Rent taxation when cost transfers are possible, but costly*

Hilde E. Halvorsen

Ministry of Transport and Communications

P.O. Box 8010, Dep., N-0030 Oslo, Norway

Internet: hilde.halvorsen@sd.dep.telemax.no

Diderik Lund

Department of Economics, University of Oslo

P.O. Box 1095, Blindern, N-0317 Oslo, Norway

Internet: Diderik.Lund@econ.uio.no

First version October 3, 1997

This Memorandum version August 1998

*We gratefully acknowledge comments from an anonymous referee. The views expressed are those of the authors, not necessarily those of their employers. Corresponding author: Diderik Lund, Department of Economics, University of Oslo, P.O. Box 1095, Blindern, N-0317 Oslo, Norway, Voice: +47 2285 5129, Fax: +47 2285 5035

Abstract

While rent taxation in some theories is neutral, and the tax rate could be set to one hundred percent to minimize the need for distortionary taxes, this does not occur in practice. An important reason for this is the transfer incentives that would result. Monitoring to prevent transfer pricing is difficult, in particular on the cost side. For corporations, monitoring implies that both transfer pricing and real transfers will be costly. Assuming a convex cost function for cost transfers, it is shown that the optimal tax system combines a cash flow tax with a royalty, i.e., a tax on gross revenues. This contributes to explaining the frequent occurence of royalties in actual rent tax systems. It is also shown that the optimal tax rates depend on the output price. This contributes to explaining the frequently observed tailoring of rent tax systems in response to output price changes, in particular in the petroleum sector. These results hold when the corporations are heterogeneous and the government only knows a probability distribution for the two cost parameters. An analogy to the problem analyzed in the present paper, is the problem of a minority shareholder in preventing a majority shareholder to withdraw profits through, e.g., transfer pricing.

KEYWORDS: Rent taxation, transfer pricing, income shifting

Rent taxation when cost transfers are possible, but costly

1 Introduction

Most of the literature on taxation regards a rent tax as neutral. A natural implication is that the tax rate should be increased as long as this reduces the need for other, distortionary taxes. But 100 per cent rent taxes do not occur in practice. Since significant amounts of revenue may be raised this way in many countries, it is important to consider why not.

The popularity of 100 per cent rent taxes in the theoretical literature, e.g., Stiglitz & Dasgupta (1971) and Guesnerie (1995), has to do with analytical convenience: Theoretical results are clearer when rents are always taxed away. These theorists should also consider why the assumption is not satisfied in the real world.

Stiglitz & Dasgupta (1971, p. 165) write, "No government has imposed on a regular basis 100 per cent taxes on profits and the income of fixed factors, in spite of the long standing advice of economists (e.g. Henry George) of the desirability of such non-distortionary taxes." They mention two reasons for this, one having to do with the problem in separating out pure profits. The present paper may be regarded as an elaboration of the consequences of that problem.

The design of many rent tax systems is poor, violating obvious requirements for a neutral tax. Most frequently there is imperfect loss offset, reducing incentives to invest, cf. Brown (1948), Mayo (1979), Ball & Bowers (1982). This alone may be sufficient to explain why tax rates are lower than one hundred percent.

Our focus here will be different, so we assume that the rent tax is a tax on non-financial cash flows with full loss offset, cf. Brown (1948). In standard models of investment decisions this tax is neutral. This is also true under uncertainty when corporations maximize their market values. From theory the rate of the tax should be about 99 percent, avoiding indifference, leaving the necessary incentive to invest. But our analysis does not lead to such a recommendation.

Less standard models are needed to explain why the rate should be set significantly below one hundred percent, even for a cash flow tax. In Haaparanta & Piekkola (1997) the reason is an entrepreneurial effort which is necessary to undertake investments. Being unobservable (or non-verifiable), this effort is not tax deductible. Although the cash flow tax formally has full loss offset, the lack of this deduction invalidates the standard neutrality result. Observe that such a mechanism might be present in a closed, one-sector economy.

The explanation given in the present paper may be complimentary to the one in Haaparanta & Piekkola. We observe that many rent taxes are applied in open economies, or in one sector of a larger economy. This creates incentives to transfer revenues out of a high-tax sector and costs into a high-tax sector.

Sansing (1996) observes the same phenomenon, and presents a model somewhat similar to ours. He shows that under some conditions it will be optimal for authorities to allow deductions for costs at a lower marginal tax rate than the marginal tax rate on revenues. This explains the widespread practice of royalties, i.e., taxes on gross revenue, often combined with standard corporate income (profits, rents) taxes. Sansing (1996) based his analysis on a tax system using the comparable profits method (CPM) to avoid transfer problems.

Our model will also lead to optimal royalties, but from a different mechanism, without the CPM. We assume that costs may be transferred, but only at a cost, and we derive the optimal royalty.

