Do Firms Maximize their Cost of Finance?

by Hans-Werner Sinn

Finanzarchiv 50, 1993, pp. 21-36

Do Firms Maximize their Cost of Finance?

by

Hans-Werner Sinn *

1. The Problem

Predictions of the allocation effects of capital income taxation are extremely sensitive to assumptions about how the tax system affects the firm's marginal investment decisions. The Harberger literature, which tends to find huge distortions, maintains that the full burden of the double taxation of dividends enters the firm's cost of capital and it is often assumed that a wedge of the size of the corporate tax rate is driven between the marginal product of capital and the market rate of interest. These assumptions contrast sharply with approaches in the tradition of King (1974a, 1977), Stiglitz (1973), and Auerbach (1979) where the size of the wedge depends crucially on the firm's marginal source of finance. For example, in the book by King and Fullerton (1985), the cost of capital is a weighted average of the costs of debt financing, new share issues, and profit retentions, and, in general, the wedge is much smaller than the corporate tax rate. In the King-Fullerton approach, a wedge of the size of the corporate tax rate would only result if capital were assumed to be financed exclusively by new share issues. However, new share issues are the most expensive source of finance in the classical system of capital income taxation which the Harberger literature typically considers. Thus it seems that the findings of the Harberger literature depend implicitly on the assumption that firms maximize their cost of finance.

This unattractive conclusion has been defended by an interesting argument put forward by Hansson and Stuart (1985). The essence of this argument is that, unlike equity finance, debt finance involves invisible costs which, in a financial optimum, just compensate for the tax advantages at the margin [1]. The invisible costs are similar to the costs of rent seeking in public choice models and can, for example, be taken to represent the cost of avoiding bankruptcy or, more generally, the differential transactions

* This work is part of the CES and NBER research programs on public economics.

[1] The argument was made verbally and has, to the best of my knowledge, not been investigated in a formal model.

costs resulting from the use of debt in lieu of equity capital. According to Hansson and Stuart, the presence of these costs implies that, although firms actually *minimize* their cost of finance, they make their real investment decisions *as if they maximized* the cost of finance with regard to the visible costs and as if they used only equity at the margin.

If correct, this argument would rehabilitate the conclusions of the Harberger literature and would constitute a strong criticism of all models that assume the firm's cost of capital to be a weighted average of the direct, visible costs of different sources of finance. This paper presents a simple model of the firm's investment and financial decisions and compares the allocative implications of three alternative specifications of the firm's financial decision making. One of these is an interpretation of Harberger's theory, the second is in the spirit of King/Fullerton (1985) and Sinn (1985), and the third takes up Hansson and Stuart's invisible cost argument. The specifications cannot pay full justice to all aspects of the financial decision process the respective authors may have had in mind. However, they result from the attempt to compare, and cover in a systematic way, the essentials of three rival approaches to a central problem of taxation theory.

2. A Model of the Firm

Consider a forward-looking neoclassical firm that optimizes its real and financial decisions for all points in time. Assume the firm is a corporation and operates under the classical system of capital income taxation or under a partial imputation system. Let τ_r denote the corporate tax rate on retained profits, τ_d the (possibly lower) corporate tax rate on distributed profits, τ_c the effective personal capital gains tax rate on accrued share appreciation, and τ_p the personal tax rate that applies to both dividends after corporate tax and interest income shareholders can earn in the capital market. For each tax rate τ_i, let Θ_i denote the corresponding tax factor: $\Theta_i \equiv 1-\tau_i, i=r,d,c,p$. Assume that $0<\Theta_i\leq 1$ and that the overall tax burden on dividends exceeds that on retained profits which, in turn, exceeds the tax burden on personal interest income[2]: $\Theta_d\Theta_p \leq \Theta_c\Theta_r \leq \Theta_p$. This assumption implies that, from a tax perspective, debt is the cheapest source of finance and that profit retentions are cheaper than new share issues.

---

[2] See Sinn (1985, ch. 3) for an extensive discussion of this assumption in the context of the tax systems currently existing in the OECD countries. In the current U.S. American tax system, for example, this chain of inequalities can be taken to be satisfied strictly due to the fact that the personal tax rate is typically below the corporate tax rate, the same corporate tax rate applies to dividends and retained profits, and realized rather than accrued capital gains are included in the personal tax base.

Note that the assumptions include the special case $\Theta_d \Theta_p = \Theta_c \Theta_r = \Theta_p$ where the tax system treats all three sources of finance equally. This is the case of a non-corporate firm whose owner only faces one personal tax rate that applies to all kinds of capital income he earns. Although the terminology used refers exclusively to corporations, it is clear that the model set up in this section applies analogously to a non-corporate firm.

