

Optimal Regional Redistribution Under Asymmetric Information[#]

by

Massimo Bordignon* **Paolo Manasse**** **Guido Tabellini*****

First Draft: January 1996

This version: May 13, 1996

Abstract: This paper studies optimal redistribution among two different regions in a federal state. Regional governments supply local public goods financed with distorting local taxes. They have better information on their tax bases than the federal government. We model this both as an adverse selection problem on the size of local tax bases and/or as moral hazard problem on local tax enforcement. Moral hazard alone does not affect the first best redistribution rule, which is a lump sum transfer from the rich to the poor region. In all other cases the optimal transfer rule involves a lump sum tax on the rich regions and a premium for fiscal effort by the poor regions, with the transfer falling short of the first-best level. In the equilibrium with moral hazard and adverse selection, tax evasion occurs only in the poor region, even though the possibility of lax tax enforcement benefits the rich and harms the poor region because it reduces equilibrium redistribution.

Key words: intergovernmental grants, regional redistribution formula, asymmetric information, fiscal federalism

Jel classification: H23

[#] We thank Jordi Brandts, Andreas Haufler, Christian Schulz and participants in seminars at the Universities of Brescia, Bologna, Copenhagen, Paris, and at Catholic University in Milan for helpful comments, and CNR and the HCM research network for financial support.

* Università Cattolica, Milan, and IGER

**Università Statale, Milan, and IGER

***Università Bocconi, Milan, and IGER

1. Introduction

Consider two regions in a federation with non-verifiable differences in per capita tax bases. What is the optimal way to redistribute among them? This is the question addressed in this paper.

Regional redistribution is common in most federal states. In Canada, the federal government transfers revenue to the poor Provinces according to a formula designed to equalise per capita tax bases. In Germany, the revenue collected from the value added tax and from the income tax is redistributed from rich to poor Landers according to a formula which is periodically re-negotiated. In Australia, an independent body, the Commonwealth Grants Commission, periodically re-estimates the parameters of a formula according to which 40% of the federal income tax revenue is distributed to the regional States. In countries such as Italy, that are considering to decentralise some functions of government and to finance them locally, it is commonly accepted that any decentralisation must be accompanied by regional redistribution, to compensate the poor regions for their smaller tax bases. Conversely, in the debate on European integration it is often argued that centralisation of some functions of government (such as monetary policy) could eventually require setting up a risk sharing or redistribute scheme among European states.

The idea that redistribution ought to take place also among regions, and not only among individuals, is quite natural. If taxes are distorting, in the absence of regional redistribution rich and poor regions would be endowed with different amounts of local public goods, even if private consumption was equalised by federal transfers directed at individuals. Thus, from a normative point of view, regional redistribution could be thought of as implementing a risk sharing contract entered into before the regions know their relative state. Alternatively, from a positive point of view, the redistribution could be regarded as a political prerequisite for decentralisation, to convince residents of the poorer region to accept further decentralisation of spending and tax collection.

There is a large theoretical literature on how to redistribute among individuals, that dates back to the seminal work of Mirrlees (1971). But not much is known about how to optimally redistribute among regional governments. Regional redistribution is more complicated because local tax rates and local tax enforcement are endogenous and likely to react to a transfer mechanism among regional governments. Furthermore, while individual taxpayers are many, regions are few, which makes strategic behaviour to affect the transfer paid or received more important. Hence, the inefficiencies of the redistribution are potentially more devastating when it is carried out at the level of the regional government rather than among individuals. The details of the redistribute mechanism are accordingly more likely to be relevant.

Under conditions of full information, unlimited commitment capacity and no spillover effects across regions, optimal redistribution is lump sum. In general, however, lump sum redistribution is not feasible. Exogenous regional circumstances change over time, and are difficult to observe or incorporate in legislation as verifiable contingencies. At the same time, any simple redistribute rule based on measures of locally collected tax revenues or on other behavioural responses is highly distorting, since it creates incentives for under-taxation by regional governments. At the root of these difficulties lies an information problem. A non-distorting redistribute rule ought to be based on regional differences in tax bases that are not manipulable by the regional government. But these parametric differences are generally non-observable or at least non-verifiable by the federal government.

This paper focuses on regional redistribution under asymmetric information over the size of regional tax bases. Local governments are assumed to know more than the federal government. We start with a pure adverse selection problem, in which the informational advantage of local governments is exogenous. We then add to that a problem of moral hazard, where local governments also control tax enforcement as one of their policy tools, and the federal government does not observe it. In both cases, local governments fix the tax rates and choose how much local public goods to provide, knowing with precision the size of the relevant tax base in their region. The only role of the federal government is to design and implement a scheme for horizontal (i.e. without contributions from the federal government) regional redistribution. In the pure adverse selection problem, the federal government observes the local statutory tax rates, but not the size of the local tax base. In particular, the federal government ignores the elasticity of the local tax base to the observed tax rate, as well as the after tax private disposable income associated with a given tax rate. In the version with moral hazard, the federal government also ignores the tax enforcement effort of the local government in fighting tax evasion. These informational constraints are likely to arise in many real world situations. For instance, statutory tax rates on real estate properties are readily observable by everybody, but local governments generally have much better information about the true market value of real estates in their jurisdiction than the federal government. Similarly, in the case of income or indirect taxes, local governments in a federation often have more information than the federal government about the size of the tax evaded or legally eluded. These informational advantages are stronger over locally administered tax bases, and in more decentralised fiscal structures.

One of the central results of the paper is that, to cope with asymmetric information, optimal regional redistribution entails a premium on fiscal effort (i.e. a high tax rate) by the poor region. This result holds with and without moral hazard. The role of this premium is to induce separation of different types and create incentives for truth telling. Fiscal effort is

less costly for the poor region, whose marginal valuation of tax revenue is higher than for a rich region. Hence high fiscal effort is a reliable signal of a small tax base.

As a result of this premium on fiscal effort by the poor, asymmetric information changes drastically the qualitative properties of optimal redistribute schemes. In the model of this paper, under full information and lump sum interregional redistribution, public spending is equalised across regions and the local tax rate is higher in the richer region, since the latter has a more inelastic tax base. With asymmetric information, on the other hand, the poor region spends less but has a higher tax rate than the rich region. Since tax effort by the poor is distorting, however, redistribution is incomplete. These results are analogous to those obtained in the theory of optimal income taxation (Mirrlees (1971), Stiglitz, (1982)), where it is shown that the richest individual is undistorted at the margin by the optimal tax and redistribution is below the first-best level. The incentive scheme studied in this paper is more complex, however, due to strategic interactions among regions.

These features of the equilibrium are preserved when we add moral hazard in the form of endogenous local tax enforcement. In this case, the extent of optimal redistribution is reduced further, because it becomes harder to separate the different types. Hence, the possibility of incomplete local tax enforcement benefits the rich and harms the poor region. Yet, in equilibrium, taxes are never evaded in the rich region, while they might be evaded in the poor region.

The theory of fiscal federalism has long studied the optimal design of intergovernmental transfers. But typically it has focused on other incentive problems, concerning spillover effects across regions due to mobility of tax bases or externalises in the provision of public goods (see the seminal contribution of Oates (1971), Boadway and Flatters, (1982) and for more recent examples Dahlby(1995) and Wildasin (1991)). This paper abstracts from these other incentive problems, not because they are unimportant, but because they are already rather well understood. More recently, other papers have also studied optimal intergovernmental grants under asymmetric information, but focused on different aspects. In particular, Cremer et alii (1995), Bucovetsky and alii (1996) and Levaggi (1991) consider asymmetric information over the local preferences for public goods, whereas Boadway et alii (1995), Raff and Wilson (1995) and Lookwood (1996) study asymmetric information over the local technology for producing local public goods. A recent paper by Laffont (1995) is closer to the present analysis. He focuses on the peculiarities of fiscal arrangements in China, however, where tax rates are set by the federal government while local governments are in charge of a tax enforcement technology with specific assumed properties.

The paper outline is as follows. Section 2 sets out the basic model. In section 3, we study the benchmark case of optimal regional redistribution with full information and full commitment by the federal government. That section also considers the consequences of

having full information but relaxing the commitment assumption. The equilibrium under adverse selection is then characterised in section 4, while section 5 adds the moral hazard problem. Section 6 concludes the paper.

2. The Model

Consider two regions, identical in all respects except in per capita income. They produce only one type of good, which can be transformed in either private or public consumption. Thus, there is no regional trade. Moreover the only input, labour, is immobile. The production technology is linear. There are two levels of government: a local government, that raises taxes and provides a local public good; and a federal government, whose only task is to redistribute among regional governments. Alternatively, the federal government could be thought of as simply enforcing a contract between the two regions.

Individuals within each region are all identical. Population size is normalised to one. Individual preferences are represented by the utility function:

$$(2.1) \quad U = c + V(x) + H(g)$$

where c and g denote (per capita) private and public consumption respectively, and x is leisure time. The functions $V(\cdot)$ and $H(\cdot)$ are well behaved concave functions. Throughout the paper the variables without superscripts refer generically to both regions. Variables referring to the rich region will be indicated with " ' ", while those referring to the poor region with " * ". To capture differences in per capita income, we assume that the "effective" time endowment of individuals differs across regions. Individuals allocate their time endowment, $1 + e$, between leisure and working time, l , so as to satisfy:

$$(2.2) \quad 1 + e = l + x$$

where e is a parameter that differs across regions. Thus, $e' > e^*$. Residents in the rich region are more productive, and we formulate this assumption by saying that they have more "effective" time.

Local taxes are proportional to labour income. Thus, the individual budget constraint is:

$$(2.3) \quad c = l(1 - t)$$

where t is the local tax rate and the real wage is unity. If we denote by τ the redistribute transfer *received* by the regional government, then the local government budget constraint can be written as:

$$(2.4) \quad g = t l + \tau$$

Finally, the federal budget constraint implies:

$$(2.5) \quad \tau' + \tau^* = 0$$

Individual decisions in this economy are very simple, and we can easily derive the individual labour supply in both regions as a function of the local tax rate and of the productivity parameter:

$$(2.6) \quad l = L(t) + e$$

where the function $L(t)$ is decreasing in t by concavity of $V(\cdot)$.¹ Thus, the productivity parameter e differentiates per capita income, and hence the tax base, of the two regions. For the same tax rate, the rich region extracts more tax revenue from the consumer.