2 The model

We use a simple, static production model with a quadratic cost function. One might argue that most rent tax systems apply to natural resource extraction, and that the effects in a dynamic exhaustible resource model would more interesting. We believe that for each producer, most decisions are not subject to Hotelling-type optimal timing considerations related to output price growth. For each deposit it is most reasonable to assume extraction at full capacity as soon as capacity is installed, cf. Campbell (1980). Thus a static model may be sufficient to capture the interesting effects.

Consider a corporation which earns rents before tax of

$$R(q) = pq - bq^2, \tag{1}$$

where $q \geq 0$ is quantity produced, $p > 0$ is output price, and $b > 0$ is a parameter of a cost function. We shall assume throughout that p is exogenously given, both for the corporation and for the tax authorities.

A tax with a high rate will be levied on this rent. The corporation also has activities in another sector. This gives an incentive to income shifting, transfers of revenues away from this sector, and costs into this sector. In line with Sansing (1996) we assume that the revenue side is more easily monitored by tax authorities than the cost side. One reason is the multitude of inputs which normally go into the production of one (or a few) output(s). Another is the fact that inputs are often tailor-made, which complicates the enforcement of arms-length pricing. For simplicity we assume that no transfers are made on the revenue side. Quantity may be measured accurately, and arms-length prices may be observed in a world market.

Cost transfers may be of two kinds: Real transfers or transfer pricing. Transfer pricing consists in charging a higher price for inputs when the seller is a related company in another sector. "Related" may mean a company with the same owners, but more complicated common interests may also prevail. Although we shall not model it explicitly, we mention that the outcome of bargaining over prices is likely to be affected by tax differences.

Real transfers consist in using more input factors than would have been the case without tax differences. This may be motivated by additional non-taxed benefits, such as benefits from testing of equipment or technology or training of personell. These benefits may be reaped in a different jurisdiction. A multinational corporation will have incentives to do testing and training in the sector in which they are decuctible against the highest tax rate.

For simplicity we assume that an amount a is deducted in the rents without having any productive effect. Thus the taxable corporate income is $pq - bq^2 - a$. Both a rent tax and a royalty are levied, and the tax revenue is

$$T = t(pq - bq^2 - a) + r(1-t)pq, \tag{2}$$

where t is the rent tax rate, r is the royalty rate, and the royalty is deductible in the base for the rent tax, which is common in practice.

We assume that the transfer of costs is costly. These additional costs may be incurred to avoid monitoring by the authorities. In case of real transfers, there may be real transport costs, or lower efficiency in testing and training than in other sectors. For simplicity we assume that these costs are independent of q, that they are not tax deductible, and we assume a quadractic cost function ca^2, where $c > 0$ is a parameter. There are alternative assumptions in the transfer pricing literature, such as the proportional profits shifting found in Weichenrieder (1996). The idea of a convex cost function for transfer costs may be found in Gordon & MacKie-Mason (1995), who stress the importance of income shifting for the design and analysis of tax policy.

Thus the after-tax corporate profits are

$$\Pi(q, a) = (pq - bq^2)(1-t) - r(1-t)pq - ca^2 + a(t-s), \tag{3}$$

where $s \in (0, t)$ is the relevant tax rate in another sector, so that the last term represents the taxes saved by the transfer.

The maximization of Π with respect to q and a leads to the following optimal values:

$$q = \frac{p(1-r)}{2b}, \tag{4}$$

and

$$a = \frac{t-s}{2c}. \tag{5}$$

We observe that q is not affected by the rent tax rate t, even when there is a royalty, which means that the rent tax is still neutral in this sense. However, a royalty rate, r, does remove q from its pre-tax optimum. We also observe the obvious effect that a is increasing in the difference in tax rates, $t-s$.

We now assume that the corporation is owned by foreigners, so that its after-tax profits do not count in the welfare function. The authorities choose tax rates t and r in order to maximize the tax revenue, T, conditional on the optimizing behavior of the corporation.

Under the assumptions made so far, the authorities know the cost function parameters as well as p and s, and is able to predict the corporation's behavior perfectly. A fixed fee would be able to capture the rent with no deduction for costs, and thus no dissipation problem. We return to a more elaborate model with asymmetric information in section 4, and show that when the rest of the model is maintained, the maximization problem is essentially equivalent. Thus the result we now derive, also hold in the case with asymmetric information.

