FUNDACION BBV

FINANCIAL INTERMEDIATION AND THE OPTIMAL TAX SYSTEM

Ramón Caminal
Institut d'Anàlisi Econòmica, CSIC
Universitat Autònoma de Barcelona

ESTUDIOS BANCARIOS

FINANCIAL INTERMEDIATION AND THE OPTIMAL TAX SYSTEM

Ramón Caminal
Institut d'Anàlisi Econòmica, CSIC
Universitat Autònoma de Barcelona

Centro de Estudios Bancarios

Director: **Luis Angel Lerena,** catedrático de Economía Internacional de la Universidad Complutense de Madrid.

Director del Programa de Investigación: **Xavier Vives.**

© FUNDACION BBV
DOCUMENTA
Plaza de San Nicolás, 4
48005 BILBAO

D. L.: BI-2044-93

RAMON CAMINAL

Nace en la Seu d'Urgell en 1955. Licenciado en Ciencias Económicas (1980) por la Universidad Autónoma de Barcelona. Realizó estudios de postgrado en la misma Universidad Autónoma de Barcelona y en la Universidad de Harvard, donde obtuvo el título de doctor en Economía en el año 1987.

Ha sido profesor Titular de la Universidad Autónoma de Barcelona y *Visiting Scholar* de la Universidad de Harvard. En la actualidad es Colaborador Científico del Instituto de Análisis Económico (CSIC).

Su investigación se inscribe fundamentalmente en las áreas de macroeconomía y organización industrial. Ha publicado, entre otras, en las siguientes revistas: *Journal of Industrial Economics, Economica, International Journal of Industrial Organization, y Journal of International Money and Finance.*

Ha sido evaluador de diversas revistas científicas nacionales e internacionales, miembro del Consejo Editorial de las revistas Moneda y Crédito y Revista Española de Economía, y consultor de la Comunidad Económica Europea.

Los Centros Permanentes de Reflexión de la Fundación Banco Bilbao Vizcaya abordan, desde una perspectiva multidisciplinar, áreas específicas de actualidad. En cada una de estas áreas se incluyen proyectos de investigación propios, a partir de los cuales se desarrolla una actividad de encuentros periódicos, generalmente en la modalidad de seminarios y conferencias anuales.

Aspiran estos Centros a que la sociedad vea en ellos puntos de referencia de calidad, en los estudios y debates de los temas encuadrados dentro de cada área.

El Centro de Estudios Bancarios de la Fundación BBV, está interesado en el estudio e investigación de los nuevos escenarios y herramientas que conformarán el mundo económico y financiero de los próximos años. Este Centro pretende convertirse en punto de referencia en el campo de los estudios financieros, favoreciendo la investigación sobre estos mercados y difundiendo sus resultados.

ABSTRACT

In this paper a stylized general equilibrium model is constructed to analyze the relative efficiency of taxing financial intermediaries. A crucial feature of the model is that tax collection costs are endogenous, that is, they result from primitive assumpyions about information and transaction costs, instead of being assumed *ad hoc.* The model provides useful insights on the welfare costs and incidence of banks' reserve requirements or, equivalently, of a tax on deposits. In particular, it is shown that a tax on bank deposits can be part of an optimal tax system, provided banks' monopoly power is significant, However, if the banking industry is perfectly competitive, the first dollar of revenue is more efficiently collected by a uniform capital income tax, although strictly positive welfare loss is incurred.

FINANCIAL INTERMEDIATION AND THE OPTIMAL TAX SYSTEM

Ramon Caminal

I. Introduction

The theory of optimal taxation deals with the problem of determining the tax structure that minimizes the efficiency costs of collecting a certain amount of revenue. Although it is clear that this line of research has provided useful insights, its practical relevance is sometimes questioned on the ground that some of the tax schemes proposed are simply unfeasible or, more generally, because the theory has overlooked the significance of different tax collection costs for alternative tax categories [1].

A common answer to such criticism has been either to rule out some types of taxes or, similarly, to assume some ad hoc collection cost function for each type of tax. The problem with such an approach is that the same factors that explain the existence of different tax collection costs can also influence private behavior, and therefore modify the incidence and efficiency losses of alternative tax schemes [2].

Consequently, some progress can be made not only by considering explicitly collection costs, but also by analyzing optimal taxation in models where these costs are endogenous. That is, the government faces different tax collection costs due to primitive assumptions about the economy, and these assumptions also affect the behavior of private agents.

This paper suggests the feasibility of this approach. We examine a particular problem in the theory of optimal taxation, where the costs of information gathering are likely to play a relevant role, and examine the effect of this informational problem on the equilibrium behavior of the private sector as well as on the optimal tax structure.

More specifically, this paper attempts to shed some light on the issue of whether taxing financial intermediaries can be part of an optimal tax scheme or, on the contrary, it is likely to be dominated by alternative taxes. This requires us to consider explicitly the particular role of financial intermediaries in the allocation of savings. Thus, the issues involved are somewhat different than those considered in standard models of optimal taxation.

In most economies, the most important tax levied on financial intermediaries are reserve requirements. Reserve requirements are an implicit form of taxation since financial intermediaries must hold a fraction of their deposits in the form of non-interest-bearing reserves. Thus, the implicit tax on deposits equals the required reserve ratio times the rate of return on interest-bearing assets (Fama, 1980). At the same time, such a requirement changes the relative demand for alternative assets (reserves versus government bonds and alternative investment opportunities) [3].

[1] See Slemrod (1990) for a discussion of these issues, and references cited there.

[2] Tax collections costs may be purely admistrative; in this case this criticism does not apply. However, in most cases tax collection costs are associated with the informational requirements for implementing the tax scheme. Clearly, informational asymmetries are often crucial to explain the characteristics of both markets and non-market institutions.

[3] In other words, reserve requirements operate as a tax by expanding the base for the inflation tax.

Therefore, a complete characterization of reserve requirements must involve a flow and a stock effect. More precisely, under certain assumptions, a reserve requirement is equivalent to a proportional tax on deposits plus an open market sale of government bonds of an amount equivalent to the volume of resources kept captive by the requirement (Romer, 1985; and Bacchetta and Caminal, 1992a).

In principle, it could be argued that the main purpose of banks' reserve requirements may not be to raise revenue but to facilitate monetary control or prevent bank runs. Three observations suggest that, in most cases, revenue-raising is the most important, or perhaps the only, rationale for existing reserve requirements. First, in many countries required reserve ratios are very high. It is difficult to argue that a reserve ratio of an order of magnitude of 20% or 30% is needed to stabilize any monetary aggregate or to compensate for the suboptimally low level of voluntary reserves [4]. Second, required reserve ratios are positively correlated with inflation rates, which is compatible with the minimization of the welfare costs of inflationary finance (Brock, 1989). Third, a negative correlation seems to hold between reserve ratios and GDP per capita, which suggests that those countries with less developed tax collection systems must rely more heavily on such form of taxation [5].

In informal discussions about the potential optimality of reserve requirements, two alternative arguments are often implicitly or explicitly used. The first one emphasizes the disintermediation effect caused by the tax, which is viewed as a negative aspect of the use of reserve ratios. The second is the relatively low collection cost in comparison with alternative taxes, which obviously militates in favor of reserve requirements.

With respect to the first argument, it is well known that taxing intermediate goods is usually dominated by taxing final goods (Diamond and Mirreless, 1971). In the case of financial intermediaries, the intuition behind this result is clear. Presumably, these intermediaries exist because they perform an efficiency-enhancing role in the allocation of savings. Thus, taxing their activity distorts an additional margin. It induces some savers to avoid intermediaries and invest their funds directly (the disintermediation effect). Therefore, in principle, taxing financial intermediaries should be dominated by taxing final investment projects [6].

Turning to the second argument, at least under certain circumstances, tax collection costs may be substantially higher for individual production units than for large financial intermediaries (the tax collection costs effect) [7]. This is how reserve requirements are usually justified in countries with less developed financial and taxation systems.

In this paper, we construct a stylized general equilibrium model with particular features which introduce the trade-off discussed above. First, entrepreneurs can choose to finance their investment projects directly from savers (issuing bonds) or through financial intermediaries (applying for a bank loan); thus, some forms of taxation may cause a disintermediation effect. Second, banks can monitor certain actions of firms which affect the probability distribution of output; since entrepreneurs are subject to a moral hazard problem, this monitoring enhances welfare. In other words, disintermediation is welfare decreasing. Third, the government faces different collection costs on different taxes but these are not assumed exogenously. The same assumptions about endowments, technology and information that endogenously give rise to a certain financial structure (debt contracts, banks, direct investment) and equilibrium behavior explain why the government faces different tax collection costs.

[4] In some countries, there seems to be a recent trend towards a reduction of required reserve ratios. However, it is likely that this reduction is an undesired consequence of the deregulation of the financial system and the liberalization of capital movements (Bacchetta and Caminal, 1992b).