Finally, the consumer indirect utility function in both regions can be written as a function of the relevant policy variables and of the productivity parameter:

$$(2.7) \quad W(t, \tau, e) = (1-t)[L(t)+e] + V(1-L(t)) + H(t[L(t)+e]+\tau)$$

The function $W(\cdot)$ is the same in both regions, but the variables t , τ and e naturally vary across regions. Equation (2.7) completely summarises the welfare effects of alternative regional and federal policies. Hence, in the remainder of the paper we use it to evaluate the welfare effects of alternative fiscal arrangements

3. Optimal Redistribution with Full Information

This section describes optimal redistribution across regions when there are no incentive constraints whatsoever. We first describe the optimal allocation when fiscal policy is fully centralised, and the same benevolent social planner chooses all policy instrument. We then show that this allocation can be decentralised, provided that the federal government can enter into binding commitments and has full information about regional differences in per capita income, that is when the productivity parameter e is observable and verifiable in both regions. This is the natural benchmark against which to contrast alternative arrangements when information becomes imperfect. Finally, we investigate the consequences of relaxing the commitment assumption, while still retaining full information.

3.1 Full Centralisation

We start with the assumption that fiscal policy is fully centralised. In this benchmark solution, a central planner chooses all policy instruments to maximise social welfare throughout the federation:

$$(3.1) \quad \text{Max} \quad W(t', \tau', e') + W(t^*, \tau^*, e^*)$$

by choice of t' , t^* , τ' and τ^* , subject to the federal budget constraint (2.5).

¹The first order condition for an optimum implies: $V_x(1 + e - l) = 1 - t$, where the subscript denotes a partial derivative. Thus, the function $L(t)$ is defined as $L(t) = 1 - V_x^{-1}(1-t)$.

The main features of the solution are summarised in

Proposition 1

The solution to problem (3.1) implies: (i) $g' = g^*$; (ii) $\tau^* = [t' l' - t^* l^*]/2 = -\tau'$.
 (iii) $t' > t^*$ if the elasticity of labour supply is increasing in the tax rate t .

Proof: Invoking the envelope theorem, the first order conditions for a maximum of (3.1) with respect to t and τ , subject to (2.5), can be written respectively as:

$$(3.2.a) \quad -l + H_g(g) [l + t L_t(t)] = 0$$

$$(3.2.b) \quad H_g(g') = H_g(g^*)$$

where (3.2.a) holds for both regions, and where l is given by (2.6). Equation (3.2.b) immediately implies (i) and, together with the budget constraints (2.4) and (2.5), it also implies (ii). To prove part (iii) of the proposition, it is useful to rewrite (3.2.a) in the following more revealing form:

$$(3.3) \quad H_g(g) = 1/[1 - \varepsilon(t, e)]$$

where $\varepsilon(t, e) = -t L_t(t)/[L(t) + e] > 0$ is the elasticity of labour supply with respect to the tax rate. Thus, the larger is this elasticity, the more distorting is the income tax, and the smaller is the optimal amount of public consumption.² Equations (3.3) and (3.2.b) imply that in equilibrium these elasticities must also be equal in the two regions:

$$(3.4) \quad \varepsilon(t', e') = \varepsilon(t^*, e^*)$$

Since the elasticity is decreasing in e , and since $e' > e^*$, part (iii) of the proposition immediately follows³. *QED*

Thus, quite intuitively, optimal interregional transfers are set so as to fully equalise local public spending in the two regions. Moreover, the rich region faces a more inelastic labour supply; hence, taxes in the rich region are less distorting, and the optimal tax rate is accordingly higher.

²This property of the equilibrium is also due to our assumption of linear and separable preferences for private consumption. If preferences for private and public consumption are non-separable, then the link between the elasticity of labour supply and the optimal amount of public spending is more complicated. Note that when $\varepsilon = 0$, that is when taxes are non distorting, (3.3) reproduces the well known rule that optimal public good provision equates the marginal rates of substitution and transformation between private and public consumption.

³ Two alternative sufficient (but not necessary) conditions for the elasticity of labour supply (as defined in the text) to be increasing in t is that either $V_{xxx} \geq 0$ or that $[V_{xx}]^2 \geq V_{xxx}$ - see also footnote 1.

3.2 Decentralisation with full information and commitment

This optimal solution can easily be decentralised if there is full information and the federal government has sufficient commitment power. Specifically, consider the following sequence of events:

- The values of the productivity parameters (e' , e^*) are observed by local and federal authorities.
- The federal government sets the regional transfer τ .
- The local governments set their policy instruments.

Suppose further that the federal government maximises social welfare in the two regions, as defined by (3.1), while the two local governments only care about social welfare in their own region. Then it is easy to show that the equilibrium of this game replicates the second best equilibrium illustrated above.⁴

Proposition 2: Under commitment and full information by the federal government, the allocation described in Proposition 1 can be decentralised by means of appropriate lump sum interregional transfers.

Proof: Note, first of all, that equations (3.2) and (3.3), together with the rest of the model laid out in section 2, fully characterise the second best equilibrium solution. In such equilibrium, the local tax rates and the regional transfer can be expressed as known functions of the productivity parameters (e' , e^*). Let $\tau^* = -\tau' = J(e', e^*)$ be the value of the interregional transfer, as a function of (e', e^*), that equates public spending in the two regions in the fully centralised equilibrium of Proposition 1. If the federal government commits to this function, then it remains optimal for both regional governments to set their tax rates according to (3.3) - or, equivalently, to (3.2.a). By construction of the function $J(\cdot)$, local public spending is fully equalised in the two regions, and the second best equilibrium still obtains. QED

Thus, under full information, the federal government commits to a lump sum transfer from the rich to the poor region. The intuition behind Proposition 2 is clear when it is recalled that we assume away any spillover effects across regions. Hence, the optimal transfer does not distort local decisions and the second best is attainable with lump sum transfers.

Alternatively, the second best equilibrium can be interpreted as the outcome of an optimal risk sharing contract between the two regions, written under a veil of ignorance about who is going to be rich and who is going to be poor, and assuming that the

⁴The equilibrium is second best because taxes are distorting.

distribution of risks is symmetric.⁵ Of course, we need to assume that the contract cannot be reneged or renegotiated once the realisation of the productivity parameters becomes known.

To redistribute in a lump sum fashion, the transfer must be a function of the exogenous endowments e . This requires that the endowments are verifiable and the federal government has sufficient commitment power. On the contrary, a redistribute rule based on collected tax revenues would be highly distorting. To illustrate this point, the remainder of this section considers what happens if binding commitments on e are not possible. Imperfect information about the realisation of e is studied in the next two sections.

3.2 No Commitment and full information

Suppose that the timing assumption is reversed: the regional transfer rate is set by the federal government *after* local governments have already committed to a tax rate. Suppose as before that the federal government still cares about the welfare of both regions and weights them identically, while the local governments only care about welfare in their own region. It is easy to show that it remains optimal, for the federal government, to equalise public spending in the two regions. Thus, the regional transfer is still determined as in part (ii) of Proposition 1. But now, the two local governments no longer take τ as given. They instead take *the reaction function* of the federal government as given. That is, they realise that τ is no longer a lump sum, but it is set as a function of collected tax revenue, as in part (ii) of Proposition 1. This distorts their choices. Repeating the previous steps and again appealing to the envelope theorem, local taxes are now set so as to satisfy:

$$(3.4) \quad H_g(g) = 2/[1 - \epsilon(t, e)]$$

Comparing (3.4) and (3.3), we see immediately that public spending, and hence local tax rates, are lower in this equilibrium than with full commitment. The reason is very intuitive. Both local governments now realise that a fraction of the tax revenue raised locally is lost, because it is compensated by smaller regional transfers received (or by larger regional transfers paid). Hence, they have an incentive to undertax relative to the second best. Under the assumed timing, there is nothing the federal government can do to remedy the situation. Since it is ex-post optimal to equalise public spending, the federal government is trapped in a third best equilibrium.

We summarise this discussion in the following:

Proposition 3: If federal commitment to a specific lump sum interregional transfer is not possible, the decentralised equilibrium entails full equalisation of public spending, but too little taxation and public consumption, and too much private consumption.

⁵Persson and Tabellini (1996a) focus on risk sharing under relevant asymmetries in the risk distribution.

Proposition 3 underlines the need for regional redistribution to be based on pre-set formulas to which the federal government can credibly commit. This creates a potential trade-off between commitment and information. Sustainable commitments have to be *simple*, in the sense that they cannot depend on too many contingencies. Simple rules for lump sum redistribution across regions can be written into federal legislation. But it is unlikely that such legislation can incorporate detailed information about *exogenous* features of regional tax bases. Regional circumstances change over time, and exogenous differences in the size of local tax bases are generally not verifiable. These regional circumstances, on the other hand, are likely to be observed by local governments before they make policy decisions. Hence, the federal legislator often operates in a situation of asymmetric information with respect to local governments. This motivates the analysis of the next two sections.

4. Adverse Selection

We now reverse the assumptions made at the end of the previous section. We assume that the federal government is able to commit to a specific transfer rule, but we drop the assumption of full information and verifiability of regional endowments. We argued above that regional differences are unlikely to be verifiable - i.e. they cannot be easily incorporated as contingencies in national legislation. But even observability by the federal government is often problematic. Local and regional governments typically rely on the taxation of a number of assets --housing, local business, tourism, environment, etc. -- where local politicians and local officials generally enjoy an informational advantage over the federal government, not least because the value of these assets depend partly on the policy decisions of local governments themselves. In confederations such the European Union or in federal countries such as China (Laffont, 1995) this problem is even more serious, as national or local governments are the primary source of statistical information and are in charge of tax enforcement decisions.

In all these examples, the federal government is unable to implement an equalising lump sum transfer. To determine the optimal size of the redistributive transfer, it must rely on locally provided information, such as the value of real estate, or the size of value added produced locally. Local governments then have an incentive to exploit their informational advantage to affect the transfer they pay or receive. The optimal transfer scheme trades-off the equalisation concern of federal government against the need to offer the regions incentives for truth telling.