Plugging (4) and (5) into (2) gives the following expression,

$$T = \left[\frac{p^2(1-r)^2}{4b} + \frac{s}{2c}\right] t - \frac{t^2}{2c} + \frac{rp^2(1-r)}{2b}. \tag{6}$$

This may be simplified by the definition of $\alpha \equiv p^2c/4b$, which gives

$$T(t, r; \alpha) = \frac{1}{c}\left\{\left[\alpha(1-r)^2 + \frac{s}{2}\right] t - \frac{t^2}{2} + 2\alpha r(1-r)\right\}. \tag{7}$$

The first-order conditions for maximizing $T(t, r; \alpha)$ with respect to t and r can be expressed in two equations, in which we for convenience define functions $f(t; \alpha)$ (for $t > s/2$)

and $g(t)$,

$$r = 1 - \sqrt{\frac{1}{\alpha}\left(t - \frac{s}{2}\right)} \equiv f(t;\alpha), \tag{8}$$

and

$$r = \frac{1-t}{2-t} \equiv g(t). \tag{9}$$

From these two equations in t and r we find solution(s) when

$$f(t;\alpha) = g(t), \tag{10}$$

We shall not write down the explicit analytical solutions, since they are messy solutions to third-order polynomial equations. The structure is sufficiently simple, however, to give straightforward analytical results.

A solution is only of interest (as tax rates) if it is in the interval $[0,1]^2$. Such (t,r) values between zero and unity, which satisfy (8) and (9), will be the relevant stationary points of T. Together with corner solutions, they are the candidates for a maximum point. For given values of α and s, the two functions f and g are continuous and strictly decreasing in this interval, and f is convex while g is concave.

Figure 1 shows an example of graphs of the two functions in a (t,r) diagram for $s = 0.3$ and $\alpha = 0.89$. The value $s = 0.3$ is chosen as an approximation to an average corporate income tax rate across countries of the world. In the subsequent discussion we keep this value of s fixed, while we want to trace the dependence of the maximum point on α.

(INSERT FIGURE 1 HERE)

In the diagram based on $\alpha = 0.89$ there are two intesections between the two curves, i.e., two stationary points. If we change α, the graph of the g function is unaltered, while the graph of f is changed. Its intersection with the horizontal $r = 1$ line is fixed, however, at $t = s/2 = 0.15$. A higher α makes it flatter (i.e., a reduced $|\partial f/\partial t|$). There is a slightly higher α for which there just one intersection between the curves (a tangency), and for even higher α values, there is none. A lower α makes the graph of f steeper, and below some α value there will be only one intersection within the relevant interval, $(t,r) \in [0,1]^2$.

For high values of α it will turn out that the maximum within this interval is not found at the stationary point, but as a corner solution, $(t, r) = (1, 0)$. This occurs when there is no intersection between the curves, but also for slightly lower α values. In the appendix we prove the following.

Lemma *Keep a fixed value of s. Within the interval $(t, r) \in [0, 1]^2$ the function $T(t, r; \alpha)$ has an interior maximum for low values of α. This occurs in the interval of low α values for which there is only one stationary point in $[0, 1]^2$, but also in an adjacent interval of higher α values for which there are two stationary points. For even higher values of α the maximum occurs at the corner $(t, r) = (1, 0)$. When there is an interior global maximum, this point has a t value which is increasing in α and an r value which is decreasing in α.*

From this follows the following:

Proposition *For a fixed value of s, the optimal rent tax rate is increasing, and the optimal royalty rate is decreasing, as we consider these partial changes: (a) The output price p is increasing, from zero to some upper limit, (b) the parameter c of the transfer cost function is increasing, from zero to some upper limit, and (c) the parameter b of the productive cost function is decreasing, from infinity to some lower limit. Above these two upper limits for p and c, and below the lower limit for b, the optimal rent tax rate is one hundred percent, while the optimal royalty rate is zero.*

3 Discussion

The Proposition shows that the model may work well as a descriptive model: While many other models prescribe neutral cash flow taxes and tax rates which do not change when prices change, this model explains why authorities may want to use royalties and may want to change tax rates when prices change.

The model is of course highly stylized. Many existing tax systems are more complicated, with depreciation schemes and less than full loss offset. But they often have the feature that the marginal tax rate which applies to revenues is higher than the marginal tax rate at which costs can be deducted. The problem of monitoring cost transfers may be a reason for such a feature.

The model is also quite simple when it comes to the objective function of authorities, and the two cost functions. We have used these in order to get a model that could be handled analytically. Another simplifying assumption is that the alternative tax rate s (abroad, or in another sector) has been assumed fixed, while in reality, there may be different relevant values for different firms and at different points in time.

We regard the model more as a descriptive than as a normative model. If we were to advise tax authorities, we might recommend something different from a cash flow tax, but the idea of letting tax rates depend on the output price seems to be flawed. In a more realistic multi-period model, one would run into the problem that output prices change and that the authorities might want to change tax rates over time. If firms realize this, there will be a dynamic game situation. With uncertainty also, the problem may be difficult to solve. We leave this for future research.

