigma^* < \sigma^{**} \leq 1$ and

a. *if $\sigma \in [0, \sigma^*)$ then TR 3 applies;*

b. *if $\sigma = \sigma^*$ then there is no welfare improving tax reform;*

c. *if $\sigma \in (\sigma^*, \sigma^{**})$ then TR 2 applies;*

d. *if $\sigma \in [\sigma^{**}, 1]$ then TR 1 applies.*

Proof. Lemma 2–(a) and $\partial\Psi_2/\partial\sigma > 0$ imply that $\Psi_2 > 0$ for all $\sigma \in (0,1)$. From lemma 2–(c), condition (2.39) implies that $\Psi_1(0, \bar{\varepsilon}, \gamma^d, \gamma^u) < 0$; then lemma 2–(b), continuity of $\Psi_1(.)$, $\partial\Psi_1/\partial\sigma > 0$ and $\Psi_1(1, \bar{\epsilon}, \gamma^d, \gamma^u) \geq 0$ imply that there exists a unique $\sigma^{**} \in (0, 1]$ such that $\Psi_1 < 0$ for $\sigma \in [0, \sigma^{**})$ and $\Psi_1 > 0$ for $\sigma \in (\sigma^{**}, 1]$. Finally, there exists a unique σ^*, $0 < \sigma^* < \sigma^{**}$, such that $\Psi = \Psi_1 + \Psi_2 < 0$ for $\sigma \in [0, \sigma^*)$ and $\Psi > 0$ for $\sigma \in (\sigma^*, 1]$. *Q.E.D.*

Corollary 1 *If $\gamma^d = 0$ then proposition 5 applies and $\sigma^{**} = 1$.*

Proposition 5 shows that when condition (2.39) holds and input substitution is relatively low ($0 < \sigma < \sigma^*$), then the welfare improving tax reform is *TR 3*. Upstream oligopolists price above marginal cost and there is aggregate production inefficiency. By taxing upstream producers and subsidizing downstream producers the government raises the price of the intermediate good and introduces additional aggregate production inefficiency: in terms of figure 2.1 the LIR line rotates in anticlockwise direction and this causes a welfare loss. Nonetheless this effect is more than compensated by the welfare gain of lowering the price of the final good, which causes the CBC line to become flatter. The intuition is the following. If σ is low then the intermediate good and labour are almost perfect complements in downstream production, thus the input mix is not distorted by much if p is set above marginal cost by upstream producers. A high proportion of the welfare loss arises from price-cost margins and only a low proportion from aggregate production inefficiency. Therefore the welfare gain of reducing q more than compensates the welfare loss of creating additional inefficiency by raising p.

The final step would be to look at the case $\sigma \in (\bar{\varepsilon}, \infty)$. However, since no analytical results are available, this is left to the numerical computations of section 2.7.

2.6 Tax Reform: specific taxes

The effect on welfare, prices and profits of a tax reform that employs specific taxes and raises no revenue is

$$dV^{\circ} = \frac{\beta}{w} mx \left(\theta_q s^d_{ts} Q + \theta_p P \right) dt_s$$

$$Q = s^u_{is} - 1 - C_p \frac{\partial p}{\partial t_s}$$

$$P = s^u_{is} - C_p \frac{\partial p}{\partial t_s}$$

$$dq^{\circ} = -s^d_{ts} Q \, dt_s, \quad dp^{\circ} = -\frac{1}{C_p} P \, dt_s, \quad d\Pi^{\circ} = mx[(\theta_q - 1) s^d_{ts} Q + \theta_p P] \, dt_s$$

$$di_s = \frac{1}{C_p} dt_s \tag{2.40}$$

The algebraic proofs are similar to those referring to ad valorem taxes and therefore are omitted. The direction of the tax reform is determined by the sign of $\theta_q s^d_{ts} Q + \theta_p P$. Notice that $\theta_q \geq 0$, $s^d_{ts} > 0$, $\theta_p \geq 0$ and $Q = P - 1$. Therefore there are four cases of tax reform to consider, which are summarized in table 2.4.

Assuming a general demand function, the only result is the following.

Table 2.4: Tax Reform: specific taxes

		dt_s	di_s	dq	dp	$d\Pi$	
$Q \geq 0 \Rightarrow P > 0$	$\theta_q s^d_{ts} Q + \theta_p P > 0$	+	−	−	−	?	*TR 1*
$0 < P < 1$ $Q < 0$	$\theta_q s^d_{ts} Q + \theta_p P > 0$	+	−	+	−	+	*TR 2*
	$\theta_q s^d_{ts} Q + \theta_p P < 0$	−	+	−	+	?	*TR 3*
$P \leq 0 \Rightarrow Q < 0$	$\theta_q s^d_{ts} Q + \theta_p P < 0$	−	+	−	−	?	*TR 4*

Proposition 6 *(Specific taxation) Assume that upstream firms price at marginal cost and downstream oligopolists price above marginal cost, i.e.* $\gamma^u = 0$ *and* $\gamma^d \in (0, 1]$. *Then there exists no welfare improving Tax Reform.*

Proof. $\gamma^u = 0$ implies $\theta_p = 0$ (because $\pi^u = 0$), $\partial p / \partial t_s = 0$ and $s^u_{is} = 1$. Thus $Q = 0$ and $P > 0$. *Q.E.D.*

Propositions 1 and 6 give an interesting result as far as the comparison between ad valorem and specific taxation is concerned. When the upstream sector prices at marginal cost then a tax reform which employs ad valorem taxes is effective, whereas specific taxes are ineffective. Ad valorem taxes manage to lower the price of the final good and downstream profits: in terms of figure 2.1 the CBC line shifts up and becomes flatter, so that the net effect is to allow the consumer to move to a lower indifference curve. Instead, specific taxes are unable to affect the consumer's budget constraint.

To get further results, one has to resort to specific functional forms for final demand. Constant elasticity was suitable when considering tax reforms with ad valorem taxes. As for specific taxes, a convenient functional form is exponential demand.

2.6.1 Exponential final demand

Assume that the utility function and the corresponding final demand are

$$U(mx) = \alpha\Big[\log A - \log(mx) + 1\Big]mx - \beta L \tag{2.41}$$

$$mx = A\exp(-bq) \tag{2.42}$$

where $\alpha > 0$, $\beta > 0$, $A > 0$ and $b = \beta/(w\alpha)$. The properties of this functional form are that (i) the elasticity is linear in q, $\varepsilon^d(q) = bq$ and (ii) the elasticity of the elasticity is constant and equal to $F^d = 1$.

The elasticity of derived demand is

$$\varepsilon^u = \frac{wC_w}{C}\sigma + \frac{pC_p}{C + t_s}(\varepsilon^d - \gamma^d) \tag{2.43}$$

From eq. (2.6)

$$q = \frac{\gamma^d}{b} + \frac{C(p,w) + t_s}{1 - t_v} \tag{2.44}$$

so that

$$\varepsilon^u(p,w,t_v) = \frac{wC_w}{C}\sigma + \frac{pC_p}{1 - t_v}b \tag{2.45}$$

Since ε^u is independent of t_s, $\partial p/\partial t_s$ vanishes in (2.40) and

$$Q = s^u_{is} - 1 \tag{2.46}$$

$$P = s^u_{is} > 0 \tag{2.47}$$

The tax reform *TR 4* cannot occur, for $P > 0$. The factor determining the direction of the tax reform is the indicator shifting coefficient s^u_{is}. If $s^u_{is} \geq 1$ then *TR 1* applies, whereas if $s^u_{is} < 1$ then either *TR 2* or *TR 3* occurs.

When $\gamma^u = 0$ then $s^u_{is} = 1$, $Q = 0$, $\theta_p = 0$ and the result of proposition 6 applies. This subsection focuses on upstream producers pricing above marginal cost.

When $\gamma^u \neq 0$, table 2.1 shows that

$$s^u_{is} > 1 \quad \text{iff} \quad F^u \equiv \frac{p\varepsilon^u_p}{\varepsilon^u} < 1 \tag{2.48}$$

Differentiating (2.45), the expression for ε^u_p is

$$\varepsilon^u_p = (1-\sigma)(\varepsilon^d - \gamma^d - \sigma)\frac{wC_wC_p}{C^2} + \frac{pC^2_p}{C}\frac{b}{1-t_v} \tag{2.49}$$

Substituting (2.43) and (2.49) into F^u and assuming $t_s = 0$, the condition for price overshifting (2.48) can be written

$$(1-\sigma)(\varepsilon^d - \gamma^d - \sigma)\frac{wC_wpC_p}{C^2} + \frac{p^2C^2_p}{C}\frac{b}{1-t_v} < \frac{wC_w}{C}\sigma + \frac{pC_p}{C}(\varepsilon^d - \gamma^d)$$

Multiplying both sides by $C^2/(wC_wpC_p)$

$$(1-\sigma)(\varepsilon^d - \gamma^d - \sigma) + \frac{pC_pC}{wC_w}\frac{b}{1-t_v} < \frac{C}{pC_p}\sigma + \frac{C}{wC_w}(\varepsilon^d - \gamma^d)$$

The right hand side can be written

$$\frac{wC_w}{pC_p}\sigma + \sigma + \frac{pC_p}{wC_w}(\varepsilon^d - \gamma^d) + \varepsilon^d - \gamma^d$$

Thus one gets

$$\sigma(\sigma - \varepsilon^d - \gamma^d - 2) < \frac{wC_w}{pC_p}\sigma + \frac{pC_p}{wC_w}(\varepsilon^d - \gamma^d) - \frac{pC_pC}{wC_w}\frac{b}{1-t_v}$$

Finally, from (2.44) $\varepsilon^d - \gamma^d = bq - \gamma^d = (bC)/(1-t_v)$, so that

$$\sigma - \varepsilon^d - \gamma^d - 2 < \frac{wC_w}{pC_p}$$

The right hand side is positive. A sufficient condition for price overshifting is therefore the following

$$\text{if} \quad \sigma \leq \varepsilon^d + \gamma^d + 2 \quad \text{then} \quad s^u_{is} \geq 1 \tag{2.50}$$

This leads to the following proposition.

Proposition 7 *(Specific taxation) Assume that final demand is exponential. Assume that* $\gamma^d \in [0,1]$ *and* $\gamma^u \in (0,1]$. *Assume that* $\sigma \in [0, \varepsilon^d + \gamma^d + 2]$. *Then the welfare improving tax reform is TR 1.*

Proof. From (2.50), if $\sigma \leq \varepsilon^d + \gamma^d + 2$ then $s^u_{is} \geq 1$. Thus $Q \geq 0$ and *TR 1* applies. *Q.E.D.*

This proposition shows that the welfare loss can be reduced by taxing downstream producers and subsidizing the upstream industry, provided that $\sigma < \varepsilon^d + \gamma^d + 2$, and notice that this condition is weak. Indeed, $\varepsilon^d > 1$ and $\gamma^d \geq 0$, thus the lowest upper bound for σ is 3.

2.7 Numerical computations

This section contains some numerical computations of the model of section 2.5.1, where final demand has constant elasticity and the government levies ad valorem taxes. The specification of the model is as follows.

Final demand. Consumer's utility function and consumption demand are those of eqs. (2.24)- (2.25). The corresponding parameters are given the following values: $\bar{\varepsilon} = 2$, $\alpha = 10$, $\beta = 1$. The numeraire is fixed at $w = 1$.

Market structure. Both markets are assumed to be Cournot oligopolies, thus conjectural variations are $v^d = 0$ and $v^u = 0$. There are eight firms in the downstream sector and four in the upstream, i.e. $m = 8$ and $n = 4$. Hence $\gamma^d = .125$ and $\gamma^u = .25$.

Upstream technology. It is assumed that $a_{Ly} = 0.1$.

Downstream technology. The average cost function is a CES

$$C(p,w) = (1/H)\left[a^\sigma p^{1-\sigma} + (1-a)^\sigma w^{1-\sigma}\right]^{1/(1-\sigma)}$$

where $\sigma \in [0,\infty)$. Three typical cases are

$\sigma = 0$	$C(p,w) = (1/H)(p+w)$	Leontief
$\sigma \to 1$	$C(p,w) = kp^a w^{1-a}, \quad k = (1/H)a^{-a}(1-a)^{a-1}$	Cobb-Douglas
$\sigma \to \infty$	$C(p,w) = \min\left\{\dfrac{1}{(1-a)H}w, \dfrac{1}{aH}p\right\}$	Perfect substit.

It is assumed that $H = 2$ and $a = 0.5$.

Recall that in section 2.5.1 the coefficients of Leontief technology are denoted $C_p = a_{yx}$ and $C_w = a_{Lx}$. The specification given here thus implies that $a_{yx} = 1/H$ and $a_{Lx} = 1/H$.

Notice that the parametrization of the model is such that

$$10 = \frac{a_{Lx}}{a_{yx}a_{Ly}} > \frac{\bar{\varepsilon}}{\bar{\varepsilon} - \gamma^d}\frac{\gamma^d}{\gamma^u} = 0.533$$

and this means that the results of proposition 5 should apply.

The numerical computations for some values of σ in the interval $[0, 2.5]$ are contained in table 2.5.

For each σ, the first row shows equilibrium prices, profits and welfare in the absence of taxation. Consumer's welfare is increasing in σ.

Both prices fall as σ increases. The explanation is simple. If the elasticity of substitution is low, derived demand is inelastic, for downstream producers cannot easily substitute labour for the intermediate good when the price of the latter increases. Thus upstream oligopolists can set a high price-cost margin, which in turn is transferred onto the price of the final good by downstream producers.

Downstream total profits are increasing in σ. The reason is the following. As σ increases, q and p go down: the first effect augments revenues, for the elasticity of final demand is greater than one; the second effect reduces unit production

Table 2.5: Optimal ad valorem taxes

σ	t_v	i_v	q	p	$m\pi^d$	$n\pi^u$	V	Ψ_1	Ψ_2	Ψ
			.6705	.2571	9.32	17.47	175.94	−.127	.000	− .127
0.0	−.3904	.7691	.6282	.6378	13.83	5.98	179.00			
			.6307	.2108	9.91	15.44	183.91	−.070	.016	− .054
0.1	−.0956	.3785	.6167	.2878	11.10	11.21	184.48			
			.5947	.1797	10.50	14.12	192.78	−.034	.032	− .003
0.2	−.0042	.0217	.5942	.1827	10.56	13.94	192.79			
			.5619	.1597	11.12	13.52	202.62	−.015	.042	.027
0.3	.0382	−.2138	.5647	.1356	10.64	15.07	202.79			
			.5311	.1468	11.76	13.51	213.56	−.005	.047	.042
0.4	.0542	−.2825	.5335	.1172	11.08	15.45	213.97			
			.5017	.1381	12.45	13.96	225.75	.001	.047	.048
0.5	.0607	−.2772	.5031	.1097	11.66	15.92	226.34			
			.4732	.1320	13.20	14.79	239.33	.004	.046	.050
0.6	.0646	−.2542	.4739	.1061	12.33	16.71	240.06			
			.4456	.1276	14.02	15.94	254.39	.005	.043	.048
0.7	.0679	−.2307	.4458	.1041	13.06	17.84	255.23			
			.4190	.1243	14.91	17.40	270.99	.006	.039	.045
0.8	.0713	−.2108	.4188	.1029	13.85	19.30	271.92			
			.3936	.1219	15.88	19.14	289.11	.007	.035	.042
0.9	.0750	−.1950	.3931	.1021	14.70	21.05	290.13			
			.3693	.1200	16.92	21.15	308.87	.007	.031	.038
1.0	.0790	−.1830	.3686	.1014	15.61	23.09	310.01			
			.3262	.1174	19.16	25.82	351.55	.007	.023	.030
1.2	.0882	−.1684	.3253	.1004	17.52	27.83	352.78			
			.2899	.1159	21.56	31.20	397.73	.006	.017	.023
1.4	.0991	−.1639	.2888	.0995	19.49	33.29	399.08			
			.2604	.1150	23.99	36.92	444.89	.006	.012	.018
1.6	.1122	−.1684	.2592	.0984	21.40	39.11	446.33			
			.2371	.1145	26.36	42.64	490.76	.005	.008	.013
1.8	.1289	−.1823	.2357	.0969	23.09	44.94	492.28			
			.2188	.1143	28.56	48.06	533.66	.005	.006	.011
2.0	.1517	−.2088	.2172	.0945	24.40	50.53	535.29			
			.1884	.1141	33.16	59.32	623.15	.004	.002	.006
2.5	.2770	−.4180	.1858	.0805	24.32	62.87	625.46			

costs. Upstream total profits are decreasing for $\sigma \leq .4$ and increasing for $\sigma \geq .5$. As σ increases there are two effects on derived demand: on one hand it becomes more elastic and this reduces upstream profits; on the other hand it shifts upwards for there is a corresponding shift in final demand and this increases profits. The former effect prevails for $\sigma \leq .4$, whereas the latter dominates for $\sigma \geq .5$.

The last three columns of the table report the values of Ψ_1, Ψ_2 and Ψ, determining the direction of the tax reform. For each σ, the second row shows the welfare maximizing tax-subsidy scheme and the associated equilibrium prices, profits and welfare.

As pointed out above, this numerical example illustrates proposition 5. It turns out that $.2 < \sigma^* < .3$ and $.4 < \sigma^{**} < .5$.

When $\sigma = 0, .1, .2$ the tax reform *TR 3* applies. When $\sigma = 0$, the optimal taxes are $t_v = -.39$ and $i_v = .77$; the price of the final good goes down from .67 to .62, the price of the intermediate good goes up from .26 to .64; profits of downstream and upstream producers increase and fall respectively. Consumer's welfare goes from 176 to 179.

When $\sigma = .3, .4$, *TR 2* applies and the government taxes the dowstream industry and subsidizes upstream producers: q goes up, p goes down and welfare increases because total profits augment.

For all $\sigma \geq .5$ *TR 1* occurs: downstream producers are taxed, whereas upstream producers are subsidized. Both prices fall.[11] Taxation reduces and raises downstream and upstream profits respectively.

Numerical computations have been attempted for $\sigma \geq 3$. It turns out that the welfare maximizing tax-subsidy scheme has a corner solution, namely $t_v \to 1$ and $i_v \to -\infty$. An explanation of this result has not been found.

[11]The fact that q slightly increases for $\sigma = .5, .6, .7$ is due to the *instability* of the numerical model at these points. In fact the values of Ψ_1 are close to zero.

2.8 Conclusions

This chapter has examined tax reforms in a model of vertically related oligopolies. The analysis has shown that there is scope for improving market performance by designing tax-subsidy schemes that raise no revenue. Ad valorem taxes manage to increase welfare also when specific taxes are ineffective.

The analysis suffers from some limitations that are now briefly mentioned. The government is assumed to have perfect knowledge of all the *ingredients* necessary for the design of tax reforms, such as consumer's preferences, firms' technology and pricing behaviour in the two markets. Obviously, most policy makers do not possess such information and the learning process may involve substantial costs.

Administrative costs have been ignored. If some revenue were to be raised in order to cover administrative costs, the tax-subsidy schemes considered here may turn out to be ineffective in raising welfare.

Taxation may create incentives to vertical integration, an issue which has not been considered.

A Appendix

A.1 Proofs of the results in table 2.2

Consumer's welfare, government's revenue and total profits are respectively

$$\begin{aligned} V(q,p,T) &= U\big(\chi(q)\big) - \frac{\beta}{w}\Big[wC_w(p,w) + wa_{Ly}C_p(p,w)\Big]\,\chi(q) \;+ \\ &\quad - \frac{\beta}{w}\,r(q,p,T)\,\chi(q) \qquad (2.51) \\ R(q,p,T) &= r(q,p,T)\chi(q) \qquad (2.52) \\ \Pi(q,p,T) &= m\pi^d + n\pi^u = \\ &= \Big[q - wC_w(p,w) - wa_{Ly}C_p(p,w)\Big]\chi(q) - r(q,p,T)\chi(q) \qquad (2.53) \end{aligned}$$

where

$$r(q,p,T) = t_v q + t_s + (i_v p + i_s)C_p(p,w)$$

Consumer's welfare. Differentiating (2.51) with respect to q

$$\frac{\partial V}{\partial q} = U'\chi' - \beta(C_w + a_{Ly}C_p)\chi' - \frac{\beta}{w}r\chi' - \frac{\beta}{w}t_v\chi$$

From the first order conditions of consumer's optimum $U' = (\beta/w)q$. From the definition of price elasticity $\chi' = -(\varepsilon^d\chi)/q$. Thus

$$\begin{aligned} \frac{\partial V}{\partial q} &= -\frac{\beta}{w}\varepsilon^d\left[1 - \frac{wC_w + wa_{Ly}C_p + r}{q}\right]\chi - \frac{\beta}{w}t_v\chi \\ &= -\frac{\beta}{w}\varepsilon^d\frac{q(1-t_v) - t_s - C + [(1-i_v)p - i_s - wa_{Ly}]C_p}{q}mx \;+ \\ &\quad - \frac{\beta}{w}t_v mx \\ &= -\frac{\beta}{w}\varepsilon^d\frac{m\pi^d + n\pi^u}{q} - \frac{\beta}{w}t_v mx \\ &= -\frac{\beta}{w}\left(\frac{m\pi^d + n\pi^u}{qmx}\varepsilon^d + t_v\right)mx \\ &= -\frac{\beta}{w}\theta_q mx \qquad (2.54) \end{aligned}$$

Differentiating with respect to p

$$\frac{\partial V}{\partial p} = -\beta(C_{wp} + a_{Ly}C_{pp})\chi - \frac{\beta}{w}\Big[(i_v p + i_s)C_{pp} + i_v C_p\Big]\chi$$

From the properties of the cost function

$$C_{wp} = -\frac{p}{w}C_{pp}, \qquad C_{pp} = -\frac{wC_wC_p}{pC}\sigma$$

Thus

$$\begin{aligned}
\frac{\partial V}{\partial p} &= \beta\left(\frac{p}{w} - a_{Ly}\right)\left(-\frac{wC_wC_p}{pC}\right)\sigma mx + \\
&\quad + \frac{\beta}{w}(i_v p + i_s)\frac{wC_wC_p}{pC}\sigma mx - \frac{\beta}{w}i_v C_p mx \\
&= -\beta\left(\frac{p - wa_{Ly}}{p}\right)\frac{C_pC_w}{C}\sigma mx + \\
&\quad + \beta\left(\frac{i_v p + i_s}{p}\right)\frac{C_pC_w}{C}\sigma mx - \frac{\beta}{w}i_v C_p mx \\
&= -\beta\left(\frac{p(1 - i_v) - i_s - wa_{Ly}}{p}\right)\frac{C_pC_w}{C}\sigma mx - \frac{\beta}{w}i_v C_p mx \\
&= -\frac{\beta}{w}\left(\frac{n\pi^u}{pny}\,\frac{wC_w}{C}\sigma + i_v\right)C_p mx \\
&= -\frac{\beta}{w}\theta_p C_p mx \qquad (2.55)
\end{aligned}$$

Finally, for any $\tau \in T$

$$\begin{aligned}
\frac{\partial V}{\partial \tau} &= \frac{\partial V}{\partial q}\frac{\partial q}{\partial \tau} + \frac{\partial V}{\partial p}\frac{\partial p}{\partial \tau} - \frac{\beta}{w}h(\tau)mx \\
&= -\frac{\beta}{w}\left(\theta_q\frac{\partial q}{\partial \tau} + \theta_p C_p\frac{\partial p}{\partial \tau} + h(\tau)\right)mx \qquad (2.56)
\end{aligned}$$

where $h(t_v) = q$, $h(i_v) = pC_p$, $h(t_s) = 1$, $h(i_s) = C_p$.