The firm's capital, K, consists of equity and debt, where debt occurs in the form of variable-interest bonds, B, issued to the public. Abstract for simplicity from all factors of production other than capital and assume the prices of the capital good and the output commodity to be constants normalized to unity[3]. Let $f(K)$, $f'>0$, $f''<0$, be the firm's output net of depreciation and $\varphi(B, K)$ its invisible cost function; the properties of this function will be specified later. The visible cost of debt finance is rB where r denotes the instantaneous market rate of interest; assume that r is bounded away from zero. The pre-tax dividend the firm pays out to its shareholders is net revenue, f, minus the visible and invisible costs of debt finance, $rB+\varphi$, minus the corporate tax on retained profits, T_r, minus the excess of net investment $\dot{K}$ over the inflowing funds from issuing debt $\dot{B}$ and new shares Q:

$$(1) \qquad D \equiv f(K) - [rB + \varphi(B, K)] - T_r - (\dot{K} - \dot{B} - Q)\,.$$

The dividend net of all taxes is

$$(2) \qquad D_n \equiv D\,\Theta_d \Theta_p\,,$$

and the corporate tax on retained profits is given by

$$(3) \qquad T_r = \tau_r [f(K) - rB - \varphi(B, K) - D]\,.$$

Inserting (3) into (1), solving for D, and using (2), one obtains

$$(4) \qquad D_n = \Theta_d \Theta_p [f(K) - rB - \varphi(B, K)] - \frac{\Theta_d \Theta_p}{\Theta_r} (\dot{K} - \dot{B} - Q)\,.$$

In line with Fisher's separation theorem, the firm chooses a policy that maximizes the market value of its shares, M. Let $M = mz$ where m is the value per share and z the number of shares. Assume that new shares are sold at their market price: $m\dot{z} = Q$, and that, at each point in time, a market equilibrium is characterized by:

$$\Theta_c \dot{m} z + D_n = r\,\Theta_p M$$

[3] The results and interpretations of the model do not change when other factors of production are introduced. However, for the Harberger problem as illustrated in Figure 1, at least one of these factors must be assumed to be immobile between the corporate and non-corporate sectors of the economy to ensure that both sectors employ positive amounts of capital even when the tax system does not treat them equally.

where $\Theta_c \dot{m} z$ is the net-of-tax capital gain on existing shares, D_n the net-of-tax dividend given by (4), and $r\Theta_p M$ the opportunity cost of holding shares in terms of foregone net interest income. Then it follows that $\dot{M} = \dot{z}m + \dot{m}z = Q + (r\Theta_p M - D_n)/\Theta_c$ which, upon integration, gives the following expression for the market value of shares at point in time t[4]:

$$(5)\qquad M(t) = \int_t^\infty \left[\frac{D_n(u)}{\Theta_c} - Q(u)\right]\left[\exp\int_t^u -\frac{\Theta_p}{\Theta_c}\, r(v)\,\mathrm{d}v\right]\mathrm{d}u\,.$$

It is assumed here that the time paths of $\{D_n\}$, $\{Q\}$, and $\{r\}$ are shaped in a way which ensures the convergence of the integrals and that $M(0)=0$ if the firm never issues new shares and never pays out any dividends.

Formally, the maximization problem of the firm can be expressed as

$$(6)\qquad \max_{\{\dot{K},\dot{B},Q\}} M(0)$$
$$\text{s.t.}$$
$$K(0) = K^0 > 0\,,$$
$$B(0) = B^0 < K^0\,,$$
$$\dot{K}, \dot{B}, Q \in C\,,$$

where C is a set of constraints on the controls $\dot{K}$, $\dot{B}$, and Q that has yet to be specified. The state variables of this control problem are K and B.

3. The Harberger Case

This section tries to reconcile Harberger's theory with the model of the firm set up in the last section, but it does not yet examine the invisible cost hypothesis. Assume, therefore, for the time being, that $\varphi = \varphi_B = \varphi_K = 0$.

Suppose that the firm is neither allowed to use debt nor profit to finance its net investment, but has to rely on equity injections through share issues:

$$\dot{B} \le 0\,,$$
$$Q \ge \dot{K}\,.$$

In this case, the current-value Lagrangean of problem (6) is

$$\mathscr{L} = \frac{D_n}{\Theta_c} - Q + \lambda_B \dot{B} + \lambda_K \dot{K} + \mu_1(Q - \dot{K}) - \mu_2 \dot{B}\,,$$

where D_n is given by (4), the λ's are the co-state variables associated with K and B, and the μ's are the Kuhn-Tucker multipliers associated with the two

[4] For a discussion of this and related market formulae see Sinn (1985, ch. III).