An interesting analogy to the problem analyzed here is the problem of a minority shareholder in protection against dilution of profits by a majority shareholder. The analysis suggests that the minority shareholder might be interested in handing in some of the shares in exchange for a claim to a fraction of gross revenue. This will be the case if monitoring is more difficult on the cost side, and cost transfers are costly to the majority shareholder.

4 Extending the model: Aymmetric information

If authorities know as much as the corporation does, it will be more efficient to charge a fixed fee for the resource, at almost one hundred percent of the maximum rent. Suppose instead that there is a fixed number of heterogeneous corporations knowing their own cost function parameters b_i and c_i (both strictly positive), but that the authorities only know distribution functions for these. In this case there is no option to charge a fixed fee.

Assume that authorities, in order to reduce political and administrative costs, consider two tax parameters only, i.e., the same tax system as in the previous sections. Observe that with the given cost functions, any pair of strictly positive values for b and c will result in strictly positive profits for any pair of positive tax rates t and r which are strictly less than unity. Thus all corporations will produce at any pair of tax rates.

The authorities want to maximize the expected tax revenue from each corporation, assuming that the corporation will maximize its profits for the given tax parameters. Plugging in from (6), we find

$$E(T) = E\left\{\frac{1}{b}\cdot\frac{p^2}{2}\left[\frac{t(1-r)^2}{2} + r(1-r)\right] + \frac{1}{c}\cdot\frac{t}{2}(s-2)\right\} \tag{11}$$

Clearly, the maximization of this is the same as the maximization of (7) with $1/E(1/b)$ and $1/E(1/c)$ substituted for b and c.

If we had included fixed costs in the model, the solution would have been different, since some of the corporations might have decided not to produce.

5 Conclusion

We have shown that it may be optimal for governments to use a combination of a cash flow tax (called a rent tax) and a tax on gross production value (called a royalty) in order to maximize tax revenue in the presence of costly cost transfers. This is one possible explanation for the frequent occurence of royalties in practice, in spite of the well-known non-neutrality of such taxes.

Furthermore, we have shown that both the optimal royalty rate and the optimal rent tax rate depend on, among other things, the output price. This is one possible explanation for the frequent occurence of tailoring of rent tax systems when large price changes occur.

The normative conclusions are less obvious. It is not desirable to have a tax system which needs tailoring in case prices change. A more elaborate model is needed to discuss what an optimal tax system is.

A Appendix: Proof of Lemma

The maximum of $T(t, r; \alpha)$ for $(t, r) \in [0, 1]^2$, and its dependence on α, for a given s.

For a fixed α, $T(t, r; \alpha)$ is twice continuously differentiable everywhere in $[0, 1]^2$. Thus it has at least one maximum there, and the maximum or maxima occur(s) at boundary or stationary points. Stationary points occur when $f(t; \alpha) = g(t)$, cf. (8) and (9).

The second-order conditions depend on

$$\frac{\partial^2 T}{\partial t^2} = -\frac{1}{c}, \tag{A.1}$$

which is always negative,

$$\frac{\partial^2 T}{\partial t \partial r} = -\frac{p^2}{2b}(1 - r), \tag{A.2}$$

which is always negative, and

$$\frac{\partial^2 T}{\partial r^2} = -\frac{2\alpha}{c}(2 - t), \tag{A.3}$$

which is always negative.

This excludes the possibility that a stationary point could be a local minimum, but a saddle point is possible. The necessary and sufficient conditions for a local maximum are satisfied if and only if

$$\frac{\partial^2 T}{\partial t^2}\frac{\partial^2 T}{\partial r^2} - \left(\frac{\partial^2 T}{\partial t \partial r}\right)^2 > 0,$$

which is satisfied if and only if

$$t < -2\alpha r^2 + 4\alpha r + 2(1 - \alpha) \equiv h(r; \alpha), \tag{A.4}$$

in which we have defined a new function h. Figure 2 shows the graph of h for $\alpha = 0.5$. The inequality is satisfied to the left of the graph, and in this area, any stationary point is a local maximum. To the right of the graph, any stationary point is a saddle point.

(INSERT FIGURE 2 HERE)

For all values of α the parabola in r, $h(r;\alpha)$, is symmetric around the line $r = 1$, and $h(1;\alpha) = 2$. A higher value of α makes the parabola narrower. For $\alpha < 0.5$, the parabola does not intersect the set $[0,1]^2$, which means that a stationary point is always a local maximum. For $\alpha \geq 0.5$, the parabola intersects the set, and we must check whether a stationary point is located to the left or to the right.

Let α^* be the value of α which gives tangency between the graphs of $f(t;\alpha)$ and $g(t)$, i.e., there is exactly one t which satisfies $f(t;\alpha^*) = g(t)$. Observe that when $\alpha < 1 - s/2$, the intersection of the graph of f with the horizontal axis is $t = \alpha + s/2$, which approaches 1 as α approaches $1 - s/2$.