Government's revenue. Differentiating (2.52) with respect to q

$$\begin{aligned}
\frac{\partial R}{\partial q} &= t_v\chi + r\chi' \\
&= \left(t_v - \frac{r}{q}\varepsilon^d\right)mx \\
&= \rho_q mx \qquad (2.57)
\end{aligned}$$

Differentiating with respect to p

$$\begin{aligned}
\frac{\partial R}{\partial p} &= \left[i_v C_p + (i_v p + i_s) C_{pp}\right]\chi \\
&= \left(i_v - \frac{i_v p + i_s}{p}\frac{wC_w}{C}\sigma\right) C_p m x \\
&= \rho_p C_p m x \qquad (2.58)
\end{aligned}$$

Finally, for any $\tau \in T$

$$\begin{aligned}
\frac{\partial R}{\partial \tau} &= \frac{\partial R}{\partial q}\frac{\partial q}{\partial \tau} + \frac{\partial R}{\partial p}\frac{\partial p}{\partial \tau} + h(\tau) m x \\
&= \left(\rho_q \frac{\partial q}{\partial \tau} + \rho_p C_p \frac{\partial p}{\partial \tau} + h(\tau)\right) m x \qquad (2.59)
\end{aligned}$$

Total profits. Differentiating (2.53) with respect to q

$$\begin{aligned}
\frac{\partial \Pi}{\partial q} &= (q - wC_w - wa_{Ly}C_p)\chi' - r\chi' - t_v\chi + \chi \\
&= -\varepsilon^d \frac{q - wC_w - wa_{Ly}C_p - r}{q}\chi - t_v\chi + \chi
\end{aligned}$$

The first two terms are identical to the first row in (2.54), excluding β/w. Thus one gets

$$\frac{\partial \Pi}{\partial q} = (1 - \theta_q) m x \qquad (2.60)$$

Differentiating with respect to p

$$\frac{\partial \Pi}{\partial p} = -(wC_{wp} + wa_{Ly}C_{pp})\chi - \left[(i_v p + i_s)C_{pp} + i_v C_p\right]\chi$$

This is equal to the first row in (2.55), excluding β/w. Thus

$$\frac{\partial \Pi}{\partial p} = -\theta_p C_p m x \qquad (2.61)$$

Finally, for any $\tau \in T$

$$\begin{aligned}
\frac{\partial \Pi}{\partial \tau} &= \frac{\partial \Pi}{\partial q}\frac{\partial q}{\partial \tau} + \frac{\partial \Pi}{\partial p}\frac{\partial p}{\partial \tau} - h(\tau) m x \\
&= \left((1 - \theta_q)\frac{\partial q}{\partial \tau} - \theta_p C_p \frac{\partial p}{\partial \tau} - h(\tau)\right) m x \qquad (2.62)
\end{aligned}$$

A.2 Tax Reform with ad valorem taxes: proofs

Recall that the tax reform has been defined as starting from the market equilibrium with no taxation: partial derivatives are evaluated at $t_v = i_v = t_s = i_s = 0$ and $r = 0$. This implies that $\rho_q = \rho_p = 0$ (see table 2.2). Thus

$$di_v = -\frac{\partial R^\circ/\partial t_v}{\partial R^\circ/\partial i_v}dt_v = -\frac{q}{pC_p}dt_v \tag{2.63}$$

The effect on the price of the consumption good is

$$\begin{aligned}
dq^\circ &= \left(\frac{\partial q^\circ}{\partial t_v} - \frac{q}{pC_p}\frac{\partial q^\circ}{\partial i_v}\right)dt_v \\
&= \left[\left(q + C_p\frac{\partial p^\circ}{\partial t_v}\right)s^d_{tv} - qs^u_{iv}s^d_{iv}\right]dt_v \\
&= \left[q\left(1 - \frac{\gamma^d}{\varepsilon^d}\right)s^d_{iv} + s^d_{iv}C_p\frac{\partial p^\circ}{\partial t_v} - qs^u_{iv}s^d_{iv}\right]dt_v \\
&= -s^d_{iv}\left[qs^u_{iv} - q\left(1 - \frac{\gamma^d}{\varepsilon^d}\right) - C_p\frac{\partial p^\circ}{\partial t_v}\right]dt_v \\
&= -s^d_{iv}Qdt_v
\end{aligned} \tag{2.64}$$

The effect on the price of the intermediate good is

$$\begin{aligned}
dp^\circ &= \left(\frac{\partial p^\circ}{\partial t_v} - \frac{q}{pC_p}\frac{\partial p^\circ}{\partial i_v}\right)dt_v \\
&= \left(\frac{\partial p^\circ}{\partial t_v} - \frac{q}{C_p}s^u_{iv}\right)dt_v \\
&= -\frac{1}{C_p}\left(qs^u_{iv} - C_p\frac{\partial p^\circ}{\partial t_v}\right)dt_v \\
&= -\frac{1}{C_p}Pdt_v
\end{aligned} \tag{2.65}$$

The effect on consumer's welfare is

$$\begin{aligned}
dV^\circ &= \left(\frac{\partial V^\circ}{\partial t_v} - \frac{q}{pC_p}\frac{\partial V^\circ}{\partial i_v}\right)dt_v \\
&= -\frac{\beta}{w}\left(\theta_q\frac{\partial q^\circ}{\partial t_v} + \theta_pC_p\frac{\partial p^\circ}{\partial t_v} + q - \frac{q}{pC_p}\theta_q\frac{\partial q^\circ}{\partial i_v} - \frac{q}{p}\theta_p\frac{\partial p^\circ}{\partial i_v} - q\right)mxdt_v \\
&= -\frac{\beta}{w}\left[\theta_q\left(\frac{\partial q^\circ}{\partial t_v} - \frac{q}{pC_p}\frac{\partial q^\circ}{\partial i_v}\right) + \theta_pC_p\left(\frac{\partial p^\circ}{\partial t_v} - \frac{q}{pC_p}\frac{\partial p^\circ}{\partial i_v}\right)\right]mxdt_v
\end{aligned}$$

Using the results in (2.64)–(2.65) one gets

$$dV^{\circ} = \frac{\beta}{w}\left(\theta_q s_{iv}^{d} Q + \theta_p P\right) dt_v \tag{2.66}$$

The derivation of $d\Pi^{\circ}$ is similar to the proof just given for consumer's welfare and therefore is omitted.

References

Myles, Gareth D., 1989, Imperfect competition and the taxation of intermediate goods, Public Finance 44, 62-74.

Chapter 3

Income Tax Enforcement Policy with Risk Averse Agents

3.1 Introduction

The standard theory of optimal income taxation (Mirrlees, 1971) assumes that there is a population of individuals endowed with different ability (productivity) levels, which in turn determine labour supply decisions and the distribution of income. Although the government does not observe abilities, it costlessly observes labour incomes. Thus individuals are taxed on the basis of labour income and the objective of the government is to select the tax schedule that maximizes social welfare while raising the required revenue. Since the income tax causes distortions in labour supply, the policy maker faces a trade off between equity and efficiency considerations.

In point of fact the taxing authority cannot directly observe taxable incomes unless it performs costly audits. Hence it usually requires taxpayers to report their own income by filling in an income declaration form. Obviously, this brings about the possibility of tax evasion. Therefore, in addition to the tax schedule, the government must design a tax enforcement policy that consists of a mechanism selecting how many and which taxpayers to audit and a system of penalties for detected tax evaders.

The theory of optimal income taxation in the presence of costly enforcement has been the focus of a seminal paper by Reinganum and Wilde (1985). They assume that taxpayers are risk neutral and are endowed with different pre-tax incomes which are exogenously given. Taxes and fines are lump sum. The audit mechanism can take two forms: in the first, termed *random audits*, a taxpayer faces a probability of audit which is independent of the amount of income he or she reports; in the second, termed *audit cutoff*, a taxpayer is investigated with probability one if his or her reported income is below some threshold level and escapes the audit by reporting an income at least as high as the threshold. The government's objective is to maximize total tax revenue net of audit costs by controlling taxes, penalties and the audit parameters. The main finding of Reinganum and Wilde is that the optimal audit cutoff policy weakly dominates the optimal random audit policy. The optimal lump sum tax and the optimal threshold are identical whereas the level of penalties is irrelevant. Moreover, the optimal tax-audit cutoff policy induces truthful reporting, or taxpayers declaring their true income to the taxing authority.

Border and Sobel (1987) and Chander and Wilde (1992) extend Reinganum and Wilde's framework by assuming general tax and audit probabilities functions. Border and Sobel (1987) first place the restriction that a taxpayer caught evading is never asked to pay a penalty higher than his or her pre-tax income and show that in this case an optimal solution to the auditor problem may not exist. This occurs when the solution would involve offering taxpayers infinitely high rewards when they are found to report truthfully at infinitesimal audit probabilities. Therefore they place a constraint on rewards to honest taxpayers and show that in general this will be a binding constraint. Optimal tax-audit schemes involve tax payments and audit probabilities which are respectively monotonically increasing and monotonically decreasing in reported income. Chander and Wilde (1992) provide a detailed analysis of the role of penalties, obtaining qualitatively similar results with three different specifications of the penalty function: the optimal tax schedule is generally concave and the optimal audit probabilities are nonincreasing and determined by marginal tax rates.

Sanchez and Sobel (1990) obtain a sharper characterization of the tax enforcement policy by solving the problem of a tax collecting agency that controls the audit policy while taking taxes and penalties as given. They assume that the tax function is strictly increasing and differentiable and that the penalty is proportional to the evaded tax. The audit policy consists of a function which gives a probability of investigation for each level of reported income. Two objectives of the tax enforcement agency are considered: one is the maximization of total tax revenue subjet to a budget constraint on audit costs; the other is the maximization of total tax revenue net of audit costs. In both cases, if the distribution of pre-tax income has an increasing hazard rate, the optimal audit policy is a step function which divides taxpayers into two groups on the basis of reported income: reports below a given threshold are audited at the probability level just sufficient to induce a risk neutral taxpayer to behave honestly whereas income reports above the threshold are never audited.

Clearly, the optimal audit policy makes the effective tax payments to be quite regressive (on this point see Scotchmer, 1987). On one hand, the relatively poor individuals (with income below the threshold) are forced to behave honestly so that they pay in full the legislated income tax. On the other hand, the relatively rich individuals (with income above the threshold) report an income just sufficient to escape the audit (the threshold) and safely evade the difference.

This point has been taken up by Cremer, Marchand and Pestieu (1990) and Sanchez and Sobel (1990). They frame their models in a hierarchical setting where there are two levels of authority. At the top, the government chooses the income tax schedule and is concerned with the maximization of social welfare. At the bottom, the tax administration (enforcement agency) controls the audit policy and is concerned with the maximization of tax revenue net of audit costs.[1] Penalties are assumed to be exogenously determined by social conventions. These models show that the government generally takes into account the behaviour of the tax enforcement agency and accordingly adjusts its instruments to trade off equity and compliance. For instance, when the income tax is linear, the optimal marginal tax rate under costly enforcement turns out to be lower than the corresponding one under costlessly enforcement.

Another undesirable feature of the optimal audit policy is time inconsistency. The tax enforcement agency knows that all income reports below the threshold are truthful and that all taxpayers reporting an income equal to the threshold are evaders. However, the former are randomly selected to be audited whereas the latter are never investigated. To sustain such an outcome one has to assume that the tax enforcement agency is able to make a credible commitment that it will not deviate ex-post from the strategy announced ex-ante. For instance, when the tax enforcement policy is instituted by law and the process of reforming the law is costly and lengthy, then the precommitment assumption seems reasonable. Reputation effects may help to sustain the precommitment solution when one considers that the game between the government and the taxpayers is repeated

[1] In addition to the tax schedule, Sanchez and Sobel (1990) assume that the government controls the provision of a public good. Also, they analyze the case of the tax enforcement agency maximizing tax revenue subject to an auditing budget constraint controlled by the government.

over time. Finally, time inconsistency can be avoided with delegation.[2]

All models described above share the limitations of assuming risk neutral taxpayers and the absence of supply side effects.[3] The purpose of this paper is to examine to what extent risk averse behaviour affects the features of the optimal tax enforcement policy. The distinctive feature of the model consists in the framing of risk aversion. Taxpayers maximize a utility function which is continuous, concave and linear with a kink at the level of after-tax income. This specification implies that the marginal utility of the prospective gains from evasion is lower (in absolute value) than the marginal disutility of the corresponding prospective losses. The merit of this approach lies in its analytical simplicity while retaining the traits of risk averse behaviour. The kinked utility function can be given two interpretations, which are illustrated in section 3.3. One is in terms of non-pecuniary costs of evasion, as in Gordon (1989). The other comes from portfolio

[2]Melumad and Mookherjee (1989) present a model where the government delegates the authority of performing tax audits to an independent agency. They show that the full commitment solution can be obtained when the government makes only a limited commitment over some policy variables and selects a suitable incentive scheme for the tax enforcement agency. Reinganum and Wilde (1986) present a model where the government is not able to make a credible commitment to an announced audit strategy.

[3]Mookherjee and Png (1989) study the problem of optimal auditing in the context of competitive insurance markets, debt contracts and tax enforcement policy. Taxpayers are risk averse and there is a moral hazard issue in that they first choose an action unobserved by the principal which then affects the level of pre-tax income. The government chooses taxes, penalties and audit probabilities to maximize the expected utility of the taxpayer subject to a tax revenue constraint. The main results are that optimal policies require all audits to be random (no income report is audited with probability one) and that the highest report is never audited. However, Mookherjee and Png's model, being derived from the principal-agent framework of Grossman and Hart (1983), is not directly comparable to Reinganum and Wilde type models described so far.

selection theory, namely the approach that takes the expected value of loss as measure of risk.

The model is adapted from Reinganum and Wilde's (1985) framework and is described in section 3.2. The principal is the tax enforcement agency. Its objective is to maximize revenue net of audit costs by controlling the audit policy, while taking taxes and penalties as given. The audit policy is restricted to the simple mechanism where the tax agency chooses an income threshold and then audits at different probabilities income reports below and above the threshold. Obviously, there is no assurance that this is the best audit policy. The main reason for this restriction is analytical tractability. Another is that in the real world governments usually rely on simple policies.[4] Taxpayers (agents) have different pre-tax incomes, exogenously given, and are assumed to choose an income report that maximizes expected utility.

In section 3.4 taxpayers are assumed to have the same degree of risk aversion. Proposition 1 shows that the nature of the optimal audit policy under risk neutrality is preserved under risk aversion. Income reports below the threshold are audited at the probability level just sufficient to induce truthful reporting, whereas those above it are not audited. As illustrated above, the outcome is that the *effective* tax schedule is quite regressive.

In section 3.5 attitudes towards risk vary across taxpayers. This enriches the model by allowing individuals with the same pre-tax income to send different reports to the tax agency. At a general level, a full characterization of the optimal audit policy cannot be given. Thus two constrained versions of the audit mechanism are considered, namely random audits and the enforcement policy where all taxpayers with income below the threshold are induced to report truthfully. In

[4]When taxpayers are risk neutral, Sanchez and Sobel (1990) show that the optimal audit mechanism is a step function with at most three probabilty levels.

the latter case it is shown that the optimal policy may involve auditing income reports above the threshold at a positive probability.

Section 3.6 contains some numerical results of the general model. Under plausible assumptions about the distribution of attitudes towards risk among taxpayers, the main finding is that the optimal tax enforcement policy causes only a *limited* regressive bias. Income reports above the threshold meet a positive probability of audit. Also, not all taxpayers with income below the threshold are induced to report truthfully. Section 3.7 concludes by suggesting lines for further research.

3.2 Description of the model

Consider a large population of taxpayers endowed with different pre-tax incomes I, exogenously given. The cumulative distribution of income is $F(I)$, $I \in [0, I^+]$, $I^+ > 0$. Assume that $F(I)$ is twice continuously differentiable and let $f(I) = F'(I)$, $I \in (0, I^+)$, be the density function.

Each taxpayer knows $F(I)$ and costlessly observes his or her own income. The tax enforcement agency knows $F(I)$ but it does not directly observe who earns which income. Taxpayers are thus required to send a report about their own income that the tax administration may then verify at a cost $\varphi > 0$ per audit. It is assumed that when performing an audit the tax agency observes true income.[5]

Consider a taxpayer with pre-tax income I and let r be reported income. If not audited, his or her transfer to the tax agency is tr, where $0 < t < 1$ is the tax

[5] A more realistic assumption about the audit technology would allow for the possibility of imperfect audits and decreasing returns to scale. Also, notice that it has been implicitly assumed that taxpayers bear no compliance costs when filling in the income report and that all taxpayers are known to the tax administration so that they cannot escape taxes by not sending a report.

rate. If audited, the taxpayer pays the income tax plus a proportional surcharge on the evaded tax (if any) at rate $s > 0$: the transfer is $tr + (1+s)t(I-r)$ if $r < I$ and tr if $r \geq I$. The assumption that the distribution of income is common knowledge implies that no taxpayer will report $r < 0$. Also, since in the event of audit there are no rewards for overreporting, no taxpayer will report $r > I$. Thus attention can be restricted to $0 \leq r \leq I$.

The tax enforcement agency decides how many and which reports to audit. Since the only information available is taxpayers' reports, it chooses a probability of audit for each level of reported income. The audit policy is restricted to the step function

$$p(r) = \begin{cases} p_1 & \text{if} \;\; r < I_c \\ p_2 & \text{if} \;\; r \geq I_c \end{cases} \tag{3.1}$$

so that all income reports below a threshold I_c are audited at a constant probability p_1 and those above at a constant probability p_2. The tax agency takes the tax and the penalty rates as given. These are fixed outside the model by a higher level of authority (say the government). It is assumed that $(1+s)t \leq 1$, which implies that a taxpayer audited and caught evading never pays taxes plus penalties exeeding his or her true income.[6] The objective of the tax enforcement agency is to maximize total revenue (taxes plus fines) net of audit costs by choosing p_1, p_2 and I_c while taking taxpayers optimal reports as given.

Each taxpayer is assumed to select an income report that maximizes the expected utility of disposable income in the two states of the world (audit / no audit), taking pre-tax income, tax and penalty rates and audit probabilities as

[6] More precisely, it implies that taxes plus fines on evaded income never exeed this amount. Chander and Wilde (1992) stress that the penalty function should be continuous in reported income and that a taxpayer caught evading must be left with a non-negative disposable income. The penalty function used herein satisfies both requisites.

given. The utility function is of the form

$$U(y;I,\beta)=\begin{cases}\beta y & \text{if } 0\leq y<(1-t)I\\ (\beta-1)(1-t)I+y & \text{if } y\geq(1-t)I\end{cases}\qquad(3.2)$$

where y is disposable income and $\beta \geq 1$ is a parameter capturing the degree of risk aversion. Eq. (3.2) is continuous, concave and linear in y with a kink at $y = (1-t)I$. Marginal utility is constant and is equal to β if $y < (1-t)I$ and 1 otherwise.

When the taxpayer reports truthfully, his or her disposable income is $(1-t)I$ in both states of the world. On the other hand, when $r \neq I$, his or her disposable income is

$$y_1 \equiv (1-t)I + t(I-r) > (1-t)I$$

if no audit occurs and

$$y_2 \equiv (1-t)I - st(I-r) < (1-t)I$$

if an audit occurs.

Eq. (3.2) implies that the prospective gain from evasion is valued, in utility terms, $t(I-r)$, whereas the corresponding loss in the event of audit is valued $\beta st(I-r)$. A higher value of β means a higher degree of risk aversion, for it increases the welfare loss when evasion is discovered. Notice that risk neutrality corresponds to $\beta = 1$.

In section 3.4 all taxpayers are assumed to have identical degree of risk aversion whereas in section 3.5 the parameter β is allowed to vary across taxpayers. To avoid repetitions, in what follows I will refer to *taxpayer with pre-tax income I and parameter of risk aversion β* simply as *taxpayer* (I,β).

The expected utility of taxpayer (I,β) is

$$EU(r;I,\beta) = \big(1-p(r)\big)U(y_1) + p(r)U(y_2)$$

Given the utility function (3.2), the taxpayer problem is

$$\max_{r\in[0,I]} \quad EU(r;I,\beta)=\beta(1-t)I+t\Big[1-p(r)(1+\beta s)\Big](I-r) \tag{3.3}$$

Notice that $EU(r;I,\beta)$ is linear in r and is discontinuous at $r=I_c$ when $p_1\neq p_2$ and $I\geq I_c$. Linearity implies that the optimal report may not be unique. The following assumption is made to rule out multiple solutions.

Assumption 1 *Whenever a taxpayer is indifferent between reporting r_0 and r_1, where $r_0<r_1$, then he or she always chooses to report r_1.*

Let the solution to (3.3) be denoted by $r(I,\beta)$, where the arguments (p_1,p_2,I_c) have been omitted to keep the notation simple. The expected revenue from a taxpayer (I,β) is $T(I,\beta)=tr(.)+p(r(.))(1+s)t(I-r(.))$ and the corresponding expected audit cost is $C(I,\beta)=\varphi p(r(.))$. The set of admissible audit policies is

$$A_0=\Big\{(p_1,p_2,I_c)\mid 0\leq p_1,p_2\leq 1, 0\leq I_c\leq I^+\Big\} \tag{3.4}$$

The next section deals with the economic interpretations of the kinked utility function. The analysis of the problem of the tax enforcement agency will resume in section 3.4.

3.3 Kinked utility function: interpretations

This section presents two economic justifications for adopting the utility function (3.2). The first comes from the consideration of non-pecuniary costs of evasion, as in Gordon (1989). The second is in terms of portfolio selection theory, namely the approach that takes the expected value of loss as measure of risk.

Using the notation of section 3.2, Gordon's utility function is of the form

$$U(y;I,v)=u(y)-v(I-r) \tag{3.5}$$

where $u(y)$ is a standard utility of disposable income, $u' > 0$, $u'' \leq 0$, and v is a constant, $v \geq 0$. The term $v(I-r)$ represents the non- pecuniary cost of evasion, incurred with certainty. Quoting Gordon (1989, p. 798): "One possibility is that evasion generates psychic costs incurred irrespective of whether the act of evasion is observed (by the authorities or anybody else). For example, a false income declaration may induce anxiety, guilt or a reduction in self-image".

Assume risk neutrality, i.e. $u(y) = y$. To ease the notation let $p(r) = p$. Then expected utility is

$$EU(r; I, v) = (1-t)I + t[1 - p(1+s)](I-r) - v(I-r) \tag{3.6}$$

The kinked linear utility function (3.2) gives

$$\begin{aligned} EU(r; I, \beta) &= \beta(1-t)I + t[1 - p(1+\beta s)](I-r) = \\ &= \beta(1-t)I + t[1 - p(1+s)](I-r) + \\ &\qquad -(\beta - 1)pst(I-r) \end{aligned} \tag{3.7}$$

It is now straightforward to spot the difference between the two models. In eq. (3.6) the non-pecuniary cost of evasion is incurred with certainty and depends on the amount of evaded income, whereas in eq. (3.7) it depends on the *expected* value of penalties on the evaded tax.

Turning to the interpretation in terms of optimal portfolio choice,[7] suppose that the taxpayer selects his or her income report on the basis of the *expected return* and a *measure of risk* of the evasion activity, where the measure of risk is the *expected value of loss*.[8] Formally, the taxpayer's return is $M(r; I) = I - tr$

[7]The framing of the tax evasion decision in terms of optimal portfolio choice dates back to the original contribution of Allingham and Sandmo (1972), where it is assumed that a risk averse taxpayer maximizes the expected value of a strictly concave utility function of disposable income. Cowell (1990) refers to this model as the *evader as gambler* model.