financial constraints. By the Maximum Principle, the following equations are among the necessary conditions for an optimum:

$$\frac{\partial \mathscr{L}}{\partial \dot{K}} = -\frac{\Theta_d \Theta_p}{\Theta_r \Theta_c} + \lambda_K - \mu_1 = 0\,, \tag{7}$$

$$\frac{\partial \mathscr{L}}{\partial Q} = \frac{\Theta_d \Theta_p}{\Theta_r \Theta_c} - 1 + \mu_1 = 0\,, \tag{8}$$

$$\dot{\lambda}_K - r\frac{\Theta_p}{\Theta_c}\lambda_K = -\frac{\Theta_d \Theta_p}{\Theta_c} f'(K)\,. \tag{9}$$

Adding (7) and (8) shows that $\lambda_K = 1$, and inserting this into (9) gives $r = \Theta_d f'$ or

$$f'(K) = e \tag{10}$$

where

$$e \equiv \frac{1}{\Theta_d}\,\mathrm{r} \tag{11}$$

is the cost of external equity finance. Equations (10) and (11) reveal that the corporate tax on dividends drives a wedge between the marginal product of capital and the market rate of interest.

When there is just one corporate tax rate for both retained and distributed profits as in the classical system of capital income taxation, then this confirms the Harberger assumption described in the introduction. In particular, (11) implies the Harberger-type of distortion between the corporate and non-corporate sectors of the economy [5].

To see this, assume that there are two sectors, X and Y, in the economy producing the same commodity, using the same production technology, and absorbing the aggregate capital stock $\bar{K}$:

$$\bar{K} = K^X + K^Y\,.$$

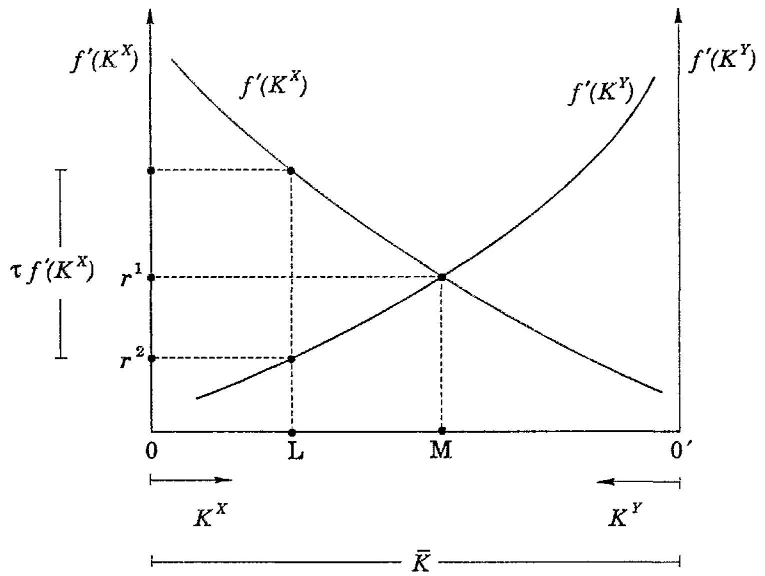
Let X indicate the corporate-sector facing the classical system of capital income taxation which has a uniform corporate tax rate: $\tau = \tau_d = \tau_r$, and let Y indicate the non-corporate sector where $\tau_d = 0$. Then (10) and (11) imply a capital market equilibrium of the kind

$$f'(K^X)(1-\tau) = r = f'(K^Y)\,.$$

This equilibrium is illustrated in Figure 1 which is the Kemp-McDougall diagram known from foreign trade theory. From left to right, this diagram shows the capital employed by the corporate sector X and from right to left

[5] See Harberger (1962, 1966).

Figure 1
The Harberger Model



that employed by the non-corporate sector Y. The corresponding marginal product curves are the downward and upward sloping curves, respectively. With a neutral tax system the equilibrium would be characterized by $K^X = \overline{\text{OM}}$ and $K^Y = \overline{\text{MO}'}$. However, when the corporate tax is introduced in addition to the personal tax, there is a distortion. The marginal product in the corporate sector rises and the marginal product in the non-corporate sector falls until a wedge of size $\tau f'(K^X)$ is created between them. The interest rate continues to equal the marginal product in the non-corporate sector and falls with it from r_1 to r_2. Capital of amount $\overline{\text{LM}}$ is driven from the corporate into the non-corporate sector to establish the new equilibrium.

4. More Financial Flexibility

An aspect of the assumptions made in the last section that could be objected to is that the firm has to use external sources of equity finance and is not allowed to build up equity capital through profit retentions. Moreover, it may be questionable to assume that no debt finance is possible at all.