We shall show the following, for $(t,r) \in [0,1]^2$:

(a): For $\alpha < 1 - s/2$ there is one stationary point, and this is the global maximum.

(b): For $1 - s/2 \leq \alpha < \alpha^*$ there are two stationary points, both along the graph of the decreasing $g(t)$. The south-eastern stationary point is a saddle point, while the north-western is a local maximum. The global maximum will in this case be one of two candidates: **(b1):** The interior local maximum, or **(b2)** the boundary point $(t,r) = (1,0)$. The first one is valid for low values of α within $[1 - s/2, \alpha^*)$, while the second one is valid for the remaining, higher values.

(c): For $\alpha > \alpha^*$ there are no stationary points, and the boundary point $(1,0)$ is the global maximum.

This suffices to prove the Lemma, since the effect of changes in α on the global maximum when it is interior, follows via the change in the graph of f, while the graph of g is fixed.

The arrows in figures 3–5 show the direction of increasing T values for each of the three cases.

(INSERT FIGURES 3–5 HERE)

Proof

(a): The stationary point is a local maximum. For $\alpha \in [0.5, 1 - s/2]$ one might worry about the stationary point being a saddle point, since the graph of the h function intersects the set $[0,1]^2$. But we show that none of the boundary points is a global maximum, so the stationary point must be. Consider the three candidates on the boundary: From $(0, 0.5)$ the T function value will increase by moving to the right. From $(\alpha + s/2, 0)$ the function value will increase by moving upwards. From $(1, s/2)$ the function value will increase by moving downwards. This implies that the global maximum is the stationary point.

(b): With two stationary points, they are separated by the graph of the h function, as illustrated in figure 4. The north-western point is a local maximum, while the south-eastern is a saddle point. Consider the three candidates on the boundary: From $(0, 0.5)$ the T function value will increase by moving to the right. From $(1, s/2)$ the function value will increase by moving downwards. Thus none of these two points is the global maximum. The remaining candidate, $(1, 0)$, cannot be ruled out, however. We resort to a numerical examination in order to determine which is the maximum between the north-western interior maximum and the corner point. It turns out that for α values close to $1 - s/2$, the interior candidate is the global maximum, while for larger α values, closer to α^*, the corner point is the global maximum.

(c): There is no interior stationary point. Consider the three candidates on the boundary: From $(0, 0.5)$ the T function value will increase by moving to the right. From $(1, s/2)$ the function value will increase by moving downwards. Thus $(1, 0)$ is the global maximum point.

This completes the proof.

References

Ball, R. & Bowers, J. (1983). Distortions created by taxes which are options on value creation: The Australian Resource Rent Tax proposal. *Australian Journal of Management,* 8(2), 1–14.

Brown, E.C. (1948). Business income, taxation, and investment incentives. In *Income, Employment , and Public Policy: Essays in Honor of Alvin H. Hansen,* New York: Norton, 300–16.

Campbell, H.F. (1980). The effect of capital intensity on the optimal rate of extraction of a mineral deposit. *Canadian Journal of Economics,* 13, 349–56.

Gordon, R.H. & MacKie-Mason, J.K. (1995). Why is there corporate taxation in a small open economy? The role of transfer pricing and income shifting. In Feldstein, M., Hines, J.R.,Jr. & Hubbard, R.G. (eds.) *The Effects of Taxation on Multinational Corporations,* Chicago: University of Chicago Press, 67–91.

Guesnerie, R. (1995). *A Contribution to the Pure Theory of Taxation,* Cambridge, UK: Cambridge University Press.

Haaparanta, P. & Piekkola, H. (1997). Taxation in an open economy and entrepreneurship. Unpublished paper, Helsinki School of Economics, Finland.

Mayo, W. (1979). Rent royalties. *Economic Record,* 55, 202–13.

Stiglitz, J.E. & Dasgupta, P. (1971). Differential taxation, public goods, and economic efficiency. *Review of Economic Studies,* 38, 151–174.

Sansing, R. (1996). Transfer pricing and the taxation of natural resource extraction. *Journal of Energy Finance & Development,* 1(1), 71–81.

Weichenrieder, A. (1996). Transfer pricing, double taxation, and the cost of capital. *Scandinavian Journal of Economics,* 98(3), 445–452.