4.1 The Problem

In this section, we study this issue as a problem of adverse selection. That is, regional differences are exogenously given, but unknown to the federal government. Since we already know that any verifiable inequality should be dealt with a lump sum inter-regional transfer, this section only focuses on unverifiable regional differences. Thus, we assume that the two regions are equal in expected value; but in some states of nature, one is rich and the other is poor, and the federal government cannot verify which is which. Specifically, suppose that *in both regions* the parameter e of the regional tax base is a random variable with two possible and equally likely realisations, e' and e^* , with $e' > e^*$. For simplicity, the realisation of e is uncorrelated across regions. Thus, there are four possible and equally likely states of nature. Both regions are rich: (e', e') . Both regions are poor: (e^*, e^*) . One region is rich and the other poor: (e', e^*) ; and viceversa: (e^*, e') . As e is known at the outset by each region and it is the only aspect over which regions can differ, e also defines the "type" of region.⁶ The realisation of e is private information: local governments observe their own type, but not that of the other region. The federal government only knows the distribution from which regional types are drawn. Thus, the federal government faces a pure "adverse selection" problem. In the next section, we add a "moral hazard" dimension to our analysis, by assuming that local governments are also in charge of tax enforcement.

The problem we discuss in this section differs from the standard problem of optimal income taxation in two respects. First, there is an intermediate level of government that interacts with individual economic agents. Second, since there are only two regions, the federal government budget constraint plays a strategic role: what one region pays, the other receives. Thus, the transfer received in equilibrium by each region depends on the behavior of both regions.

Throughout the rest of the paper, we also assume that the federal government observes the regional statutory tax rates, t , but not local public spending, g , nor the total tax revenue collected, tl .⁷ We need this assumption because otherwise the federal government could recover the unobservable parameter e from the budget constraint (2.4.), by its observation of the tax rate and of tax revenue. This assumption could be relaxed, by adding another unobservable parameter in the local tax base, but this would only complicate the model without changing the nature of the problem. Ultimately, what is needed is that the

⁶ Since the two regions are ex-ante identical, we only need to distinguish among them ex-post.

⁷ Cremer et alii. (1995) take the opposite route, and assume that only g is observable; in their model this choice is unavoidable, however, as regional income is exogenously given and taxation is lump sum.

size of the local tax base and its elasticity to the tax rate be unverifiable by the federal government. This feature is captured most simply by the assumption that only the statutory tax rate is verifiable.

Under these assumptions, the problem of finding an optimal transfer rule can then be represented as the equilibrium outcome of the following five-stage game:

1. The federal government designs an inter-regional transfer rule as a function of whatever can be verified.
2. Nature moves and determines whether each region receives a high (e') or low (e^*) endowment of effective labour time. Only local governments observe the realisation of e in their own region.
3. Local governments are asked to simultaneously communicate their type. These regional announcements are verifiable (i.e., they can enter as arguments of the transfer rule set in stage 1), and are observed by both regions.
4. Both regions simultaneously choose their tax rates. Since tax rates are verifiable, they too can enter as arguments of the transfer rule set in stage 1.
5. Local governments pay or receive the transfer according to the rule itself.

Given the coercive powers of the federal government, it is natural to assume that both regions are forced to play this game and thus we do not impose any participation constraint.

The stages of this game capture the basic elements of the interregional redistribution schemes operating in several real world federations (see Shah, 1994 for an institutional analysis of several countries). These schemes typically determine the interregional transfer as a function of an assessment of the size of local tax bases, given locally provided information (stage three of this game), and of regional tax rates (stage four of this game).

By invoking the direct revelation principle (see for instance, Fudenberg and Tirole (1991) chapter 7), if an equilibrium of the complex game stated above exists, it is equivalent (for the principal) to that of a simpler game. In the latter game, the regions directly declare their own type and the federal government chooses a contract in the tax rates and in the transfer, given both regional types, such that each region has an incentive to reveal its type truthfully. We can then solve this problem analytically by letting the federal government directly choose the local tax rates t and the transfer τ in each state of nature, subject to the incentive compatibility constraints of the two regions and to the balanced budget constraint (2.5).

To set out the notation, we need to define a local tax rate and a transfer in each of the possible states of nature. Note that regions differ only by type. By symmetry, the same types ought to be treated equally. The balanced budget constraint then implies that in the states (e', e') and (e^*, e^*) the interregional transfer must equal zero. Let the local tax rates in these

two states of nature be denoted by t'' and t^{**} respectively.⁸ Similarly, let t'^* and $t^{*'}$ be the local tax rate in the rich and poor region respectively, and let τ^* denote the transfer received by the poor region, in state (e', e^*) . By symmetry, the notation in state (e^*, e') is the same, except that the identity of the regions is reversed.

Exploiting symmetry, the optimal transfer rule maximises the expected utility of the federal government, under the relevant incentive compatibility conditions:

$$\begin{aligned} \max \quad & W(t'', 0, e') + W(t^{**}, 0, e^*) + W(t'^*, -\tau^*, e') + W(t^{*'}, \tau^*, e^*) \\ & t'', t^{**}, t'^*, t^{*'}, \tau^* \\ \text{s.t.} \quad & \end{aligned}$$

$$I.C.1 \quad W(t'', 0, e') \geq W(t^{*'}, \tau^*, e')$$

$$I.C.2 \quad W(t'^*, -\tau^*, e') \geq W(t^{**}, 0, e')$$

(4.1)

$$I.C.3 \quad W(t^{*'}, \tau^*, e^*) \geq W(t'', 0, e^*)$$

$$I.C.4 \quad W(t^{**}, 0, e^*) \geq W(t'^*, -\tau^*, e^*)$$

where I.C.1 - I.C.4 indicate the incentive compatibility constraints for dominant strategy implementation. That is, in equilibrium each agent is induced to tell the truth whatever the other type reports.⁹ Thus, I.C.1 says that in state (e', e') a rich region must prefer to tell the truth, in which case it gets the tax rate t'' and a transfer equal to zero, rather than pretending to be poor, in which case it receives the transfer τ^* but its tax rate must equal $t^{*'}$. I.C.2 also applies to a rich region, but in state (e', e^*) . Here the rich region must prefer truth telling, in which case it pays the transfer τ^* and sets the tax rate t'^* , rather than pretending it is poor, in which case it pays nothing but sets the tax rate at t^{**} . I.C.3 and I.C.4 induce truth telling by the poor region, in the two other possible states.

4.2 The Single Crossing Condition

A necessary condition for the existence of a solution to problem 4.1 is that preferences obey a familiar “single crossing” property (cf. Fudenberg and Tirole (1991)): the

⁸ By symmetry, both regions have the same tax rates in equilibrium in these states, even though $t'' \neq t^{**}$.

⁹ This incentive constraint is more demanding than that under Bayesian implementation, and thus forces the principal to leave more informational rents to the agents. It has however the advantage that it is not sensitive to the beliefs about the reports of the other agents in the game - see Fudenberg and Tirole (1991), Chapter 7.

marginal rate of substitution between the transfer and the regional tax rate must be a monotone function of the type. Figure 1 illustrates the indifference curve of a generic type e . The slope of this curve is the marginal rate of substitution between the transfer and the tax rate:

$$(4.2) \quad \left. \frac{d\tau}{dt} \right|_{W=\bar{W}} = -W_t / W_\tau ;$$

Note that as $W_\tau = H_g > 0$ and $W_t > 0$ at low levels of t and $W_t < 0$ at high level of t (as $W_{tt} < 0$), the indifference curve has the U-shape indicated in the picture. This is because at “low” levels of t , the marginal utility of spending exceeds the marginal welfare cost of distortions, and viceversa at “high” levels of t . The indifference curve reaches a minimum in τ at the point $W_t=0$. Intuitively, the transfer needed to keep utility constant at \bar{W} is smallest at the optimal tax rate. Regional welfare is increasing in τ for a given t ; thus, utility increases as we move upwards. Note, for future reference, that the minimum point of the indifference curves also moves upwards and to the left as utility increases; this occurs because the optimal tax rate is decreasing in τ .¹⁰

The single crossing condition implies that the indifference curves of the rich and poor types intersect only once. This could happen if the rich region indifference curve is everywhere either steeper or flatter than that of the poor region: in both cases the necessary condition for the existence of a solution would be fulfilled. The properties of the equilibrium, however, depend on which assumption we make. Throughout the paper, we assume that the rich region indifference curve is everywhere steeper than that of the poor region, so that at the optimum the intersection between the indifference curves of the two regions occurs in the upwards sloping segment of the two curves, as illustrated in Figure 2. More specifically, we assume:

S.C. Single crossing condition: $-W_t(t, \tau, e') / W_\tau(t, \tau, e') > -W_t(t, \tau, e^*) / W_\tau(t, \tau, e^*)$, $\forall t, \tau$.

To understand the implications of this condition, consider the area to the right of an intersection point, such as the darkened area to the right of point E in Figure 2. In this area, the utility of the poor region is higher and the utility of the rich region is lower than at point E . This is what allows separation and makes truth telling incentive compatible. By increasing the tax rate along with the transfer, we can find allocations where the poor region

¹⁰This can be obtained by simply applying the implicit function theorem to the first order condition in equation (3.2.a). This result is due to our assumption of constant marginal utility for private consumption, but decreasing marginal utility for public consumption.

is better off without the rich region wanting to mimic it¹¹. As we will see below, the SC condition implies that in equilibrium only the incentive constraints of the rich region, I.C.1 and I.C.2, bind - see the proof of Proposition 4 in Appendix 1.

To gain a better understanding of condition SC, suppose that e is a continuously distributed variable, differentiate (4.2) with respect to e and simplify. SC reduces to:

$$(4.3) \quad -(g - \tau)H_{gg} / H_g > H_g - 1$$

Equation (4.3) can be shown to imply that the optimal tax rate is decreasing in e for a given τ : the optimal tax rate is higher in the poor region for the same transfer τ (see also figure 2). As e increases, the optimal tax rate (given τ) is affected by two opposite forces. On the one hand, an "income" effect pushes the optimal tax rate downwards: as the tax base increases, with diminishing marginal utility of public consumption the optimal tax rate falls. On the other hand, a "substitution effect" pushes the optimal tax rate in the opposite direction. A higher value of e reduces the elasticity of labor supply and makes the tax rate less distorting (see the Proof of Proposition 1 in section 3). Hence the optimal tax rate increases. Condition (4.3) implies that the income effect dominates over the substitution effect, so that the optimal t is decreasing in e .¹² This is more likely to happen the more concave is the utility function $H(\cdot)$ compared to the distorting effects of the tax.