[5] Additionally, it is often argued that a reserve requirement is likely to be a redundant instrument of stabilization policy. See, for instance, Horrigan (1988).

[6] A particular application of this principle is contained in Kimbrough (1989), which is the only attempt we know to formally analyze the potential optimality of reserve requirements. He assumes that bank deposits (as well as cash) reduce shopping time, and hence they are an argument of the production function. In his framework, the reserve requirement falls exclusively on depositors, and therefore it is a tax on an intermediate good and dominated by a tax on final consumption.

[7] In the language of Slemrod (1990), when tax collection costs are considered the optimal taxation problem becomes an optimal tax system problem.

We assume that all agents are risk-neutral, derive utility only from final consumption, and are ex ante identical. Thus, ex-ante expected utility maximization is equivalent to maximizing aggregate consumption. Ex post, agents are heterogeneous and output is a function of two inputs: capital and entrepreneurial activity.

A key ingredient of the model is the asymmetry of information on the ex post realization of output. The entrepreneur costlessly observes this but all other agents (including the government) must pay a fixed cost to monitor output. Like in most of the literature on optimal financial contracting under costly state verification (Townsend, 1979; Diamond, 1984; Gale and Hellwig, 1985) the optimal arrangement between borrowers and lenders is a standard debt contract: the borrower pays a fixed amount and output is not monitored, unless the borrower defaults in which case the lender pays the monitoring cost and collects all the output [8].

Since entrepreneurs are ex post heterogeneous, the moral hazard problem affects them differently. In equilibrium, the entrepreneurs with a low probability of bankruptcy choose to borrow directly from savers (to issue bonds), while the more risky entrepreneurs choose to apply for a bank loan (and let their actions monitored).

An interesting property of the model is that, even when the bond and loan markets are perfectly competitive and there is no government intervention, the equilibrium is ex-ante inefficient. The reason is that precommitting to ex-post monitoring makes private contracts incentive-compatible but reduces aggregate consumption. In other words, the interest rate can be thought of as playin g two simultaneous roles. On the one hand, it is the relative price that guides the allocation of resources. On the other hand, it determines the frequency of bankruptcies. With respect to the second role, competitive interest rates are inefficiently high.

We show that, with a competitive banking sector, the most efficient way of collecting a small amount of revenue is through a uniform capital income tax. A tax on output (that is on the return of final investment projects) creates no allocative distortion, because it falls proportionally on both inputs (capital and entrepreneurial activity), but involves a fixed tax collection cost. Alternatively, a uniform capital income tax distorts the allocation of resources and moreover raises the interest rate, which increases the frequency of bankruptcies. The first effect vanishes for the first dollar collected but the second is strictly negative even for negligible tax rates [9]. Finally, a tax on bank deposits [10] has similar negative effects than the uniform capital income tax and, additionally, it induces disintermediation [11].

Summarizing, with a competitive banking sector, endogenous tax collection costs may explain why final investment should not be taxed, but still the model argues in favor of a uniform capital income tax and against taxing financial intermediaries.

Next, we ask whether the competitive structure of the banking industry plays any role in the determination of the optimal tax system. The answer is affirmative. In the extreme case in which banks are local monopolists, they charge entrepreneurs their reservation price (the rate they would obtain in the bond market). Thus, a tax on bank deposits can not be passed on to the loan rates, and is paid exclusively by the bank. The only effect of this tax is now the disintermediation effect, but the welfare cost of this effect is zero for the first dollar collected, and therefore in this case the most efficient tax scheme involves a positive tax on bank deposits.

[8] This is precisely the origin of the moral hazard problem. The limited liability aspect of the contract induces the entrepreneur to act as a risk-lover and choose an action that involves an inefficiently high probability of bankruptcy.

[9] This is precisely where the inefficiency of the competitive equilibrium without government intervention matters.

[10] For simplicity, we consider a direct tax on deposits instead of a reserve requirement. It is unlikely that we learn much here by worrying about the composition of government's liabilities.

[11] One of the results of the analysis is that the reserve requirement is effectively paid by both depositors and loan applicants.

Clearly, the real world lies somewhere in between these two extreme market structures. Thus, the conjecture is that in a more general model of banking (imperfect) competition the reserve requirement will fall partially on banks, and partially on depositors and loan applicants. But, as long as banks' monopoly power is large enough, the optimal tax system is likely to involve a positive tax on bank deposits.

In fact, empirical evidence available on the incidence of reserve requirements seems to corroborate this view. For instance, Osborne and Zaher (1992) [12] find that the stock prices of large banks change with announcements of changes on the reserve requirement, which is consistent with the hypothesis that the banks bear part of the implicit tax. In addition, the authors report further evidence suggesting that demand depositors and borrowers bear part of the tax as well.

The paper is organized as follows. In the next section we describe the model. Section 3 characterizes the planner's problem. The properties of the competitive equilibrium are analyzed in Section 4. The following section deals with the effect of linear taxes with a competitive banking sector. Section 6 cares about the effect of monopoly power in banking. Finally, some concluding remarks close the paper.

2. A simple general equilibrium model with financial intermediation

In order to discuss the relative optimality of taxing banks we need a general equilibrium model with an explicit role for these financial intermediaries. However, to make the problem tractable we must neglect many issues which are usually the main focus of the theory of optimal taxation. In this respect, our model differs from the standard ones.

The model is one-shot. There are three goods: labor, an intermediate good (capital) and a consumption good. There is a continuum of agents, indexed by i. All individuals are ex-ante identical. They derive utility only from the consumption good and are risk neutral. Each individual i is endowed with one unit of labor and has potential access to an investment project that produces y_i units of the consumption good from one unit of the intermediate good and one unit of own labor (applied to a specific entrepreneurial activity). Everyone also has free access to a constant returns to scale technology that produces one unit of the intermediate good per unit of labor.

The return on the individual i's project, y_i, is the sum of two variables:

$$y_i = w_i + x_i \qquad (1)$$

The timing is the following (See Figure 1). At the beginning of period 0, a signal of w_i, s_i, is publicly observed. At this point each agent has to choose an occupation: either to employ the whole labor endowment in the intermediate good sector (become a worker) or to become an entrepreneur. Entrepreneurial activity takes place in two stages. In the first stage, the agent uses half unit of labor in developing the investment project and, as a result, w_i becomes public information. At the end of this stage, the entrepreneur may abandon the project and use the remaining half unit of labor in producing the intermediate good. In the second stage, the entrepreneur uses the second half unit of labor and chooses an action a_i that influences the distribution of x_i. Next, investment projects are funded or not, i.e. the intermediate good is allocated among the investment projects, production takes place and the random variables x_i are realized.

In our simple world the only potential exchanges are of the intermediate good for claims to final output (we call this credit). On the supply side, are the workers and on the demand side the entrepreneurs. The occupational choice implies that agents select which side of the credit market to be on. On the one hand, this reflects a real world situation: when the size of available investment projects exceeds personal funds, individuals must choose whether to borrow and directly manage the investment project or to lend to other potential entrepreneurs. On the other hand, our modeling strategy is an attempt to illustrate the distortions caused by capital income taxation in a way

[12] See also other references cited in that paper.

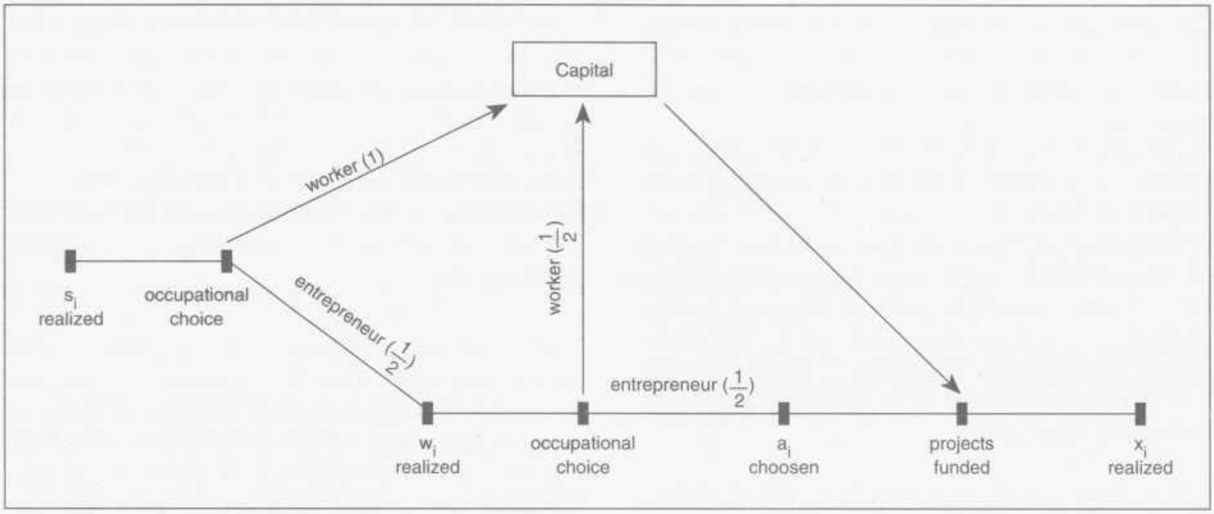


Figure 1

that allows us to carry out efficiency assessments without worrying about interpersonal comparisons or alternative factor income tax policies [13].