[8]This measure of risk is used by Domar and Musgrave (1944) to analyse the impact of taxation on risk taking.

with probability $1-p$ and $M(r;I) = I - tr - (1+s)t(I-r)$ with probability p. The loss is $L(r;I) = 0$ with probability $1-p$ and $L(r;I) = st(I-r)$ with probability p. Markowitz (1959, p. 291) shows that "If an investor (a) maximizes the expected value of some utility function, and (b) chooses among utility functions solely on the basis of expected return and expected loss, then his utility function must be of the form $U = c + aM + bL$." Without loss of generality let $U = M + vL$, $v \geq 0$.

Applying this theorem one gets

$$\begin{aligned} EU(r;I,v) &= (1-t)I + t[1 - p(1+s) - vps](I-r) = \\ &= (1-t)I + t[1 - p(1+s)](I-r) - vpst(I-r) \end{aligned} \tag{3.8}$$

and comparing eqs. (3.7) and (3.8) one immediately notices that the two models are equivalent when $v = \beta - 1$.

3.4 Identical degree of risk aversion

Assume that all taxpayers have the same degree of risk aversion $\bar{\beta} \geq 1$. Then the problem of the tax enforcement agency is to solve

$$\max_{(p_1,p_2,I_c)\in A_0} \quad R(p_1,p_2,I_c) = \int_0^{I^+} \left[T(I,\bar{\beta}) - C(I,\bar{\beta})\right] dF(I) \tag{3.9}$$

The following lemma states that the optimal audit policy is contained in a subset of A_0.

Lemma 1 *Let $(p_1,p_2,I_c) \in A_0$ be some audit policy. Then there exists an audit policy $(p_1',p_2',I_c') \in A_1$ such that $R(p_1',p_2',I_c') \geq R(p_1,p_2,I_c)$, where*

$$A_1 = \left\{(p_1,p_2,I_c) \mid 0 \leq p_2 < p_1 \leq (1+\bar{\beta}s)^{-1},\ 0 \leq I_c \leq I^+\right\}$$

Proof. See appendix A.1.

Optimal probabilities do not exceed $(1+\bar{\beta}s)^{-1}$, the level of enforcement just sufficient to induce truthful reporting, for a higher probability would increase audit costs without raising more tax revenue. Also, reports below the threshold must meet a higher probability of audit than those above it, for otherwise the audit policy would not work as a screening device, separating low and high income taxpayers according to their reports.

The following lemma characterizes taxpayers' optimal reports.

Lemma 2 *For any* $(p_1, p_2, I_c) \in A_1$*, taxpayers' optimal reports are*

If $p_1 < (1+\bar{\beta}s)^{-1}$ *then*

$$r(I,\bar{\beta}) = \begin{cases} 0 & \text{if } I < \min(\bar{\kappa} I_c, I^+) \\ I_c & \text{if } I \geq \min(\bar{\kappa} I_c, I^+) \end{cases}$$

If $p_1 = (1+\bar{\beta}s)^{-1}$ *then*

$$r(I,\bar{\beta}) = \begin{cases} I & \text{if } I < I_c \\ I_c & \text{if } I \geq I_c \end{cases}$$

where

$$\bar{\kappa} = \frac{(1+\bar{\beta}s)^{-1} - p_2}{p_1 - p_2}$$

Proof. See appendix A.1.

Taxpayers are divided into two groups. When $p_1 < (1+\bar{\beta}s)^{-1}$, then $\bar{\kappa} > 1$ and individuals with $I < \bar{\kappa} I_c$ report no income whereas those with $I \geq \bar{\kappa} I_c$ report the threshold. Notice that if the difference between p_1 and p_2 is small, then it may occur that $\bar{\kappa} I_c > I^+$ and all taxpayers report zero income. When $p_1 = (1+\bar{\beta}s)^{-1}$, then $\bar{\kappa} = 1$ and individuals with $I \leq I_c$ report truthfully, whereas those with $I > I_c$ report the threshold.

Let $v \equiv \min(\bar{\kappa} I_c, I^+)$. Applying lemma 2 the net revenue function is

$$R(p_1, p_2, I_c) = \int_0^v \Big[t p_1 (1+s) I - \varphi p_1\Big] dF(I) + \\ + \int_v^{I^+} \Big[t I_c + t p_2 (1+s)(I - I_c) - \varphi p_2\Big] dF(I), \quad \text{if } p_1 < \frac{1}{1+\bar{\beta}s} \tag{3.10}$$

$$R\Big((1+\bar{\beta}s)^{-1}, p_2, I_c\Big) = \int_0^{I_c} \Big[tI - \varphi(1+\bar{\beta}s)^{-1}\Big] dF(I) +$$
$$+ \int_{I_c}^{I^+} \Big[tI_c + tp_2(1+s)(I - I_c) - \varphi p_2\Big] dF(I), \quad \text{if } p_1 = \frac{1}{1+\bar{\beta}s} \qquad (3.11)$$

The parameter β has an effect equivalent to an increase in s when the taxpayer is risk neutral; so it may appear that this model can be reduced to the model of Reinganum and Wilde by a suitable redefinition of variables. But this is not the case and the reason is the following. When $\bar{\beta} = 1$ (taxpayers are risk neutral, i.e. the standard Reinganum and Wilde model) the net revenue function is continuous in all its arguments. In particular, notice that although from lemma 2 optimal reports are discontinuous in p_1 the revenue function is continuous in p_1: when $p_1 < (1+s)^{-1}$ taxpayers with $I < \bar{\kappa} I_c$ report $r(.) = 0$ and the gross revenue is $\int_0^v tp_1(1+s)IdF(I)$, whereas when $p_1 = (1+s)^{-1}$ taxpayers with $I < I_c$ ($\bar{\kappa} = 1$) report $r(.) = I$ and the gross revenue is $\int_0^v tIdF(I)$. On the other hand, when $\bar{\beta} > 1$ (taxpayers are risk averse) the net revenue function is continuous in p_2 and I_c, whereas it shows a discontinuity in p_1 at $p_1 = (1+\bar{\beta}s)^{-1}$ which takes the form of an upward jump: optimal reports are discontinuous at $p_1 = (1+\bar{\beta}s)^{-1} < (1+s)^{-1}$ so that gross revenue (for $I < v$) jumps from $\int_0^v t(1+\bar{\beta}s)^{-1}(1+s)IdF(I)$ to $\int_0^v tIdF(I)$. About the differences between the two models ($\bar{\beta} = 1$ and $\bar{\beta} > 1$) see also the comment to eq. (3.14) below.

The following assumption is made to ensure that the solution of the tax enforcement problem is well behaved, namely that the optimal threshold is unique.

Assumption 2 *$F(I)$ has a strictly increasing hazard rate. That is, $h'(I) > 0$ for all $I \in (0, I^+)$, where $h(I) \equiv f(I)/[1 - F(I)]$.*

Since $h(I) \geq 0$, assumption 2 implies $f(I^+) > 0$. The following assumption is somewhat arbitrary and not strictly necessary, its purpose being to ease the presentation, for it ensures that the optimal threshold is an interior point of its support set.

Assumption 3 $f(0) = 0$.

Proposition 1 *Under assumptions 2-3, the optimal audit policy is unique and is given by* $p_1^* = (1+\bar{\beta}s)^{-1}$, $p_2^* = 0$ *and* $I_c^* \in (0, I^+)$ *that solves*

$$\frac{1}{h(I_c)} = \frac{\varphi}{(1+\bar{\beta}s)t} \tag{3.12}$$

Proof. The proof is given in two steps. The first shows that eq. (3.10) does not have an interior maximum in p_1 and that its optimal value is $p_1 \to (1+\bar{\beta}s)^{-1}$. Since the revenue function has an upward jump at $p_1 = (1+\bar{\beta}s)^{-1}$, this proves that $p_1^* = (1+\bar{\beta}s)^{-1}$. The second step shows that eq. (3.11) has a maximum at $p_2^* = 0$ and determines I_c^*.

Step 1. After some algebra involving integration by parts one finds that, for any $I_c \in [0, I^+/\bar{\kappa}]$, eq. (3.10) can be written as

$$R(p_1, p_2, I_c) = t(1+s)\frac{\kappa}{\bar{\kappa}}\left\{p_1 \int_0^{\bar{\kappa}I_c} \Psi(I, p_2)dI + p_2 \int_{\bar{\kappa}I_c}^{I^+} \Psi(I, p_2)dI\right\} \tag{3.13}$$

where

$$\Psi(I, p_2) = 1 - F(I) - \frac{\kappa - \bar{\kappa}}{\kappa} If(I) - \frac{\bar{\kappa}}{\kappa}\frac{\varphi}{(1+s)t}f(I), \tag{3.14}$$

$$\kappa \equiv \frac{(1+s)^{-1} - p_2}{p_1 - p_2} \quad \text{and} \quad \bar{\kappa} \equiv \frac{(1+\bar{\beta}s)^{-1} - p_2}{p_1 - p_2} \tag{3.15}$$

Notice that $\Psi(.)$ is independent of p_1 and I_c and that $\kappa \geq \bar{\kappa}$. Also, when $\bar{\beta} = 1$ then $\kappa = \bar{\kappa}$ and the third term of eq. (3.14) vanishes.

Differentiating with respect to I_c, the first order condition for an interior solution is

$$\frac{\partial R}{\partial I_c} = t(p_1 - p_2)(1+s)\kappa\Psi(\bar{\kappa}I_c, p_2) = 0 \tag{3.16}$$

Let $z \equiv \bar{\kappa}I_c$. Since $I_c \in [0, I^+/\bar{\kappa}]$, then $z \in [0, I^+]$. After this change of variable, the first order condition (3.16) is simply $\Psi(z, p_2) = 0$. From assumption 3,

$\Psi(0,p_2) > 0$ and from assumption 2, $\Psi(I^+,p_2) < 0$. Continuity of $\Psi(.)$ and assumption 2 then imply that there exists a unique $z^*(p_2) \in (0, I^+)$ such that $\Psi(z,p_2) = 0$. In particular, $z^*(p_2)$ solves

$$\frac{1}{h(z)} - \frac{\kappa - \bar{\kappa}}{\kappa} z = \frac{\bar{\kappa}}{\kappa} \frac{\varphi}{(1+s)t} \tag{3.17}$$

The revenue function (3.13) can now be written as

$$\begin{aligned} R\left(p_1, p_2, \frac{z^*(p_2)}{\bar{\kappa}}\right) &= tp_1(1+s)\frac{\kappa}{\bar{\kappa}} \int_0^{z^*(p_2)} \Psi(I,p_2)dI + \\ &+ tp_2(1+s)\frac{\kappa}{\bar{\kappa}} \int_{z^*(p_2)}^{I^+} \Psi(I,p_2)dI \end{aligned} \tag{3.18}$$

Differentiating with respect to p_1

$$\frac{\partial R}{\partial p_1} = t(1+s)\frac{\kappa}{\bar{\kappa}} \int_0^{z^*(p_2)} \Psi(I,p_2)dI > 0 \tag{3.19}$$

Since $\Psi(.) > 0$ for all $z \in [0, z^*)$ and $\Psi(.) < 0$ for all $z \in (z^*, I^+]$, the derivative (3.19) is positive for any p_2. Thus $p_1^* \to (1+\bar{\beta}s)^{-1}$. However, since the revenue function has a positive jump at $p_1 = (1+\bar{\beta}s)^{-1}$, this implies that $p_1^* = (1+\bar{\beta}s)^{-1}$.

Step 2. After some algebra, eq. (3.11) can be written as

$$\begin{aligned} R(p_1^*, p_2, I_c) = t \int_0^{I_c} &\left[1 - F(I) - \frac{\varphi}{(1+\bar{\beta}s)t} f(I)\right] dI + \\ + tp_2(1+s) \int_{I_c}^{I^+} &\left[1 - F(I) - \frac{\varphi}{(1+s)t} f(I)\right] dI \end{aligned} \tag{3.20}$$

The first order condition for an interior solution with respect to I_c is

$$\begin{aligned} \frac{\partial R}{\partial I_c} &= t\big[1 - p_2(1+s)\big]\left\{1 - F(I_c) - \frac{\varphi}{(1+s)t} f(I_c)\right\} + \\ &+ \varphi\left(\frac{1}{1+s} - \frac{1}{1+\bar{\beta}s}\right) f(I_c) = 0 \end{aligned} \tag{3.21}$$

From assumptions 2-3 there exists a unique $I_c(p_2) \in (0, I^+)$ that solves eq. (3.21). In particular, $I_c(p_2)$ solves

$$\frac{1}{h(I_c)} = \frac{\varphi}{(1+s)t} \frac{(1+\bar{\beta}s)^{-1} - p_2}{(1+s)^{-1} - p_2} \tag{3.22}$$

The revenue function (3.20) can now be written as

$$R\big(p_1^*, p_2, I_c(p_2)\big) = t\int_0^{I_c(p_2)} \left[1 - F(I) - \frac{\varphi}{(1+\bar{\beta}s)t} f(I)\right] dI + \\ + tp_2(1+s)\int_{I_c(p_2)}^{I^+} \left[1 - F(I) - \frac{\varphi}{(1+s)t} f(I)\right] dI \tag{3.23}$$

Differentiating with respect to p_2

$$\frac{\partial R}{p_2} = t(1+s)\int_{I_c(p_2)}^{I^+} \left[1 - F(I) - \frac{\varphi}{(1+s)t} f(I)\right] dI < 0 \tag{3.24}$$

From eq. (3.22), $h(I) < \varphi/[(1+s)t]$ for all $I \in (I_c(p_2), I^+]$. Therefore the derivative (3.24) is negative for all p_2, which implies $p_2^* = 0$. Also, I_c^* solves eq. (3.12). *Q.E.D.*

Proposition 1 shows that the introduction of risk aversion ($\bar{\beta} > 1$) does not affect the qualitative properties of the optimal audit policy under risk neutrality ($\bar{\beta} = 1$). Income reports below the threshold are audited at the probability level just sufficient to induce truthful reporting, whereas those above it are not audited. This causes the *effective* tax schedule to be regressive: the effective average tax rate of a taxpayer with $I < I_c^*$ is t (the *legislated* tax rate), whereas that of an individual with $I \geq I_c^*$ is tI_c^*/I, which is less than t and is decreasing in I.

However, for given distribution of pre-tax incomes $f(I)$, a higher degree of risk aversion $\bar{\beta}$ implies a lower p_1^* and a higher I_c^*. Thus more taxpayers are forced to report truthfully and this reduces the size of the regressive bias.

The assumption that taxpayers have the same degree of risk aversion is quite restrictive, for it implies that all individuals with the same pre-tax income make the same report. This point is taken up in the following section, where the concavity of the utility function is allowed to vary across taxpayers.

3.5 Different attitudes towards risk

Suppose that taxpayers have different degrees of risk aversion. Let $G(\beta)$, $\beta \in [1,\infty)$, be the cumulative distribution of β. Assume that $G(\beta)$ is twice continuously differentiable with density function $g(\beta) = G'(\beta)$, $\beta \in (1,\infty)$. Notice that it has been implicitly assumed that I and β are independently distributed. The independence assumption has been made for the sake of analytical tractability even though intuition would suggest as plausible a negative correlation between the level of pre-tax income and the degree of risk aversion.

The tax enforcement agency solves

$$\max_{(p_1,p_2,I_c)\in A_0} R(p_1,p_2,I_c) = \int_0^{I^+}\int_1^{\infty}\Big[T(I,\beta) - C(I,\beta)\Big]dG(\beta)dF(I) \qquad (3.25)$$

The following lemma states that the optimal audit policy is contained in a subset of A_0.

Lemma 3 *Let $(p_1,p_2,I_c) \in A_0$ be some audit policy. Then there exists an audit policy $(p_1',p_2',I_c') \in A_2$ such that $R(p_1',p_2',I_c') \geq R(p_1,p_2,I_c)$, where*

$$A_2 = \Big\{(p_1,p_2,I_c) \mid 0 \leq p_2 < p_1 \leq (1+s)^{-1},\ 0 \leq I_c \leq I^+\Big\}$$

Proof. See appendix A.2.

The following lemma characterizes taxpayers' optimal reports.

Lemma 4 *For any $(p_1,p_2,I_c) \in A_2$, taxpayers' optimal reports are*

$$r(I,\beta) = \begin{cases} 0 & \text{if } \langle I,\beta\rangle \in [0,I_c) \times [1,\beta_1) & a_1 \\ I & \text{if } \langle I,\beta\rangle \in [0,I_c) \times [\beta_1,\infty) & a_2 \\ 0 & \text{if } \langle I,\beta\rangle \in [I_c,\min(\kappa I_c,I^+)) \times [1,\tilde{\beta}) & a_3 \\ I_c & \text{if } \langle I,\beta\rangle \in [I_c,\min(\kappa I_c,I^+)) \times [\tilde{\beta},\beta_2) & a_4 \\ I_c & \text{if } \langle I,\beta\rangle \in [\min(\kappa I_c,I^+),I^+] \times [1,\beta_2) & a_5 \\ I & \text{if } \langle I,\beta\rangle \in [I_c,I^+] \times [\beta_2,\infty) & a_6 \end{cases}$$

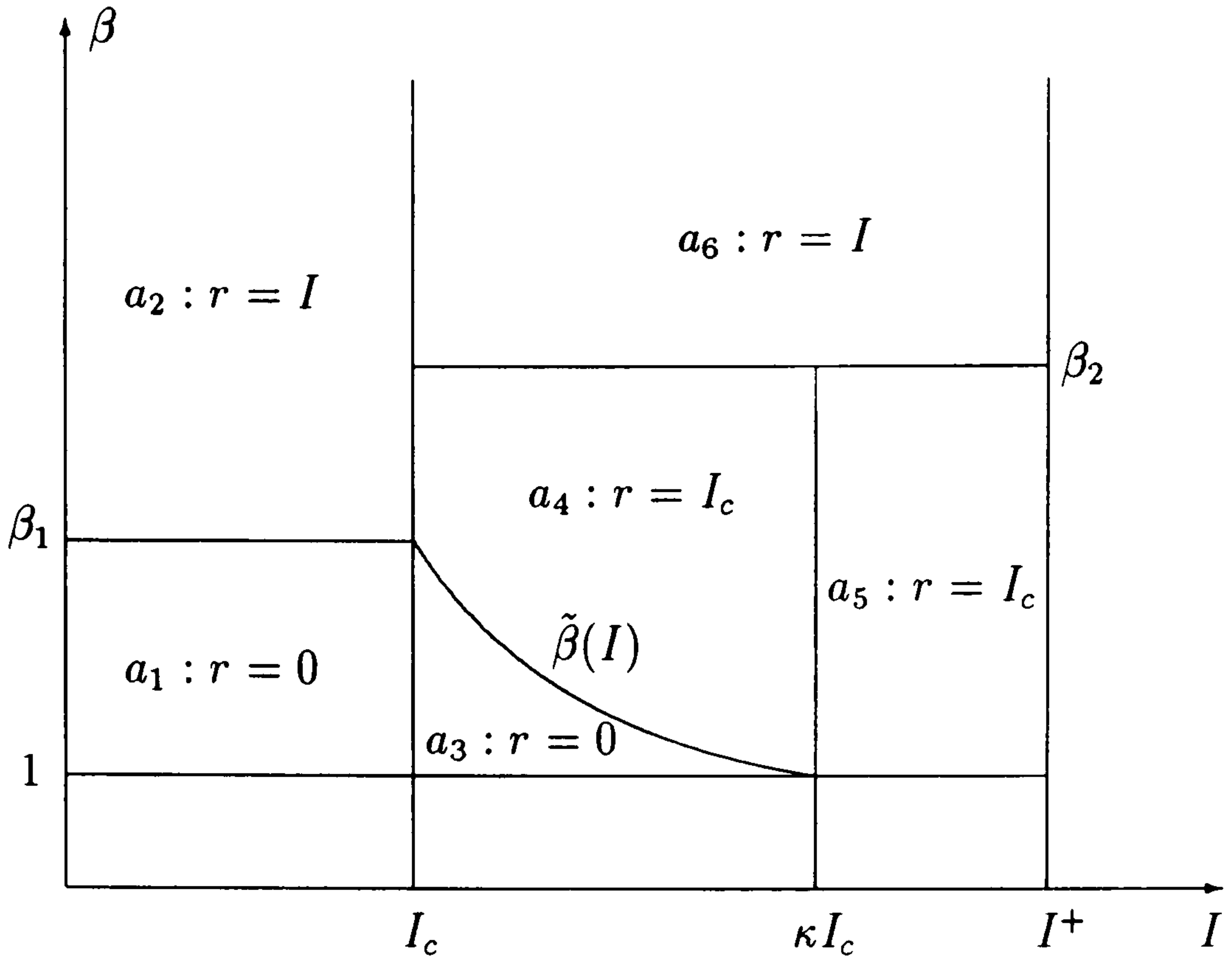

Figure 3.1: Illustration of lemma 4

where

$$\beta_1 \equiv \frac{1-p_1}{sp_1}, \qquad \beta_2 \equiv \frac{1-p_2}{sp_2},$$

$$\tilde{\beta}(I) \equiv \frac{(1-p_2)I_c - (p_1-p_2)I}{s[p_2 I_c + (p_1-p_2)I]}, \qquad \kappa \equiv \frac{(1+s)^{-1}-p_2}{p_1-p_2}$$

Proof. See appendix A.2.

Notice that β_1 and β_2 are decreasing functions of p_1 and p_2 respectively and that from lemma 3, $1 \leq \beta_1 < \beta_2 < \infty$. Also, $\kappa \geq 1$.

Lemma 4 is illustrated in figure 3.1. Taxpayers (I, β) with $I \in [0, I_c)$ are separated into two groups: those with $\beta \in [1, \beta_1)$ reporting no income and those with $\beta \in [\beta_1, \infty)$ reporting truthfully. An increase in p_1 lowers β_1 so that the number of individuals reporting truthfully increases.

Taxpayers (I, β) with $I \in [I_c, I^+]$ report zero income if $\beta \in [1, \tilde{\beta})$, report the threshold if $\beta \in [\tilde{\beta}, \beta_2)$ and report truthfully if $\beta \in [\beta_2, \infty)$. The separation of individuals declaring zero income from those declaring the threshold is not clearcut, for $\tilde{\beta}(I)$ is a strictly decreasing convex function of I such that $\tilde{\beta}(I_c) = \beta_1$ and $\tilde{\beta}(\kappa I_c) = 1$. When $\kappa I_c \leq I^+$, see figure 3.1, $r(.) = 0$ if $\langle I, \beta\rangle \in [I_c, \kappa I_c) \times [1, \tilde{\beta})$ and $r(.) = I_c$ if $\langle I, \beta\rangle \in [I_c, \kappa I_c) \times [\tilde{\beta}, \beta_2)$ and if $\langle I, \beta\rangle \in [\kappa I_c, I^+] \times [1, \beta_2)$. When $\kappa I_c > I^+$, $r(.) = 0$ if $\langle I, \beta\rangle \in [I_c, I^+] \times [1, \tilde{\beta})$ and $r(.) = I_c$ if $\langle I, \beta\rangle \in [I_c, I^+] \times [\tilde{\beta}, \beta_2)$, where $\tilde{\beta}(I_c) = \beta_1$ and $1 < \tilde{\beta}(I^+) < \beta_1$.

Lemma 4 shows that the interaction between pre-tax income and degree of risk aversion in determining optimal reports is quite rich. Taxpayers with the same pre-tax income but differing attitude towards risk may have distinct optimal reports. In turn, individuals with the same degree of risk aversion but distinct pre-tax income may send different reports.