4.3 The optimal transfer rule

We are now ready to characterise the solution to problem (4.1). We assume throughout that the second best equilibrium described in section 3 is not implementable (i.e., it violates at least one of the incentive constraints I.C.1 - I.C.4).¹³ We show that in this case, to induce

¹¹This is where our formulation of the SC matters. Had we assumed that the indifference curve of the rich region is everywhere *flatter* (rather than steeper) than that of the poor, at the optimum the single crossing would have occurred in the downward segment of the indifference curve. In this case, to obtain separation, the choice of t by the poor region should be distorted *downwards*. This is less plausible. To increase the poor region's consumption of g and still induce separation, the federal government would have to induce the poor region to choose a lower tax rate than it would have chosen otherwise. This would make it even more difficult to reach the equalisation in public consumption the federal government is seeking in the first place. Moreover, there are no examples in the real world of redistributive mechanisms with such a provision.

¹² Note that by (3.3), the right hand side of (4.3) also measures the tax distortion: $\varepsilon/(1-\varepsilon)$.

¹³ Clearly, this is the only case in which asymmetric information matters. If the second best were implementable, it would also be the solution to problem (4.1). We cannot rule out that this is indeed the case: the optimal tax rates of the two types could be sufficiently far apart to induce separation at the second best, without any further distortion.

separation, the tax rate in the poor region must be distorted upwards in all states of nature, while the tax rate in the rich region is left undistorted, also in all states of nature. In more concrete terms, the optimal transfer mechanism puts a premium on “fiscal effort” by the poor regions. Since these distortions are costly, redistribution is incomplete. Thus, as usual, asymmetric information forces the principal to leave some informational rents to the agents: the rich region is better off under asymmetric information, since the transfer paid is smaller and the choice of the tax rate by this region is not distorted.

Let the equilibrium solution to this adverse selection problem be denoted with the subscript A . Then we have:

Proposition 4. Under condition SC, the solution to problem (4.1) entails:

$$(i) t_A'' = \arg \max W(t'', 0, e'); \quad t_A'^* = \arg \max W(t'^*, -\tau_A^*, e');$$

$$(ii) t_A^{**} \geq \arg \max W(t^{**}, 0, e^*), \text{ with strict inequality if I.C.2 binds;} \\ t_A^{*'} \geq \arg \max W(t^{*'}, \tau_A^*, e^*), \text{ with strict inequality if I.C.1 binds;}$$

$$(iii) g_A'^* > g_A^{*'} ;$$

$$(iv) t_A'' < t_A^{*'}; \quad t_A'^* < t_A^{**}.$$

The proof is provided in the Appendix.

Figure 3 illustrates the equilibrium when both I.C.1 and I.C.2 bind. By the SC condition, the rich region indifference curves intersect the poor's indifference curves from below. The rich region can either face another rich region, in which case it neither pays nor receives a transfer (point A), or it can face a poor region, in which case it pays the transfer $-\tau_A^*$ (point B). Similarly, the poor region can either face a rich counterpart and receive τ_A^* (point C), or it can face another poor region, and receive zero (point D).

To verify that, say, point A is an equilibrium, suppose that the true state of nature is (e', e') : both regions are rich. If they tell the truth, both rich regions are at point A. If one of the two regions unilaterally deviates from the equilibrium and claims to be poor, it would receive the positive transfer τ_A^* , but it would also have to distort its tax rate upwards to t^{*} ; the deviation would thus put it at point C in Figure 3. By construction, points A and C lie on the same indifference curve, and the rich is therefore indifferent between lying and truth-telling. By a similar argument, no deviation is worth it at point B, where the rich faces a poor counterpart. If it lies, the rich avoids paying the transfer, but it is forced to set a higher tax rate. By construction, the rich is again indifferent between lying and truth telling. It is also clear from Figure 3 that a poor region would never want to claim that it is rich, since

its indifference curve at the poor equilibrium allocations (points C and D) are strictly higher than the indifference curves passing through the rich allocations (points A and B respectively). Hence the incentive constraints of the poor regions do not bind in equilibrium.

As stated in part (i) of the Proposition, a rich region sets its tax rate optimally (and thus is at the minimum point of its indifference curve), irrespective of which counterpart it faces. Thus, the interregional transfer paid is a lump sum tax for this agent. This is a typical result of optimal separating mechanisms (the "no distortions at the top" condition). It appears also in the literature on optimal income taxation, where it is shown that the marginal tax rate on the richest individual is always zero at the optimum (see for instance Stiglitz (1982)). Here the set up is different, because a balance budget constraint complicates the strategic interaction between the agents, and because the agents (the local governments) are themselves principals in another game *vis à vis* the private sector. But the basic intuition is similar. It is never optimal to distort the rich region, as the only reason for distortion is to make the rich unwilling to mimic the poor. This can always be achieved at a lower welfare cost by distorting the poor region's allocation only.

Indeed, as stated in part (ii) of the Proposition, a poor region is always induced to set a higher tax rate than individually optimal, and thus it is in the rising segment of its indifference curve, also irrespective of which counterpart it faces. In particular, even if both regions are poor, and thus cannot receive any positive transfer, it may be necessary to distort their tax rates upwards to prevent imitation from a rich type trying to avoid a negative transfer. The key to this result is that fiscal effort is less costly for a poor than for a rich region. The reason is that, by SC, tax revenue has a higher marginal value for a poor region than for a rich one. Hence, fiscal effort is a credible signal of a small tax base.¹⁴

Since these distortions reduce expected national welfare, optimal inter-regional redistribution is incomplete. As stated in part (iii) of the Proposition, public consumption is not fully equalised across regions. Thus, asymmetric information "bites", and forces the federal government to leave informational rents to the rich region.

Finally, it can be shown that if the redistribute motives of the federal government become stronger, either because the federal government puts more weight on the poor region welfare, or because e^* drops so that the two regions get further apart, the size of the equilibrium transfer increases. But since separation becomes more difficult, the premium

¹⁴ This suggests a precise definition of "fiscal effort" by a region. Fiscal effort (FE) is the distortion that the poor region suffers in order to receive a transfer, or to avoid paying one. Thus, in the case of a transfer received: $FE \equiv t_A^{*'} - \arg \max_{t_A^{*'}} W(t_A^{*'}, +\tau_A^*, e^*)$. That is, it is the difference between the tax rate that the poor region would have chosen optimally at the equilibrium level of the transfer, and the tax rate the poor region must set in order to receive that same transfer.

on fiscal effort also increases, and the poor region tax rate is distorted upwards even more in both states of nature.

Proposition 4 characterises the optimal allocation, under the informational constraints on the regional tax bases. But how can such an allocation be implemented in practice? The most straightforward mechanism has both regions announcing their type, and their *joint* announcement determining an assignment of both a tax rate and a transfer. But other mechanisms can also sustain the equilibrium allocation, as long as they have the feature that the transfer paid or received by each regions depends on the announcements (or actions) of *the other* region as well, and not only on their own announcement (or action).

For instance, the following alternative mechanism could be used. First, both regions simultaneously announce their type. Then, they are offered a simple step-wise transfer rule, determining the size of the transfer paid or received as a function of *their own* tax rate. The transfer rule they face depends on *the other region* announced type. If, say, region 1 announced that it is rich, then region 2 is offered the following transfer rule:

$$(4.4R) \quad \begin{aligned} \tau &= \tau_A^* \text{ if } t_2 \geq t_A^* \\ \tau &= 0 \text{ if } t_2 < t_A^* \end{aligned}$$

where t_2 is the tax rate set by region 2. If on the other hand region 1 announced that it is poor, then region 2 is offered a different transfer rule:

$$(4.4P) \quad \begin{aligned} \tau &= 0 \text{ if } t_2 \geq t_A^{**} \\ \tau &= -\tau_A^* \text{ if } t_2 < t_A^{**} \end{aligned}$$

Region 1 is offered a similar transfer rule, also as a function of the other region announcement. Facing these transfer rules, both regions find it optimal to bring about the equilibrium allocation, given that the other region has announced its type truthfully. Moreover, neither region has an incentive to announce a false type, since the announcement only affects the transfer rule faced by the other region and not by itself.¹⁵

The transfer rule specified in (4.4R) and (4.4P) is somewhat reminiscent of the redistributive schemes operating in some federal states. Interregional redistribution often

¹⁵ This indifference between truth telling and lying in the announcement stage could be easily turned into a strict preference for truth telling, by imposing a small penalty on both regions if their selected tax rate is not consistent with their own announced type, given the other region announcement. Also note that under this implementation mechanism, the federal government budget constraint could be violated out of equilibrium. By slightly changing the implementation mechanism, however, the out of equilibrium violation of the federal government budget constraint could take the form that not all the money paid by one region is always received by the other region. That is, out of equilibrium some of the money could go from the regions to the federal government, but not viceversa.

entails a premium on “fiscal effort” for the receiving region. This provision is usually justified with reference to a "moral hazard" argument, as preventing the poor region from choosing a suboptimally low tax rate. Here we derive it also as the result of "adverse selection", that is, to induce truthful revelation of information by the rich type.

A good example of such a scheme is the formula for regional redistribution among the Provinces in Canada, the so called Representative Tax System (see Bird and Slack, 1990). According to this system, on each potential tax base, a poor Province receives a transfer proportional to the difference between its per capita tax base and the nationwide average. The coefficient of proportionality is the "standard" tax rate, namely a weighted sum of the tax rates in all the Provinces.¹⁶ This system creates an incentive for a poor Province to raise its tax rate: by doing so, it raises the “standard” tax rate and therefore the transfer it receives. This premium for fiscal effort is larger for more elastic tax bases, as the higher tax rates also increase the difference between its per capita tax base and the average.¹⁷

It is interesting to note that, as predicted by the equilibrium of this model, poor provinces in Canada in fact tend to set higher tax rates than the richer one - see **Table 1**.¹⁸

5. Moral Hazard and Adverse Selection

In this section we add a moral hazard dimension to our previous analysis by letting regional governments also choose the enforcement of local taxes, in addition to the tax rates. This case is important, because the optimal transfer rule entails a premium on fiscal effort. Thus, regions may be tempted to reduce tax pressure by reducing unobservable tax enforcement, while keeping a tax rates high. In several countries, local governments are indeed wholly or partially responsible for monitoring their own tax bases¹⁹. More generally, the value of local tax bases is to some extent under local governments control, even if they are not formally in charge of tax enforcement. For instance, by enforcing

¹⁶ Thus, the fiscal capacity of a Province is computed with reference to the totality of "potential" regional taxes, although only few of them are simultaneously in use in all Provinces. As a consequence, a potential tax base has to be computed for each potential tax for each Province on the basis of local information. Clearly, this is a situation where asymmetric information problems are likely to be relevant.