The random variables w_i and s_i are jointly distributed according to the density function $f(w, s)$ on $[0,1] \times [0,1]$, and x_i is distributed according to $h(x/a)$ on $[0,1]$. The action a_i can take two values: a_0 and a_1.

Let capital letters denote distribution functions. We make the following assumptions:

A.I. Positive density in support, i.e. $f(w,s) > 0$ for all values of w and s in $[0,1]$, and $h(x/a) > 0$ for all x in $[0,1]$ and a_j, $j = 0,1$.

A.II. The signals s order the distributions of w according to first order stochastic dominance; in particular $\frac{dF(w|s)}{ds} < 0$ for all s.

A.III. The actions a_j order the distributions of x according to second order stochastic dominance; in particular: $E(x|a_0) = E(x|a_1) = \frac{1}{2}$ and $h(x|a)$ are symmetric around $\frac{1}{2}$.

$$\int_0^z H(x|a_1)\, dx < \int_0^z H(x|a_0)\, dx \text{ for all } z \in (0,1).$$

Since all random variables are independent and there is a continuum of agents, there is no aggregate uncertainty.

Assumption II implies that the higher the value of s the higher the conditional expected value of w [14]. Assumption III implies that with action a_1, x is distributed with lower dispersion than with action a_0 around

[13] The model could be further simplified by assuming that the occupational choice is made at a single decision node. Footnotes 22 and 24 discuss the costs paid by such a change.

[14] Integrating by parts, we can write the conditional expected value of w as:

$$E(w|s) = 1 - \int_0^1 F(w|s)\, dw$$

Thus,

$$\frac{dE(w|s)}{ds} = - \int_0^1 \frac{dF(w|s)}{ds}\, dw > 0$$

the same mean. In particular, with a_1 the frequency of low outcomes is lower than with a_0. Assumptions I plays a minor technical role in the analysis.

The structure of information is the following. The realizations of s_i and w_i are public information, but the action a_i is chosen by entrepreneur i and is only observable by outsiders by using an ex-ante monitoring technology. Such a technology requires spending C units of the intermediate good in observing a_i. The realization of x_i is costlessly observed by the entrepreneur, and outsiders can monitor it only if an amount D of the consumption go od is paid (using an ex-post monitoring technology).

In order to make ex-ante monitoring an economically relevant possibility, it must be the case that the ex-ante monitoring cost C is relatively low with respect to the information obtained; i.e. with respect to the difference between the two conditional distributions of x_i. For simplicity, we simply assume that C is arbitrarily low. Finally, those institutions with the ability of using such an ex-ante monitoring technology will be called banks [15].

3. The planner's problem

In this section we characterize the optimal allocation. As usually the goal is simply to have a benchmark to compare the market solution and asses the distortions caused by taxes.

The first problem is to choose a notion of optimality. Given the structure of the model, the obvious social welfare criterion is the maximization of ex-ante (before learning s_i) expected utility of the representative agent. With risk neutrality, this is equivalent to maximizing aggregate consumption, i.e. the distribution of consumption among individuals with different (ex-post) characteristics is irrelevant.

A feasible allocation must satisfy three aggregate conditions: first, the amount of intermediate good produced can not be lower than the total amount of investment (the amount of intermediate good used in the production of the consumption good); second, aggregate consumption can not be higher than total output minus ex-post monitoring costs, and third, total output is the sum of the output of completed investment projects.

Given the information asymmetries, a feasible allocation must also satisfy the incentive compatibility constraints. Thus, characterizing the optimal allocation is a problem of optimal mechanism design. However, we will see immediately that, provided the planner does not face ex-post participation constraints, the incentive problems can be trivially solved [16].

The planner's instruments are the following. For each agent i:

a) the occupational choice as a function of s_i,

b) if the project was developed in the first stage, whether to continue developing the project and allocate one unit of the intermediate good to it, or abandon the project and produce half unit of the intermediate good with the remaining labor, as a function of w_i,

c) if the project is to be carried out to the second stage, choose whether to monitor a_i or not, as a function of w_i,

d) if the project is to be carried out to the second stage, choose whether to monitor (ex-post) the realization of x_i, in principle as a function of w_i, a_i (if ex-ante monitoring has taken place) and perhaps as a function of what the agent voluntarily declares,

[15] In fact, we assume that only banks, but not individuals, have acces to the ex-ante monitoring technology. This could be justified by the existence of sunk costs in monitoring. In a similar model to the one we analyze in this paper, where the size of investment projects exceeds the personal funds of a single individual, even if everyone has acces to the monitoring technology, it would be optimal to delegate the monitoring to financial intermediaries. See Diamond (1984).

[16] In particular, the main difference between the planner's problem and the decentralized market solution arises from the fact that the planner is assumed to have the power to force a particular agent to allocate her unit of labor in the production of the intermediate good, and give the unit of the intermediate good to an entrepreneur, with no compensation for the worker. In a market economy, the worker must be compensated with a share of output (the interest rate).

e) the level of consumption, in principle as a function of s_i, w_i, a_i (if ex- ante monitored), x_i (if ex-post monitored), and perhaps what the entrepreneur has declared.

In fact the social planner's problem can be simplified after the following observations that partially characterize the optimal plan.

Observation 1:

Monitoring a project either ex-ante or ex-post is never part of an optimal plan.

Since ex-post monitoring is costly, it can only be socially valuable if it solves an incentive problem: either because the planner needs to learn the realization of output, or because it provides incentives to the entrepreneur to take the right action. However, the planner maximizes aggregate consumption and hence does not need to redistribute output ex-post. Moreover, if projects are not monitored ex-post, the choice of action is completely irrelevant (by assumption A.III, actions do not affect the expected level of output), and thus projects should not be monitored ex-ante either. Note that, if the choice of a_i is irrelevant, then the absence of ex-post monitoring can not have any incentive effects.

A corollary of Observation 1 is that the consumption of entrepreneur i must be higher or equal than the realization of x_i, otherwise the planner will need to monitor x_i.

Observation 2:

In the optimal plan, the set of agents that should allocate their unit of labor to the production of the intermediate good are those who receive a signal belonging to the closed interval $[0, s_0]$.

If an agent s_j is a worker but an agent s_k is an entrepreneur, with $s_j > s_k$, then by trading places expected output increases, which yields a contradiction.

Observation 3:

In the optimal plan, the set of agents that should allocate their entire unit of labor to the development of the investment project are those with a realization of w_i in the closed interval $[w_0, 1]$.

Again this is the only possibility consistent with the maximization of aggregate output.

Given Observations 1 to 3, the planner's problem is simply to choose s_0 and w_0 in order to maximize aggregate output, Y,

$$Y = \int_{w_0}^{1} \int_{w_0}^{1} \left(w + \frac{1}{2}\right) f(w,s)\, ds\, dw \qquad (2)$$

subject to the feasibility constraint:

$$\int_{w_0}^{1} \int_{s_0}^{1} f(w,s)\, ds\, dw \leq \int_{0}^{s_0} \int_{0}^{1} f(w,s)\, dw\, ds + \int_{0}^{w_0} \int_{s_0}^{1} f(w,s)\, dw\, ds \qquad (3)$$

The left hand side of equation (3) represents the level of investment. The first term of the right hand side is the amount of the intermediate good produced by those agents who devoted their entire unit of labor to such activity, while the second term is the amount produced by agents who changed occupation after learning the realization of w. The solution to this problem will be compared with the allocation resulting from decentralized optimizing behavior.

4. The competitive allocation

Let us now turn to the decentralized allocation. In this section, and in the next one, markets will be assumed to be perfectly competitive, i.e. agents take prices as given and markets clear.

The only possible market in this model is the credit market. Workers (savers) supply the intermediate good (capital) in exchange for a share of the entrepreneurs' final output. Since output is a random variable and there is the possibility of ex-ante and ex-post monitoring, financial contracts can be in principle quite complicated.

Let us first focus on the characteristics of the optimal contract neglecting the possibility of ex-ante monitoring. A contract specifies the payment to the lender in each state of nature, and the circumstances under which ex-post monitoring takes place.

The literature on financial contracting with costly state verification (Townsend, 1979; Diamond, 1984; Gale and Hellwig, 1985) shows that, under those circumstances, the optimal contract is a debt contract; i.e. the lender receives a fixed payment R, unless output is below R (bankruptcy), in which case monitoring takes place and all the output goes to the lender. In our case, if we denote by r the market return for a unit of the intermediate good, and S^b is the lender's payoff, then the following proposition holds.