Let $v \equiv \min(\kappa I_c, I^+)$. Applying lemma 4 the net revenue function is

$$R(p_1, p_2, I_c) = T_1 + T_2 + T_3 - C \tag{3.26}$$

where

$$\begin{aligned}
T_1 &= tp_1(1+s)\left\{G(\beta_1)\int_0^{I_c} If(I)dI + \int_{I_c}^{v} IG(\tilde{\beta})f(I)dI\right\} \\
T_2 &= \int_{I_c}^{v}\Big[tI_c + tp_2(1+s)(I - I_c)\Big]\Big[G(\beta_2) - G(\tilde{\beta})\Big]f(I)dI \; + \\
&\qquad + \int_{v}^{I^+}\Big[tI_c + tp_2(1+s)(I - I_c)\Big]G(\beta_2)f(I)dI \\
T_3 &= t\Big[1 - G(\beta_1)\Big]\int_0^{I_c} If(I)dI + t\Big[1 - G(\beta_2)\Big]\int_{I_c}^{I^+} If(I)dI \\
C &= \varphi p_1\left\{F(I_c) + \int_{I_c}^{v} G(\tilde{\beta})f(I)dI\right\} + \\
&\qquad + \varphi p_2\left\{1 - F(I_c) - \int_{I_c}^{v} G(\tilde{\beta})f(I)dI\right\}
\end{aligned}$$

The term T_1 contains penalties collected from individuals reporting zero income, the size of the group being

$$G(\beta_1)F(I_c) + \int_{I_c}^{v} G\big(\tilde{\beta}(I)\big)f(I)dI$$

The term T_2 is gross revenue collected from taxpayers reporting I_c, the number of them being

$$\int_{I_c}^{v} \left\{G(\beta_2) - G\big(\tilde{\beta}(I)\big)\right\} f(I)dI + G(\beta_2)\big[1 - F(v)\big]$$

The term T_3 is the tax revenue collected from truthful taxpayers, the size of the group being

$$\big[1 - G(\beta_1)\big]F(I_c) + \big[1 - G(\beta_2)\big]\big[1 - F(I_c)\big]$$

Finally, C represents audit costs. The first term in braces is the number of taxpayers whose income report is below I_c. The second term is the number of individuals with an income report at least as high as the threshold.

Before examining the general model, two constrained versions of the tax enforcement policy are considered.

3.5.1 Random audits

Consider the constrained version of problem (3.25) where $I_c = 0$, so that all income reports meet the audit probability p_2. From lemma 4, taxpayers optimal reports are

$$r(I,\beta) = \begin{cases} 0 & \text{if } \langle I,\beta\rangle \in [0,I^+] \times [1,\beta_2) \\ I & \text{if } \langle I,\beta\rangle \in [0,I^+] \times [\beta_2,\infty) \end{cases}$$

Individuals with $\beta < \beta_2$ report zero income, whereas those with $\beta \geq \beta_2$ report truthfully.

The tax enforcement agency solves

$$\max_{p_2 \in [0,(1+s)^{-1}]} R(p_2) = p_2(1+s)t\bar{I}G(\beta_2) + t\bar{I}\big[1 - G(\beta_2)\big] - \varphi p_2 \tag{3.27}$$

where $\bar{I} = \int_0^{I^+} IdF(I)$ is average income. The first term in the revenue function represents fines collected from taxpayers reporting $r = 0$, the size of the group being $G(\beta_2)$. The second term represents taxes collected from truthful taxpayers, the number of them being $1 - G(\beta_2)$. Problem (3.27) is well defined, for $R(p_2)$ is continuous on a compact set.

The first order necessary condition for an interior solution is

$$\frac{dR}{dp_2} = -t\big[1 - p_2(1+s)\big]\bar{I}g(\beta_2)\frac{d\beta_2}{dp_2} + t(1+s)\bar{I}G(\beta_2) - \varphi = 0 \tag{3.28}$$

The first term, which is positive, shows the gain in gross revenue that springs from the fact that as p_2 increases more taxpayers are induced to report truthfully. The second term, as well positive, shows the increase in fines collection as more audits occur. The third term is negative and is equal to minus unit audit costs.

The second order condition for a maximum is

$$\frac{d^2R}{dp_2^2} = \frac{2t\bar{I}}{p_2}g(\beta_2)\frac{d\beta_2}{dp_2} - t\big[1 - p_2(1+s)\big]\bar{I}\left(\frac{d\beta_2}{dp_2}\right)^2 g'(\beta_2) < 0 \tag{3.29}$$

The first term is negative whereas the sign of the second is determined by $g'(\beta_2)$, the slope of the density function of β.

For the sake of comparison, suppose that all taxpayers are risk neutral. Then the revenue function is $R(p_2) = tp_2(1+s)\bar{I} - p_2\varphi$, which is linear in p_2. Thus the optimal random audit policy is $p_2 = (1+s)^{-1}$ if $(1+s)t\bar{I} > \varphi$ and $p_2 = 0$ otherwise. In the first case the taxing authority induces all taxpayers to report truthfully, whereas in the second gives up the task of enforcing compliance. Of the two, it seems more plausible to assume that $(1+s)t\bar{I} > \varphi$.

Eqs. (3.28)-(3.29) clearly show that the optimal random audit probability crucially depends on the distribution of β. From (3.28), $dR/dp_2 = -\varphi$ at $p_2 = (1+s)^{-1}$ and $dR/dp_2 = t(1+s)\bar{I} - \varphi$ at $p_2 = 0$, provided that $\lim_{p_2 \to 0} g(\beta_2)\frac{d\beta_2}{dp_2} = 0$. Therefore, if $t(1+s)\bar{I} > \varphi$, there exists at least a $p_2^* \in (0, (1+s)^{-1})$ such that $dR/dp_2 = 0$. Notice that $t(1+s)\bar{I} > \varphi$ is sufficient but not necessary for the existence of an interior solution.

Sufficient conditions for uniqueness cannot be given, for d^2R/dp_2^2 is unlikely to be everywhere negative. However, suppose that $g(\beta)$ has a unique mode $\hat{\beta} \geq 1$,

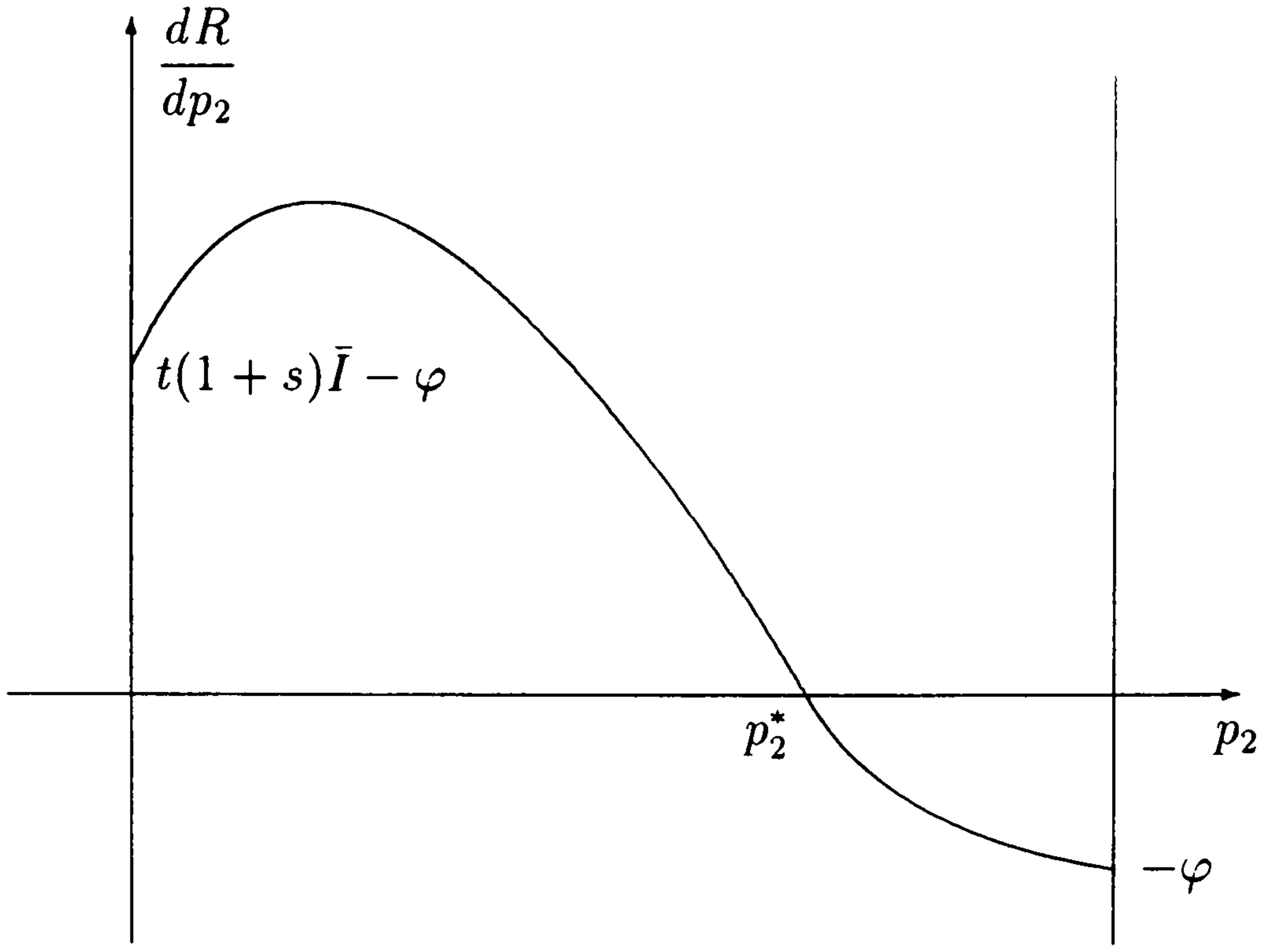

Figure 3.2: Random audits

so that $g'(\beta) > 0$ for all $\beta \in [1, \hat{\beta})$ and $g'(\beta) < 0$ otherwise. Then there exists a $\hat{p}_2$ such that $g'(\beta_2) < 0$ for all $p_2 < \hat{p}_2$ and $g'(\beta_2) > 0$ for all $p_2 > \hat{p}_2$. In this case the first order derivative (3.28) may behave as illustrated in figure 3.2, where a unique solution occurs.[9]

[9]Suppose that the distribution of β is exponential with standard deviation σ. Then $g'(\beta) < 0$ for all β so that the second term in (3.29) is always positive. However, one can show that $d^2R/dp_2^2 > 0$ for all $p_2 < (2s\sigma + 1 + s)^{-1}$ and $d^2R/dp_2^2 < 0$ otherwise, so that when $t(1+s)\bar{I} > \varphi$ the graph of $\partial R/\partial p_2$ is as illustrated in figure 3.2 and the optimal random audits policy is unique. Numerical computations show that the optimal random audit probability is unique also when $g(\beta)$ is normally distributed. See section 3.6 for the numerical results when $g(\beta)$ is normal or exponential and $f(I)$ is lognormal.

3.5.2 Another constrained problem

Suppose that the tax enforcement agency fixes $p_1 = (1+s)^{-1}$. From lemma 4, optimal reports are

$$r(I,\beta) = \begin{cases} I & \text{if } \langle I,\beta\rangle \in [0,I_c) \times [1,\infty) \\ I_c & \text{if } \langle I,\beta\rangle \in [I_c,I^+] \times [1,\beta_2) \\ I & \text{if } \langle I,\beta\rangle \in [I_c,I^+] \times [\beta_2,\infty) \end{cases}$$

Taxpayers with $I < I_c$ are forced to report truthfully regardless of their attitude towards risk. Those with $I \geq I_c$ report truthfully if $\beta > \beta_2$ and report the threshold if $\beta \leq \beta_2$.

The tax agency solves

$$\begin{aligned} \max_{p_2,I_c} R\left(\frac{1}{1+s}, p_2, I_c\right) &= \int_0^{I_c} \left(tI - \frac{\varphi}{(1+s)}\right) f(I)dI + \\ &+ G(\beta_2)\int_{I_c}^{I^+} \Big[tI_c + tp_2(1+s)(I-I_c) - \varphi p_2\Big] f(I)dI + \\ &+ \Big[1 - G(\beta_2)\Big]\int_{I_c}^{I^+} (tI - \varphi p_2) f(I)dI \end{aligned} \tag{3.30}$$

After some algebra, the net revenue function can be expressed as

$$\begin{aligned} R\left(\frac{1}{1+s}, p_2, I_c\right) &= t\int_0^{I_c} \left[G(\beta_2)\Big(1 - F(I)\Big) - \frac{\varphi}{(1+s)t} f(I)\right] dI + \\ &+ tp_2(1+s)\int_{I_c}^{I^+} \left[G(\beta_2)\Big(1 - F(I)\Big) - \frac{\varphi}{(1+s)t} f(I)\right] dI + \\ &+ t\Big[1 - G(\beta_2)\Big]\bar{I} \end{aligned} \tag{3.31}$$

The first order condition for an interior solution with respect to I_c is

$$\frac{\partial R}{\partial I_c} = t\Big[1 - p_2(1+s)\Big]\left[G(\beta_2)\Big(1 - F(I_c)\Big) - \frac{\varphi}{(1+s)t} f(I_c)\right] = 0 \tag{3.32}$$

Under assumptions 2-3, for any $0 \leq p_2 < (1+s)^{-1}$, there exists a unique $I_c(p_2) \in (0, I^+)$ that solves eq. (3.32). In particular, it is defined by

$$\frac{1}{h(I_c)} = \frac{\varphi}{(1+s)t}\,\frac{1}{G(\beta_2)} \tag{3.33}$$

Notice that as p_2 increases, $I_c(p_2)$ decreases and that $\lim_{p_2 \to (1+s)^{-1}} I_c(p_2) = 0$.

To obtain the optimality condition for p_2, first replace I_c with $I_c(p_2)$ into the revenue function (3.31) and then differentiate to get

$$\begin{aligned}\frac{\partial R}{\partial p_2} &= t(1+s)\int_{I_c(p_2)}^{I^+}\left[G(\beta_2)\big(1-F(I)\big)-\frac{\varphi}{(1+s)t}f(I)\right]dI\ + \\ &\quad - t\Big[1-p_2(1+s)\Big]g(\beta_2)\frac{d\beta_2}{dp_2}\int_{I_c(p_2)}^{I^+}\Big[1-F(I)\Big]dI = 0 \qquad (3.34)\end{aligned}$$

The first term is negative because, from the definition of $I_c(p_2)$, see eq. (3.33),

$$\frac{1}{h(I)} < \frac{\varphi}{(1+s)tG(\beta_2)}$$

for all $I \in (I_c(p_2), I^+]$. The second term is positive.

At a general level, the only result is that

$$\left.\frac{\partial R}{\partial p_2}\right|_{p_2=0} = \int_{I_c(0)}^{I^+}\left[1-F(I)-\frac{\varphi}{(1+s)t}f(I)\right]dI < 0 \qquad (3.35)$$

$$\left.\frac{\partial R}{\partial p_2}\right|_{p_2\to(1+s)^{-1}} = -\varphi < 0 \qquad (3.36)$$

Eq. (3.35) implies that $p_2^* = 0$ is a local maxima of problem (3.30). Notice that this corresponds to the optimal audit policy under risk neutral taxpayers.

Clearly, the graph of $\partial R/\partial p_2$ subject to $I_c = I_c(p_2)$ is shaped by the distributions of β and I. Numerical computations, see section 3.6 for the results, show that this graph is as depicted in figure 3.3, when $f(I)$ has a lognormal distribution and $g(\beta)$ is normal or exponential. As illustrated, there are two local maxima: one at $p_2^* = 0$, $I_c^* = I_c(p_2^*)$ and one at $p_2^{**} < (1+s)^{-1}$, $I_c^{**} = I_c(p_2^{**})$. The numerical results show that the second audit policy raises more revenue. Also, notice that it is markedly more equitable, for $p_2^* < p_2^{**}$ and $I_c^* > I_c^{**}$.

3.5.3 The general problem

The analysis of the two constrained problems has shown that the characterization of the optimal audit policy is not straightforward even when attention is restricted to simple audit mechanisms.

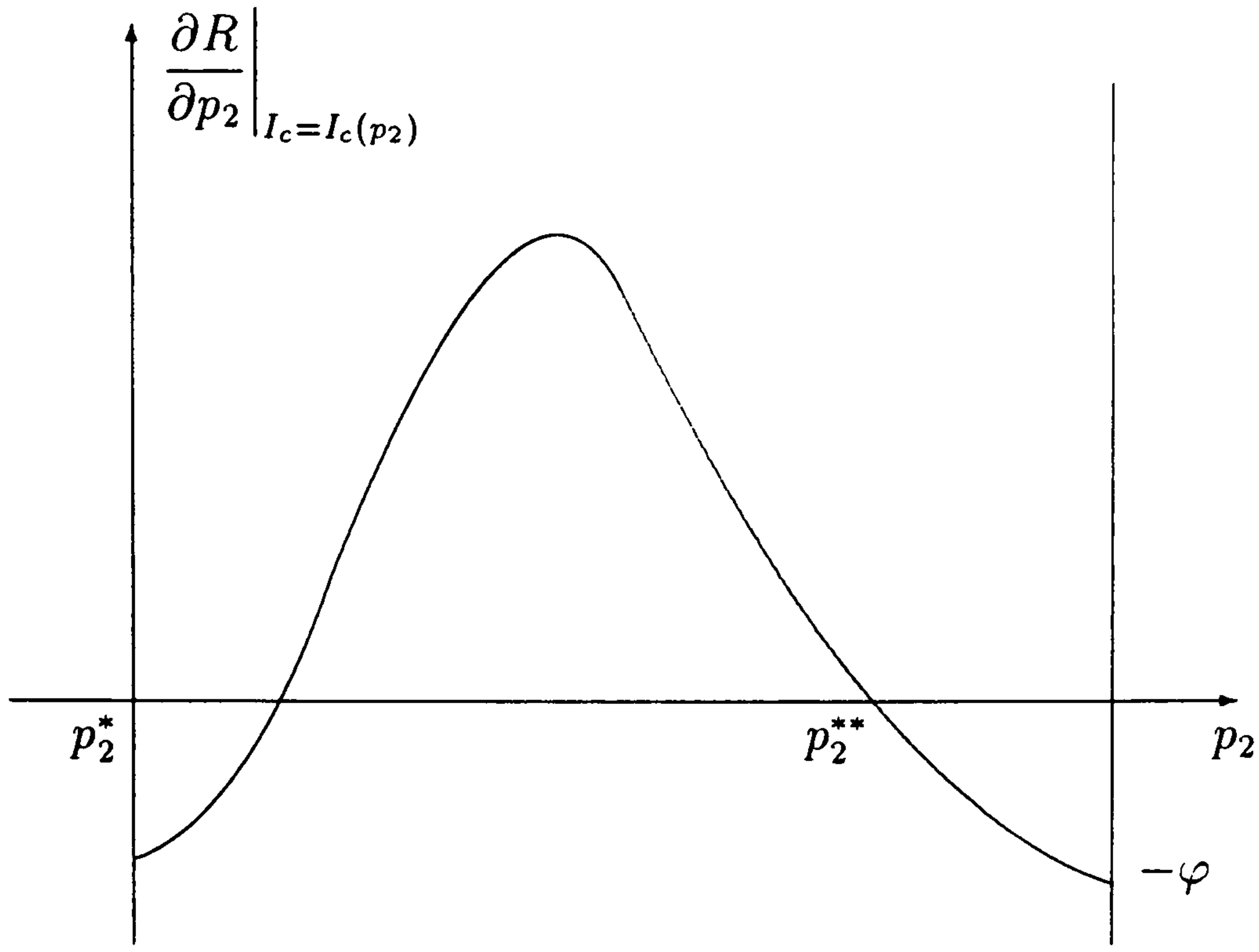

Figure 3.3: Another constrained problem

Consider the unconstrained problem. Differentiating eq. (3.26) with respect to I_c, the first order condition for an interior solution is

$$\begin{aligned}\frac{\partial R}{\partial I_c} &= t(1+s)(p_1-p_2)\kappa\left\{G(\beta_2)\left[1-F(I_c)\right]-\int_{I_c}^{v} G(\tilde{\beta})f(I)dI\right\}+\\&-\varphi(p_1-p_2)\left\{\left[1-G(\beta_1)\right]f(I_c)+\int_{I_c}^{v} g(\tilde{\beta})\frac{\partial\tilde{\beta}}{\partial I_c}f(I)dI\right\}+\\&-t(1+s)(p_1-p_2)\int_{I_c}^{v}(\kappa I_c-I)g(\tilde{\beta})\frac{\partial\tilde{\beta}}{\partial I_c}f(I)dI=0 \qquad (3.37)\end{aligned}$$

The first term in braces is the number of taxpayers reporting $r = I_c$. The second term in braces is the derivative, with respect to I_c, of the number of individuals reporting an income below the threshold. No interpretation can be given to the third term.

Eq. (3.37) gives a grasp of the complexity of the problem of solving for the optimal audit policy. The first order conditions with respect to the audit probabilities are not presented because of their length and because no useful economic interpretations can be given. At this level of generality, the following proposition gives two limited results.

Proposition 2 *Under assumptions 2–3, for any* $0 \leq p_2 < p_1 \leq (1+s)^{-1}$ *there exists at least a* $I_c(p_1, p_2) \in (0, I^+)$ *such that* $\partial R/\partial I_c = 0$. *Also,* $p_1^* < (1+s)^{-1}$.

Proof. From (3.37) and assumptions 3 and 2 one gets respectively

$$\left.\frac{\partial R}{\partial I_c}\right|_{I_c=0} = t\big[1 - p_2(1+s)\big]G(\beta_2) > 0$$
$$\left.\frac{\partial R}{\partial I_c}\right|_{I_c=I^+} = -\varphi(p_1 - p_2)\big[1 - G(\beta_1)\big]f(I^+) < 0$$

Then continuity of $\partial R/\partial I_c$ implies that there exists at least one $I_c \in (0, I^+)$, which depends on p_1, p_2, such that $\partial R/\partial I_c = 0$. To prove the second part of the proposition, simply differentiate the revenue function with respect to p_1 to obtain

$$\left.\frac{\partial R}{\partial p_1}\right|_{p_1=(1+s)^{-1}} = -\varphi F(I_c) < 0$$

which is strictly negative for any $I_c \in (0, I^+)$. *Q.E.D.*

As for the first part of the proposition, notice that the assumption that the distribution of income has a strictly increasing hazard rate is no longer sufficient to ensure that the optimal threshold is unique. The second part shows that it does not pay to induce all taxpayers with $I < I_c$ to report truthfully. Nothing can be said about the probability p_2.

3.6 Numerical results

In this section the results of some numerical computations are presented. The purpose is to explore the relation between the distribution of attitudes towards risk among taxpayers and the properties of the optimal audit policy.

The specification of the model is as follows.

Distribution of income. The density function $f(I)$ is a truncated lognormal with parameters of the associated normal distribution $\mu = 1.533$, $\sigma = 0.4$ and truncation point at $I^+ = 15$. The resulting distribution has mode $\hat{I} = 3.95$, median $I_{50} = 4.63$ and mean (national income) $\bar{I} = 5.00$. The poorest 5- percent of the population has incomes below 2.39, whereas the richest 5-percent has incomes above 8.96. The poorest 25- percent of the population possesses 14.2% of national income, the second 25-percent possesses 20.4%, the third 26.4% and finally the richest 25-percent owns 39.0% of national income.

Attitudes towards risk. The numerical computations differ about the specification of the distribution of β.

Model 1. Taxpayers are risk neutral. That is, $\bar{\beta} = 1$.

Model 2. All taxpayers have the same degree of risk aversion $\bar{\beta} = 5$.

Model 3. The density function $g(\beta)$ is a truncated normal with parameters of the associated normal distribution $\mu = 5$, $\sigma = 1.33$ and truncation point at $\beta = 1$. The resulting distribution has mode $\hat{\beta} = 5$ and mean $\bar{\beta} = 5.006$. Notice that 54.7 percent of taxpayers have $\beta \in (4,6)$ and 86.8 percent have $\beta \in (3,7)$. Only 1.08 percent of them are *almost* risk neutral, with $\beta \in (1,2)$.