¹⁷ Rich provinces do not face such incentives, since the transfer is not paid by them directly, but by the Federal government.

¹⁸ We are grateful to Andreas Haufler for pointing this out to us and supplying us with the data and sources.

¹⁹ For example, local taxes on housing are audited by local and regional governments in Denmark, Germany, Australia, the US and the United Kingdom and jointly with the central state in Spain and Italy. Only very centralised states, such as France, leave this task entirely to the central government (see Messere, 1993).

changes in local policies (e.g. local transports, rent regulations and so on) local governments can raise the value of housing or land and thus the tax base for the property tax, a typical local tax. It would be clearly difficult for the federal government to condition the transfer on these local policies. Thus, the optimal transfer rule must take into account the effect of the transfer itself on local policies .²⁰

In this section, we retain the assumption that only the tax rates are verifiable. In particular, the federal government cannot observe local effort in tax enforcement, and thus cannot condition the transfer upon it. We ask three basic questions. How does moral hazard affect the optimal redistribution scheme we derived in the previous section? Does tax evasion occur in equilibrium? In which state of nature? And in which region ?

5.1 Tax evasion

Each worker may either be a tax-payer or a tax-evader. The former is identical to the representative consumer of the previous section. The latter is a full evader and so faces a tax rate $t=0$. When a tax -evader is "caught", he instantly becomes a tax-payer, and no extra penalty is imposed upon him . A tax-evader works more, as $L(0) > L(t)$, $t>0$, and can afford a higher level of private consumption than a tax-payer. Consumption of the local public good is the same for tax-payers and tax-evaders. Hence, if possible, everybody would prefer to free ride and evade taxes. We assume that only local governments can enforce tax payments. Let $a \in [0,1]$ represent the intensity of enforcement, as well as the probability of finding a potential tax-evader. With population in each region normalised to one, a and $1-a$ therefore also indicate the number of tax-payers and tax-evaders, respectively. For simplicity we assume that tax enforcement is costless for the local authority. Thus the local government budget constraint is now:

$$(5.1) \quad g = a t (L(t) + e) + \tau.$$

For the rest, the model is as described in section 2.

In order to characterise the optimal transfer scheme we follow this strategy. First, we look at the problem with only moral hazard and no adverse selection. Then we add adverse selection and study the optimal transfer scheme when the new incentive constraints are taken into account.

5.2 The choice of tax enforcement.

Assume for the moment that the federal government can observe type e (but not tax enforcement, a). The timing of events is as follows: first the e 's are realised; then the

²⁰ Persson and Tabellini (1996a) also study moral hazard and redistribution , though in a different settings.

federal government, having observed the e 's, commits to a lump-sum inter-regional transfer, τ , conditional on the realisation of e in the two regions; finally local governments choose the local tax rate, t , and tax enforcement, a , given τ .

Let the tax-payers' and evaders' indirect utility be denoted by superscripts "P" and "E", respectively. Going through the same steps as in section 2, we have:

$$(5.2) \quad W^P = (1-t)(L(t) + e) + V(1-L(t)) + H(g);$$

$$W^E = L(0) + e + V(1-L(0)) + H(g);$$

In order to preserve the comparability with previous sections, we assume that local governments weight consumers equally, irrespective of whether they evade taxes or not. Hence local governments maximise the social welfare function²¹:

$$(5.3) \quad W(a, \tau, t, e) = aW^P + (1-a)W^E$$

subject to (5.1), (5.2) and $0 \leq a \leq 1$.

Let $A(t, \tau, e)$ denote the optimal choice of tax enforcement for region e , given t and τ . Since the non negativity constraint for a never binds, we can substitute for (5.1) and (5.2) into (5.3), and write the first-order condition for the choice of a as:

$$(5.4) \quad W_a(a, t, \tau, e) = [(H_g(\cdot) - 1) t (L(t) + e)] + [L(t) + V(1-L(t)) - L(0) - V(1-L(0))] \geq 0$$

$$(5.5) \quad \text{if } W_a(A(\cdot), \tau, t, e) > 0 \text{ then } A(t, \tau, e) = 1$$

Equation (5.4) states that the social benefit of an additional monitoring effort must be at least equal to its cost. The former, given by the first square bracket in the RHS of (5.4), is the marginal utility of the extra units of public goods financed by the new contributor, net of (the utility of) consumption foregone. The latter, the second bracket, is the utility loss for the tax-evader who becomes a tax-payer. Equation (5.5) is the standard Kuhn-Tucker condition. We now prove:

Lemma 5.1: Suppose $t \leq \arg \max_{t'} W(a, \tau, t', e)$. Then, the optimal enforcement technology is $A(t, \tau, e) = 1$.

²¹Note that in this section, with a slight abuse of notation, the indirect welfare function $W(\cdot)$ now also depends on a .

Proof: see Appendix

Simply stated, the lemma says that in this model tax evasion does not occur unless the local tax rate is distorted upwards. Consider first the case in which t is optimally chosen. A region would never want to allow for tax evasion for the following reason. Not fully enforcing the tax law is equivalent to taxing identical individuals at different tax rates. But this policy can never be efficient when tax rates are optimally chosen. Given that monitoring is costless and tax rates are optimally set, by increasing a (when it is less than one), the local government can transfer leisure from the tax-payer (who works less and enjoys more leisure) to the tax-evader, whose marginal valuation is higher²². Tax evasion is even less desirable when the tax rate is below the optimal level, since allowing for tax evasion would further reduce the supply of public goods when it is already inefficiently low.

Lemma 5.1 allows us to derive the following conclusion:

Proposition 5: When types are verifiable and the federal government can commit, to a lump sum transfer rule, tax evasion never occurs in equilibrium: $A(t, \tau, e) = 1$ for all τ and e ; hence the optimal transfer scheme is the same as that derived under full information and commitment in Section 3.

Proposition 5 follows immediately by recalling, from Section 3, that under full information and federal commitment, the interregional transfer is lump sum. Hence, both regions optimally choose their tax rates, $t = \arg \max_t W(a, t', \tau, e)$, (cf. Proposition 2). As all types are undistorted in their choice of the tax rates, Lemma 5.1 implies that tax payments are fully enforced. Thus, even if the federal government cannot monitor the choice of a , it knows that by choosing a lump sum transfer, no tax evasion occurs. Note that this result holds even without federal commitment, as long as there is full information. As shown in Section 3.2, in this case we would still have $t = \arg \max_t W(a, t', F(t', t^*, a, a^*), e)$, where $F(\cdot)$ denotes equilibrium transfers (see equation (3.4)). Hence, Lemma 5.1 could still be invoked, and Proposition 5 would follow as before, although tax rates would be lower. Nor the result depends on the assumption of costless monitoring. If enforcement costs were increasing in a , the optimal monitoring technology A would be between zero and one, but it still would not be affected by interregional transfers.

Hence, moral hazard on tax enforcement *per se* is irrelevant for the design of optimal interregional redistribution. As long as the interregional transfer is lump sum or entails no

²²Note that since utility is linear in consumption, private consumption reallocation between consumers is irrelevant for aggregate welfare.

premium on fiscal effort, local tax rates are not distorted upwards. Hence, regional governments would always fully enforce local tax payments

Before turning to a discussion of the model with both moral hazard and adverse selection, it is useful to derive the following lemma, obtained by totally differentiating (5.4):

Lemma 5.2: *single crossing (SC) implies $A_t, A_\tau, A_e \leq 0$, with strict inequalities at an interior solution, $A < 1$.*

The proof is in the Appendix

Suppose that local tax rates are "too high" and some tax evasion is optimal ($A < 1$). Then a further increase in the tax rate enhances the tax distortions even more, and strengthens the incentive to allow for tax evasion. Furthermore, as the transfer received increases, the smaller is the marginal benefit of tax revenue for a given tax rate, and thus again the stronger is the incentive to allow for tax evasion. Finally, at the *same* level of t and τ the rich type always evades more than the poor: $A(t, \tau, e') \geq A(t, \tau, e^*)$ for all t, τ . Intuitively, this follows from our SC assumption, which implies that the optimal tax rate is decreasing in the type e (see the discussion in section 4). In terms of our previous diagrams, whenever a type allows for tax evasion, the right branch of its indifference curve becomes flatter, the flatter the lower is A . This happens because the welfare cost of a higher (than optimal) tax rate is reduced by allowing some workers to evade taxes. Hence the region is willing to accept a smaller transfer to compensate for the higher tax rate²³.