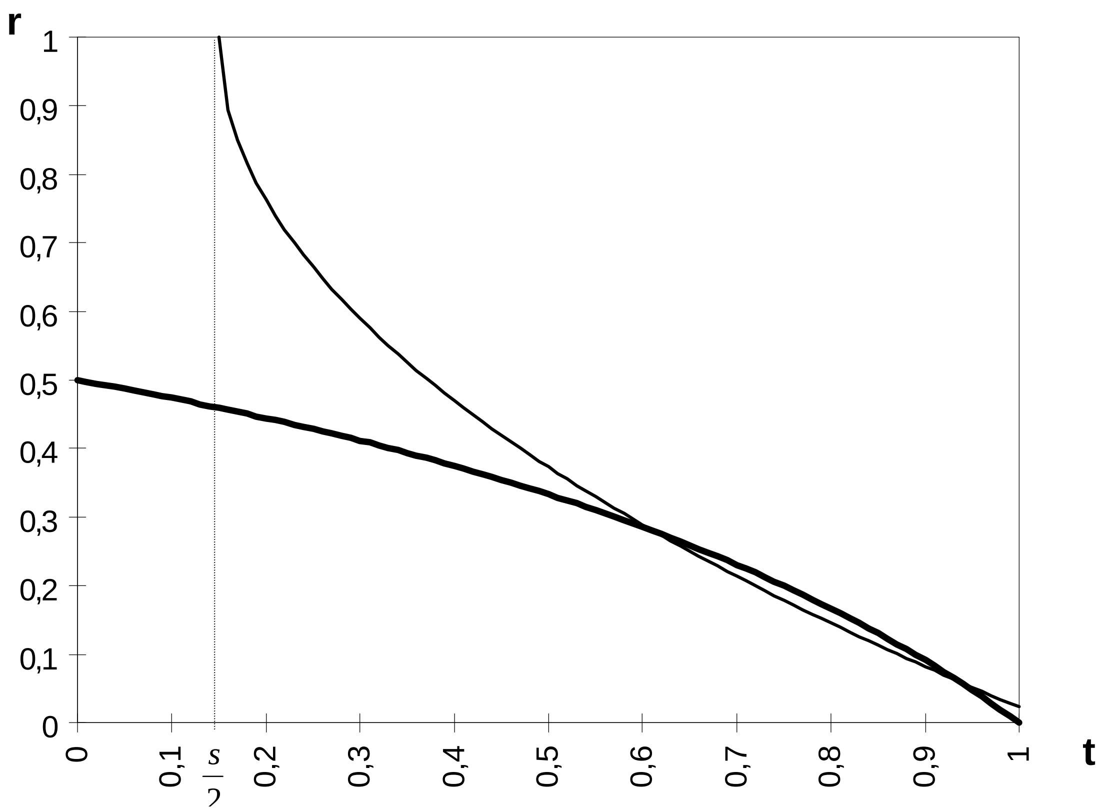

Figure 1: Graphs of g(t) (thickly drawn) and f(t; a) when a = 0.89 and s = 0.3.