Model 4. Attitudes towards risk are distributed according to the exponential density $g(\beta) = \sigma^{-1} \exp(-(\beta - 1)/\sigma)$, with $\sigma = 4$. The mode is $\hat{\beta} = 1$, the median $\beta_{50} = 3.77$ and the mean $\bar{\beta} = 5$. Although the average degree of risk aversion is the same as model 3, notice that 22.1 percent of taxpayers are *almost* risk neutral, with $\beta \in (1,2)$. The share of taxpayers with $\beta \in (4,6)$ is 18.6 percent, whereas those with $\beta \in (3,7)$ are 38.3 percent.

Tax and penalty rates and unit audit costs. These are fixed at $t = 0.2$, $s = 4$, $\varphi = 2$. Notice that average income is 2.5 times unit audit costs.

Let T be gross revenue, C total audit costs and $R = T - C$ net revenue. The numerical results are as follows.

Model 1. When taxpayers are risk neutral, the optimal audit policy is $p_1^* = .2$, $p_2^* = 0$, $I_c^* = 5.544$, $R = .6139$, $T = .8835$, $C = .2696$. The optimal threshold is just above average income $\bar{I} = 5$, the percentage of taxpayers with $I < I_c^*$ and reporting truthfully is 67.41%, whereas the remaining 32.59% report the threshold and evade the difference between I and I_c. In terms of net revenue collected, the optimal cutoff policy does just better than optimal random audits, where $I_c = 15$, $p_1 = .2$, $R = .6$, $T = 1.0$, $C = .4$. However, audit costs are much higher with random audits. As for the equity issue, with random audits all taxpayers are forced to behave honestly and therefore the *effective* and the *legislated* tax schedules coincide: the average tax rate is equal to 0.2 for all income levels. Instead, the audit cutoff policy causes a considerable regressive bias: for $I \leq I_c^*$ the effective and legislated average tax rates are equal, whereas for $I > I_c^*$ the effective rate falls below the legislated rate and is decreasing in I. To illustrate, the average tax rate is .184 for a taxpayer with $I = 6$, .138 when $I = 8$, .092 if $I = 12$ and .074 for the highest income level $I = 15$.

Model 2. When all taxpayers have $\bar{\beta} = 5$, the optimal audit policy is $p_1^* = .0476$ (the level of enforcement just sufficient to induce truthful reporting), $p_2^* = 0$, $I_c^* = 14.45$, $R = .9048$, $T = .9999$, $C = .0951$. Since the optimal threshold is close to the top income level $I^+ = 15$, the proportion of taxpayers with $I < I_c^*$ and reporting truthfully is 99.94%. Thus the optimal policy is almost that of optimal random audits. The regressive bias is kept to a minimum.

Next consider models 3 and 4. The quest for the optimal policy has been made using a grid-search procedure. The net revenue function and the associated partial derivatives have been computed for the selected values of the instruments (p_1, p_2, I_c). The visual inspection of both grids suggests the following statements.

Fact 1 *For any (p_1, p_2), there exists a unique $I_c(p_1, p_2)$ such that $\partial R/\partial I_c > 0$ for all $I_c < I_c(p_1, p_2)$ and $\partial R/\partial I_c < 0$ for all $I_c > I_c(p_1, p_2)$.*

This allows to extract a reduced grid by choosing, for each (p_1, p_2), the value of I_c that maximizes net revenue. A second round of visual inspection suggests the following.

Fact 2 *For any p_2, there exists a unique $p_1(p_2)$ such that $\partial R/\partial p_1 > 0$ for all $p_1 < p_1(p_2)$ and $\partial R/\partial p_1 < 0$ for all $p_1 > p_1(p_2)$.*

A further reduction of the grid is thus obtained by selecting, for each p_2, the value of p_1 that maximizes net revenue. The remaining data show the following.

Fact 3 *The graph of $\partial R/\partial p_2$ on p_2 exhibits the pattern illustrated in figure 3.3 in section 3.5.2 above.*

This means that there are two local maxima, one at $p_2 = 0$ and one for $p_2 \in (0, (1+s)^{-1})$. The numerical results show that the latter is the global maximum, which gives the optimal audit policy.

Model 3. Table 3.1 shows the results when $g(\beta)$ is normally distributed. The optimal audit policy is under the label *gm* (global maximum). Notice that (i) the optimal threshold is *low*, $I_c^{gm} = 2.70$, so that the proportion of taxpayers with $I < I_c^{gm}$ is 8.87% and (ii) $p_1^{gm} = .0992$ and $p_2^{gm} = .0728$ are alike, implying that the regressive bias caused by $p_1^{gm} > p_2^{gm}$ is not strong. Also, the table shows that 92.05% of individuals report truthfully, whereas .90% report no income and 7.05% declare an income equal to the threshold and are thus *evaders*.

Table 3.1: Normal distribution of β.

	I_c	p_1	p_2	R	T	C	$r=0$	$r=I_c$	$r=I$	$F(I_c)$
gm	2.70 (0)	.0992 (0)	.0728 (0)	.8219	.9726	.1507	.90	7.05	92.05	8.87
lm	13.69 (0)	.0815 (0)	.0000 (-)	.8078	.9705	.1627	4.96	0.17	94.88	99.83
ra	15.00 (-)	.0815 (0)		.8076	.9706	.1630	4.96	0.00	95.04	100.00
bb	5.54 (0)	.2000 (-)	.0000 (-)	.6140	.8836	.2696	.00	32.59	67.41	67.41
bb'	1.81 (0)	.2000 (-)	.0759 (0)	.8182	.9723	.1541	.00	6.93	93.07	.94

Columns 2-4: The sign in parentesis refers to the sign of the corresponding first order derivative.
Columns 8-10: $r=0,I_c,I$, is the percentage of taxpayers reporting $0,I_c,I$ respectively.

The second row of the table shows that the local maximum (lm) is at $I_c^{lm} = 13.69$, $p_1^{lm} = .0815$, $p_2^{lm} = 0$. Notice that $T^{gm} > T^{lm}$ and $C^{gm} < C^{lm}$. The local maximum is virtually equal to the optimal random audits (ra), where $I_c^{ra} = 15$ and $p_1^{ra} = .0815$.

Consider the issue of the regressive bias. Taxpayers declaring $r = I$ pay the legislated tax liability so that the average tax rate is t. For those reporting $r = 0$ the *expected* average tax rate (EATR) is $p_1 t(1+s)$, whereas for those declaring $r = I_c$ is

$$\frac{tI_c}{I} + \frac{p_2 t(1+s)(I-I_c)}{I}$$

With the optimal audit policy gm, the EATR for an individual reporting $r = 0$ is .0992, which is less than a half of the tax rate $t = .2$. The EATR for taxpayers declaring $r = I_c^{gm}$ varies with I, few values are the following (income in parenthesis): .187 (3), .130 (6), .115 (8), .101 (12), .095 (15). This pattern is similar to that of the optimal audit policy with risk neutral taxpayers, see model 1. However, the regressive bias is higher in model 1, where 32.59% of

Table 3.2: Exponential distribution of β.

	I_c	p_1	p_2	R	T	C	$r=0$	$r=I_c$	$r=I$	$F(I_c)$
gm	2.97 (0)	.1593 (0)	.0935 (0)	.7245	.9309	.2064	2.45	24.53	73.02	13.34
lm	11.61 (0)	.1273 (0)	.0000 (-)	.6865	.9390	.2525	16.31	0.80	82.89	99.08
ra	15.00 (-)	.1271 (0)		.6860	.9402	.2542	16.41	0.00	83.59	100.00
bb	5.54 (0)	.2000 (-)	.0000 (-)	.6139	.8836	.2697	.00	32.59	67.41	67.41
bb'	2.55 (0)	.2000 (-)	.0982 (0)	.7200	.9304	.2104	.00	25.77	74.23	6.83

Columns 2-4: The sign in parentesis refers to the sign of the corresponding first order derivative.

Columns 8-10: $r=0,I_c,I$, is the percentage of taxpayers reporting $0,I_c,I$ respectively.

individuals are tax evaders, whereas with *gm* the evaders add up to 7.95%.

The last two rows pertain to the constrained model of section 3.5.2, where $p_1 = .2$, so that all taxpayers with income below the threshold are induced to report truthfully. The first local maxima is called *bang- bang* policy (*bb*),[10] where $I_c^{bb} = 5.54$, $p_1^{bb} = .2$ and $p_2^{bb} = 0$. The share of taxpayers with $I < I_c^{bb}$ is 67.41%. This policy raises the lowest gross and net revenues and is the most expensive in terms of audit costs. Notice that *bb* is the optimal audit policy when taxpayers are risk neutral, see model 1. Finally, *bb'* is the second local maxima, confirming that the graph of $\partial R/\partial p_2$ given $I_c(p_2)$ is as illustrated in figure 3.3. Notice that the threshold is down to $I_c^{bb'} = 1.81$ and that $p_2^{bb'} = .0759$. Also, *bb'* raises more net revenue than *lm* and *ra*.

Model 4. The examination of table 3.2 shows that qualitatively similar results are obtained when $g(\beta)$ is exponentially distributed. As for the quantitative figures,

[10]The term *bang- bang* is taken from Cowell (1990). Notice however that he applies this term to the case $p_1 = 1$.

audit probabilities are higher in model 4 than in model 3. The explanation is simple: the proportion of taxpayers with $\beta < 5$ (the average) is 50.1% in model 3 and 63.2% in model 4.

The global maximum is $I_c^{gm} = 2.97$, $p_1^{gm} = .1593$, $p_2^{gm} = .0935$. The EATR for an individual reporting $r = 0$ is .1593, which is about 75% of $t = .2$. The EATR for taxpayers declaring $r = I_c^{gm}$ is (income in parenthesis): .198 (3), .146 (6), .133 (8), .119 (12), .114 (15). Since the proportion of evaders is only 26.98%, this fall in the EATR does not cause a strong regressive bias.

To conclude, the numerical computations have shown that the form of the distribution of attitudes towards risk among taxpayers is a crucial factor in determining the nature of the optimal tax enforcement policy. When all taxpayers have the same degree of risk aversion, the optimal audit policy may cause the effective tax schedule to be quite regressive. Instead, when attitudes towards risk vary across individuals, the optimal audit policy introduces only a *limited* regressive bias. The optimal threshold is generally *low* and $p_2^{gm} < p^{ra} < p_1^{gm}$.

3.7 Conclusions

This final section suggests few lines for further research.

The assumption of a kinked linear utility function gives a weak model of individual behaviour. Taxpayers with pre-tax income below the threshold operate a zero-one decision in response to the audit policy, the optimal choice being either to conceal all income or to report truthfully. Taxpayers with income greater than the threshold restrict their choice to declaring zero income, declaring the threshold or reporting truthfully. Also, the tax rate does not affect the evasion decision. The obvious generalization is to assume a strictly concave, twice continuously differentiable utility function.

The audit policy has been restricted to the one-threshold, two-probabilities form. Although it has the merit of being operationally simple, there is no assurance that it is the optimal audit probability function.

A proportional income tax has been assumed. One could add the government to the model, its objective being to maximize a social welfare function by choice of the tax schedule, while taking into account the behaviour of the tax agency. On this point see Sanchez and Sobel (1990) and Cremer *et al.* (1990).

Efficiency aspects could be added to the analysis, by letting the audit policy, taxes and penalties to affect pre-tax incomes.

Finally, the assumption that the only information available to the tax agency is taxpayers reports is quite restrictive. The efficacy of the audit policy can be improved by using signals other than taxpayers reports, which are correlated with taxable income. On dividing taxpayers into audit classes, see Scotchmer (1987) and Cremer *et al.* (1990).

A Appendix

A.1 Identical degree of risk aversion

Proof of lemma 1. The proof is given by establishing two properties. Let $b \equiv (1+\bar{\beta}s)^{-1}$. Let p_1^*, p_2^* denote optimal audit probabilities.

P1. *For any* I_c, $\max(p_1^*, p_2^*) \leq b$.

Proof. The arguments are taken from Cremer *et al.* (1990, appendix, properties 1-3) and adapted to the model herein. The property is proven by contradiction.

Consider first a taxpayer with $I < I_c$. His or her expected utility is

$$EU(r; I, \bar{\beta}) = \bar{\beta}(1-t)I + t(1-p_1/b)(I-r), \qquad r \in [0, I] \tag{3.38}$$

Assume that $p_1^* > b$ (the value of p_2^* is irrelevant). Then the optimal report is $r(.) = I$. Replacing p_1^* with $p_1' = b$ leaves the optimal report unchanged (from assumption 1) whereas reduces audit costs. Thus $p_1^* > b$ is not optimal.

Next consider a taxpayer with $I \geq I_c$. His or her expected utility is

$$EU(r; I, \bar{\beta}) = \begin{cases} \bar{\beta}(1-t)I + t(1-p_1/b)(I-r), & r \in [0, I_c) \\ \bar{\beta}(1-t)I + t(1-p_2/b)(I-r), & r \in [I_c, I] \end{cases} \tag{3.39}$$

Assume that P1 does not hold. Then there are five cases to consider: (i) $p_1^* > b$, $p_2^* > b$, (ii) $p_1^* = b$, $p_2^* > b$, (iii) $p_1^* < b$, $p_2^* > b$, (iv) $p_1^* > b$, $p_2^* = b$, (v) $p_1^* > b$, $p_2^* < b$. Let $p_i' = \min(p_i^*, b)$, $i = 1, 2$. One can show that replacing p_1^*, p_2^* with p_1', p_2' leaves taxpayers optimal reports unchanged whereas decreases audit costs. Thus p_1^*, p_2^* are not optimal. To illustrate, two cases are examined. Consider first case (iii). Then $r(.) = 0$ for all I. Switching to $p_1' = p_1^* < b$, $p_2' = b$ leaves both optimal reports and audit costs unchanged. Next take case (v). Then $r(.) = 0$ for all $I < I_c$ and $r(.) = I_c$ for all $I \geq I_c$. Switching to $p_1' = b$, $p_2' = p_2^* < b$ does not affect optimal reports whereas reduces audit costs. *Q.E.D.*

P2. *For any* I_c, $p_2^* < p_1^*$.

Proof. The arguments are taken from Scotchmer (1987, p. 231). Let $0 \leq p_1' < p_2' \leq b$, $I_c' \in [0, I^+]$ be some audit policy. Then $r(.) = 0$ for all I and all reports are audited at the probability p_1'. Next consider the audit policy $0 \leq p_1' = p_2'' < b$, I_c'. This change of policy does not affect optimal reports as well as audit costs. Finally, provided that $p_1' > 0$, consider the audit policy $0 \leq p_2''' < p_1' < b$, $I_c = I^+$. Again, both optimal reports and audit costs are not affected and this proves the property. *Q.E.D.*

Proof of lemma 2. Linearity of expected utility and assumption 1 imply that a taxpayer with $I < I_c$ chooses his or her optimal report $r(.)$ from the discrete set $\{0, I\}$, whereas a taxpayer with $I \geq I_c$ chooses from the set $\{0, I_c, I\}$. Let $EU(r)$ denote the expected utility when the report is r. Let $b \equiv (1 + \bar{\beta}s)^{-1}$.

Consider first a taxpayer with $I < I_c$. Eq. (3.38) implies that $EU(I) \geq EU(0)$, thus $r(.) = I$, iff $p_1 \geq b$. Therefore, for all $I \in [0, I_c)$, $r(.) = 0$ if $p_1 < b$ and $r(.) = I$ if $p_1 = b$.

Next consider a taxpayer with $I \geq I_c$. Eq. (3.39) implies that (i) $EU(I) \geq EU(0)$ iff $p_1 \geq b$ and $EU(I) \geq EU(I_c)$ iff $p_2 \geq b$, (ii) $EU(I_c) > EU(I)$ iff $p_2 < b$ and $EU(I_c) \geq EU(0)$ iff $I \geq \bar{\kappa} I_c$, (iii) $EU(0) > EU(I)$ iff $p_1 < b$ and $EU(0) > EU(I_c)$ iff $I < \bar{\kappa} I_c$. From lemma 1, $p_2 < b$, thus case (i) never occurs. When $p_1 < b$, then $r(.) = I_c$ if $I \geq \min(\bar{\kappa} I_c, I^+)$ and $r(.) = 0$ if $I < \min(\bar{\kappa} I_c, I^+)$. When $p_1 = b$, then $\bar{\kappa} = 1$ and $r(.) = I_c$ for all $I \in [I_c, I^+]$. *Q.E.D.*

A.2 Different attitudes towards risk

Proof of lemma 3. As for lemma 1, the proof is given by establishing two properties. Let $b \equiv (1 + s)^{-1}$ and let p_1^*, p_2^* denote optimal audit probabilities.

P1. *For any* I_c, $\max(p_1^*, p_2^*) \leq b$.

Proof. The property is proven by contradiction.

Consider first a taxpayer (I,β) with $I < I_c$. His or her expected utility is

$$EU(r; I,\beta) = \beta(1-t)I + t[1 - p_1(1+\beta s)](I-r), \quad r \in [0, I] \tag{3.40}$$

Assume that $p_1^* > b$ (the value of p_2^* is irrelevant). Then the optimal report is $r(.) = I$ for all $\beta \in [1,\infty)$. Replacing p_1^* with $p_1' = b$ leaves the optimal report unchanged (from assumption 1) whereas reduces audit costs. Thus $p_1^* > b$ is not optimal.

Next consider a taxpayer (I,β) with $I \geq I_c$. His or her expected utility is

$$EU(r; I,\bar{\beta}) = \begin{cases} \beta(1-t)I + t[1 - p_1(1+\beta s)](I-r), & r \in [0, I_c) \\ \beta(1-t)I + t[1 - p_2(1+\beta s)](I-r), & r \in [I_c, I] \end{cases} \tag{3.41}$$

Assume that P1 does not hold. Then there are five cases to consider: (i) $p_1^* > b$, $p_2^* > b$, (ii) $p_1^* = b$, $p_2^* > b$, (iii) $p_1^* < b$, $p_2^* > b$, (iv) $p_1^* > b$, $p_2^* = b$, (v) $p_1^* > b$, $p_2^* < b$. Let $p_i' = \min(p_i^*, b)$, $i = 1,2$. One can show that replacing p_1^*, p_2^* with p_1', p_2' leaves taxpayers optimal reports unchanged for all $\beta \in [1,\infty)$ whereas decreases audit costs. Thus p_1^*, p_2^* are not optimal.

To illustrate, two cases are examined. Consider first case (iii). For all I, $r(.) = 0$ if $\beta \in [1,\beta_1)$ and $r(.) = I$ if $\beta \in [\beta_1,\infty)$, where $\beta_1 \equiv (1-p_1^*)/(sp_1^*) > 1$. Switching to $p_1' = p_1^* < b$, $p_2' = b$ leaves optimal reports unchanged whereas audit costs are reduced. Next take case (v). For all $I \in [0, I_c)$, $r(.) = I$. For all $I \in [I_c, I^+]$, $r(.) = I_c$ if $\beta \in [1,\beta_2)$ and $r(.) = I$ if $\beta \in [\beta_2,\infty)$, where $\beta_2 = (1-p_2^*)/(sp_2^*) > 1$. Switching to $p_1' = b$, $p_2' = p_2^* < b$ does not affect optimal reports whereas reduces audit costs. *Q.E.D.*

P2. *For any I_c, $p_2^* < p_1^*$.*

Proof. Let $0 \leq p_1' < p_2' \leq b$, $I_c' \in [0, I^+]$ be some audit policy. For all I, $r(.) = 0$ if $\beta \in [1,\beta_1)$ and $r(.) = I$ if $\beta \in [\beta_1,\infty)$. Notice that optimal reports are independent of p_2. Next consider the audit policy $0 \leq p_1' = p_2'' < b$, I_c'. This change of policy does not affect optimal reports whereas reduces audit costs. Finally, provided that $p_1' > 0$, consider the audit policy $0 \leq p_2''' < p_1' < b$,

$I_c = I^+$. Both optimal reports and audit costs are not affected and this proves the property. *Q.E.D.*

Proof of lemma 4. Linearity of expected utility and assumption 1 imply that a taxpayer with $I < I_c$ chooses his or her optimal report $r(.)$ from the discrete set $\{0, I\}$, whereas a taxpayer with $I \geq I_c$ chooses from the set $\{0, I_c, I\}$. Let $EU(r)$ denote the expected utility when the report is r. Let $b \equiv (1+s)^{-1}$.

Consider first a taxpayer with $I < I_c$. Eq. (3.40) implies that $EU(I) \geq EU(0)$, thus $r(.) = I$, iff $\beta \geq \beta_1$.

Next consider a taxpayer with $I \geq I_c$. Eq. (3.41) implies that (i) $EU(I) \geq EU(0)$ iff $\beta \geq \beta_1$ and $EU(I) \geq EU(I_c)$ iff $\beta \geq \beta_2$, (ii) $EU(I_c) > EU(I)$ iff $\beta < \beta_2$ and $EU(I_c) \geq EU(0)$ iff $\beta \geq \tilde{\beta}$, (iii) $EU(0) > EU(I)$ iff $\beta < \beta_1$ and $EU(0) > EU(I_c)$ iff $\beta < \tilde{\beta}$. From lemma 3, $p_2 < p_1$, thus $\beta_2 > \beta_1$. Also, $\tilde{\beta}$ is decreasing in I and $\tilde{\beta} = \beta_1$ at $I = I_c$, $\tilde{\beta} = 1$ at $I = \kappa I_c$. Hence (i) $r(.) = I$ if $\beta \geq \beta_2$, (ii) $r(.) = I_c$ if $\beta \in [\tilde{\beta}, \beta_2)$, (iii) $r(.) = 0$ if $\beta < \tilde{\beta}$. *Q.E.D.*

References

Allingham, Michael G. and Agnar Sandmo, 1972, Income tax evasion: A theoretical analysis, Journal of Public Economics 1, 323–338.

Border, Kim C. and Joel Sobel, 1987, Samurai accountant: a theory of auditing and plunder, Review of Economic Studies 54, 525–540.

Chander Parkash and Louis Wilde, 1992, A general characterization of optimal income taxation and enforcement, California Institute of Technology, Social Science Working Paper 791.

Cowell, Frank A., 1990, Cheating the government: The economics of evasion, The MIT Press.

Cremer, H., Marchand, M. and P. Pestieu, 1990, Evading, auditing and taxing: the equity-compliance tradeoff, Journal of Public Economics 43, 67–92.

Domar, E. D. and R. Musgrave, 1944, Proportional income taxation and risk-taking, Quarterly Journal of Economics 57, 387–422.

Gordon, James P.F., 1989, Individual morality and reputation costs as deterrents to tax evasion, European Economic Review 33, 797–805.

Grossman, Sanford J. and Oliver D. Hart, 1983, An analysis of the principal-agent problem, Econometrica 51, 7–45.

Markowitz, Harry M., 1959, Portfolio selection, Cowles Foundation Monographs 16, John Wiley & Sons, New York.

Melumad, Nahum D. and Dilip Mookherjee, 1989, Delegation as commitment: the case of income tax audits, Rand Journal of Economics 20, 139–163.

Mirrlees, James, 1971, An exploration in the theory of optimal income taxation, Review of Economic Studies 38, 175–208.

Mookherjee, Dilip and Ivan Png, 1989, Optimal auditing, insurance and redistribution, Quarterly Journal of Economics 104, 399–415.

Reinganum, Jennifer F. and Louis L. Wilde, 1985, Income tax compliance in a principal-agent framework, Journal of Public Economics 26, 1–18.

Reinganum, Jennifer F. and Louis L. Wilde, 1986, Equilibrium verification and reporting policies in a model of tax compliance, International Economic Review 27, 739–760.