5.3 Optimal redistribution under adverse selection and moral hazard

Suppose now that neither the local endowment e nor the enforcement effort a are verifiable. Like in section 4, regional endowments are uncorrelated, and there are four states of nature. The timing of events is as in section 4, except that now t and a are simultaneously chosen by the regions in the fourth stage. The choice of tax enforcement by local authorities is still characterised by (5.4) and (5.5). In order to describe how this choice affects the possibility of mimicking, we have to introduce some new notation. We adopt the convention of denoting by " $_$ " the (optimal) choice of one type when imitating the other type. Thus:

²³This can be seen by taking the partial derivative of the indifference curve's slope wrt a :

$$d(d\tau/dt|_{W=cost})/da = d(-W_t/W_\tau)/da = (-W_{ta}W_\tau + W_tW_{\tau a})/W_\tau^2.$$

Since $W_t < 0$ to the right of the point of minimum, $W_\tau = H_g > 0$, and the cross derivatives are negative from SC, the previous expression is positive (i.e the slope decreases when a falls).

$$(5.6) \quad \begin{array}{ll} A^{*'} = A(t^{*'}, -\tau^*, e'); & A^{*'} = A(t^{*'}, \tau^*, e^*); \\ A'' = A(t'', -\tau^*, e'); & A^{**} = A(t^{**}, \tau^*, e^*) \\ \underline{A}^{**} = A(t^{**}, 0, e'); & \underline{A}^{*'} = A(t^{*'}, 0, e^*) \\ \underline{A}^{*'} = A(t^{*'}, \tau^*, e'); & \underline{A}^{**} = A(t^{**}, -\tau^*, e^*) \end{array}$$

Under this convention, the optimal redistribution scheme with adverse selection (AS) and moral hazard (MH) solves:

$$(5.7) \quad \begin{array}{l} \text{Max} \quad W(A'', t'', 0, e') + W(A^{**}, t^{**}, 0, e^*) + W(A^{*'}, t^{*'}, -\tau^*, e') + \\ t'', t^{**}, t^{*'}, t^{*'}, \tau^* \\ + W(A^{*'}, t^{*'}, \tau^*, e^*) \end{array}$$

s.t (5.6) and

$$I.C.5 \quad W(A'', t'', 0, e') \geq W(\underline{A}^{*'}, t^{*'}, \tau^*, e')$$

$$I.C.6 \quad W(A^{*'}, t^{*'}, -\tau^*, e') \geq W(\underline{A}^{**}, t^{**}, 0, e')$$

where we have used the fact that the constraints for the poor region are not binding at the solution (as in section 4, this follows from the single crossing condition, (MSC), stated below). Note that, in contrast with section 4, the incentive constraints I.C.5 and I.C.6 now allow for the possibility that when the rich type chooses the allocation designed for the poor, he may also want to modify his choice of tax enforcement. The moral hazard constraints in (5.6) show how this is done optimally.

A single crossing condition analogous to (SC) is needed. As in section 4, we want the indifference curve of the rich region to be everywhere steeper than that of the poor region. But now, this requires:

$$\underline{\text{Modified single crossing:}} \quad - [(g - \tau) H_{gg} (1 + A_e (L+e)/A)] / H_g > H_g - 1 \quad (MSC)$$

The interpretation of this condition is similar to that of (SC). As the productivity parameter e increases, the labour supply function becomes less elastic, calling for higher taxes in the rich region (RHS); on the other hand, as the level of public good increases, its marginal utility is reduced, calling for lower taxes in the rich region (LHS). With moral hazard this second effect is weaker than before: at high tax rates, the rich region potentially allows for more tax evasion (as $A_e \leq 0$ from Lemma 5.2), so that only a share of the larger tax base generates tax revenue and ends up in more public goods. This implies that (MSC) is more restrictive than (SC).

Under condition MSC, the following proposition can be proved:

Proposition 6: In the equilibrium with MH and AS: (a) the properties (i)-(iv) of Proposition 4 still hold; (b) in equilibrium tax evasion can occur in the poor but not in the rich region ($A'' = A'^* = 1$, $A^* \leq 1$, $A^{**} \leq 1$).

Proof: See Appendix

Part (a) of the proposition shows that the qualitative features of the optimal contract are not affected by moral hazard. Whenever the second best cannot be achieved, the region who declares to be rich is left undistorted; the region who declares to be poor continues to be distorted upwards, and redistribution falls below the full equalising level of section 3. Part (b) follows immediately from Lemma 5.1. As the rich type is always undistorted at the optimal allocation, it would never allow for tax evasion.

While the general properties of the contract are not affected, moral hazard changes the equilibrium level of redistribution and the equilibrium tax rates, and therefore affects regional and national welfare. Specifically, let t''_M , t'^*_M , t^*_M , t^{**}_M and τ^*_M denote the solutions to the moral hazard and adverse selection problem in (5.8), and W''_M , W'^*_M , W^*_M , W^{**}_M the corresponding levels of welfare of the different types at the optimal contract. Let $W_M = (W''_M + W'^*_M + W^*_M + W^{**}_M)$ be the total welfare of the federal government. Finally, let the suffix A indicate, as in proposition 4, the corresponding solutions and welfare levels under adverse selection alone. Our results are summarized in the following:

Proposition 7: Under standard regularity conditions on $A(\cdot)$ (see Appendix),

- (i) $\tau^*_M \leq \tau^*_A$, $t''_M = t''_A$, $t'^*_M \leq t'^*_A$;
- (ii) $W''_M = W''_A$, $W'^*_M \geq W'^*_A$;
 $W^*_M \leq W^*_A$, $[W^{**}_M + W^*_M] \leq [W^{**}_A + W^*_A]$.
- (iii) $W_M \leq W_A$

with strict inequalities provided either $\underline{A}^{**} < 1$ or $\underline{A}^{*' } < 1$ (or both).

Proof. See Appendix

Clearly, when $\underline{A}'' = \underline{A}'^* = 1$ (implying that all As are one, from Lemma 5.2) no tax evasion occurs, and the objective function and the constraints of problem (5.8) are identical to the AS problem, and so is the solution. In turn, Part (i) of the proposition says that provided the region who has the positive shock finds it optimal to allow tax evasion when mimicking the poor, at either possible out of equilibrium allocation, then the optimal level

of redistribution falls unambiguously compared to the AS contract. This makes the rich region better off and the poor worse off in expected terms.

More in detail, when both regions happen to be rich, they levy the same tax rate than in AS, while the transfer is always zero from the budget constraint. So their welfare is the same as in AS. Conversely, when one is rich and one poor, the poor is always worse off and the rich better off, because the transfer is reduced. The only ambiguity concerns welfare in the poor region in state e^*, e^* . On the one hand, for a given transfer, the poor region tax rate t^{**} must increase to induce self revelation of the rich (who now can evade taxes), causing a welfare loss. On the other hand, t^{**} must decrease relative to AS, since the lower transfer already makes the rich better off, and thus reduces the need to distort the poor for revelation purposes. It is easy to show, however, that the expected utility of the poor is lower than in AS, while the rich is better off. Finally, part (iii) shows that the poor region loses more, on average, than what is gained by the rich, so that moral hazard reduces national welfare²⁴. This result follows from the fact that, under moral hazard, public good consumption is even more unequal in the two regions than with AS alone.

In order to grasp the intuition for these results, consider the particularly simple case where tax evasion does not occur in the poor region but may be chosen by the rich when mimicking the poor in state e', e' (although not in $e', e^* : \underline{A}'' < I = \underline{A}'^* = A^{**} = A^*$). As we have seen, the indifference curves of the rich type in the upper quadrant become flatter (the dotted curves in Figure 4). As is clear from the Figure (and by inspection of the first IC constraint), now the rich can break the AS contract, point C, and be better off by pretending to be poor. Of course this follows from the fact that the welfare loss due to excessive taxation is reduced by evading taxes, and is more than compensated by the welfare gain from receiving, rather than paying, the transfer. In order to induce the rich type to separate, the poor must be made relatively worse off. Note, however, that in the new optimal contract t'' continues to be undistorted, and, in order to satisfy the budget constraint, must be chosen at the point of tangency between the rich indifference curve and the $\tau=0$ line, point A. This effectively pins down this tax rate together with the rich's level of welfare W'' . It then follows that in order to prevent misreporting of the rich, the transfer τ^* must fall in order to make the poor worse off in absolute terms ($W^{*'} falls).$

²⁴In fact, were this not true, national welfare would increase, which cannot be true since we have added a binding constraint to the original

Moving to the lower quadrant of the figure, this means that when state e',e^* is effectively realised, the region who truthfully declares to be poor now receives a lower transfer from the rich, simply because the contract must prevent the rich from misreporting in state e',e' ! Hence, the region who experiences a positive shock in state e',e^* benefits for two reasons: it pays a lower transfer, and, as a consequence, it chooses a lower tax rate (t'^* falls).

This particular case illustrates the possibility that the poor, although losing in expected terms, may actually benefit from moral hazard in the state of the world where the rich does *not* evade taxes out of equilibrium. Consider what happens, in our example, when e^*,e^* is realised. In this state, the Figure shows that the poor is better off under MH (W^{**} rises). The reason is simply that, as t'^* is reduced by the lower transfer, truth telling can be enforced with a lower tax rate, t^{**} . This example is an illustration of an interesting property of the optimal contract: relative to AS, moral hazard redistributes the poor welfare across states of nature, from the state where tax evasion occurs (state e',e' in our case), to the state where it does not (state e^*,e^*).

Proposition 7 holds irrespective of whether the poor evades taxes at his equilibrium allocation (i.e. $A^{*'} > I$, or $A^{**} < I$). In terms of our last Figure, when also the poor finds it optimal to evade, say in state e',e' ($A^{*'} > I$), also his indifference curves become flatter. However, we know from the previous discussion that, at the same allocation, the rich always evades more ($\underline{A}'' < A^{*'} < I$), so that his indifference curve rotate even more down to the right. This implies that the new intersection between the "dotted" indifference curves must occur at a point where the poor is worse off than in the AS equilibrium.

We have shown that moral hazard does not alter the qualitative properties of the AS contract. In particular, the poor region signals its bad luck by imposing a higher tax rate than would find otherwise optimal. However, tax evasion reduces the effectiveness of fiscal effort as a signal of bad luck. The rich finds it easier to mimic the poor by increasing its tax rate (which is observable) while reducing tax enforcement (which is not). To enforce separation, the transfer must fall, which implies that even less redistribution is possible with MH than with AS alone. This harms the poor and benefits the rich. Quite strikingly, this occurs even though it is the poor and not the rich region which may allow tax evasion in equilibrium.

6. Concluding Remarks

We started with a simple question: how to optimally redistribute among regions with different per capita tax bases. The paper suggests the following answer. Any known exogenous differences in the size of local tax bases should be compensated by equalising interregional lump sum transfers. There are likely to remain unknown or unverifiable regional differences, however, since the relative position of regions is imperfectly observed and changes over time. Furthermore, there is a trade-off between the ability of federal government to commit to a redistributive formula and the number of contingencies this formula can rely upon. To deal with unverifiable differences, an optimal transfer mechanism should then induce truth telling by means of a premium on fiscal effort (i.e. on high tax rates) by regions that claim to be poorer than expected. Since statutory tax rates are observable, and under the assumption that poorer regions value tax revenue more highly, fiscal effort is less costly for a poor region. Hence, fiscal effort is a credible signal of bad luck. High fiscal effort is distorting, however. For this reason, optimal redistribution is incomplete, and the unexpectedly rich regions get some informational rents.