Suppose, therefore, that the firm has somewhat more financial flexibility in that a maximum marginal debt-asset ratio σ, $0 \leq \sigma = \text{const.} \leq 1$, is imposed, while it is left open whether the required amount of equity finance is generated internally through profit retentions or externally through issuing new shares:

$$\dot{B} \leq \sigma \dot{K} .$$

Assume that the firm's earnings are large enough to provide for the required minimum equity formation $(1-\sigma)\dot{K}$. This assumption must be satisfied for a firm that pays out dividends, because such a firm could always generate more internal funds by not distributing any dividends if it wished. Assume, moreover, that

$$Q \geq 0$$

to exclude the possibility of firms avoiding the double taxation of dividends by distributing their profits through share repurchases. With these new constraints, the current-value Lagrangean of problem (6) is

$$\mathscr{L} = \frac{D_n}{\Theta_c} - Q + \lambda_B \dot{B} + \lambda_K \dot{K} + \mu_1 (\dot{K}\sigma - \dot{B}) + \mu_2 Q$$

where λ's and μ's have the same meaning as before except that there are now different constraints. Among other things, the Maximum Principle implies the following necessary optimality conditions:

$$\text{(12)} \qquad \frac{\partial \mathscr{L}}{\partial \dot{K}} = -\frac{\Theta_d \Theta_p}{\Theta_r \Theta_c} + \lambda_K + \mu_1 \sigma = 0 ;$$

$$\text{(13)} \qquad \frac{\partial \mathscr{L}}{\partial \dot{B}} = \frac{\Theta_d \Theta_p}{\Theta_r \Theta_c} + \lambda_B - \mu_1 = 0 ;$$

$$\text{(14)} \qquad \dot{\lambda}_K - r \frac{\Theta_p}{\Theta_c} \lambda_K = -\frac{\Theta_d \Theta_p}{\Theta_c} f' ;$$

$$\text{(15)} \qquad \dot{\lambda}_B - r \frac{\Theta_p}{\Theta_c} \lambda_B = \frac{\Theta_d \Theta_p}{\Theta_c} r .$$

As it can easily be established that $\dot{\lambda}_B = 0$[6], (15) implies that $\lambda_B = -\Theta_d$.

[6] Suppose, contrary to this contention, that $\dot{\lambda}_B < 0$ for some point in time t^*. Then, as $r >$ const. > 0, it follows from (15) that $\dot{\lambda}_B <$ const. < 0 for all $t \geq t^*$ such that $\lim_{t\to\infty} \lambda_B(t) = -\infty$. Because of (13), this violates the requirement that the Kuhn-Tucker multiplier μ_1 be non-negative. Thus, only the possibility $\dot{\lambda}_B \geq 0$ remains. Try now, alternatively, the case $\dot{\lambda}_B > 0$ for some point in time t^*. Here, (15) implies that $\lambda_B > 0$ in finite time. This, however, is impossible as, by the definition of a co-state variable, $\lambda_B(t) = \partial M(t)/\partial B(t)$ and it is obvious from (4) and (5) that $\partial M/\partial B < 0$ for all t. Thus, $\dot{\lambda}_B = 0$ for all t remains as the only possible case.

Inserting this into (13), solving for μ_1, and substituting the result for μ_1 in (12) gives

$$\lambda_K = \sigma\,\Theta_d + (1-\sigma)\,\frac{\Theta_d\,\Theta_p}{\Theta_r\,\Theta_c}\,.$$

Together with (14), this equation implies that

(16) $f'(K) = \sigma\, r + (1-\sigma)\, i$

where

(17) $i \equiv \dfrac{\Theta_p}{\Theta_c\,\Theta_r}\, r$

is the cost of internal equity finance.

Note that the basic assumption $\Theta_d\,\Theta_p \leq \Theta_r\,\Theta_c \leq \Theta_p$ is equivalent to

(18) $e \geq i \geq r\,.$

This chain of inequalities reveals a substantial difference between the Harberger equation (10) and equation (16). While the former says that the cost of capital equals the cost of the most expensive source of finance, the latter says that the cost of capital is a weighted average of the costs of the cheapest and the second cheapest sources of finance where the weights are the marginal debt-asset and equity-asset ratios. It is obvious that (16) will typically imply lower distortions in the economy than those depicted in Figure 1.

In fact, the difference can be dramatic. When firms enjoy full financial flexibility in the sense that $\sigma=1$: that is, if they can minimize their cost of finance under tax aspects, then $f'(K^X)=f'(K^Y)$ and there is no distortion at all. Similarly, when $\Theta_p=\Theta_c\,\Theta_r$ (and hence $i=r$), there is no distortion even when equity capital must be used at the margin ($\sigma<1$) and there is an unmitigated double taxation of dividends ($e=r/\Theta_d>i=r$). This demonstrates the sensitivity of the Harberger model with regard to a relaxation of the implicit assumption that firms maximize their cost of finance.