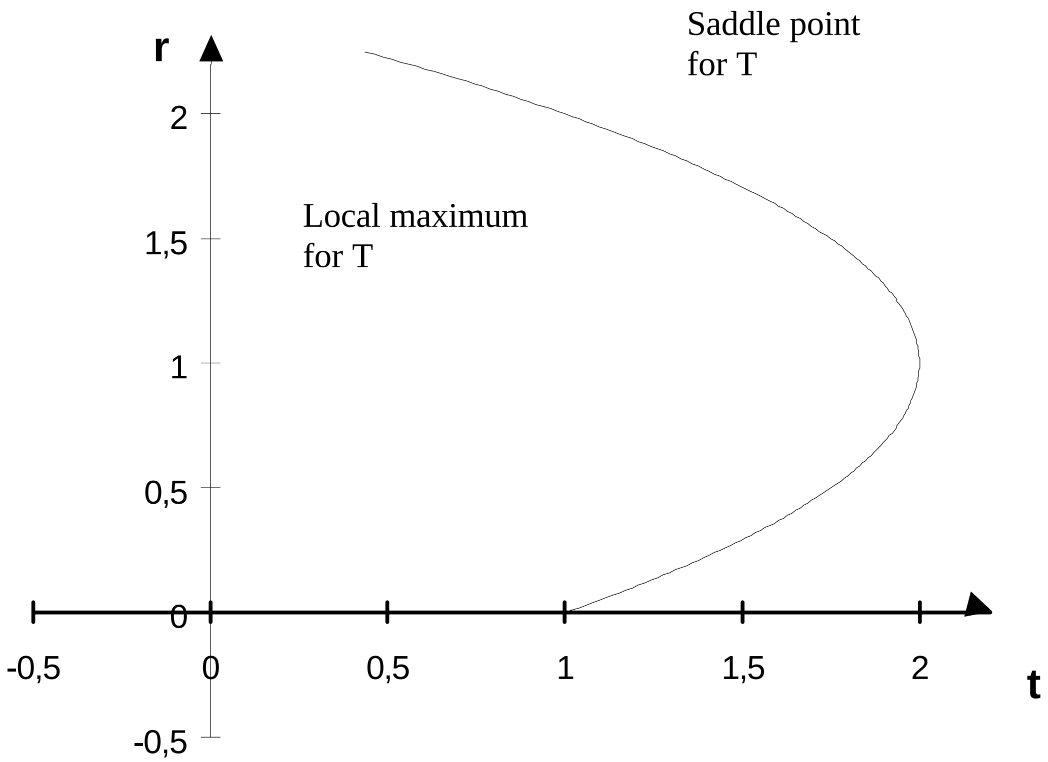

Figure 2: Graph of h(r; a) when a = 0.5, separating possible stationary points of T(t,r; a) by type.

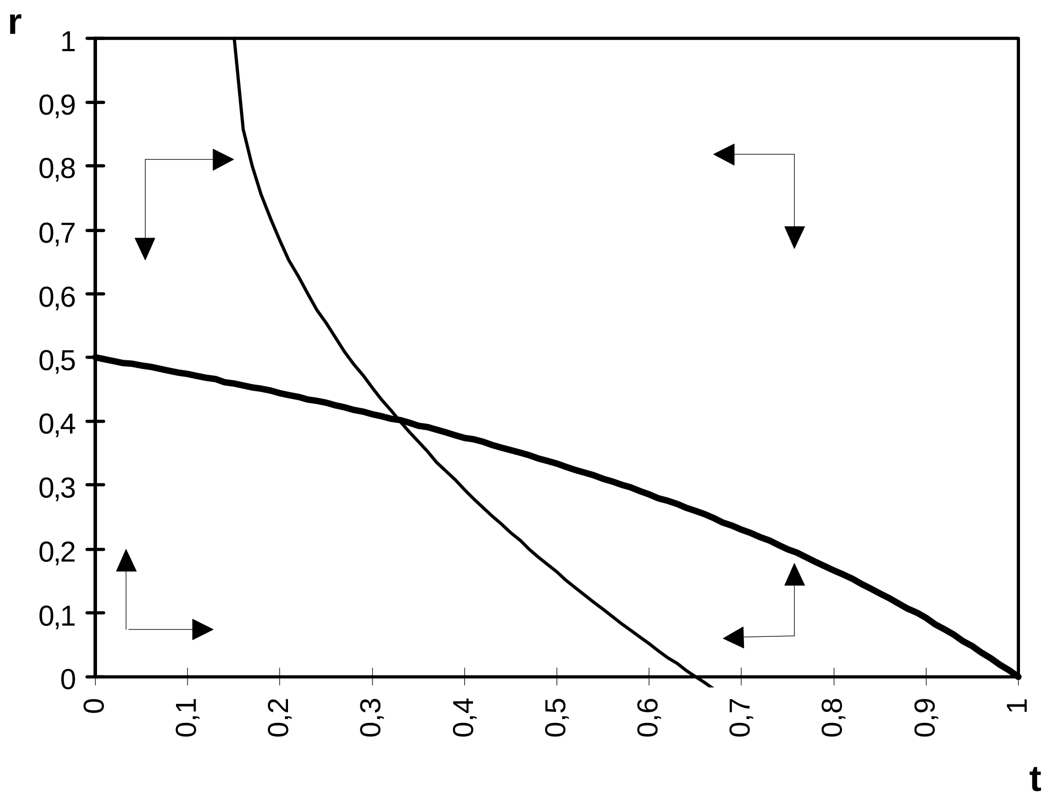

Figure 3: Graphs of g(t) (thickly drawn) and f(t; a) when there is only one intersection, a = 0.5.