Proposition 1

Without ex-ante monitoring, the equilibrium contract takes the following form (See Graph 1):

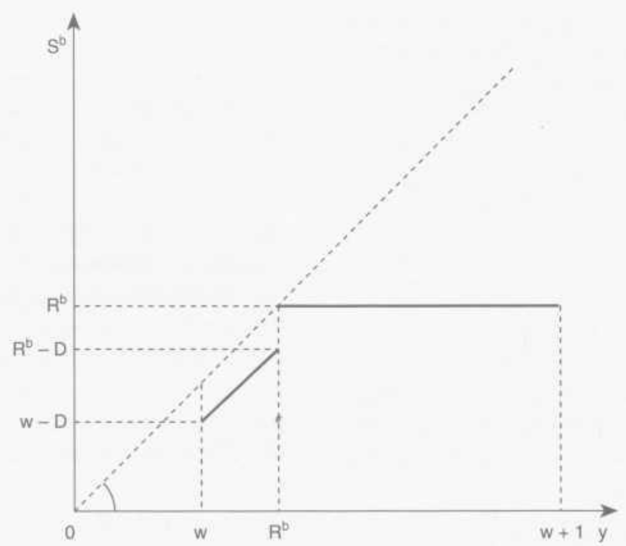
$$S^b = \begin{cases} R^b & \text{if } w + x \geq R^b \\ w + x - D & \text{otherwise} \end{cases}$$

and monitoring takes place if and only if $w + x - \leq R^b$. Finally, R is given by the lowest solution to:

$$\int_0^{R^b-w} (w+x-D)\, h(x|a_0)\, dx + R^b\,[1-H(R^b-w|a_0)] = r \qquad (4)$$

A contract without ex-ante monitoring will be called a bond.

Clearly, if $w \geq r$ then $R^b = r$ and the lender never monitors output. Also if $R^b \geq w+1$ then the project can not generate any surplus above r and thus can not be funded. Hence, the only non-trivial case involves $w < r$ and $R^b < w+1$.



Graph 1.

The proof of the first part of the proposition is standard and therefore omitted. At an intuitive level, for a contract to be incentive compatible the lender must monitor output in those states that specify a payment strictly below the maximum level. Also, optimality requires the minimization of monitoring costs. The consequence of both observations is that the optimal contract involves a fixed payment unless default takes pl ace in which case the lender monitors the realization of x and receives all the output [17].

The second part of the proposition indicates that under this contract the borrower is expected to choose the action a_0, that is the one that implies the highest output dispersion; at least, as long as $w<r$ and the project generates a positive surplus for the borrower. To see that we can write the borrower's payoff under this contract, π^b (a), as follows:

$$\pi^b(a) = \int_{R^b-w}^{1} (w+x-R^b)\, h(x|a)\, dx \qquad (5)$$

Thus, the incentives to choose a_0 are easily signed:

$$\pi^b(a_0) - \pi^b(a_1) =$$

$$= \int_{R^b-w}^{1} (w+x-R^b)\, [h(x|a_0) - h(x|a_1)]\, Dx =$$

$$= (R^b-w)\, [H(R^b-w|a_0) - H(R^b-w|a_1)] -$$

$$- \int_{0}^{R^b-w} x\, [h(x|a_0) - h(x|a_1)]\, dx>0$$

This sign is positive because $h(x|a_1)$ second order stochastically dominates $h(x|a_0)$. To check that define f (z) as follows:

$$\phi(z) = \equiv z\, [H(z|a_0) - H(z|a_1)] -$$

$$\int_{o}^{z} x\, [h(x|a_0) - h(x|a_1)]\, dx$$

Notice that $\phi(0)=\phi(1)=0$, and

$$\phi'(z) = H(z|a_0) - H(z|a_1)]>0 \quad \text{if } o<z<\frac{1}{2}$$

$$= \quad z=\frac{1}{2}$$

$$< \quad \frac{1}{2}<z<1$$

Consequently, nsequently, $\phi(z)>0$ for $z\in(0,1)$.
QED

Thus, it has been shown that, for a given R^b, the entrepreneur will choose the action a_0, but obviously lenders anticipate it and they take it into account in computing the face value of the bond, R^b, to set the expected return of the bond equal to r. As a result, for a given r, the choice of a_i affects only the borrower's expected profits. Using equation (4) we can write (5) as:

$$\pi^b(a) = w + \frac{1}{2} - r - D\, H(R^b-w|a) \qquad (6)$$

Hence, if $0<R^b-w<\frac{1}{2}$, bond contract involves an inefficient choice of action, since the characteristics of the bond induces the entrepreneur to choose a_0, which implies a high probability of bankruptcy, while the action a_1 implies lower output dispersion and thus lower bankruptcy costs [18].

[17] The optimal contract can be thought of as being implemented by a mechanism that involves the borrower announcing the level of output, and the contract specifying the payment to the lender and the probability of monitoring as a function of the announcement. Notice that if cheating has no exogenous cost, incentive compatibility requires the probability of monitoring to be one for those states that involve payments strictly below the maximum level. Hence nothing is gained by allowing random monitoring. However, if there are exogenous cheating costs (established, for instance, by the legal system) then the efficiency of the contract improves by allowing random monitoring. In the simple case of these costs being a proportion γ of the difference between the true output level and the announcement, then the optimal contract involves a probability $\frac{1}{1+\gamma}$ of monitoring in the default states, which reduces the absolute magnitude of monitoring costs without altering any of the qualitative results.

[18] It is well know that the limited liability aspect implicit in the debt contract implies that the firm acts as risk-lover agent and thus chooses the riskiest action. The novel aspect here is that despite of the risk neutrality assumption but because there are ex-post monitoring costs, the choice of the riskiest action is inefficient.

Therefore, provided C is negligible, the efficient financial contract when the bond involves positive expected ex-post monitoring costs and the probability of bankruptcy is lower than one half, involve ex-ante monitoring. We call the contract with ex-ante monitoring costs a bank loan, along with the interpretation that banks are the institutions that perform such a monitoring activity. Thus, we denote by S^l the borrower's payoff in this type of contract, and assume that banks behave perfectly competitive in such a way that the rate of return on a loan is r [19], then

Proposition 2

As long as C is sufficiently small, $r>w$ and $R^b(w)<w+\frac{1}{2}$,, where $R^b(w)$ is given by (4), the equilibrium contract consists of providing the unit of the intermediate good if and only if $a=a_l$ (ex-ante monitoring), and

$$S^l = \begin{cases} R^l & \text{if } w + x \geq R^l \\ w + x - D & \text{otherwise} \end{cases}$$

and ex-post monitoring takes place if and only if $w+x \leq R^l$. Finally, R^l is given by the lowest solution to [20]:

$$\int_0^{R^l-w} (w+x-D)\ h(x|a_l)\ dx + R^l[1-H(R^l-w|a_l)] = r \tag{7}$$

Thus, in our model entrepreneurs will have to choose between two alternative financing schemes: loans or bonds [21]. In a more sophisticated model the choice could include other ways of raising capital, like shares or partnership.

Entrepreneurs choose bank loans only if the probability of bankruptcy under the bond contract is between 0 and 1/2. An entrepreneur with a probability of bankruptcy higher than 1/2 has incentives to choose the safest action and thus would rather issue bonds and avoid the ex-ante monitoring costs. However, Lemma 1 in the Appendix shows that, provided ex-post monitoring costs are not too large, in equilibrium no entrepreneur faces a probability of bankruptcy higher than one half. In this case, Propositions 1 and 2 indicate that safe entrepreneurs (those with $w \geq r$) issue bonds (sign a financial contract without ex-ante monitoring), with $R^b=r$, while the risky ones (those with $w<r$) apply for loans (which involve ex-ante monitoring) with characteristics given by Proposition 2 [22]. Graph 2 illustrates how R^l and R^b depend on w and r. Therefore, an entrepreneur has an expected payoff equal to:

$$\pi^b(w) = w + \frac{1}{2} - r \qquad \text{if } w \geq r \tag{8}$$

$$\pi^l(w) = \int_{R^l-w}^{1} (w+x-R^l)\ h(x|a_l)\ dx = w + \frac{1}{2} - r - D\ H(R^l-w|a_l) \qquad \text{if } w<r \tag{9}$$

with R^l given by equation (7).