One of the reasons for the reduction or even disappearance of the tax wedge is that the double taxation of dividends does not affect the internal cost of equity finance as given by (17). This aspect of the model can be attributed to the trapped-equity view put forward by King (1974b), Auerbach (1979), and, most pointedly, Bradford (1981). The way the model is set up implies that it is not possible to get money out of the pockets of the firm into the pockets of the public except by payment of dividends subject to dividend tax. The opportunity cost of one dollar available for internal investment therefore is less than one dollar and this just compensates for the taxation of the returns of this investment.

An objection to the "trapped equity" view could be that the firm's profit is not large enough to provide for the required amount of internal equity formation $(1-\sigma)\dot{K}$. This objection is true and it may be relevant for new or rapidly growing firms. Note, however, that these firms must then choose not to pay dividends. Dividend paying firms can increase their internal funds by curtailing their dividend payments; they cannot be in a situation where they are forced to use expensive external equity sources.

The case of insufficient profits and firm growth from birth to maturity has been extensively studied in Sinn (1990, 1991 a). There it was shown that, with immature firms, the cost of capital exceeds the costs of internal (i) and external (e) equity funds as specified with (11) and (17). Thus dividend taxes create more distortions than both the Harberger and the King-Fullerton literature assume when they are *not* paid. However, these taxes are definitely neutral when they are paid, and this is the case to which (16) refers.

Another objection against the trapped equity view is that it is inconsistent with the frequent observation of share repurchases in the United States. Share repurchases, it is maintained, rehabilitate the old Harberger view since reductions in such repurchases, which are equivalent to new share issues, are available as a marginal source of finance. As explained in Sinn (1991 b, 1991 c), this objection is wrong. Published models with share repurchases tend to rehabilitate the old view simply because they tacitly assume that reduced share repurchases are the only marginal source of finance while dividends are the (or a) marginal use of profits. If reduced share repurchases and dividend cuts contribute to funding an investment project in the same proportions as share repurchases and dividends participate in profit distributions, the new view and the cost of capital formulae it generates stay perfectly valid.

5. The Invisible Cost Hypothesis

Turn now to the invisible cost hypothesis and forget about all previous constraints except for the constraint

$$Q \geq 0$$

which excludes the case of corporate distributions that evade the personal dividend tax[7]. Assume that $\varphi, \varphi_B, \varphi_{BB} \geq 0$, that $\varphi_K \leq 0$, and that φ and all of its derivatives are zero when $B=0$. Retain the assumption of sufficient profits for an internal creation of equity capital.

The Lagrangean of problem (6) is

$$\mathscr{L} = \frac{D_n}{\Theta_c} - Q + \lambda_K \dot{K} + \lambda_B \dot{B} + \mu Q\,,$$

[7] As mentioned above this constraint could be changed to $Q \geq -c \cdot D$, where c is a positive constant and D is dividends, without changing the results of this paper.

and an application of the Maximum Principle gives a list of necessary optimization conditions including

(19) $$\frac{\partial \mathscr{L}}{\partial \dot{K}} = -\frac{\Theta_d \Theta_p}{\Theta_r \Theta_c} + \lambda_K = 0\,,$$

(20) $$\frac{\partial \mathscr{L}}{\partial \dot{B}} = \frac{\Theta_d \Theta_p}{\Theta_r \Theta_c} + \lambda_B = 0\,,$$

(21) $$\dot{\lambda}_K - r\frac{\Theta_p}{\Theta_c}\lambda_K = -\frac{\Theta_d \Theta_p}{\Theta_c}(f' - \varphi_K)\,,$$

(22) $$\dot{\lambda}_B - r\frac{\Theta_p}{\Theta_c}\lambda_B = \frac{\Theta_d \Theta_p}{\Theta_c}(r + \varphi_B)\,.$$

Inserting λ_K and λ_B as given by (19) and (20) into (21) and (22), and using the definition (17), one obtains

(23) $$f'(K) = i + \varphi_K(B, K)$$

and

(24) $$i = r + \varphi_B(B, K)\,.$$

Equation (24) captures the idea of an interior solution of the debt-equity choice due to the existence of invisible costs. At the margin, the sum of the visible and invisible cost of debt financing just equals the cost of equity finance. However, equation (23) shows that this does not imply that the cost of capital equals the cost of the most expensive source of finance. On the one hand, as in (16), the cost of equity that appears in this equation is the internal cost of equity i rather than the higher external cost e. On the other hand, it is obvious from (23) that the cost of capital deviates from the cost of equity i if the invisible cost of debt finance depends on the firm's stock of capital K or, equivalently, its stock of equity $K-B$. If we make the plausible assumption that an increase in equity reduces the invisible cost of debt finance, i.e. $\varphi_K < 0$, then, contrary to the "as if maximization hypothesis",

(25) $$f'(K) < i\,.$$

Obviously, the fact that the overall costs of debt and equity are equal at the margin does not imply that firms choose their investment projects as if they maximized their visible cost of finance.