Sanchez, Isabel and Joel Sobel, 1990, Hierarchical design and enforcement of income tax policies, Southern European Economics Discussion Series, Universidad del Pais Vasco, D.P. 86.

Scotchmer, Suzanne, 1987, Audit classes and tax enforcement policy, American Economic Review 77, 229–233.

Chapter 4

Presumptive Income Coefficients and Tax Enforcement Policy: a Theoretical Analysis

4.1 Introduction

In 1989 the Italian government introduced a new tax enforcement mechanism, called Presumptive Income Coefficients (PIC), with the objective of reducing tax evasion in the non-corporate sector. The system applies to small size firms, professionals and freelance, where it is a common practice to evade the income tax and the value added tax by under-reporting revenues.

Consider a homogeneous group of professionals or firms operating in the same sector of activity. The rationale of the mechanism is simple and is based on two considerations. The first is that sales revenues are easily concealed to the taxing authority whereas production costs are not. On one hand, selling *under the counter* for cash is a common way of not reporting revenues and it is costly, if not impossible, for the tax agency to discover the true amount when performing an audit. On the other hand, expenditures such as wages to regular employees, telephone and electricity bills are difficult to hide and can be monitored at a low cost. The second consideration is that when examining various firms, one observes a similar relation between production costs and the corresponding revenues, the discrepancies being due to different technologies, locations and demand conditions. Hence it is possible to estimate actual revenues and income by applying simple algorithms on observed production costs. The choice of the Italian tax administration has been that of determining the PIC by regressing (ordinary least squares) revenues on production costs over a sample of taxpayers.

The PIC audit policy works as follows. For each group of taxpayers (classified according to the sector of economic activity) the taxing authority publishes the list of PIC, one for each type of expenditure. The application of the coefficients to production costs then determines the so called *presumptive income.* If reported income is at least as high as the presumptive income, the taxpayer is subjected to random audits. Otherwise the taxpayer is expected to pay the tax on presumptive income, unless he or she demonstrates that the actual income is lower than presumptive income. Section 4.2 illustrates in more detail the PIC audit mechanism implemented by the Italian government.[1]

The purpose of this paper is to provide a theoretical analysis of the PIC tax enforcement policy. The basic model, described in section 4.3, consists of a large population of professionals producing and selling a homogeneous commodity in a competitive market. Each professional supplies labour services into his or her own business and makes earnings from sales into the market. Pre-tax incomes and production costs vary across individuals, for they are endowed with different

[1] A similar system, called *tachshiv,* is used in Israel since 1954, see Lapidoth (1977). Procedures to estimate taxable incomes in the agricultural sector have been used in France since 1935, see Kelly and Oldman (1973).

technologies.

Tax compliance and tax enforcement policy are introduced in section 4.4. Following Reinganum and Wilde (1985), the *game* between the taxing authority and taxpayers is modelled in a principal-agent framework. Taxpayers choose an income report that maximizes expected utility, taking taxes, penalties and the audit policy as given. The tax agency observes production costs and income reports and chooses the audit policy that maximizes the tax revenue net of audit costs. The audit policy is restricted to the simple mechanism where income reports below and above a given mark-up on production costs are audited at different probabilities. When the mark-up coefficient is meant to give an *estimate* of taxable income, its interpretation is that of *presumptive income coefficient.* Two types of presumptive coefficients are considered: the weighted average of individual mark-up coefficients and the ordinary least squares coefficient with no intercept term.

Section 4.5 examines the case of exogenous labour supply. Proposition 1 characterizes the nature of the optimal tax enforcement policy when the mark-up coefficient is set at the net revenue maximizing level. Income reports below the corresponding mark-up on production costs are audited at the probability level just sufficient to induce honest behaviour, whereas those above it are not audited. This result mirrors those obtained in Reinganum and Wilde (1985) type models,[2] where the tax agency observes only income reports and audits at different probabilities reports below and above a given income threshold. Proposition 2 deals with optimal probabilities of audit when the mark-up is constrained to be a presumptive income coefficient. It is shown that either the result of proposition 1 applies *or* random audits is the best policy. In the latter case the notion of presumptive coefficient becomes meaningless.

---

[2]See for instance Cremer, Marchand and Pestieu (1990) and Sanchez and Sobel (1990).

In section 4.6 labour supply and pre-tax incomes are endogenously determined by the tax enforcement instruments.[3] An analytical characterization of the optimal audit policy cannot be given. Thus some numerical computations are examined. It is shown that the properties of the optimal audit policy depend on the size of unit audit costs relative to the size of average pre-tax income. If audit costs are relatively low, the optimal audit policy is random audits, and thus the mark-up coefficient is a redundant instrument. If audit costs are relatively high, income reports below the optimal mark-up on production costs are audited at the probability level just sufficient to induce truthful reporting, whereas those above it meet a lower but positive probability of audit. Finally, when unit audit costs are of considerable size relative to average income, income reports above the mark-up level should be audited at probability zero. The numerical computations also show that the net revenue maximizing mark-up coefficient is lower than the presumptive income coefficients. This is an important result, especially when one considers that in taxpayers' minds a high value of the mark-up is likely to be perceived as *unfair*. Conclusions are given in section 4.7.

4.2 Presumptive income coefficients

The PIC audit policy is a recent innovation of the Italian tax code. Not surprisingly, the provisions of the tax law are numerous and it would be lenghty to describe them all. Also, recently the original system has been modified. The purpose of this section is to give a short illustration of the basic features.

[3] Mookherjee and Png (1989) allow for pre- tax income to be endogenous. However their model is derived from Grossman and Hart's (1983) framework and it is not directly comparable to the model presented herein.

4.2.1 The 1989-92 regime

The PIC mechanism applies to firms in the non-corporate sector with revenues, or invoiced sales proceeds, lower than a given threshold (about 180,000 English pounds). These taxpayers have the opportunity to opt for two alternative book-keeping systems. One is the Ordinary Book-Keeping (OBK), which is expensive to run because the entrepreneur must fill in numerous accounting books. When opting for the OBK, the taxpayer is subjected to a pure random audits mechanism.[4] The other is the Simplified Book-Keeping (SBK), which consists of keeping a book of invoiced sales proceeds and a book of invoiced expenditures. Thus book-keeping costs are low. Clearly, the opportunites to evade the income tax and the value added tax are higher under SBK than OBK. Taxpayers opting for SBK are subjected to the PIC audit mechanism.

Although the system applies to both the income and the value added tax, in what follows the focus will be only on the income tax. Under SBK, each year the taxpayer fills in an Income Tax Declaration Form (ITDF) by entering the amount of the Invoiced Sales Revenues and a list of Invoiced / Registered Expenditures, the difference being taxable income. The PIC audit mechanism operates in two steps. The first uses the Standard Revenues Coefficients (SRC) to determine whether reported revenues are *suspiciously* low. In case of an affirmative answer, then in the second step the tax authority computes the presumptive income using the Presumptive Income Coefficients.

Step 1. Let R^d be declared revenues. Let C_i^d, $i = 1, \ldots, k$, be the list of expenditures the taxpayer is required to enter into the ITDF and let ϑ_i be the corresponding SRC. The tax authority then computes $R_i^s = \vartheta_i C_i^d$, the standard revenues associated to expenditure i. If the number of R_i^s for which $R_i^s > R^d$

[4]The OBK is compulsory for non-corporate firms with revenues above the threshold and for all corporations.

is greater than two then the tax agency suspects that revenues are too low and proceeds to step 2. Otherwise the taxpayer is subjected to the ordinary random audits.

Step 2. The taxing authority determines the presumptive revenues and income on the basis of the PIC α_i, $i = 1, \ldots, k$. Presumptive revenues are equal to $R^p = \sum_1^k (1 + \alpha_i) C_i^d$, and the presumptive income is $I^p = \sum_1^k \alpha_i C_i^d$. If $R^p < R^d$, the taxpayer is subjected to the ordinary random audits. When $R^p > R^d$, the taxpayer is expected to pay the income tax on the basis of presumptive income, unless he or she demonstrates that the declared income is equal to the actual income.

4.2.2 The 1993 regime

In 1993 the system has been modified as follows:

- the Standard Revenue Coefficients have been suppressed (so that *step 1* described above no longer applies);
- the revenue threshold below which the PIC audit mechanism operates has been increased (for firms in the manufacturing industry the level is about 450,000 English pounds);
- In some instances the PIC apply also to firms opting for the OBK system;
- taxpayers are expected to pay a *minimum tax,* which is computed on the *shadow labour income* of the entrepreneur, that is the part of profits that are attributed to labour services supplied by the taxpayer into his or her own business.

The most relevant innovation concerns the *minimum tax.* The estimate of the shadow labour income is based on several elements: sector of activity, location, age, expertise (measured by the years of activity) and number of employees. Presumptive income is $I^p = I^m + \sum_1^k \alpha_i C_i^d$, where the shadow labour income I^m is the intercept term in the equation determining I^p. The analysis that will follow is conducted assuming $I^m = 0$, the main reason for doing so being analytical tractability.

4.3 The basic model

This section describes the basic framework. Tax compliance and tax enforcement policy issues will be introduced in the next section.

The model consists of a large population of professionals producing and selling a homogeneous commodity, or a service, in a competitive market. The term *professional* is meant to characterize an independent producer who provides labour services into his or her own business and derives earnings from the sale of his or her products into the market.[5]

Assume that the professionals have identical preferences over disposable income y and labour supply L. The utility function is of the form

$$U(y, L) = y - \frac{1}{2}L^2 \tag{4.1}$$

The government levies a proportional income tax at rate t, $0 < t < 1$, on pre-tax income I, so that $y = (1-t)I$.

Let π be the price of the service and assume that inverse demand is horizontal so that π is fixed. The key feature of the model is that individuals are endowed with different production technologies, which are characterized by the pair of parameters (c, w), $0 \leq c \leq \pi$, $w \geq 0$. The first is an *efficiency* parameter and represents the cost of producing one unit of output (materials, telephone calls,

[5] Alternative terms for *professional* are *freelance* and *self-employed worker*.

wages to employees). The second is a *skill* parameter and gives the units of output for each unit of labour services supplied by the professional. Thus gross income is $I = (\pi - c)x$ and labour supply is $L = (1/w)x$, where x is individual output. Notice that c and w are constant so that technology is constant returns to scale.

To avoid repetitions, in what follows I will refer to *professional with efficiency parameter c and skill parameter w* simply as *professional* (c, w).

Professional (c, w) solves

$$\max_x U = (1-t)(\pi - c)x - \frac{1}{2w^2}x^2 \tag{4.2}$$

From the first order condition of this problem one obtains $x(c, w) = (1-t)w^2(\pi - c)$ and $I(c, w) = (1-t)w^2(\pi - c)^2$. Clearly, gross income is increasing in w and decreasing in c. Also, notice that the income tax is distortionary, for gross income is decreasing in t.

The distribution of the efficiency and skill parameters across individuals completes the description of the model. Assume that c and w are independently distributed with cumulative distributions $F(c)$, $c \in [0, \pi]$, and $G(w)$, $w \in [0, \infty)$. Assume that $F(c)$ and $G(w)$ are twice continuously differentiable and let $f(c) = F'(c)$, $c \in (0, \pi)$, and $g(w) = G'(w)$, $w \in (0, \infty)$, be the corresponding density functions.

4.4 The problem of the tax enforcement agency

The information set of the tax enforcement agency is as follows. The agency knows $F(c)$ and $G(w)$ but it is unable to observe the characteristics (c, w) of each individual, as well as the output x. Therefore pre-tax income $I = (\pi - c)x$ is not directly observed.

Professionals are thus required to fill in an income declaration form, showing separately total revenues and total production costs. It is assumed that taxpayers are unable to over or under-report production costs, or alternatively that the agency costlessly observes cx. To evade the income tax, professionals must then under-report total revenues. The cost of verifying an income report is equal to $\varphi > 0$ and it is assumed that when performing an audit the tax agency observes the true revenues πx.

Consider a professional (c, w) and let

$$r = R - cx$$

be reported income, where R is the amount of declared revenues.

If not audited, his or her transfer to the tax agency is

$$\begin{array}{ll} tr & \text{if } \; r \geq 0 \\ 0 & \text{if } \; r < 0 \end{array}$$

When audited, the professional pays the income tax plus a proportional surcharge on the evaded tax (if any) at rate $s > 0$, so that the transfer is

$$\begin{array}{ll} tr + (1+s)t(I-r) & \text{if } \; 0 \leq r < I \\ tr & \text{if } \; r \geq I \\ (1+s)tI & \text{if } \; r < 0 \end{array}$$

Since in the event of no audit the income tax is zero for negative reports, no taxpayer will report $r < 0$. Also, since in the event of audit there are no rewards for over-reporting, no taxpayer will report $r > I$. Thus attention can be restricted to $0 \leq r \leq I$, or $cx \leq R \leq \pi x$, where professionals declare an amount of revenues at least as high as total production costs.

The tax enforcement agency decides how many and which reports to audit. The available information is taxpayers' reports, showing r, R and cx. The agency fixes a mark-up coefficient $\alpha \geq 0$ and the audit policy is restricted to the step function

$$p(r) = \begin{cases} p_1 & \text{if } \; r < \alpha cx \\ p_2 & \text{if } \; r \geq \alpha cx \end{cases} \tag{4.3}$$

A professional is audited at a constant probability p_1 if reported income is less than α times the amount of production costs, whereas the probability is p_2 if reported income is at least as high as αcx. Notice that the way in which an audit is triggered differs from the real PIC mechanism described in section 4.2. In particular, the two systems coincide when reported income is at least as high as presumptive income: taxpayers face a probability p_2 of audit and the tax agency can verify an income report at a cost φ per audit. However, the model and the real PIC mechanism differ when reported income is less than presumptive income. In the former, taxpayers meet a probability p_1 of audit and the taxing authority may verify an income report at the same unit cost φ. In the latter, the taxpayer is expected (with probability one) to pay the tax on presumptive income, unless she or he demonstrates (and this means costly effort) that the actual income is lower than presumptive income. The tax agency bears no audit costs.

The tax agency takes the tax and the penalty rates as given, these being fixed outside the model by a higher level of authority. Its objective is to maximize total revenue (taxes plus fines) net of audit costs by choosing the tax enforcement instruments p_1, p_2 and α, while taking taxpayers optimal reports as given. A few words about the instrument α are in order. When α is unconstrained, its interpretation is that of a *mark-up threshold*, where the level of enforcement changes from p_1 to p_2. However, the actual value of α may bear some importance in a taxpayer's mind, especially when $p_1 > p_2$. In other words, a taxpayer may feel that reporting an income below αcx makes the tax agency to believe that the income declaration is suspiciously low. In this case a high value of α is considered *unfair*, for it attaches a kind of stigma to individuals with actual income below αcx. Hence the tax agency may opt for constraining α to an *average* mark-up, and in this case α is interpreted as a *presumptive income coefficient*, so that αcx is the *presumptive income* and $(1 + \alpha)cx$ are the corresponding *presumptive* *revenues.*

Each professional is assumed to select an income report and an output level that maximize expected utility in the two states of the world (audit / no audit), taking tax and penalty rates and the audit policy as given. Labour supply is the same in both states of the world. As for disposable income, when the taxpayer reports truthfully, his or her income is $(1-t)I$ in both states of the world. On the other hand, when $r \neq I$, disposable income is

$$\begin{aligned} y_1 &\equiv (1-t)I + t(I-r) = \\ &= (1-t)(\pi - c)x + t[(\pi - c)x - r] \end{aligned}$$

if no audit occurs and

$$\begin{aligned} y_2 &\equiv (1-t)I - st(I-r) = \\ &= (1-t)(\pi - c)x - st[(\pi - c)x - r] \end{aligned}$$

if an audit occurs.

Hence the expected utility of professional (c, w) is

$$EU(x, r; c, w) = \big(1 - p(r)\big)y_1 + p(r)y_2 - \frac{1}{2w^2}x^2$$

Notice that the utility function (4.1), which is linear in y, implies that professionals are risk neutral over random realizations of income.

Professional (c, w) solves

$$\begin{aligned} \max_{x,r} EU(x, r; c, w) &= (1-t)(\pi - c)x + \\ &\quad + t\big[1 - p(r)(1+s)\big]\big[(\pi - c)x - r\big] - \frac{1}{2w^2}x^2 \qquad (4.4) \end{aligned}$$

where $0 \leq r \leq (\pi - c)x$. Notice that $EU(.)$ is linear in r and is discontinuous at $r = \alpha cx$ when $p_1 \neq p_2$ and $(\pi - c)x \geq \alpha cx$. Linearity implies that the optimal report may not be unique. The following assumption is made to rule out multiple solutions.

Assumption 1 *Whenever a professional is indifferent between reporting r_0 and r_1, where $r_0 < r_1$, then he or she always chooses to report r_1.*

Let the solution to (4.4) be denoted by $x(c,w)$, $r(c,w)$, where the arguments (p_1, p_2, α) have been omitted to keep the notation simple. Gross income is $I(c,w) = (\pi - c)x(c,w)$.

The set of admissible audit policies is

$$A_0 = \Big\{(p_1, p_2, \alpha) \mid 0 \leq p_1, p_2 \leq 1,\ \alpha \geq 0\Big\} \tag{4.5}$$

When α is interpreted as a mark-up threshold, the problem of the tax enforcement agency is

$$\begin{aligned} \max_{(p_1, p_2, \alpha) \in A_0} T(p_1, p_2, \alpha) &= \int_0^\infty \int_0^\pi \Big\{tr(.) + p\big(r(.)\big)(1+s)t\big[I(.) - r(.)\big] + \\ &\qquad -\varphi p\big(r(.)\big)\Big\} dF(c) dG(w) \end{aligned} \tag{4.6}$$

When the mark-up coefficient is constrained to be a presumptive income coefficient, the maximization is over the audit probabilities with α fixed.

The following lemma states that the optimal audit policy is contained in a subset of A_0.

Lemma 1 *Let $(p_1, p_2, \alpha) \in A_0$ be some audit policy. Then there exists an audit policy $(p_1', p_2', \alpha') \in A_1$ such that $T(p_1', p_2', \alpha') \geq T(p_1, p_2, \alpha)$, where*

$$A_1 = \Big\{(p_1, p_2, \alpha) \mid 0 \leq p_2 < p_1 \leq (1+s)^{-1},\ \alpha \geq 0\Big\}$$

The proof is omitted.[6] Optimal probabilities do not exceed $(1+s)^{-1}$, the level of enforcement just sufficient to induce truthful reporting, for a higher probability would increase audit costs without raising more tax revenue. Also, income

[6] The proof follows the same lines of the proof of lemma 1 in chapter 3, to which the reader is referred to.

reports below αcx must meet a higher probability of audit than those above it, for otherwise the audit policy would not work as a screening device, separating low and high income taxpayers according to their reports.

The next section examines a constrained version of this model, where labour supply and therefore pre-tax incomes are assumed to be exogenously fixed. The case of endogenous labour supply is postponed to section 4.6.

4.5 Exogenous labour supply

Assume that labour supply is fixed at $L = L^{\circ}$. In other words, all professionals provide the same amount of labour services into their own business. The output of a professional (c, w) is $x(c, w) = wL^{\circ}$. Hence total production costs and pre-tax income are respectively

$$cx(c, w) = cwL^{\circ} \tag{4.7}$$

$$I(c, w) = (\pi - c)wL^{\circ} \tag{4.8}$$

4.5.1 The information set of the tax enforcement agency

Although the tax enforcement agency observes the individual production costs, this piece of information is not sufficient to infer the corresponding individual pre-tax income, for two professionals with the same level of production costs will have different incomes whenever they have different skill and efficiency parameters.

To illustrate, consider professionals (c_1, w_1) and (c_2, w_2). Let $L^{\circ} = 1$ and suppose they have identical production costs $c_1w_1 = c_2w_2$. Without loss of generality, suppose that $c_1 < c_2$ and $w_1 > w_2$. Thus $I_2 = (\pi - c_2)w_2 = \pi w_2 - c_1w_1$ and $I_1 = (\pi - c_1)w_1 > I_2$, so that the more skilled and efficient professional has a higher income.[7] Graph (a) in figure 4.1 shows production costs and pre-tax

[7]Conversely, professionals with the same gross income may have different total production

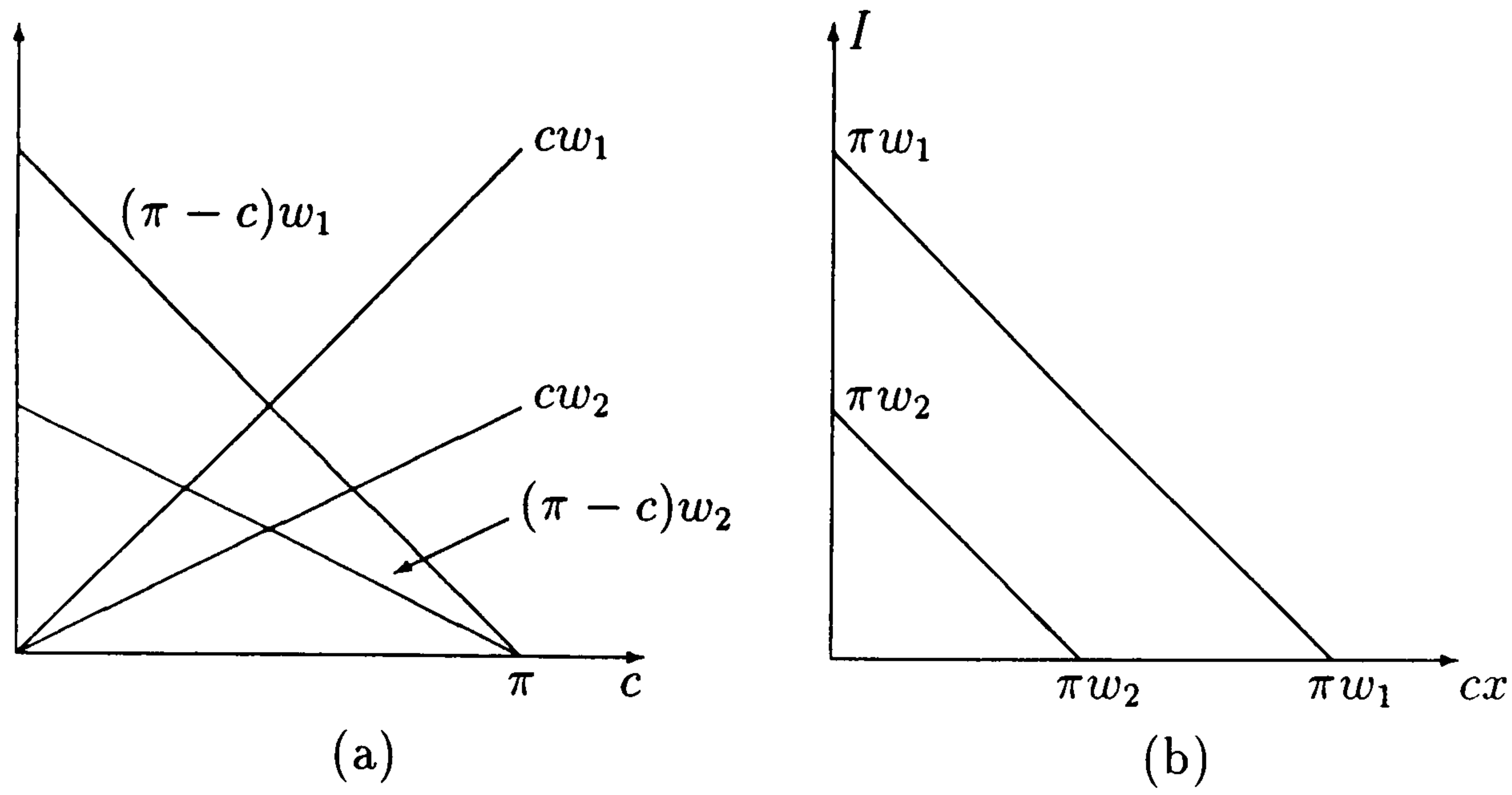

Figure 4.1: Exogenous income and production costs

incomes as a function of c for two different levels $w_1 > w_2$ of the skill parameter. Graph (b) shows the resulting relation between production costs and pre-tax income.