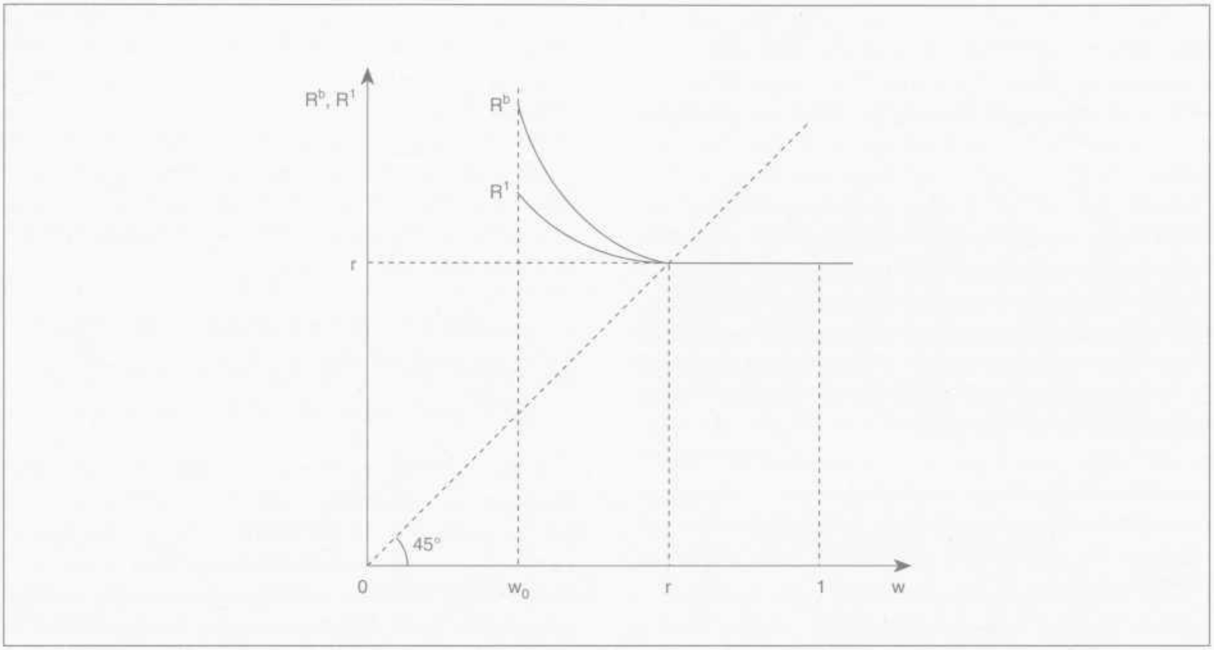
With this background information, the definition of equilibrium becomes very simple.

[19] It is implicitly assumed that there is a large number of banks and each one holds a perfectly diversified portfolio. Thus, the law of large numbers is assumed to hold not only for the whole economy but also for each individual bank.

[20] In general, the right hand side of equation (7) would be $r(1+C)$. However, as discussed above, for positive values of C ex-ante monitoring will be used only if the effect of a_i on the distribution of x_i is significant (if the information obtained is economically relevant). To avoid notational complexity and imposing a non-trivial condition on those distributions we simply assume that C is arbitrarily small.

[21] Diamond (1991) also discusses the determinants of choosing between loans and bonds.

[22] If the entrepreneur can not reallocate the second half unit of labor after learning the realization of w, the probability of bankrupcy of a funded project could be higher than one half. In this case the entrepreneur would take the action a_l even if not monitored ex-ante. Thus, the equilibrium would contain the undesirable feature of very risky entrepreneurs issuing bonds, instead of applying for a bank loan.



Graph 2.

— *Definition*

A competitive equilibrium is a vector (s_0, w_0, r) such that [23]:

(1) Given r and w_0, agents with $s < s_0$ choose to become workers (while those with $s > s_0$ become entrepreneurs). s_0 is given by:

$$\frac{r}{2}\int_0^{w_0} f(w|s_0)\, dw + \int_{w_0}^{r} \pi^l(w)\, f(w|s_0)\, dw + \int_r^1 \pi^b(w)\, f(w|s_0)\, dw = r \qquad (10)$$

(2) Given r, entrepreneurs with $w < w_0$ give up the project at the end of the first stage and use their second half unit of labor in the production of the intermediate good (those with $w \geq w_0$ go on with the project into their second stage). w_0 is given by:

$$\pi^l(w_0) = \frac{r}{2} \qquad (11)$$

(3) The market clears, i.e.

$$\int_{w_0}^{1}\int_{s_0}^{1} f(w, s)\, ds\, dw = \int_0^{s_0}\int_0^{1} f(w, s)\, dw\, ds + \frac{1}{2}\int_0^{w_0}\int_{s_0}^{1} f(w, s)\, dw\, ds \qquad (12)$$

Remember that $\pi^b(w)$ and $\pi^l(w)$ are given by equations (8) and (9), and R^l is given by equation (7). Equation (10) defines the threshold of the signal that leaves the agent indifferent between becoming a worker or

[23] Notice that from the point of view of workers, bank deposits and bonds are perfect substitutes.

an entrepreneur. Notice that the expected payoff of becoming an entrepreneur, the left hand side, is increasing in s. Thus, all agents with a higher signal will strictly prefer to develop their investment projects, while agents with a lower signal rather become workers.

Similarly, equation (11) defines the level of w that leaves the entrepreneur at the end of the first stage indifferent between quitting or carrying on the project. Again the expected payoff of developing the second stage of the project is increasing in w, so that all entrepreneurs with lower w will quit and those with higher w will go on [24].

Sufficient conditions for existence and uniqueness of equilibrium are given in the Appendix. Now we can compare the competitive equilibrium with the social planner's solution.

— *Proposition 3*

In competitive equilibrium the levels of consumption, investment and output are inefficiently low, with respect to the solution of the social planner's problem. As D goes to 0, the market allocation approaches the ex-ante optimal allocation.

The proof can also be found in the Appendix. The intuition goes as follows. The market return performs two jobs at the same time. On the one hand, it drives the allocation of (labor) resources between the production of the intermediate good and the entrepreneurial activities, and second it determines the frequency of bankruptcies. On the second ground the lower the return the better (the lower ex-post monitoring costs). Unfortunately, the interest rate that clears the market involves a positive frequency of bankruptcies [25].

In the limit case of $D=0$, the market allocation is efficient because the second job vanishes as bankruptcies are costless. But as D increases, an increasing amount of output goes to monitoring and thus consumption falls. On the top of that higher monitoring costs imply lower return for the marginal entrepreneur, thus lower credit demand and lower r. Consequently, s_0 decreases while w_0 increases, and investment also falls.

5. Linear taxes with perfectly competitive banks

In this section we investigate the effect of linear taxes on the competitive equilibrium allocation. Two shortcomings should be pointed out at the outset. First, in the presence of indivisibilities, like in our case, to lim it ourselves to a menu of linear taxes may be suboptimal. Second, the statements about optimal taxation are made for the case of an arbitrarily small revenue. Both limitations are imposed by the tractability of the problem.

5.1. *A tax on output*

Let us consider a proportional tax on output. Thus, the tax revenue from entrepreneur i is given by:

$$T_i = \tau y_i$$

where τ is the uniform tax rate.

The analog of Proposition 2 indicates that with a loan contract ex-post monitoring occurs if and only if:

$$(1-\tau)(w+x) \leq R^l$$

[24] If the occupational choice can only be made after learning w, then an equilibrium is a pair (w_0, r) satisfying conditions analogous to (11) and (12). In this case, the system is dichotomic with w_0 determined exclusively by the market clearing condition and the private payoff maximizing condition would determine the market return. In this simpler world the role of intermediaries and the entrepreneur's financial decisions would similar. The main difference would be that in this case taxes can not distort the allocation of labor.

[25] The inefficency of the competitive equilibrium can be explained in terms of incomplete markets. If agents can trade before they learn their characteristics the market allocation would be efficient. However, in this model only spot markets exist, i.e. agents are allowed to trade only after they learn their characterisitics.

and R^l will be given by[26]:

$$\int_{0}^{\frac{R_l}{1-\tau}-w} (1-\tau)\,(w+x-D)\,h(x|a_l)\,dx + \\ + R^l\left[1-H\left(\frac{R^l}{1-\tau}-w|a_l\right)\right] = r \tag{13}$$

Similarly, the entrepreneur's payoff is given by:

$$\pi^l(w) = \int_{\frac{R_l}{1-\tau}-w}^{1} [(1-\tau)\,(w+x) - R^l \\ h(x|a_l)\,dx \tag{14}$$

Also, entrepreneurs choose to issue bonds if and only if $(1-\tau)w \geq r$.

It is immediate to see that if $\frac{r}{1-\tau}$ is constant with t , then equation (13) implies that $\frac{R^l}{1-\tau}$ remains also constant. Therefore, by equation (14) $\frac{\pi^l(w)}{1-\tau}$ is also unchanged. Also, notice that the set of entrepreneurs that choose to issue bonds remains invariant and trivially $\frac{R^b}{1-\tau}$ as well as $\frac{\pi^b(w)}{1-\tau}$ are also constant.