While (25) demonstrates the fallacy of the "as if maximization hypothesis" it does not clarify how the cost of capital is related to the composition of marginal sources of finance, which, as shown by (16), is very important when the amount of debt is effectively determined through constraints rather than through an interior solution with invisible costs such as (24). The use of the debt constraint was seen to imply that the cost of capital is i when only equity can be used at the margin, is r when it is admissible to use only debt at the margin, and is between i and r when some, but not full debt financing is allowed. Does a similar result hold in the presence of invisible costs or will these costs induce a fundamental change of the cost of capital formula? For example, will the cost of capital necessarily exceed the interest rate even in the case where only debt is used at the margin?

To investigate these questions it is necessary to see what (24) implies for the marginal sources of funds and how this relates to (23). Let

$$\sigma(B,K) \equiv \left.\frac{\mathrm{d}B}{\mathrm{d}K}\right|_{i-r=\varphi_B(B,K)=\text{const.}}$$

be the firm's optimal marginal debt-asset ratio at a particular point in (B,K) space assuming that the wedge between i and r that corresponds to this point is being kept constant. Obviously, (24) implies that

$$\sigma(B,K) = -\frac{\varphi_{BK}(B,K)}{\varphi_{BB}(B,K)}. \tag{26}$$

Using this equation and noting that $\int_0^B \varphi_{BK}(u,K)\,\mathrm{d}u = \varphi_K(B,K)$, condition (23) can be rewritten as

$$\begin{aligned} f'(K) &= i + \int_0^B \varphi_{BK}(u,K)\,\mathrm{d}u \\ &= i - \int_0^B \varphi_{BB}(u,K)\,\sigma(u,K)\,\mathrm{d}u\,. \end{aligned}$$

This expression, in turn, can be transformed to

$$f'(K) = i - \int_0^B \varphi_{BB}(u,K)\,\mathrm{d}u \cdot \sigma^*(B,K)\,,$$

where σ^* is a weighted average of the marginal debt-asset ratios for alternative values of B in the range between zero and its actual value:

$$\sigma^* \equiv \int_0^B \frac{\varphi_{BB}(u,K)}{\int_0^B \varphi_{BB}(v,K)\,\mathrm{d}v}\,\sigma(u,K)\,\mathrm{d}u\,. \tag{27}$$

To simplify further, use the fact that $\int_0^B \varphi_{BB}(u,K)\,\mathrm{d}u = \varphi_B(B,K)$ and substitute $i-r$ for φ_B accordingly to (24). This gives

$$f'(K) = \sigma^* r + (1-\sigma^*)\,i\,. \tag{28}$$

Equation (28) shows that the composition of the firm's marginal sources of funds matters in the "invisible cost model". As in (16), the cost of capital is a weighted average of the cost of debt and the cost of internally created equity where the weights are marginal debt-asset and equity-asset ratios.

Note, however, that these ratios are not necessarily equal to the *actual* marginal debt-asset and equity-asset ratios. By definition, σ^* is an average of the hypothetical marginal debt-asset ratios for alternative values of B (or alternative values of $i-r$), given the value of K. Thus σ^* cannot, in general, be equated to the marginal debt-asset ratio observed at one particular value of the wedge between the cost of equity and the cost of debt, let alone with the average debt-asset ratio B/K. On the other hand, it is obvious that, even in the invisible cost model, it *is* the composition of marginal sources of finance that determines the cost of capital. The invisible cost assumption implies that the statistical procedure of estimating the composition of marginal sources of funds would be somewhat more complicated than it has been anyway, but it does not call for a fundamental revision of approaches that assume this composition to be determined by exogenous constraints.