A difficulty with inducing high fiscal effort on the unexpectedly poor regions is that it could discourage local tax enforcement and lead to tax evasion. Allowing for this moral hazard problem does not change the nature of the result: the optimal redistributive mechanism still entails a premium on fiscal effort by the poor regions. The extent of optimal redistribution is further reduced, however, since it becomes easier to lie. An unexpectedly rich region now has a stronger incentive to pretend that it is poor, by pushing up its (observable) tax rate while at the same time reducing its (unobservable) tax enforcement. In equilibrium, tax evasion could be allowed in the unexpectedly poor regions, but never occurs in the unexpectedly rich ones. Yet, the possibility of evading taxes harms the poor and benefits the rich regions, since it leads to smaller interregional redistribution.

The optimal redistributive scheme we derived in this paper bears some similarities with the actual schemes operating in some federal countries. In Canada in particular, the redistributive scheme induces a fiscal effort on the tax bases on which a Province is relatively poor, as predicted by our model.

This paper has focused on incentive constraints that arise from informational asymmetries, abstracting from two other incentive problems that are likely to arise in these situations: mobility of tax bases, and politically motivated policymakers. It would be interesting to know whether mobility of the tax base reinforces or relaxes the informational constraints we discussed in this paper. Furthermore, while the literature (i.e. Keen, 1993) has largely studied problems of regional tax competition, little effort has been made to connect this theme with the optimal redistributive scheme. How should the latter be organised to cope with both informational and tax competition problems?

Politics is also likely to play a large role in the setting up of redistribution schemes (see in particular Persson and Tabellini (1996a)). Indeed, a strong argument, for regional redistribution is that it "levels the playing field" between competing regions, thus allowing citizens to better compare the relative performance of local politicians (Salmon, 1994). As local politicians are likely to enjoy an informative advantage vis-a-vis both the center and their citizens, it could be interesting examining the optimal redistributive mechanism in this context. Extending the analysis to these other issues would be an exciting direction for future research.

Appendix

1. Proof of Proposition 4

"Consider the problem (4.1). Under SC, at most two of the incentive compatibility conditions I.C.1 - I.C.4 can bind at the optimum. If all four were binding, then both types would be just indifferent between truth telling and lying, in all states of nature. But this cannot happen if the indifference curve of the two types only intersect once. Intuition suggests that the incentive constraints that bind in equilibrium are those of the rich region, that may have to pay the transfer, while the poor region is strictly better off under truth telling than if it pretends to be rich. We conjecture that this is indeed the case, and verify at the end of the proof that the conjecture is right.

Thus, ignoring I.C.3 and I.C.4, we can write the Lagrangean of Problem (4.1) as:

$$(A.1) \quad L = W(t', 0, e') + W(t^{**}, 0, e^*) + W(t^*, -\tau^*, e') + W(t^{*'}, \tau^*, e^*) \\ + \lambda [W(t', 0, e') - W(t^{*'}, \tau^*, e')] + \mu [W(t^*, -\tau^*, e') - W(t^{**}, 0, e^*)]$$

where λ and μ are the Lagrange multipliers of I.C.1 and I.C.2 respectively. If both constraints bind, the first order conditions of this optimisation problem are:

$$(A.2) \quad \begin{aligned} a) & (1 + \lambda) W_t(t', 0, e') = 0 \\ b) & (1 + \mu) W_t(t^*, -\tau^*, e') = 0 \\ c) & W_{t^{**}}(t^{**}, 0, e^*) - \mu W_{t^{**}}(t^{**}, 0, e^*) = 0 \\ d) & W_{t^{*'}}(t^{*'}, \tau^*, e^*) - \lambda W_{t^{*'}}(t^{*'}, \tau^*, e^*) = 0 \\ e) & -(1 + \mu) W_{\tau^*}(t^*, -\tau^*, e') + W_{\tau^*}(t^*, -\tau^*, e') - \lambda W_{\tau^*}(t^*, -\tau^*, e') = 0 \end{aligned}$$

Part (i) of the Proposition follows directly from equations (A.2a) and (A.2b).

Part (iii): Note that $W_{\tau^*}(t^*, -\tau^*, e') = H_g(g_A^*)$, and similarly for the other partial derivatives in (A.2e). If either λ or μ are positive (that is, if at least one incentive constraint binds), then, by concavity of $H(g)$, equation (A.2e) implies $g_A^* > g_A^*$.

Part (ii): Equations (A.2d) and (A.2c) directly imply that, if the corresponding Lagrange multiplier is positive, then t^{**} and $t^{*'}$ must be distorted. The reason is that, since $e' > e^*$, the two partial derivatives in the left hand sides of (A.2d) and of (A.2c) cannot both be zero at the same values of t and τ . Moreover, by (A.2d) and (A.2c), these two tax rates must be distorted in the same direction for both regions: for the poor that tells the truth, and for the rich that pretends to be poor. It remains to prove that the distortion is upwards, as in Figure

3. But this follows from the SC condition, which says that at the same allocation the rich region indifference curve is always steeper than that of the poor. To see why, suppose by contradiction that at the poor equilibrium allocation, the rich and the poor indifference curves intersect on the downward sloping segment. By SC, the equilibrium allocation of the poor region would have to be drawn as in Figure A.1. But then, I.C.3 and I.C.4 would be violated, since the poor region would be better off by mimicking the rich. Hence, in both states of the world, at the poor equilibrium allocation the indifference curves of the two types must intersect in the upwards sloping part. Hence, the poor region tax rate is distorted upwards, as in Figure 3. This argument also immediately implies part (iv) of the proposition.

It remains to prove that our initial conjecture is right, namely that I.C.3 and I.C.4 do not bind in equilibrium. Suppose by contradiction that this was the case. Recall our initial argument, that I.C.1 and I.C.2 could not also bind in this case. Repeating the same steps as above, it would then be possible to show that, at the poor equilibrium allocation, the tax rate for the poor would be undistorted. By mimicking the rich, the poor would then have to pay a transfer (or give up a transfer it receives, depending on the state of nature), and on top of that also distort its tax rate (in this case downwards, by SC). But this would result in a net welfare loss, contradicting the initial conjecture that the poor was indifferent between mimicking or accepting the equilibrium allocation." QED

Proof of Lemma 5.1

"Consider the following, general, problem. There are n individuals of "type 1" and $(1-n)$ of "type 2", where $0 < n < 1$, who differ only to the extent that they may be taxed at different tax rates t_1 and t_2 , and who enjoy the same level of public good. Optimal tax rates must solve

$$\begin{aligned} & \text{Max}_{t_1, t_2} \{ (1-t_1) [L(t_1) + e] + V(1-L(t_1)) \} + (1-n) \{ (1-t_2) [L(t_2) + e] + V(1-L(t_2)) \} + \\ & + H \{ n t_1 [L(t_1) + e] + (1-n) t_2 [L(t_2) + e] + \tau \end{aligned}$$

Using the Envelope theorem, the first order conditions for t_1, t_2 are

$$-n [L(t_1) + e] + n H_g(.) [L(t_1) + e + t_1 L_t(t_1)] = 0$$

$$-(1-n) [L(t_2) + e] + (1-n) H_g(.) [L(t_2) + e + t_2 L_t(t_2)] = 0$$

which clearly imply that for all n strictly between zero and one, $t_1 = t_2$ at the optimum.

This result rules out $A < 1$, because that would imply $t_1 > 0$, $t_2 = 0$ (or viceversa) for n strictly between zero and one, contradicting our result. This proves $A=1$ at $t = \text{argmax } W(.)$. It remains to be proved that $A=1$ for $t < \text{argmax } W(.)$, and $A < 1$ is possible for $t > \text{argmax } W(.)$. Both statements follows immediately from the fact that $W_{ta} < 0$. See lemma 5.2. "QED

Proof of Lemma 5.2

"Differentiating equation (5.4) with respect to t , e and τ respectively, using equation (5.1) and the consumer's first-order condition for the choice of l , we get, after some simplifications

$$A.3 \quad W_{at} = l \{ (H_{gg} [g - \tau] + H_g - 1)(1 - \varepsilon(t, e)) - \varepsilon(t, e) \}$$

$$A.4 \quad W_{ae} = t \{ (H_{gg} [g - \tau] + H_g - 1) \}$$

$$A.5 \quad W_{a\tau} = t l H_{gg}$$

A.5 is certainly negative due to the concavity of $H(.)$; A.4 is negative because of assumption S.C; and A.3 is negative because of assumption S.C. and $0 < \varepsilon(t, e) < 1$. As $W_{aa} < 0$ by assumption, the signs of A.3, A.4, A.5 entail Lemma 5.2" QED

Proof of Proposition 6

"Part (a). The first order conditions for the problem in (5.7) are

$$(1 + \lambda) W_{t''} T(A'', t'', 0, e') = 0$$

$$W_{t'^*} T(A'^*, t'^*, -\tau^*, e') (1 + \mu) = 0$$

$$W_{t^*} T(A^*, t^*, \tau^*, e^*) = \lambda W_{t^*} T(\underline{A}^*, t^*, \tau^*, e')$$

$$W_{t^{**}} T(A^{**}, t^{**}, 0, e^*) = \mu W_{t^{**}} T(\underline{A}^{**}, t^{**}, 0, e')$$

$$- W_{\tau^*} T(A'^*, t'^*, -\tau^*, e') + W_{\tau^*} T(A^*, t^*, \tau^*, e^*) - \lambda W_{\tau^*} T(\underline{A}^*, t^*, \tau^*, e') -$$

$$- \mu W_{\tau^{**}} T(A'^*, t'^*, -\tau^*, e') = 0$$

where the total derivative wrt the generic variable x is defined as

$$W_x^T \equiv W_x(A, t, \tau, e) + W_d(A, t, \tau, e) A_x(t, \tau, e) \text{ for } x = t, \tau.$$

From a straightforward application of the envelope theorem, $W_x^T = W_x(A, t, \tau, e)$ so that the first order conditions become formally identical to those under AS (see equation A.2 above). Therefore the properties (i)-(vi) are derived exactly as under AS. Part (b) follows because t'' and t'^* are undistorted (cf. Lemma 5.1)". *QED.*

Proof of Proposition 7

"The strategy of the proof is as follows. We first simplify the federal government's problem, using the results of Proposition 6 to focus on the optimal choice for τ^* . We then show that if one of the moral hazard constraints for the rich mimicking the poor binds, then τ^* is lower under MH and AS than under AS alone. Finally, we derive the other results of proposition 7 from this fact. We start by imposing the following condition, derived by substituting for (5.6) in (5.1) and by differentiating (5.1) with respect to t :

$$(R) \quad dg / dt = \frac{(g - \tau)}{t} (-\varepsilon(t, e) - \sigma(t, e, \tau)) \quad \forall t, e, \tau$$

where $\sigma(t, e, \tau) = -t(dA/dt)/A$ is the elasticity of the optimal tax enforcement parameter with respect to a change in t . Condition (R) is a simple regularity assumption; it implies that even allowing for tax evasion, g is always an increasing function of t .