Thus, by checking the equilibrium conditions (10) to (12), we realize that if $\frac{r}{1-\tau}$ is constant with τ, then the allocation of labor is invariant to t , and so is output. Consequently, $\frac{r}{1-\tau}$ is invariant with τ. Despite of the fact that the allocation of resources is unaffected, there is an efficiency loss given that implementing such a tax scheme implies monitoring ex-post all projects. This result is summarized in the following proposition:

— *Proposition 4*

A proportional tax on output creates no allocative distortion but involves a fixed collection cost.

The intuition is immediate. The incidence of a tax on output is proportional to all inputs, the intermediate good and entrepreneurial activity, so that no decision is distorted but all expected returns fall proportionally to the tax. However, since monitoring output is costly the government must pay a cost D in all contingencies, except when bankruptcy occurs in which case the government pays t D. Thus, for small tax rates, revenues will not be able to cover the collection cost.

5.2. *A tax on the intermediate good*

Suppose now that taxes are proportional to the returns on the intermediate good (a proportional capital income tax). Thus, the tax revenue from worker i is given by:

$$T_i = \gamma\, r$$

where γ is the uniform tax rate.

Now the equilibrium conditions are affected in a non-trivial way. In fact, we must rewrite (10) and (11) by:

$$(1-\gamma)\,\frac{r}{2}\int_{0}^{w_0} f(w\{s_0)\,dw + \\ + \int_{w_0}^{r} \pi^l(w)\,f(w\{s_0)\,dw + \\ + \int_{r}^{1} \pi^b(w)\,f(w\int_{w_0}^{1}\{s_0)\,dw = (1-\gamma)\,r \tag{15}$$

$$\pi^l(w_0) = (1-\gamma)\,\frac{r}{2} \tag{16}$$

[26] It is assumed that when bankruptcy occurs the tax collection agency bears a proportion τ of the ex-post monitoring costs.

By totally differentiating (8), (9), (12), (15), and (16):

$$\frac{dr}{d\gamma} < 0, \qquad \frac{d(1-\gamma)r}{d\gamma} < 0,$$
$$\frac{ds_0}{d\gamma} < 0, \qquad \frac{dw_0}{d\gamma} > 0$$

The intuition about the allocative effects of such a tax is the following. Other things equal, a tax on the return from the intermediate good reduces the incentives to become a worker. To eliminate the excess demand for the intermediate good r must increase, although the after-tax return $(1-\gamma)r$ falls. Higher r reduces the expected profits of the marginal entrepreneur, and thus w_0 increases. However, at the moment of observing the signal, the increase in r reduces the expected return from entrepreneurial activity but less than the after-tax return on the intermediate good, and hence s_0 falls [27].

The main idea is that a proportional capital income tax distorts the allocation of resources because only one input is directly taxed. Since the distortion is reflected in the market return on the intermediate good, entrepreneurs also bear part of the tax.

Let us now compute the welfare losses caused by such a tax. Let us first write the formula for aggregate consumption (which includes private and government consumption):

$$c = \int_{s_0}^{1} \int_{w_0}^{1} \left(w + \frac{1}{2}\right) f(w,s)\, dw\, ds - \\ - D \int_{s_0}^{1} \int_{w_0}^{r} \int_{0}^{R^1-w} f(w,s)\, h(x|a_1)\, dx\, dw\, ds \qquad (17)$$

The first term is aggregate output, while the second term is the aggregate ex-post monitoring costs. Using the equilibrium conditions (8), (9), (12), (15) and (16) we can compute the marginal welfare losses of a change in the tax rate:

$$\frac{dc}{d\gamma} = \gamma \frac{r}{3} f_s(s_0) \frac{ds_0}{d\gamma} - D \int_{s_0}^{1} \int_{w_0}^{r} f(w,s) \\ \frac{h(R^1-w|a_1)}{1 - H(R^1-w|a_1) - D\, h(R^1-w|a_1)}\, dw\, ds\, \frac{dr}{d\gamma} \qquad (18)$$

The first term is zero for the first dollar ($\gamma = 0$) and negative in general, and it represents the allocative distortion caused by the tax. The second is always negative (even when $\gamma = 0$) and represents the increase in ex-post monitoring costs caused by higher interest rates.

Notice that the first dollar collected has also a strictly negative welfare cost. The reason is a pre-existing distortion demonstrated in Proposition 3, i.e. the interest rate in a competitive equilibrium is too high because it induces too many bankruptcies. A tax on the intermediate good raises interest rates and thus increases the frequency of bankruptcies [28]. We summarize these results in the following proposition:

— *Proposition 5*

A proportional tax on the intermediate good (capital) has strictly negative welfare costs, even for the first dollar collected, since it distorts the allocation of resources and increases the frequency of bankruptcies.

5.3. *A tax on bank loans*

Finally, consider a proportional tax on bank deposits. Since bank deposits and bonds are perfect substitutes this implies that the before-tax return on deposits is $\frac{r}{1-\phi}$ and the tax revenue from worker i is:

[27] The market clearing condition (12) implies that s_0 and w_0 must move in opposite directions.
[28] If $D = 0$ then the welfare costs of the first dollar collected are null.

$$T_i = \frac{\phi\, r}{1 - \phi}$$

Thus, a bank loan must provide a return equal to the before-tax return on deposits $\frac{r}{1-\phi}$. Let us rewrite the face value of a loan:

$$\int_0^{R^l-w} (w+x-D)\, h(x|a_1)\, dx + R^l\,[1-H(R^l-w|a_1)] = \frac{r}{1-\phi} \qquad (19)$$

Hence it is no longer true that the optimal contract involves ex-ante monitoring if $w<r$. Now we can define w_1, $w_0<w_1<r$, as the value of w such that the entrepreneur is indifferent between applying to a bank loan or issuing bonds:

$$\sigma^l(w_1) = \pi^b(w_1) \qquad (20)$$

Thus, the equilibrium condition (10) is replaced by:

$$\frac{r}{2}\int_0^{w_0} f(w|s_0)\, dw + \int_{w_0}^{w_1} \pi^l(w))\, f(w|s_0)\, dw + \int_{w_1}^{1} \pi^b(w)\, f(w|s_0)\, dw = r \qquad (21)$$

where

$$\pi^b(w) = w + \frac{1}{2} - r - D\, H9R^b - w|a_0) \qquad (22)$$

and R^b is the lowest solution to (4) (i.e. $R^b = r$ if $w \geq r$).

Summarizing, the equilibrium conditions are given by (4), (9), (11), (12), and (19) to (22).

Totally differentiating these equilibrium conditions:

$$\frac{dr}{d\phi} < 0, \qquad \frac{d\,\frac{r}{1-\phi}}{d\phi} > 0,$$

$$\frac{dw_1}{d\phi} < 0 \qquad \frac{ds_0}{d\phi} < 0, \qquad \frac{dw_0}{d\phi} > 0$$

Most of the intuition is analogous to the previous tax scheme. A tax on bank loans have a different impact on different investment projects, so this explains the distortion in the allocation of labor (the changes in w_0 and s_0). Moreover, a tax on bank loans makes issuing bonds relatively more attractive than applying for a loan, so w_1 falls with the tax. Finally, since the profitability of entrepreneurial activity falls this lowers r, although the cost of loans increase $\left(\left(\text{i.e. } \frac{r}{1-\phi} \text{ increase}\right).\right)$

Let us now turn to the welfare loss caused by the tax. We can first rewrite the formula for aggregate consumption:

$$c = \int_{s_0}^{1}\int_{w_0}^{1} \left(w + \frac{1}{2}\right) f(w,s)\, dw\, ds - D\int_{s_0}^{1}\int_{w_0}^{w_1}\int_0^{R^l-w} f(w,s)\, h(x|a_1)\, dx\, dw\, ds - \int_{s_0}^{1}\int_{w_1}^{r}\int_0^{R^b-w} f(w,s)\, h(x|a_0)\, dx\, dw\, ds \qquad (23)$$

The first term is aggregate output, the second term the bankruptcy costs associated with bank loans, and the third the bankruptcy costs associated with bonds. The marginal welfare losses are given by:

$$\frac{dc}{d\phi} = \phi\, k_1 \frac{ds_0}{d\phi} + K_2 \frac{dw_1}{d\,\Phi} - D\, k_3 \frac{d\frac{r}{1-\phi}}{d\phi} \qquad (24)$$

where

$$k_1 = f_s(s_0)\ \frac{4}{1-\phi}\left[\frac{4}{3} - F(w_1|s_0)\right] > 0$$

$$k_2 = \frac{r}{1-\phi}\int_{s_0}^{1} f(w_1, s)\ ds > 0$$

$$k_3 = \int_{s_0}^{1}\int_{w_1}^{r} f(w,s)\ \frac{h(R^1-w|a_1)}{1-H(R^1-w|a_1)-D\ h(R^1-w|a_1)}\ dw\ ds\ +$$

$$+\int_{s_0}^{1}\int_{w_1}^{r} f(w,s)\ \frac{h(R^1-w|a_0)}{1-H(R^1-w|a_0)-D\ h(R^1-w|a_0)} dw ds > 0$$

The first term captures the distortion in the allocation of labor, which as usual is zero for the first dollar collected. The second term captures the distortion in the firm's financing decision between bonds and bank loans; again, this distortion is zero for the first dollar of taxes. The third term represents the increase in ex-post monitoring costs caused by the increase in the loan interest rate. As in the previous subsection this term is strictly positive even for the first dollar.