Consider a special example to illustrate the similarity between the two rival approaches. Assume that

$$\varphi(B,K) = \Phi(B-\alpha K),\ \Phi', \Phi'' \geq 0,\ 0 \leq \alpha \leq 1,\ \Phi = \Phi' = \Phi'' = 0 \quad \text{for} \quad B = 0\,.$$

With $\alpha=0$, this example captures the case where the invisible cost is only a function of the firm's stock of debt and, with $\alpha=1$, the case where it is only a function of its stock of equity. In general, the formulation allows the invisible cost of debt financing to depend on a linear combination of debt and equity as $B-\alpha K = B(1-\alpha)-(K-B)\alpha$. Because $\varphi_{BK} = -\Phi'' \cdot \alpha$ and $\varphi_{BB} = \Phi''$, it follows from (26) and (27) that

$$\sigma = \sigma^* = \alpha$$

and hence (28) reduces to equation (16) that was derived in the previous section:

$$f'(K) = \sigma r + (1-\sigma)\, i\,.$$

The case implicitly addressed by the "as if maximization hypothesis" is $\alpha=\sigma=0$. The firm uses both debt and equity as implied by (24), but at the margin it uses only equity. The marginal cost of capital therefore equals the cost of equity i. (Remember that the latter is the lower of two possible costs of equity.)

An equally illuminating extreme is $\alpha=\sigma=1$. Again the firm employs both debt and equity capital, but at the margin it only employs debt. Despite an interior solution of the debt-equity choice the cost of capital equals the rate of interest on debt r.

Between the extremes, when $0<\alpha<1$, the specific cost function assumed generates more reasonable intermediate values of σ and a cost of capital between i and r. This is what models that employ exogenous constraints on the debt-equity choice predict.

The second case, $\alpha = \sigma = 1$, is important in so far as it is the case where no Harberger distortion occurs between the corporate and non-corporate sectors of the economy. For the corporate sector, it holds that $f'(K^X) = r$ and, for the non-corporate sector, the similar condition $f'(K^Y) = r$ holds true at any rate as for this sector $i = r$. In Figure 1, point M characterizes the allocation of resources and the market rate of interest is r_1, as in the absence of taxation. All distortions show up in the financial sphere in that the corporate sector is driven into debt financing while the non-corporate sector uses only little or no debt. If $i > r$ holds in the corporate sector, there is a welfare loss in that this sector chooses to bear the invisible cost of debt financing in order to avoid the tax discrimination against profit retentions. This welfare loss, however, results exclusively from financial distortions and is clearly not the one Harberger described.

6. Concluding Remarks

The conclusion of this paper is that the invisible cost hypothesis does not imply that firms choose their investment projects as if they maximized their cost of finance. Instead, this hypothesis was shown to imply that firms behave as if they minimized their cost of finance subject to the constraint that at least some given proportion of equity capital must be used at the margin.

Given the tax systems considered, this implied on the one hand that the cost of equity capital was lower than that assumed in the Harberger model. On the other hand, the cost of capital as such was seen to be a weighted average of the visible costs of equity and debt where, by assumption, the latter was the lowest cost of all.

The reason for the second part of this result was that the invisible cost hypothesis will, in general, imply an interior solution not only to the firm's average, but also to its marginal debt-equity choice. This fact is responsible for the similarity between the solutions resulting from the invisible cost hypothesis and the imposition of constraints on the firm's marginal sources of funds. The Hansson-Stuart conjecture implicitly assumes an invisible cost function with which equity is the only marginal source of finance.

The reason for the first part – the comparatively lower cost of equity – was that the cost of capital turned out not to be affected by the double taxation of dividends, and this in turn depended on the assumption that firms are able to pay out dividends. As has been clarified in earlier literature, firms that pay dividends and dividend taxes prove that they have enough profits to generate equity internally. They cannot be forced to rely on equity injections from the public against which the double taxation discriminates so heavily. Dividend taxes are neutral when they are paid.

To illustrate the allocative meaning of the tax wedges, the paper discussed Harberger's original problem of distortions between the corporate and non-corporate sectors of the economy. However, the analysis also has implications for other problems such as the international and intertemporal allocation of resources. Here, too, the wedge between the marginal product of capital and the market rate of interest can bring about significant distortions. If the residence principle is applied to the taxation of border-crossing interest income flows, there is a tendency to equalize the national interest rates and capital movements similar to those illustrated in Figure 1 can be induced between countries by unilateral tax reforms. Moreover, the wedge between the marginal product of capital and the market rate of interest is one of two components of the overall wedge between the marginal product of capital and the consumer rate of time preference which is essential for the intertemporal allocation of resources in general and the volume of savings in particular. For all of these problems, the invisible cost model as formulated in this paper, would predict distortions that will probably not be comparable to those resulting from approaches in the Harberger tradition.