Now let $T'(\tau)$ be that function which gives the optimal tax rate for the rich region as a function of τ (i.e. $T'(\tau) = \text{argmax}_{t'} W(A(t', \tau, e'), t', \tau, e')$). From Proposition 6 we know $t'' = T'(0)$ and $t'^* = T'(-\tau^*)$, and $A(T'(0), 0, e') =$

$A(T'(-\tau^*), -\tau^*, e') = I$. Furthermore, note that by substituting for $t'' = T'(0)$ and $t'^* = T'(-\tau^*)$ in , respectively , constraints I.C.5 and I.C.6 we can get $t^{*'}$ and t^{**} as a function of the transfer (assuming both constraints bind at the optimum and recalling that to enforce separation $t^{*'} > t''$ and $t'^* > t^{**}$ at the optimum). Let $t^{*'} = T^{*'}(\underline{A}^*, \tau^*, e')$ and $t^{**} = T^{**}(\underline{A}^{**}, -\tau^*, e')$ indicate these two functions. By implicitly differentiating I.C.5 and I.C.6 with respect to τ^* and using the envelope theorem we find the slope of these functions to be $dT^{*'} / d\tau^* = (MRS(\underline{A}^*, t^{*'}, \tau^*, e'))^{-1} > 0$

and $dT^{**} / d\tau^* = - (W_\tau(A^*, T'(\tau^*), -\tau^*, e') / W_t(\underline{A}^{**}, t^{**}, 0, e')) > 0$, where

$MRS = -W_t / W_\tau$ is the slope indifference curves of the rich type. From $A_t \leq 0$, see Lemma 5.2, it follows that

$T^{*'}(\underline{A}^{*'}, \tau^*, e') \geq T^{*'}(1, \tau^*, e')$ and $T^{**}(\underline{A}^{**}, -\tau^*, e') \geq T^{**}(1, -\tau^*, e')$, with inequalities which become strict if $\underline{A}^{*'} < 1$ and $\underline{A}^{**} < 1$; if the rich evades at the allocations of the poor, a larger tax rate must be imposed on the poor to allow for separation.

Using the functions defined above, and Proposition 6, it is readily seen that the optimal transfer which solves problem (5.7), also solves the following problem:

(A.6)

$$\text{Max } \Omega(\tau^*) = W(1, T'(0), 0, e') + W(1, T'(-\tau^*), \tau^*, e') + W(A^{**}, T^{**}(\underline{A}^{**}, -\tau^*, e'), 0, e^*) + W(A^{*'}, T^{*'}(\underline{A}^{*'}, \tau^*, e'), \tau^*, e^*)$$

Upon manipulation, the F.O. condition for this problem can be written as:

$$\Omega_{\tau^*}(\tau^*) = 0 =$$

$$= -H_g(1, t^{*'}, -\tau^*, e') \left\{ 1 + \frac{W_t(A^{**}, t^{**}, 0, e^*)}{W_t(\underline{A}^{**}, t^{**}, 0, e')} \right\} + H_g(A^{*'}, t^{*'}, \tau^*, e^*) \left\{ 1 - \frac{MRS(A^{*'}, t^{*'}, \tau^*, e^*)}{MRS(\underline{A}^{*'}, t^{*'}, \tau^*, e')} \right\}$$

where $\Omega_{\tau\tau} < 0$ to satisfy the Second Order conditions. Suppose now $\underline{A}^{**} = \underline{A}^{*'} = 1$, which also implies by lemma 5.2. Then, it is clear that τ_A^* , the optimal transfer under AS alone, solves problem (A.6) as the objective function and the constraints are the same as in Problem 4.1. Suppose now instead that either $\underline{A}^{**} < 1$ or $\underline{A}^{*'} < 1$ or both, (in which case A^{**} and/or $A^{*'} < 1$ is also possible) and evaluate (A.7) at τ_A^* . Then compare the expression so obtained with the case with no tax evasion. Going through each element in equation (A.7) in turn, we see that, in *all* possible cases: 1) the first element to the LHS is unchanged, as t^{*} is only a function of τ^* ; 2) the second element, in curly brackets, is at least as large because $\underline{A}^{**} \geq A^{**}$ (see lemma 5.2); 3) the third element is unchanged or smaller, because t^{*} is now equal or larger and condition (R) above holds; and (4) the last element in curly brackets is also equal or smaller because of lemma 5.2 (i.e the ratio of the two MRS's is now at least as large). We then conclude that if either $\underline{A}^{**} < 1$ or $\underline{A}^{*'} < 1$ or both, then $\Omega_{\tau}(\tau_A^*) < 0$. From $\Omega_{\tau\tau} < 0$, this implies that $\tau_M^* \leq \tau_A^*$ with the inequality which becomes strict if either $\underline{A}^{**} < 1$ or $\underline{A}^{*'} < 1$ or both. This, together with the properties of function $T'(\cdot)$ prove part (i) in the proposition.

To prove part (ii), note first that as the rich is undistorted at this allocation, $\tau_M^* \leq \tau_A^*$ immediately implies $W_M^* \geq W_A^*$, and $t_A^* = t_M^*$ implies $W_A^* = W_M^*$. Second, note from Lemma 5.2, that the following is true:

$$\begin{aligned}
W_A^{*'} &\equiv W(1, T^{*'}(1, \tau_A^*, e'), \tau_A^*, e^*) \geq W(A^{*'}, T^{*'}(\underline{A}^{*'}, \tau_A^*, e'), \tau_A^*, e^*) \geq \\
&\geq W(A^{*'}, T^{*'}(\underline{A}^{*'}, \tau_M^*, e'), \tau_M^*, e^*) \equiv W_M^{*'}
\end{aligned}$$

Part (iii) follows because the objective function is unchanged with respect to the case with AS alone, while the constraints are stricter. Finally note that Part (iii) together with $W_A^{*'} + W_A^{**} \geq W_M^{*'} + W_M^{**}$ entails $W_M^{*'} + W_M^{**} \leq W_A^{*'} + W_A^{**}$. This concludes the proof. QED"

Table 1

Local Tax Rates in Canada

Province	Average Family Income Pre-tax and Transfers* (1993)	Provincial Sales Tax Rate (1994)
Ontario	51.5	8
Alberta	50.9	0
British Columbia	49.6	7
Manitoba	43.9	7
Saskatchewan	41.3	9
Quebec	40.3	8
Nova Scotia	38.5	11
New Brunswick	38.1	11
Prince Edward Island	34.6	10
Newfoundland	31.3	12

* x 1000 dollars

Sources: Statistics Canada 1995, Table VI; Mennel and Foerster (1995)

References

- Bird, R.M. and Slack, E. 1990 "Equalization: the representative tax system revised", Canadian Tax Journal, 38, 913-927
- Boadway R. and Flatters F. 1982 "Efficiency and equalization payments in a federal system of government" Canadian Journal of Economics 15, 121-143
- Boadway R., Horiba I. and Raghbendra J. 1995 "The design of conditional grants as a principal agent problem" Paper presented at the IIPF meeting, Lisboa, August 1995
- Bucovetsky S., Marchand M. and Pestieu P. (1996) "Tax Competition and Revelation of Preferences for Expenditures", mimeo
- Cremer H., Marchand M, and Pestieu P. 1994 "International redistribution through tax surcharge" Core Discussion Paper, 9469
- Fudenberg H. and Tirole J. 1991 Game theory MIT Press
- Keen M. 1993 "Structure of fiscal and social charges according to their degree of mobility" *mimeo* Brussels, EU Commission
- Laffont J.J. 1995 "Incentives in China's federal tax system" *mimeo* Essex University
- Levaggi R. 1991 Fiscal federalism and grants in aid Aldershot: Avebury
- Lockwood, Ben (1996) "Inter-regional Insurance with Asymmetric Information", mimeo
- Mirrlees J.A. 1971 "An exploration in the theory of optimal income taxation" Review of Economic Studies 38, 175-208
- Mennel, A. and J. Foerster (1995) Steuern in Europa, USA, Kanada und Japan, Verlag Neue Wirtschaftsbrieft, Berlin
- Oates, W.E. 1972 Fiscal federalism New York: Harcourt Brace Jovanovich
- Persson T. and Tabellini G. 1996a "Federal fiscal constitutions: risk sharing and redistribution" *forthcoming*, Journal of Political Economy
- Persson T. and Tabellini G. 1996b "Federal Fiscal Constitutions: risk sharing and redistribution", Econometrica, May
- Raff H. and Wilson J.D. 1995 "Income redistribution with well informed local governments" *mimeo* Indiana University
- Rosen H.S. (ed) 1986 Studies in State and Local Public Finance, Chicago U.P, Chicago

Salmon, P. 1987 "Decentralization as an incentive scheme" Oxford Review of Economic Policy 3, 24-43

Shah, A. 1994 "The practice of federalism in four industrial countries" in The reform of intergovernmental fiscal relations in developing and emerging market economies, The world Bank, Washington D.C.

Statistics Canada 1995 Income after tax: distribution by size in Canada, 1993, Ottawa, Minister of Supply and Services

Stiglitz, J.E. 1982 "Self selection and Pareto-efficient Taxation" Journal of Public Economics 17, 213-240

Tirole, J. 1994 "The internal organisation of government" Oxford Economic Papers

Wildasin D.E. 1991 "Income redistribution in a Common Market" American Economics Review 81, 757-774