Thus, in principle a tax on bank loans creates similar distortions as a tax on the intermediate good but in addition distorts firms' financial decisions. These results are summarized in the following proposition:

— *Proposition 6*

A proportional tax on bank deposits has strictly negative welfare costs, even for the first dollar collected, since it distorts the allocation of resources, the entrepreneurs' choice between loans and bonds, and increases the frequency of bankruptcies.

Now we turn to the issue of comparing the relative efficiency of these alternative forms of taxation. This can be easily done for the first dollar, but unfortunately it becomes quite complicated to keep track of the evolution of the different tax bases as tax rates become strictly positive. The next proposition reports the result for the first dollar of taxes.

— *Proposition 7*

With a perfectly competitive banking industry, the most efficient way of collecting the first dollar is through the tax on the intermediate good. In other words, for small tax revenues a tax on the intermediate good strictly dominates the two other forms of taxation.

— *Proof*

(a) Proposition 4 already shows that the tax on output involves a fixed cost and is therefore unable to raise any positive revenue if the tax rate is small.

(b) From Propositions 5 and 6 we see that the welfare loss of taxation for the first dollar have the same coefficient and only depend on the difference between

$$\frac{dr}{d\gamma} \text{ and } \frac{d\,\frac{r}{1-\phi}}{d\phi} \quad \text{(evaluated at } \gamma = \phi = 0\text{)}$$

It can be shown that the second is larger than the first. Moreover, the base of the tax on the intermediate good is larger than the one on bank loans.

QED

It is quite complicated to extend the result to discrete values of the tax rates. However, the insight provided by the analysis suggests that the same result is likely to hold for large tax revenues, since the tax on bank loans distorts in the same direction the labor allocation decisions, it similarly increases the frequency of bankruptcies and on the top of that it distorts the entrepreneurs' financial decisions.

6. Taxes with a monopolistic banking industry

Suppose entrepreneurs and workers live in different islands. In each island there is a single bank (a local monopolist), but all workers and entrepreneurs have access to a centralized bond market, but not to banks located in other islands. This implies banks must offer to their depositors at least the same rate of return on capital than the one prevailing in the bond market (no market power in the deposit market), but entrepreneurs can only choose between the bond market or the local bank, which can exert some degree of monopoly power in the loan market.

Banks are assumed to maximize profits, which are distributed at the end of the period. For simplicity, the distribution of bank ownership is assumed to be exogenous, and since the social welfare criterion is simply aggregate consumption, need not be specified.

The monopoly power of the local bank is limited by the bond market. That is, it can not charge an interest rate that implies lower profits than those achieved by the entrepreneur in the bond market, i.e. $R^l(w)$ must satisfy:

$$\int_{R^b-w}^{1} (w+x-R^b)\, h(x|a_0)\, dx \leq \int_{R^l-w}^{1} (w+x-R^l)\, h(x|a_1)\, dx \quad (25)$$

where R^b is given by (4).

However, the bank has incentives to charge the highest possible interest rate. To check that, we can write bank's profits as a function of the interest rate:

$$S^l(w) = \int_{0}^{R^l-w} (w+x-D)\, h(x|a_0)\, dx + R^l[1-H(R^l-w|a_0)] \quad (26)$$

and

$$\frac{dS^l}{dR^l} = 1 - H(R^l-w|a_1) - D\, h(R^l-w|a_0) > 0$$

Therefore, in equilibrium R^l will be determined by equation (25) with equality.

An important remark is that the monopolist can perfectly discriminate across borrowers, but this is not an ad hoc assumption but the natural consequence of the structure of the model. In fact, this feature is likely to be quite robust to changes in the environment, as long as banks perform a role in ex-ante monitoring, since they acquire private information about firms' return distribution.

A second remark is that, despite of the fact that banks are able to extract all the surplus from firms, the existence of banks is welfare enhancing. The reason is that by monitoring firms' actions they are able to reduce ex-post monitoring costs. In fact, banks profits are equal to the reduction of ex-post monitoring costs achieved by ex-ante monitoring.

Thus, entrepreneurs' expected payoffs are given by:

$$\pi^b(w)=w+\frac{1}{2}-r \qquad \text{if } w>r \quad (27)$$

$$\pi^l(w)=w+\frac{1}{2}-r-D\, H(R^b-w|a_0) \qquad \text{if } w<r \quad (28)$$

where R^b is given by (4). Notice that the payoff function of a loan applicant, equation (28), depends on R^b since such an entrepreneur is indifferent between loans and bonds.

The remaining equilibrium conditions are (10) to (12), as before.

Notice that the effect of output and intermediate good taxation would be very similar to the one described in the previous section. However, taxing bank deposits will have qualitatively different effects.

Figure 2

	Welfare loss through			
	Distorsion in the labor allocation	*Increase in the frequency of bankruptcies*	*Distorsion in financial decisions*	*Tax collection cost*
Tax on output	—	—	—	yes (∞)
Tax on capital	yes (0)	yes (+)	—	—
Tax on deposits (perfect competition)	yes (0)	yes (+)	yes (0)	—
Tax on deposits (monopoly)	—	—	yes (0)	—

Nota: Each entry indicates whether the particular tax causes a welfare loss through that particular channel. In the case of a yes, it is indicated (between brackets) whether the distorsion is null, positive but finite, or infinite, for the first dollar collected.

As usual, when taxing banks we need to worry about the determination of w_1. To an entrepreneur with $w > w_1$ the bank is unable to make positive profits. Thus, the entrepreneur with $w = w_1$ is indifferent between issuing bonds and applying for a loan and the bank makes zero profits; i.e. w_1 is given by:

$$w_1 + \frac{1}{2} - r - D\,H(R^b - w_1 | a_0) = w_1 + \frac{1}{2} - \frac{r}{1-\phi} - D\,H(R^l - w_1 | a_1) \qquad (30)$$

where R^l is given by (25) with equality.

Notice that the system is decomposable, and that w_0, s_0, and r (and R^l) are independent of ϕ, while the only variable affected by the tax is w_1 (and R^b for $w_1 < w < r$). In fact, from equation (30):

$$\frac{dw_1}{d\phi} < 0$$

Aggregate consumption is given by equation (23), where R^l is implicitly defined by equation (25) when it holds with equality. The marginal welfare loss from taxing bank deposits, using equation (30), is:

$$\frac{dc}{d\phi} = \frac{dw_1}{d\phi}\frac{\phi\, r}{1-\phi}\int_{s_0}^{1} f(w_1, s)\, ds \leq 0 \qquad (31)$$

Thus, the following proposition holds:

— *Proposition 8*

With a monopolistic banking industry, the first dollar collected by taxing bank deposits has zero welfare costs. Hence, taxing deposits strictly dominates the other alternative forms of taxation. In fact the only welfare loss comes from the disintermediation effect.

A tax on bank deposits increases the minimum return that the bank must raise from a loan. *Ceteris* *paribus*, the bank will exclude only safe borrowers, those who can actually borrow from the bond market at rates that incorporate a low risk premium, while it will keep the rates constant to the rest of the pool of applicants. Since in equilibrium loan applicants are indifferent between issuing bonds or getting a bank loan, disintermediation has no effect on the profitability of entrepreneurial activity. Thus, the allocation of labor is unaffected, and so is the market return. The only effect will be on the amount of ex-ante monitoring performed and thus on the bankruptcy costs. The first dollar collected has no welfare cost since the marginal borrower excluded by the bank is that with no risk ($w = r$), and hence this fact has no effect on ex-post monitoring costs.

In other words, the incidence of the tax goes entirely to banks' profits, and since these profits are equal to the amount of bankruptcy costs saved by monitoring entrepreneurs' actions, taxing bank deposits will generally have welfare costs, except for the first dollar.

In fact, the empirical evidence available [29] indicates that banks do bear a significant part of the deposit taxes. Thus, our model provides a possible rationale for this observation, and suggests the relative magnitude of the welfare costs associated [30].

7. Concluding remarks

In this paper we have analyzed the relative efficiency of taxing financial intermediaries in a stylized general equilibrium model. A particular feature has been emphasized: the optimal tax system is partly determined by the presence of different tax collection costs, but these costs are not ad hoc but arise from primitive assumptions about the economy. However, the particular modeling strategy chosen is unsatisfactory in several other respects. Let us close the paper by making explicit at least two of these shortcomings.