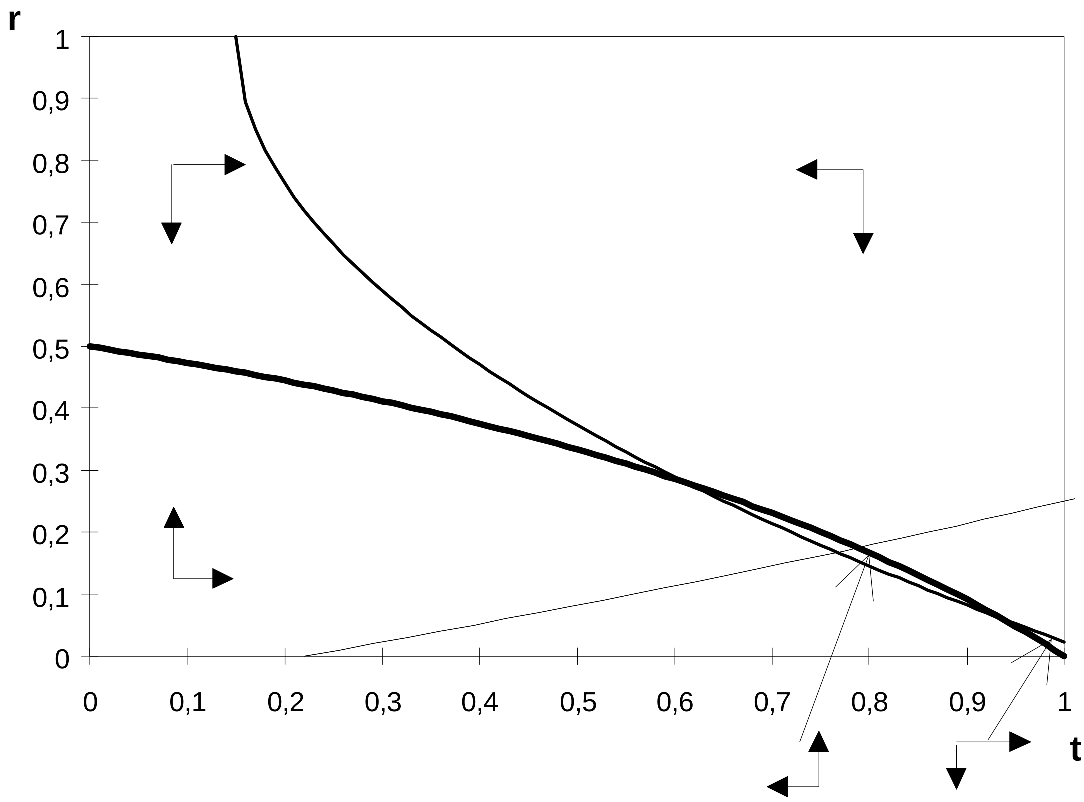

Figure 4: Graphs of g(t) (thickly drawn) and f(t; a) when there are to intersections, a = 0.89. The thinly drawn, increasing curve is the graph of h(r; a), indicating the types of the stationary points.

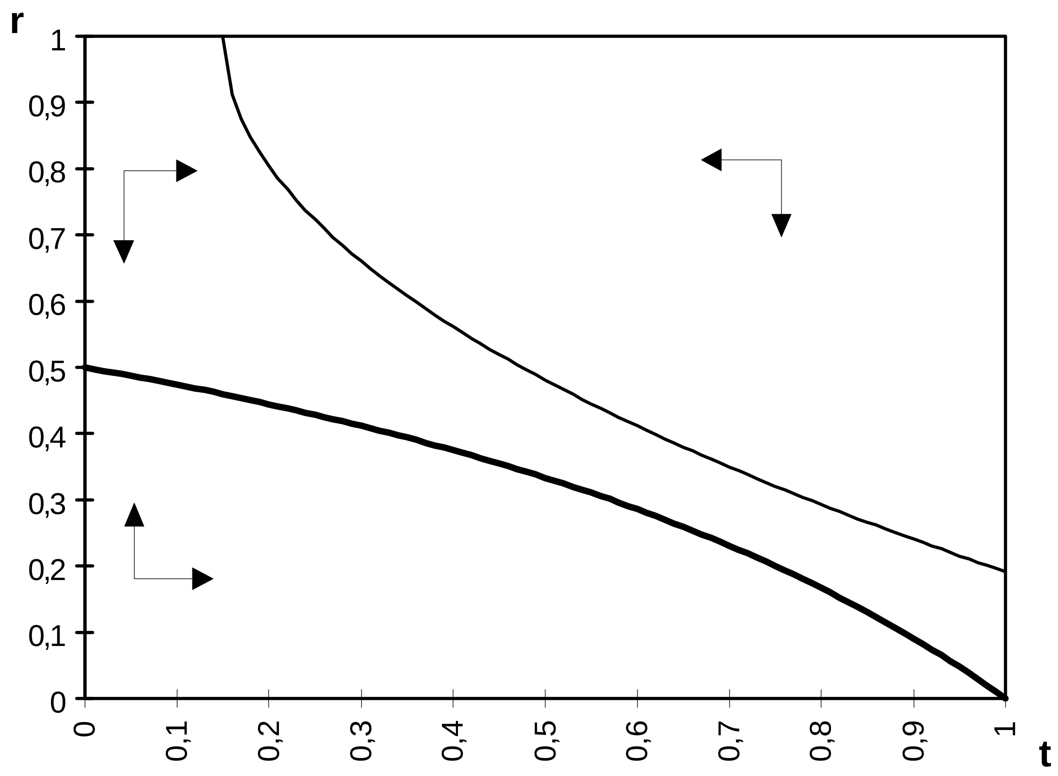

Figure 5: Graphs of g(t) (thickly drawn) and f(t; a) when there is no intersection, a = 1.3.