One point worth mentioning is that the restriction to the audit function with no intercept term and positive α given in (4.3) is potentially quite strong, for the correlation between total production costs and income may, in some cases, be negative. When the correlation is negative, a more reasonable audit function would be one with an intercept term (minimum income) which allows for a negative mark up threshold $\alpha < 0$. For the sake of illustration, suppose that the efficiency and the skill parameters are uniformily distributed with support $c \in [c_l, c_h]$ and $w \in [0, w_h]$ respectively; also let $L^\circ = 1$. After some tedious computations, one finds that the covariance between $I = (\pi - c)w$ and $C = cw$ is $\sigma_{IC} = (1/144)w_h^2[6\pi(c_l + c_h) - 3(c_l + c_h)^2 - 4(c_h - c_l)^2]$. If $c_l = 0$ and $c_h = \pi$ (as assumed at the end of section 4.3) then $\sigma_{IC} = -(1/144)w_h^2\pi^2$ so that average income, conditional on C, decreases with C. On the other hand, if $c_l = 0$ and $c_h < \pi$ then $\sigma_{IC} = (1/144)w_h^2(6\pi - 7c_h)c_h$ so that $\sigma_{IC} > 0$ iff $\pi > (7/6)c_h$.

---

costs.

The mark-up of a professional (c, w) is independent of the skill parameter w and is defined by

$$\beta(c) \equiv \frac{I(c,w)}{cx(c,w)} = \frac{\pi - c}{c} \tag{4.9}$$

The presumptive income coefficient, or presumptive mark-up, can be defined in different ways. Here two alternatives are considered. The first approach is to define the presumptive income coefficient as the average of individual mark-up coefficients

$$\alpha^{am} = \int_0^{\pi} \frac{\pi - c}{c} f(c)dc \tag{4.10}$$

The second approach is to define it as the least square coefficient obtained when regressing pre-tax incomes over total production costs with a line constrained to pass through the origin. Thus α solves

$$\min_{\alpha} \int_0^{\infty} \int_0^{\pi} \left[(\pi - c)wL^{\circ} - \alpha cwL^{\circ}\right]^2 f(c)g(w)dc\,dw$$

and the solution gives

$$\alpha^{ls} = \frac{\pi \bar{c}}{\sigma_c^2 + \bar{c}^2} - 1 \tag{4.11}$$

where $\bar{c}$ and σ_c^2 are respectively the mean and the variance of unit costs.

The next subsection looks at the unconstrained problem of the tax agency, where net revenue is maximized over p_1, p_2 and α. The search for the optimal probabilities with α constrained to be a presumptive income coefficient is examined in subsection 4.5.3.

4.5.2 Optimal audit policy

The following lemma characterizes professionals' optimal reports.

Lemma 2 *For any $(p_1, p_2, \alpha) \in A_1$, taxpayers' optimal reports are*

If $p_1 < (1+s)^{-1}$ *then*

$$r(c,w) = \begin{cases} 0 & \textit{if} \;\; \langle c,w\rangle \in (v,\pi] \times [0,\infty) \\ \alpha c w L^{\circ} & \textit{if} \;\; \langle c,w\rangle \in [0,v] \times [0,\infty) \end{cases}$$

If $p_1 = (1+s)^{-1}$ *then*

$$r(c,w) = \begin{cases} (\pi - c) w L^{\circ} & \textit{if} \;\; \langle c,w\rangle \in (v,\pi] \times [0,\infty) \\ \alpha c w L^{\circ} & \textit{if} \;\; \langle c,w\rangle \in [0,v] \times [0,\infty) \end{cases}$$

where

$$v = \frac{\pi}{1+\kappa\alpha}, \qquad \kappa = \frac{(1+s)^{-1} - p_2}{p_1 - p_2}$$

Proof. See appendix A.1.

Taxpayers are divided into two groups and the separation is related to the efficiency parameter c. When $p_1 < (1+s)^{-1}$, then $\kappa > 1$ and the relatively inefficient professionals, those with $c > \pi/(1+\kappa\alpha)$, report no income, whereas the relatively efficient, those with $c \leq \pi/(1+\kappa\alpha)$, report an income equal to the mark-up threshold on production costs. When $p_1 = (1+s)^{-1}$, then $\kappa = 1$ and individuals with $c > \pi/(1+\alpha)$ report true income, whereas those with $c \leq \pi/(1+\alpha)$ report $\alpha c x$.

The decision on whether to report no income, the mark-up threshold on production costs, or true income is independent of the skill parameter. In other words, the *qualitative* response to the audit policy is independent of w. The reason is that the audit policy sets two different levels of enforcement on the basis of the coefficient α which is a mark-up on production costs, and individual mark-up coefficients are independent of w, see eq. (4.9). However, the skill parameter affects the *quantity* of reported income, for pre-tax income and production costs are greater the higher is w.

Applying lemma 2 the net revenue function is

$$T(p_1,p_2,\alpha) = \int_v^\pi \Big\{tp_1(1+s)(\pi-c)\bar{w}L^\circ - \varphi p_1\Big\}dF(c) + \\ + \int_0^v \Big\{t\alpha c\bar{w}L^\circ + tp_2(1+s)\Big[(\pi-c)-\alpha c\Big]\bar{w}L^\circ - \varphi p_2\Big\}dF(c) \tag{4.12}$$

where $\bar{w} = \int_0^\infty wg(w)dw$ is the average skill level.

Notice that $T(.)$ is continuous in all its arguments. Making the change of variable $\alpha = (\pi - v)/(\kappa v)$ and rearranging eq. (4.12) one gets

$$\frac{T(p_1,p_2,v)}{t(1+s)\bar{w}L^\circ} = p_1\int_v^\pi \Big\{(\pi-c)-\gamma\Big\}f(c)dc + \\ + p_2\int_0^v \Big\{(\pi-c)-\gamma\Big\}f(c)dc + (p_1-p_2)\frac{\pi-v}{v}\int_0^v cf(c)dc \tag{4.13}$$

where $\gamma = \varphi/(t(1+s)\bar{w}L^\circ)$.

From the formula of integration by parts

$$\begin{aligned}\int_0^v cf(c)dc &= vF(v) - \int_0^v F(c)dc \\ \int_v^\pi cf(c)dc &= \pi - vF(v) - \int_v^\pi F(c)dc\end{aligned}$$

Using these expressions, and after some algebra, eq. (4.13) can be written as

$$\frac{T(p_1,p_2,v)}{t(1+s)\bar{w}L^\circ} = p_2\int_0^v \left\{\frac{\pi}{v}F(c) - \gamma f(c)\right\}dc + \\ + p_1\int_v^\pi \left\{\frac{\pi}{v}F(c) - \gamma f(c)\right\}dc - p_1\frac{\pi-v}{v}\int_0^\pi F(c)dc \tag{4.14}$$

Differentiating with respect to v

$$\frac{\partial T}{\partial v} = t(1+s)\bar{w}L^\circ(p_1-p_2)\left\{\frac{\pi}{v^2}\int_0^v F(c)dc - \frac{\pi}{v}F(v) + \gamma f(v)\right\} \tag{4.15}$$

Let $\mathcal{F}(v) = \int_0^v F(c)dc$. From eq. (4.15) the first order necessary condition with respect to v is

$$\Psi(v) \equiv \frac{\pi}{v^2}\mathcal{F}(v) - \frac{\pi}{v}F(v) + \gamma f(v) = 0 \tag{4.16}$$

Unfortunately this expression does not have a straightforward interpretation.[8] Consider the values of $\Psi(v)$ at the extreme points of its support set

$$\lim_{v\to 0}\Psi(v) = \gamma f(0) > 0 \tag{4.17}$$

$$\lim_{v\to \pi}\Psi(v) = -\frac{\bar{c}}{\pi} + \gamma f(\pi) \tag{4.18}$$

where $\bar{c} = \int_0^\pi cf(c)dc$ is the average unit production cost. The second limit is likely to be negative if $f(\pi)$ is sufficiently small.

When eq. (4.18) is negative the following assumption is sufficient to ensure that the optimal audit policy is unique.

Assumption 2 *There exists a $v^* \in (0,\pi)$ such that $\Psi(v) > 0$ for $v < v^*$ and $\Psi(v) < 0$ for $v > v^*$.*

Proposition 1 *Under assumption 2, the optimal audit policy is unique and is given by $p_1^* = (1+s)^{-1}$, $p_2^* = 0$ and $\alpha^* = (\pi - v^*)/v^*$, where $v^* \in (0,\pi)$ solves $\Psi(v) = 0$.*

Proof. Assumption 2 implies

$$\int_0^{v^*} \Psi(v)dv = -\int_0^{v^*}\left\{\frac{\pi}{v^*}F(v) - \gamma f(v)\right\}dv > 0 \tag{4.19}$$

$$\begin{aligned}\int_{v^*}^{\pi} \Psi(v)dv &= -\int_{v^*}^{\pi}\left\{\frac{\pi}{v^*}F(v) - \gamma f(v)\right\}dv + \\ &\quad + \frac{\pi - v^*}{v^*}\int_0^{\pi} F(v)dv < 0\end{aligned} \tag{4.20}$$

Let $v = v^*$ into eq. (4.14). Since v^* is independent of p_1 and p_2, one immediately notice that the revenue function is linear in the audit probabilities. Differentiating

$$\frac{\partial T/\partial p_2}{t(1+s)\bar{w}L^\circ} = \int_0^{v^*}\left\{\frac{\pi}{v^*}F(c) - \gamma f(c)\right\}dc < 0 \tag{4.21}$$

$$\frac{\partial T/\partial p_1}{t(1+s)\bar{w}L^\circ} = \int_{v^*}^{\pi}\left\{\frac{\pi}{v^*}F(c) - \gamma f(c)\right\}dc - \frac{\pi - v^*}{v^*}\int_0^{\pi} F(c)dc > 0 \tag{4.22}$$

[8] In Reinganum and Wilde's (1985) model, the corresponding first order condition for the income threshold can be expressed in terms of the hazard rate of the distribution of pre-tax incomes. Under the assumption of increasing hazard rate the optimal threshold is then unique.

so that optimal probabilities are $p_1^* = (1+s)^{-1}$ and $p_2^* = 0$. *Q.E.D.*

The optimal tax enforcement policy is such that income reports below the corresponding mark-up on production costs are audited at the probability level just sufficient to induce truthful reporting, whereas those above it are not audited. The relatively inefficient professionals, those with $c \geq v^*$, are forced to behave honestly and pay in full the legislated income tax. On the other hand, the relatively efficient business, those with $c < v^*$, report $\alpha c x$ and this is sufficient to escape the audit, so that they evade the difference between true income and $\alpha c x$.

4.5.3 Presumptive income coefficients

Suppose that the taxing authority constrains the mark-up coefficient to one of the presumptive income coefficients defined in eqs. (4.10)–(4.11) and chooses audit probabilities to maximize net revenue. Let α^{pic} be the presumptive income coefficient, where α^{pic} is either α^{am} or α^{ls}. Let $v^{pic} = \pi/(1+\kappa\alpha^{pic})$.

Proposition 2 *Let the tax enforcement agency fix a presumptive income coefficient α^{pic}. Under assumption 2 optimal probabilities are*

If $\alpha^{pic} < \alpha^*$ then $p_1^* = (1+s)^{-1}$ and

$$\text{if} \int_0^{v^{pic}} \Psi(v)dv > 0 \quad \text{then} \quad p_2^* = 0$$
$$\text{if} \int_0^{v^{pic}} \Psi(v)dv < 0 \quad \text{then} \quad p_2^* \to (1+s)^{-1}$$

If $\alpha^{pic} > \alpha^*$ then $p_2^* = 0$ and

$$\text{if} \int_{v^{pic}}^{\pi} \Psi(v)dv < 0 \quad \text{then} \quad p_1^* = (1+s)^{-1}$$
$$\text{if} \int_{v^{pic}}^{\pi} \Psi(v)dv > 0 \quad \text{then} \quad p_1^* \to 0$$

Proof. From eqs. (4.19) and (4.21) and eqs. (4.20) and (4.22)

$$\int_0^{v^{pic}} \Psi(v)dv = -\frac{\partial T/\partial p_2}{t(1+s)\bar{w}L^\circ} \tag{4.23}$$

$$\int_{v^{pic}}^{\pi} \Psi(v)dv = -\frac{\partial T/\partial p_1}{t(1+s)\bar{w}L^\circ} \tag{4.24}$$

If $\alpha^{pic} < \alpha^*$ then $v^{pic} > v^*$. From assumption 2 $\int_{v^*}^{\pi} \Psi(v)dv < 0$. Thus $\int_{v^{pic}}^{\pi} \Psi(v)dv < 0$ and $\partial T/\partial p_1 > 0$ from eq. (4.24), so that $p_1^* = (1+s)^{-1}$. Eq. (4.23) can be either positive or negative.

If $\alpha^{pic} > \alpha^*$ then $v^{pic} < v^*$. From assumption 2 $\int_0^{v^*} \Psi(v)dv > 0$. Thus $\int_0^{v^{pic}} \Psi(v)dv > 0$ and $\partial T/\partial p_2 < 0$ from eq. (4.23), so that $p_2^* = 0$. Eq. (4.24) can be either positive or negative. *Q.E.D.*

When the presumptive income coefficient is lower than the optimal mark-up threshold ($\alpha^{pic} < \alpha^*$), then the best audit policy may collapse into random audits at probability $(1+s)^{-1}$, and in this case the notion of presumptive income becomes meaningless. Of course, if α^{pic} is sufficiently *close* to α^*, then optimal probabilities are $p_1^* = (1+s)^{-1}$, $p_2^* = 0$. A sufficient condition for ruling out random audits can be obtained from the following expression

$$-\int_0^{\pi} \Psi(v)dv = -(\pi - \bar{c}) - \gamma = -(\pi - \bar{c}) - \frac{\varphi}{t(1+s)\bar{w}L^\circ} \tag{4.25}$$

If eq. (4.25) is positive, that is $t(1+s)(\pi - \bar{c})\bar{w}L^\circ < \varphi$, then random audits never occur, because in this case $\partial T/\partial p_2 < 0$ for all v^{pic}, see eq. (4.23). To interpret this condition, notice that $(\pi - \bar{c})\bar{w}L^\circ$ is the average pre-tax income.

When $\alpha^{pic} > \alpha^*$, then the tax agency may opt for no enforcement, thus raising no revenue. However, if α^{pic} is sufficiently *close* to α^*, then optimal probabilities are $p_1^* = (1+s)^{-1}$, $p_2^* = 0$. Also, from eq. (4.25), the no enforcement solution never occurs if $t(1+s)(\pi - \bar{c})\bar{w}L^\circ > \varphi$.

4.6 Endogenous labour supply

In this section labour supply and output levels are endogenously determined from the utility maximization problem (4.4). The following lemma characterizes optimal reports, equilibrium outputs and pre-tax incomes.

Lemma 3 *For any $(p_1, p_2, I_c) \in A_1$, professionals' optimal reports are*

If $p_1 < (1+s)^{-1}$ then

$$r(c,w) = \begin{cases} 0 & \text{if } \langle c,w\rangle \in (v,\pi] \times [0,\infty) \\ \alpha c w^2 x^{**}(p_2,\alpha;c) & \text{if } \langle c,w\rangle \in [0,v] \times [0,\infty) \end{cases}$$

If $p_1 = (1+s)^{-1}$ then

$$r(c,w) = \begin{cases} w^2 I^*(p_1;c) & \text{if } \langle c,w\rangle \in (v,\pi] \times [0,\infty) \\ \alpha c w^2 x^{**}(p_2,\alpha;c) & \text{if } \langle c,w\rangle \in [0,v] \times [0,\infty) \end{cases}$$

where

$$v = \frac{\pi}{1+\kappa\alpha}, \qquad \kappa = \frac{(1+s)^{-1} - p_2}{p_1 - p_2}$$

Equilibrium outputs are

$$\begin{aligned} w^2 x^*(p_1;c) &= w^2\Big[1 - tp_1(1+s)\Big](\pi - c), \quad c \in (v,\pi] \\ w^2 x^{**}(p_2,\alpha;c) &= w^2\Big[1 - tp_2(1+s)\Big](\pi - c) + \\ &\qquad -w^2 t\alpha c\Big[1 - p_2(1+s)\Big], \quad c \in [0,v] \end{aligned}$$

and pre-tax incomes are

$$\begin{aligned} w^2 I^*(p_1;c) &= (\pi - c)w^2 x^*(p_1;c) \\ w^2 I^{**}(p_2,\alpha;c) &= (\pi - c)w^2 x^{**}(p_2,\alpha;c) \end{aligned}$$

Proof. See appendix A.2.

Professionals with $c > v$ report no income if $p_1 < (1+s)^{-1}$ and report true income when $p_1 = (1+s)^{-1}$. Professionals with $c \leq v$ report α times production costs. Notice that the tax enforcement policy affects equilibrium pre-tax incomes. In particular $\partial x^*/\partial p_1 < 0$, $\partial x^{**}/\partial p_2 < 0$: an increase in audit probabilities lowers equilibrium outputs because it causes an increase of the expected tax rate on concealed income. Also, $\partial x^{**}/\partial\alpha < 0$, for taxpayers with $c \leq (\pi - \alpha)/\alpha$ meet the higher probability p_1 over a larger range of income reports r.

Applying lemma 3, the problem of the tax agency is

$$\max_{(p_1,p_2,\alpha)\in A_1} T(p_1,p_2,\alpha) = \int_v^\pi \Big\{tp_1(1+s)\omega I^* - \varphi p_1\Big\} f(c)dc + \\ + \int_0^v \Big\{t\alpha c\omega x^{**} + tp_2(1+s)\omega(I^{**} - \alpha c x^{**}) - \varphi p_2\Big\} f(c)dc \qquad (4.26)$$

where $\omega = \sigma_w^2 + \bar{w}^2$ and σ_w^2 is the variance of w.

4.6.1 The information set of the tax enforcement agency

When labour supply is endogenous, the relation between production costs and pre-tax incomes is affected by the tax enforcement instruments. As for the case of exogenous labour supply, the knowledge of individual production costs is not sufficient to determine the corresponding individual pre-tax income, for identical levels of production costs will result in different incomes whenever the skill and efficiency parameters differ.

To illustrate, suppose that $t = 0$, so that incomes and production costs are those in the absence of taxation. Graph (a) in figure 4.2 shows production costs and pre-tax incomes as a function of c for a given value of the skill parameter. Graph (b) shows the relation between production costs and pre-tax income for two different levels $w_1 > w_2$ of the skill parameter.

As for the case of exogenous incomes, individual mark-up coefficients are independent of w. They are defined by

$$\beta^*(c) \equiv \frac{w^2 I^*(p_1;c)}{cw^2 x^*(p_1;c)} = \frac{\pi - c}{c} \quad \text{for} \quad c \in (v, \pi] \qquad (4.27)$$

$$\beta^{**}(p_2,\alpha;c) \equiv \frac{w^2 I^{**}(p_2,\alpha;c)}{cw^2 x^{**}(p_2,\alpha;c)} =$$

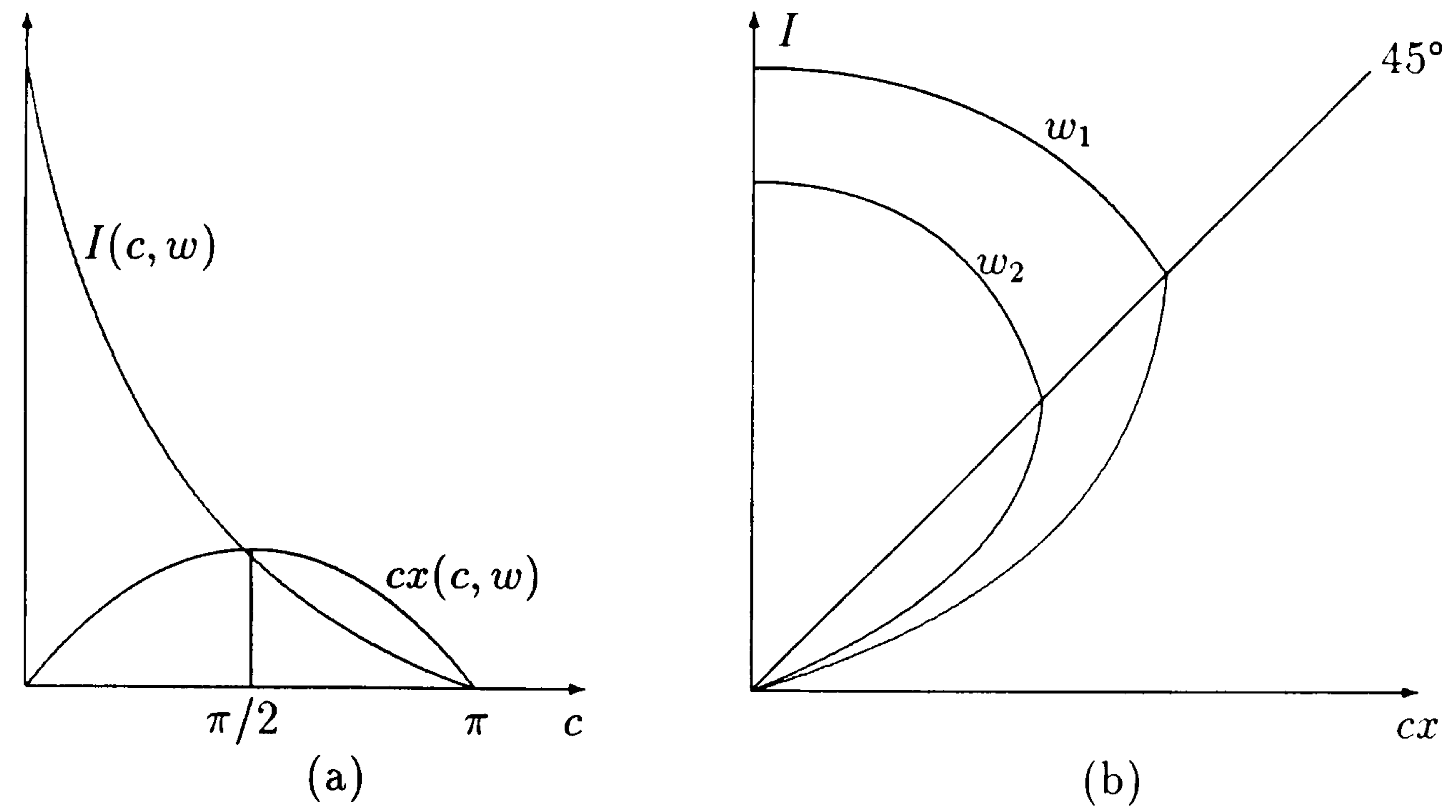

Figure 4.2: Endogenous income and production costs

$$= \frac{[1 - tp_2(1+s)](\pi - c)^2 - t\alpha[1 - p_2(1+s)](\pi - c)c}{[1 - tp_2(1+s)](\pi - c)c - t\alpha[1 - p_2(1+s)]c^2} \quad \text{for} \quad c \in [0, v] \tag{4.28}$$

The computation of the presumptive income coefficients is complicated by the fact that β^{**} is endogenously determined by p_2 and α. However, suppose that the tax agency considers the individual mark-up coefficients under the hypothesis of a random audits policy with probability p_1, in which case one can show that $\beta^{**} = \beta^{*}$.