A comparison of the results of sections 5 and 6 indicate that whether taxing banks' activities is dominated or not by alternative taxes, largely depends on whether banking is better characterized by perfect competition or monopoly. Obviously, it would be desirable to perform a similar experiment in intermediate market structures (oligopoly or monopolistic competition). However, modeling bank interaction in an imperfectly competitive framework, when the monitoring role of banks is explicitly recognized, turns out to be quite complicated [31].

The second comment is related to the existence of pure economic profits in banking. It is obvious that a proportional profit tax involves no efficiency loss. In particular, in the model of section 6, a tax on bank deposits is dominated a profit tax. One can probably argue that in the real world the observation of banks' economic profits is at least as difficult as the observation of firms' outcomes. However, in the spirit of the paper this implies that a different (and much more complex) model of banking is needed.

Both issues indicate that this paper is only a first step in the right direction, but that clearly further research is needed.

[29] See again Osborne and Zaher (1992).

[30] The prescription that taxing the monopolist rather than the competitive fringe is optimal is not a general result. In fact, the opposite is more likely. See, for instance, Mintz and Seade (1989).

[31] See, for instance, Bhattacharya (1992).

References

Bacchetta, P. and R. Caminal (1992a): A note on reserve requirements and public finance, working paper #176-92, Universitat Autònoma de Barcelona, forthcoming in the *International Review of Economics and Finance.*

Bacchetta, P. and R. Caminal (1992b): Optimal seigniorage and financial liberalization, *Journal of International Money and Finance*, 11, 6 (December), 518-538.

Bhattacharya, S. (1992): Financial intermediation with proprietary information, CEPR-ESF working paper #21.

Brock, P. (1989): Reserve requirements and the inflation tax, *Journal of Money, Credit and Banking*, 21, 106-121.

Diamond, D. (1991): Monitoring and reputation: the choice of bank loans and directly placet debt, *Journal of Political Economy*, 99, 689-721.

Diamond, D. (1984): Financial intermediation and delegated monitoring, *Review of Economic Studies*, 51, 393-414.

Fama, E. (1980). Banking and the theory of finance, *Journal of Monetary Economics*, 6, 39-57.

Gale, D, and M. Hellwig (1985): Incentive-compatible debt contracts: the one-period problem, *Review of Economic Studies*, 52, 647-663.

Horrigan, B. (1988): Are reserve requirements relevant for economic stabilization?, *Journal of Monetary Economics*, 21, 97-105.

Kimbrough, K. (1989: Optimal taxation in a monetary economy with financial intermediaries, *Journal of Macroeconomics*, 11, 493-511.

Mintz, J. and J. Seade (1989): Small enterprise tax policy: effects on market price, industry profits and allocative efficiency in a dominant firm model, mimeo.

Osborne, D. and T. Zaher (1992): Reserve requirements, banks share prices and the uniqueness of bank loans, *Journal of Banking and Finance*, 16, 799-812.

Romer, D. (1985): Financial intermediation, reserve requirements and inside money: a general equilibrium analysis, *Journal of Monetary Economics*, 16, 175-194.

Siemrod, J. (1990): Optimal taxation and optimal tax systems, *The Journal of Monetary Economics*, 16, 175-194.

Siemrod, J. (1990): Optimal taxation and optimal tax systems, *The Journal of Economic Perspectives*, 4, 157-178.

Townsend, R. (1979): Optimal contracts and competitive markets with costly state verification, *Journal of Economic Theory*, 21, 1-29.

Appendix

LOW PROBABILITY OF BANKRUPTCY

Lemma 1

$$w_0 < r < 1$$

$$R^1(w) - w < \frac{1}{2} \text{ for all } w \geq w_0$$

Proof of Lemma 1

If $D=0$, then in equation (11) becomes:

$$r = \frac{2}{3}\left(w_0 + \frac{1}{2}\right)$$

Since $0 \leq w_0 < 1$ then $w_0 < r < 1$.

Also $R^1(w) = r$, so that

$$R^1(w) - w \leq r - w_0 = \frac{1 - w_0}{3} \leq \frac{1}{3}$$

By continuity, the Lemma is correct for D small enough. QED

EXISTENCE AND UNIQUENESS OF EQUILIBRIA

First, let us introduce some notation:

$$H(w_0, s_0) \equiv \int_{w_0}^{1} wf(w|s_0)\, dw +$$

$$+ w_0 F(w_0|s_0) - \frac{1}{6} - \frac{4w_0}{3}$$

$$J(w_0, s_0) \equiv \int_{w_0}^{1}\int_{s_0}^{1} f(w, s)\, ds\, dw -$$

$$- \int_{0}^{s_0}\int_{0}^{1} f(w, s)\, ds\, dw - \frac{1}{2}\int_{0}^{w_0}\int_{s_0}^{1} f(w, s)\, ds\, dw$$

$$H(\hat{w}_0, 0) = 0$$

$$H(0, \hat{s}_0) = 0$$

$$J(\bar{w}_0, 0) = 0$$

$$J(0, \bar{s}_0) = 0$$

Lemma 2 (Sufficient conditions)

If and $H(0, \bar{s}_0) > 0$ and $H(\bar{w}_0, 0) < 0$, then the competitive equilibrium exists and is unique.

Proof of Lemma 2

The equilibrium conditions (equations (10) to (12)) for $D=0$ can be written as:

$$H(w_0, s_0) = J(w_0, s_0) = 0$$

It is clear that

$\bar{s}_0$, $\bar{w}_0$ are strictly less than 1 and that at least one of them is strictly positive.

The slope of these two conditions can be easily signed:

$$\frac{dw_0}{ds_0}\Big|_{J=0} < 0,$$

$$\frac{dw_0}{ds_0}\Big|_{H=0} > 0,$$

This implies that if the equilibrium exists, then it is unique.

Existence will be guaranteed if

$$\hat{s}_0 < \hat{s}_0$$

$$\hat{w}_0 < \hat{w}_0$$

but since

$$\frac{\delta H}{\delta s_0}(0, s_0) > 0$$

$$\frac{\delta H}{\delta w_0}(w_0, 0) < 0$$

these conditions are equivalent to $H(0, \bar{s}_0) > 0$ and $H(\bar{w}_0, 0) < 0$.

By continuity, these conditions guarantee existence and uniqueness provided D is small enough. QED

OPTIMALITY OF THE COMPETITIVE EQUILIBRIUM

Proof of Proposition 3

The planner's problem consists of maximizing (2) subject to (3) and subject to the non-negativity constraints:

$$w_0 \geq 0$$

$$s_0 \geq 0$$

since it is clear that w_0 and s_0 are strictly less than 1.

Let us denote by λ, δ μ the Lagrange multiplier associated with the feasibility and the non-negativity constraints respectively.

The first order conditions will be given by:

$$\int_{w_0}^{1}\left(w+\frac{1}{2}\right) f(w|s_0)\, dw =$$

$$= \lambda\left[2 - \frac{3}{2} F(w_0|s_0)\right] + \frac{\mu}{f_s(s_0)}$$

$$\lambda = \frac{2}{3}\left(w_0 + \frac{1}{2}\right) - \frac{\delta}{\frac{3}{2}\int_{s_0}^{1} f(w_0, s)\, ds}$$

plus the feasibility constraint.

Only one negativity constraint can be binding. Suppose, first that $\mu > 0$, i.e. $\delta = s_0 = 0$, and $\bar{w}_0 > 0$. Then the first order conditions imply that $H(\bar{w}_0, 0) > 0$.

Next, suppose that $\delta > 0$, i.e. $\mu = w_0 = 0$, $s_0 > 0$. Then the first order conditions imply that $H(0, \bar{s}_0) < 0$.

Finally, $\mu = \delta = 0$, then the first order conditions imply that $H(w_0, s_0) = J(w_0, s_0) = 0$.

Consequently, as long as the sufficient conditions (given in Lemma 2) for existence and uniqueness of the competitive equilibria are satisfied, the competitive equilibrium with $D = 0$ is ex-ante efficient.

What if $D > 0$? Totally differentiating (10), (11), and (12) it turns out that:

$$\frac{dr}{dD} < 0, \quad \frac{dw_0}{dD} > 0, \quad \frac{ds_0}{dD} < 0$$

Since the allocation of labor is distorted from the output maximizing levels, it follows that with $D > 0$, output is lower than in the first best. And so is consumption: lower output and higher ex-post monitoring costs.

Finally, let us check on the effect of D on the level of investment. Investment , I, is given by:

$$I \equiv \int_{w_0}^{1}\int_{s_0}^{1} f(w, s)\, ds\, dw$$

Thus,

$$\frac{dI}{ds_0} = -\int_{w_0}^{1} f(w, s_0)\, dw - \frac{dw_0}{ds_0}\int_{s_0}^{1} f(w_0, s)\, ds =$$

$$= \frac{1}{3}\int_{0}^{1} f(w, s_0)\, dw > 0.$$

Thus as s_0 falls with D, so does investment. QED

FUNDACION BBV

Gran Vía, 12 - 48001 BILBAO
Alcalá, 16 - 28014 MADRID