The rigid and non-optimal financial behavior, which the Harberger theory assumes, forces all distortions to show up as distortions in the allocation of resources among ordinary productive uses. This is not only objectionable from a theoretical point of view, it also contradicts the empirical facts. After all, retained profits rather than new share issues are the dominant channel of equity formation, and debt undoubtedly contributes to financing investment projects. Firms obviously do enjoy some financial flexibility. This flexibility serves as a kind of buffer that helps cushion the real economy against the blows that have traditionally been believed to be imposed by the tax system. However, such flexibility may not be costless. Models that assume the financial decisions to be governed by constraints neglect these costs. It is an advantage of the invisible cost approach that it makes them explicit and indicates a way they can, in principle, be incorporated into the welfare analysis of tax reforms.

References

Auerbach, A. J., 1979, Wealth Maximization and the Cost of Capital, Quarterly Journal of Economics, 94, pp. 433–436.

Bradford, D., 1981, The Incidence and Allocation Effects of a Tax on Corporate Distributions, Journal of Public Economics, 15, pp. 1–22.

Hansson, I., and Ch. Stuart, 1985, The Taxation of Income from Capital. A Comparative Study of the United States, the United Kingdom, Sweden, and West Germany (review article), Journal of Political Economy, 93, pp. 826–831.

Harberger, A. C., 1962, The Incidence of the Corporation Income Tax, Journal of Political Economy, 70, pp. 215–240.

–, 1966, Efficiency Effects of Taxes on Income from Capital, in: M. Krzyzaniak (ed.), Effects of Corporation Income Tax, Detroit, pp. 107–117.

King, M. A., 1974a, Taxation and the Cost of Capital, Review of Economic Studies, 41, pp. 21–35.

–, 1974b, Dividend Behavior and the Theory of the Firm, Economica, 41, pp. 25–34.

–, 1977, Public Policy and the Corporation, London-New York.

King, M. A., and D. Fullerton (eds.), 1985, The Taxation of Income from Capital: A Comparative Study of the United States, the United Kingdom and West Germany, Chicago.

Sinn, H.-W., 1985, Kapitaleinkommensbesteuerung, Tübingen. English edition: Capital Income Taxation and Resource Allocation, Amsterdam-London-New York, 1987.

–, 1990, The Vanishing Harberger Triangle, Journal of Public Economics, 45, pp. 271–300.

–, 1991a, Taxation and the Birth of Foreign Subsidiaries, Woodrow Wilson School, Princeton, Discussion paper No. 66. Forthcoming in: H. Herberg and N. V. Long (eds.), Current Issues in International Trade Theory: Papers in Honour of Murray C. Kemp, Ann Arbor, Michigan, 1993.

–, 1991b, Taxation and the Cost of Capital: The "Old" View, the "New" View, and another View, in: D. Bradford (ed.), Tax Policy and the Economy, 5, Cambridge, Mass., pp. 25–54.

–, 1991c, Share Repurchases, the 'New' View, and the Cost of Capital, Economics Letters, 36, pp. 187–190.

Stiglitz, J., 1973, Taxation, Corporate Financial Policy, and the Cost of Capital, Journal of Public Economics, 2, pp. 1–34.

Abstract

It is a popular criticism of the weighted average formulation of the cost of capital under distortionary taxation that this formulation underestimates the true cost because it does not take account of the endogeneity of the firm's financial decisions. It is argued that, with an interior financial choice, all sources of funds are equally expensive and that, therefore, only the source of funds against which the tax system discriminates most heavily determines the cost of capital. Using an explicit model of intertemporal optimization with an endogenous financial choice this paper demonstrates the fallacy of that view. It is shown that the firm's cost of capital can be expressed as a weighted average of the direct, tax-inclusive costs of the used sources of finance although the financial decisions are endogenous.

Kurzfassung

In der Literatur zur Kapitaleinkommensbesteuerung werden die Kapitalkosten bisweilen als gewogenes Mittel der Kosten verschiedener Finanzierungswege dargestellt. Diese Formulierung wurde kritisiert, weil sie die Endogenität der Finanzierungsentscheidung nicht berücksichtigt, und es wurde vermutet, daß bei einer endogenen Finanzierungsentscheidung die Kapitalkosten nur durch die teuerste Finanzierungsart bestimmt werden, weil im Optimum alle marginalen Finanzierungskosten gleich sind. Auf der Basis eines expliziten intertemporalen Optimierungsmodells mit endogener Finanzplanung wird in diesem Aufsatz gezeigt, daß diese Vermutung falsch ist. Es wird nachgewiesen, daß die Kapitalkosten der Firma als gewogenes Mittel der direkten, unter Einschluß der Steuereffekte berechneten Kosten der gewählten Finanzierungswege ausgedrückt werden können, obwohl die Finanzierungsentscheidung endogen bestimmt wird.

Prof. Dr. Hans-Werner Sinn
Universität München
Volkswirtschaftliche Fakultät
Schackstraße 4
W-8000 München 22
Deutschland