Then the presumptive income coefficient computed as the average of individual mark-up coefficients is

$$\alpha^{am} = \int_0^{\pi} \frac{\pi - c}{c} f(c) dc \tag{4.29}$$

and this is equal to the corresponding coefficient in the case of exogenous labour supply.

The minimum least square coefficient (of a line passing through the origin) is obtained from

$$\min_{\alpha} \int_0^{\infty}\int_0^{\pi} \Big[(1-t)w^2(\pi-c)^2 - \alpha(1-t)w^2(\pi-c)c\Big]^2 f(c)g(w)dc\,dw$$

and the solution is

$$\alpha^{ls} = \frac{\displaystyle\int_0^{\pi}(\pi-c)^3 c f(c)dc}{\displaystyle\int_0^{\pi}(\pi-c)^2c^2 f(c)dc} \tag{4.30}$$

The rest of this section is organized as follows. The next subsection examines the case of random audits policy. The first order necessary conditions for solving the general problem are presented in subsection 4.6.3. Since an analytical solution cannot be found, subsection 4.6.4 considers some numerical computations of the model.

4.6.2 Random audits

Suppose that the tax enforcement agency audits all professionals at a constant probability p independent of the level of income reports. Random audits are a constrained version of problem (4.26), where $\alpha = 0$ and $p_2 = p$.

Under random audits, government's net revenue and taxpayers' equilibrium pre-tax incomes are respectively

$$T(p) = \int_0^{\infty}\int_0^{\pi} \Big\{tp(1+s)I(p;c,w) - \varphi p\Big\}dF(c)dG(w) \tag{4.31}$$

$$I(p;c,w) = \Big[1 - tp(1+s)\Big]w^2(\pi-c)^2 \tag{4.32}$$

Aggregate, or average, income is

$$\begin{aligned}\bar{I}(p) &= \int_0^{\infty}\int_0^{\pi} I(p;c,w)f(c)g(w)dc\,dw \\ &= [1-tp(1+s)]A \end{aligned} \tag{4.33}$$

where

$$A = (\sigma_w^2 + \bar{w}^2)\Big[(\pi-\bar{c})^2 + \sigma_c^2\Big] \tag{4.34}$$

is aggregate income in the absence of taxation ($t = 0$).

The problem of the taxing authority is to solve

$$\max_p T(p) = tp(1+s)\bar{I}(p) - \varphi p \tag{4.35}$$

The first order condition is

$$\frac{dT}{dp} = t(1+s)\bar{I}(p) - \varphi + tp(1+s)\frac{d\bar{I}}{dp} = 0 \tag{4.36}$$

The first term is positive and shows the increase in fines collection as more audits occur. The second is negative and represents the marginal cost of increasing the audit probability. The third term is negative because an increase in the level of enforcement reduces aggregate pre-tax income. The first order condition is also sufficient, for $d^2T/dp^2 < 0$.

It is immediate to show that the optimal random audits probability is

$$p^* = \begin{cases} 0 & \text{if} \quad t(1+s)A - \varphi < 0 \\ p^{**} & \text{if} \quad t(1+s)A - \varphi > 0 \\ & \text{and} \quad t(1+s)A(1-2t) - \varphi < 0 \\ (1+s)^{-1} & \text{if} \quad t(1+s)A(1-2t) - \varphi > 0 \end{cases} \tag{4.37}$$

where

$$p^{**} \in \left(0, (1+s)^{-1}\right), \quad p^{**} = \frac{t(1+s)A - \varphi}{2t^2(1+s)^2 A}$$

When labour supply and pre-tax incomes are exogenous, the optimal random audit probability is either $p = 0$ or $p = (1+s)^{-1}$, whereas with endogenous incomes also interior solutions may occur, the reason being that an increase in p causes aggregate income to fall.

4.6.3 Optimal audit policy: first order conditions

Returning to the general problem (4.26), the first order condition with respect to the mark-up threshold α is

$$\frac{\partial T}{\partial \alpha} = t\omega\left[1 - p_2(1+s)\right] \int_0^v \left\{cx^{**} + \alpha c\frac{\partial x^{**}}{\partial \alpha}\right\} f(c)dc +$$

$$+ t\omega p_2(1+s)\int_0^v \frac{\partial I^{**}}{\partial \alpha} f(c)dc +$$
$$+ \varphi(p_1 - p_2)f(v)\frac{\partial v}{\partial \alpha} = 0 \qquad (4.38)$$

The first and the second term represent the net benefit, in terms of tax and fines collection, from a marginal increase in the mark-up coefficient. The second term is negative and shows that as α increases the aggregate pre-tax income of professionals with $c < v$ falls. The third term is the marginal cost of increased tax enforcement.

The corresponding first order conditions for the audit probabilities are

$$\begin{aligned}\frac{\partial T}{\partial p_1} &= t(1+s)\omega\int_v^\pi \left\{I^* + p_1\frac{\partial I^*}{\partial p_1}\right\} + \\ &\quad + \varphi(p_1 - p_2)f(v)\frac{\partial v}{\partial p_1} - \varphi\big[1 - F(v)\big] = 0 \qquad (4.39)\end{aligned}$$

$$\begin{aligned}\frac{\partial T}{\partial p_2} &= t\omega\int_0^v \left\{\big[1 - p_2(1+s)\big]\alpha c\frac{\partial x^{**}}{\partial p_2} + p_2(1+s)\frac{\partial I^{**}}{\partial p_2}\right\} f(c)dc + \\ &\quad + t(1+s)\omega\int_0^v (I^{**} - \alpha c x^{**})f(c)dc + \\ &\quad + \varphi(p_1 - p_2)f(v)\frac{\partial v}{\partial p_2} - \varphi F(v) = 0 \qquad (4.40)\end{aligned}$$

To solve for the optimal tax enforcement instruments from these first order conditions is not an easy task. Hence some numerical computations will be considered.

4.6.4 Numerical results

The specification of the model is as follows. The price of the service is $\pi = 4$, the tax rate $t = .35$ and the surcharge rate $s = 4$. The density function $f(c)$ of unit production costs is a truncated normal with parameters of the associated normal distribution $\mu = 2$, $\sigma = .8$ and truncation points at $c = 0$, $c = 4$. The resulting distribution has mean $\bar{c} = 2$ and standard deviation $\sigma_c = .764$. There is no need to specify the form of the distribution $g(w)$ of labour productivity. It is assumed that $\bar{w} = \sqrt{.5}$ and $\sigma_w = \sqrt{.5}$.

Table 4.1: Audit policies.

	Optimal policy				r.a.	$\alpha^{ls} = 1.235$		$\alpha^{am} = 1.730$	
φ	α^*	p_1^*	p_2^*	$F(v^*)$	p^*	p_1^{ls}	p_2^{ls}	p_1^{am}	p_2^{am}
1.0	any α	.2000	.2000		.2000	.2000	.2000	.2000	.2000
1.5	any α	.2000	.2000		.2000	.2000	.2000	.2000	.2000
2.0	.745	.2000	.1929	.644	.2000	.2000	.2000	.2000	.2000
2.5	.638	.2000	.1618	.712	.1967	.1967	.1967	.1967	.1967
3.0	.588	.2000	.1318	.745	.1789	.2000	.1258	.1789	.1789
3.5	.557	.2000	.1025	.765	.1610	.2000	.0903	.1610	.1610
4.0	.537	.2000	.0733	.778	.1432	.2000	.0548	.1432	.1432
4.5	.523	.2000	.0443	.787	.1254	.2000	.0195	.1254	.1254
5.0	.511	.2000	.0157	.794	.1076	.2000	.0000	.1076	.1076
6.0	.442	.2000	.0000	.837	.0720	.0720	.0720	.0720	.0720
7.0	.379	.2000	.0000	.875	.0363	.0363	.0363	.0363	.0363

The optimal audit policies corresponding to various values of the unit audit cost φ are shown in table 4.1. To give a scale to the level of φ, notice that average (aggregate) income is 4.583 in the absence of taxation and 2.979 with perfect enforcement (the income tax is distortionary and affects the incentive to work). When $\varphi \leq 1.5$ the optimal audit policy is random audits at probability $(1+s)^{-1} = .2$, the level of enforcement just sufficient to induce honest reports. For higher values of φ, there is a unique optimal mark-up threshold separating taxpayers into two groups. Income reports below $\alpha^* cx$ are audited at probability $p_1^* = .2$, whereas those above at a lower probability p_2^*. Notice that α^* and p_2^* decrease as φ increases and that for $\varphi \geq 6$ $p_2^* = 0$.

The value of $F(v^*)$ gives the proportion of taxpayers audited at probability p_2^*. For instance, when $\varphi = 4$ the mark-up threshold is .537, so that 77.8% of professionals are audited at probability .0733 and 22.2% at probability .2. Notice that for all values of φ the mark-up threshold α^* is relatively low: in fact, the proportion of professionals audited at the higher probability p_1^* ranges from 36% to 13%. This is a good property of the tax enforcement policy, for it is reasonable to expect that taxpayers would not tolerate a high value of the mark-up threshold.

The column labelled *r.a.* shows the optimal random audits probability p^*. Notice that $p_2^* < p^* < p_1^*$ and that p^* decreases as φ increases.

The last four columns contain the optimal probabilities when α is constrained to be a presumptive income coefficient. The *minimum least squares* coefficient, defined in eq. (4.30), is equal to $\alpha^{ls} = 1.235$; 39.48% of professionals have a mark-up higher than α^{ls}, whereas 60.52% have a lower mark-up. The *average mark-up* coefficient, defined in eq. (4.29), is equal to $\alpha^{am} = 1.730$ and leaves 24.88% of professionals with a higher mark-up. Notice that α^{ls} and α^{am} are larger than α^* for all values of φ. As for the audit probabilities, consider α^{ls} first. When $\varphi \leq 2.5$ or $\varphi \geq 6$, optimal probabilities collapse into random audits and the notion of presumptive coefficient becomes meaningless. When $3 \leq \varphi \leq 5$, p_1^{ls} and p_2^{ls} differ: notice that $p_1^{ls} = p_1^* = .2$ and $p_2^{ls} < p_2^*$. Turning to α^{am}, optimal probabilities collapse into random audits for all values of φ.

Let T^i be government's net revenue and I^i be aggregate income under policy i, where $i = op, ra, ls$ means respectively *optimal policy*, *random audits* and *least squares* presumptive income coefficient. As expected, table 4.2 shows that T^{op} is decreasing in φ, for tax enforcement becomes less profitable as audits are more expensive. However I^{op} increases, for a lower level of enforcement has a positive incentive effect on labour supply. The table also shows the loss in net revenue and aggregate income associated to policies *ra* and *ls*. For instance, when $\varphi = 4$, $T^{ra} = .84T^{op}$ and $I^{ra} = .95I^{op}$. The loss of revenue is large for high values of φ, whereas the incentive effect on average income is not relevant.

Let GT^i and C^i be respectively gross revenue and audit costs associated to policy i. For each audit policy, table 4.3 shows GT^i/I^i, the *effective* average tax rate. Recall that the government levies a proportional income tax at rate $t = .35$, which is the *legislated* average tax rate. With costly audits, tax enforcement is

Table 4.2: Net revenue and aggregate income.

φ	T^{op}	$\frac{T^{ra}}{T^{op}}$	$\frac{T^{ls}}{T^{op}}$	I^{op}	$\frac{I^{ra}}{I^{op}}$	$\frac{I^{ls}}{I^{op}}$
1.0	.8410	1.000	1.000	2.974	1.000	1.000
1.5	.7410	1.000	1.000	2.974	1.000	1.000
2.0	.6373	1.000	1.000	2.992	1.994	0.994
2.5	.5466	0.990	0.990	3.139	0.956	0.956
3.0	.4667	0.959	0.960	3.293	0.955	0.954
3.5	.3983	0.910	0.919	3.447	0.953	0.938
4.0	.3418	0.838	0.852	3.603	0.952	0.922
4.5	.2972	0.739	0.752	3.759	0.950	0.908
5.0	.2645	0.610	0.612	3.915	0.949	0.885
6.0	.2260	0.318	0.318	4.068	0.983	0.983
7.0	.1980	0.091	0.091	4.138	1.035	1.035

Table 4.3: Gross revenue and audit costs.

φ	$\frac{GT^{op}}{I^{op}}$	$\frac{GT^{ra}}{I^{ra}}$	$\frac{GT^{ls}}{I^{ls}}$	$\frac{C^{op}}{I^{op}}$	$\frac{C^{ra}}{I^{ra}}$	$\frac{C^{ls}}{I^{ls}}$	$\frac{C^{op}}{GT^{op}}$	$\frac{C^{ra}}{GT^{ra}}$	$\frac{C^{ls}}{GT^{ls}}$
1.0	.350	.350	.350	.067	.067	.067	.192	.192	.192
1.5	.350	.350	.350	.101	.101	.101	.288	.288	.288
2.0	.344	.349	.349	.131	.134	.134	.380	.386	.386
2.5	.312	.344	.344	.138	.164	.164	.441	.476	.476
3.0	.278	.313	.306	.136	.171	.163	.490	.545	.533
3.5	.243	.282	.283	.127	.171	.170	.524	.609	.600
4.0	.207	.251	.259	.113	.167	.172	.543	.667	.662
4.5	.172	.219	.235	.093	.158	.170	.540	.720	.722
5.0	.136	.188	.221	.068	.145	.175	.503	.769	.789
6.0	.103	.126	.126	.048	.108	.108	.463	.857	.857
7.0	.090	.064	.064	.042	.059	.059	.470	.934	.934

not perfect and the effective tax rate falls short of the legislated rate. The table also shows audit costs in percentage of aggregate income and in percentage of gross tax collection. Notice that audit costs are much higher with *ra* and *ls* than with *op*.

The analysis of the numerical computations of the model with endogenous labour supply is summarized as follows. The first result is that the nature of the optimal tax enforcement policy crucially depends on the size of unit audit costs relatively to the average income of the group. When audit costs are relatively low, see table 4.1, the best policy is random audits. When audits are relatively expensive, then the two-probabilities policy is effective. The relatively inefficient professionals, those with an individual mark-up lower than α^*, meet the probability of audit just sufficient to induce truthful reporting, $p_1^* = (1+s)^{-1}$, whereas the relatively efficient individuals, those with a mark-up higher than α^*, face the lower probability p_2^*. Notice that $p_2^* > 0$ for *intermediate* values of φ, whereas $p_2^* = 0$ when φ is very large.

The second finding is that the optimal mark-up coefficient α^* is generally low and is lower than the presumptive income coefficients α^{ls} and α^{am}. This means that constructing presumptive income coefficients with statistical measures that underestimate the actual coefficients is probably a good policy in terms of net revenue collected. In addition, this normative prescription is supported by the consideration that taxpayers favours a low value of the mark-up threshold.

4.7 Conclusions

This paper has examined in a principal-agent setting the Presumptive Income Coefficients tax enforcement policy implemented by the Italian government since 1989.

When labour supply is fixed, the net revenue maximizing audit policy is such that income reports below the the corresponding mark-up on production costs are audited at the probability level just sufficient to induce truthful reporting, whereas those above it are never audited. Also, when the mark-up coefficient is constrained to be a presumptive income coefficient, the best audit policy may collapse into random audits.

The numerical computations of the model with endogenous labour supply show that the properties of the optimal policy crucially depend on the size of taxpayers' average income relative to unit audit costs. If average income is relatively high the best policy is random audits. If average income is relatively low, the two probability policy is effective: income reports below the corresponding mark-up on production costs meet the probability of audit just sufficient to induce honest behaviour, whereas those above it face a lower, but positive, level of enforcement. Also, the optimal mark-up coefficient is generally low and is lower than the presumptive income coefficients.

The rest of this section suggests few lines for further research. The professional's utility function, which is linear in income and quadratic in labour supply, could be generalized to allow for risk aversion and substitution between income and labour. Also, the homogeneity assumption about professionals' output could be relaxed, allowing for product variety and different demand conditions.

A simple audit technology has been assumed: the cost of verifying an income report is constant and when performing an audit the tax agency observes true income. Instead, the audit mechanism applied by the Italian government is asymmetric. If reported income is at least as high as presumptive income, the taxpayer is subjected to random audits and the task of demonstrating any discrepancy between reported and actual income is on the tax agency. If reported income is lower than presumptive income, the taxpayer is expected to pay the tax on presumptive income, unless he or she demonstrates that the actual income is lower than presumptive income. This implies that monitoring reports below the corresponding presumptive income is less costly than auditing those above it. Also, a taxpayer with actual income lower than presumptive income bears the cost of demonstrating that his or her report is truthful.

A Appendix

A.1 Proof of lemma 2

For a given audit policy, professionals are divided into two groups.

Those with $(\pi - c)wL^\circ < \alpha cwL^\circ$ solve

$$\begin{aligned}\max_r EU &= (1-t)(\pi - c)wL^\circ + t\big[1 - p_1(1+s)\big]\big[(\pi - c)wL^\circ - r\big] + \\ &\quad -\frac{1}{2w^2}(wL^\circ)^2, \quad 0 \le r \le (\pi - c)wL^\circ \end{aligned} \tag{4.41}$$

Linearity in r and assumption 1 imply that $r(c,w) = 0$ if $p_1 < (1+s)^{-1}$ and $r(c,w) = (\pi - c)wL^\circ$ if $p_1 = (1+s)^{-1}$.

Professionals with $(\pi - c)wL^\circ \ge \alpha cwL^\circ$ solve

$$\max_r EU = \begin{cases} (1-t)(\pi - c)wL^\circ + t\big[1 - p_1(1+s)\big]\big[(\pi - c)wL^\circ - r\big] + \\ \qquad -\frac{1}{2w^2}(wL^\circ)^2, \quad 0 \le r < \alpha cwL^\circ \\ (1-t)(\pi - c)wL^\circ + t\big[1 - p_2(1+s)\big]\big[(\pi - c)wL^\circ - r\big] + \\ \qquad -\frac{1}{2w^2}(wL^\circ)^2, \quad \alpha cwL^\circ \le r \le (\pi - c)wL^\circ \end{cases} \tag{4.42}$$

Let $EU(r)$ be the expected utility when the report is r. Linearity in r and assumption 1 imply that the professional chooses his or her optimal report from the discrete set $\{0, \alpha cwL^\circ, I\}$, where $I = (\pi - c)wL^\circ$. From eq. (4.42) one obtains (i) $EU(I) \ge EU(0)$ iff $p_1 \ge (1+s)^{-1}$ and $EU(I) \ge EU(\alpha cwL^\circ)$ iff $p_2 \ge (1+s)^{-1}$; (ii) $EU(\alpha cwL^\circ) > EU(I)$ iff $p_2 < (1+s)^{-1}$ and $EU(\alpha cwL^\circ) \ge EU(0)$ iff $c \le \pi/(1+\kappa\alpha)$; (iii) $EU(0) > EU(I)$ iff $p_1 < (1+s)^{-1}$ and $EU(0) > EU(\alpha cwL^\circ)$ iff $c > \pi/(1+\kappa\alpha)$. From lemma 1, $p_2 < (1+s)^{-1}$, thus case (i) never occurs. When $p_1 < (1+s)^{-1}$, then $r(.) = \alpha cwL^\circ$ if $c \le \pi/(1+\kappa\alpha)$ and $r(.) = 0$ if $c > \pi/(1+\kappa\alpha)$. When $p_1 = (1+s)^{-1}$, then $\kappa = 1$ and $r(.) = \alpha cwL^\circ$ for all $c \le \pi/(1+\alpha)$. This completes the proof.

A.2 Proof of lemma 3

For a given audit policy, professionals are divided into two groups, according to the level of the efficiency parameter c. The relatively inefficient entrepreneurs, those with $(\pi - c) < \alpha c$, solve

$$\max_{x,r} EU = (1-t)(\pi - c)x + t\Big[1 - p_1(1+s)\Big]\Big[(\pi - c)x - r\Big] + \\ -\frac{1}{2w^2}x^2, \quad 0 \le r \le (\pi - c)x \tag{4.43}$$

whereas the relatively efficient, those with $(\pi - c) \ge \alpha c$, solve

$$\max_{x,r} EU = \begin{cases} (1-t)(\pi - c)x + t\Big[1 - p_1(1+s)\Big]\Big[(\pi - c)x - r\Big] + \\ \qquad -\frac{1}{2w^2}x^2, \quad 0 \le r < \alpha c x \\ (1-t)(\pi - c)x + t\Big[1 - p_2(1+s)\Big]\Big[(\pi - c)x - r\Big] + \\ \qquad -\frac{1}{2w^2}x^2, \quad \alpha c x \le r \le (\pi - c)x \end{cases} \tag{4.44}$$

For given x, eqs. (4.43) and (4.44) are identical to eqs. (4.41) and (4.42) respectively. Thus applying the results of the previous proof the optimal reports are

$$r(c,w) = \begin{cases} 0 & \text{if} \quad \langle c, w \rangle \in (v, \pi] \times [0, \infty) \\ \alpha c x & \text{if} \quad \langle c, w \rangle \in [0, v] \times [0, \infty) \end{cases}$$

if $p_1 < (1+s)^{-1}$ and

$$r(c,w) = \begin{cases} (\pi - c)x & \text{if} \quad \langle c, w \rangle \in (v, \pi] \times [0, \infty) \\ \alpha c x & \text{if} \quad \langle c, w \rangle \in [0, v] \times [0, \infty) \end{cases}$$

if $p_1 = (1+s)^{-1}$.

To determine optimal outputs, substitute the optimal reports into eqs. (4.43)–(4.44). Thus professionals with $c > v$ solve

$$\max_{x} EU = \begin{cases} (1-t)(\pi - c)x + t\Big[1 - p_1(1+s)\Big](\pi - c)x + \\ \qquad -\frac{1}{2w^2}x^2, \quad \text{if} \quad p_1 < (1+s)^{-1} \\ (1-t)(\pi - c)x - \frac{1}{2w^2}x^2, \quad \text{if} \quad p_1 = (1+s)^{-1} \end{cases} \tag{4.45}$$

and it is straightforward to show that the solution is

$$X^*(p_1;c,w) = w^2x^*(p_1;c) = w^2\Big[1 - tp_1(1+s)\Big](\pi - c)$$

Professionals with $c \leq v$ solve

$$\max_x EU = (1-t)(\pi - c)x + t\Big[1 - p_2(1+s)\Big]\Big[\pi - (1+\alpha)c\Big]x - \frac{x^2}{2w^2} \quad (4.46)$$

and the solution is

$$\begin{aligned} X^{**}(p_2,\alpha;c,w) = w^2x^{**}(p_2,\alpha;c) &= w^2\Big[1 - tp_2(1+s)\Big](\pi - c) + \\ &\quad -w^2t\alpha c\Big[1 - p_2(1+s)\Big] \end{aligned}$$

This completes the proof.

References

Cremer, H., Marchand, M. and P. Pestieu, 1990, Evading, auditing and taxing: the equity-compliance tradeoff, Journal of Public Economics 43, 67–92.

Grossman, Sanford J. and Oliver D. Hart, 1983, An analysis of the principal-agent problem, Econometrica 51, 7–45.

Kelly, P.L. and O. Oldman eds., 1973, Readings on income tax administration, New York, The Foundation Press.

Lapidoth, Arye, 1977, The Israeli experience of using the Tachshiv for estimating the taxable income, Bulletin for International Fiscal Documentation 31, International Bureau of Fiscal Documentation, Amsterdam, 99–106.

Mookherjee, Dilip and Ivan Png, 1989, Optimal auditing, insurance and redistribution, Quarterly Journal of Economics 104, 399–415.

Reinganum, Jennifer F. and Louis L. Wilde, 1985, Income tax compliance in a principal-agent framework, Journal of Public Economics 26, 1–18.

Sanchez, Isabel and Joel Sobel, 1990, Hierarchical design and enforcement of income tax policies, Southern European Economics Discussion Series, Universidad del Pais Vasco, D.P. 86.