A PARADOX IN THE THEORY OF SECOND BEST

by

Roger L. Faith
Virginia Polytechnic Institute and State University

and

Earl A. Thompson
University of California, Los Angeles

Discussion Paper No. 161
September 1979

A PARADOX IN THE THEORY OF SECOND BEST

It is now conventional to conclude that when an uncontrollable monopolist exists in an otherwise competitive economy, positive taxes or subsidies on the controllable, competitive sectors of the economy are generally required to move the economy to a social optimum (e.g., Lipsey-Lancaster (1956), Davis-Whinston (1965), Bohm (1967), Negishi (1972) and Allingham-Archibald (1975)). This basic result of the general theory of second best is generated from models that take the existence of uncontrollable monopoly as given and unrelated to the ability of the government to determine policy over the rest of the economy. Yet why should there be a constraint on monopoly policy to begin with? And, if monopoly policy is in fact constrained, does the institutional environment that permits the constraint imply additional policy restrictions that are not picked up by the formal mathematics of "second best" constrained maximization?

After specifying a political model generating an uncontrollable monopolist, this paper shows that conventional second best theory implies a logical paradox that can be removed only by imposing additional restrictions on the feasibility of controlling other, competitive sectors of the economy. Moreover, the additional restrictions imply the optimality of a laissez faire policy. The argument generalizes to all forms of market failure; it is not at all restricted to monopoly distortions. Generalizing further, when some of the controllable sector contain distortions, the result of adding the implicit political restrictions is still that the standard optimality conditions apply to the controllable sectors. This means that classical, piecemeal policy can once again be employed, thereby eliminating the conventional second-best implication that an economist must model every imperfection in the economy in order to logically derive a single policy recommendation.

I. The General Model

Consider an n-industry economy generating an equilibrium set of outputs, $x=(x_1,\ldots,x_n)$, in which the output of each industry depends on the set of per unit tax (or subsidy) rates on the various outputs, $(t_1,\ldots,t_n)$. Outputs here include "leisure" so that x does not vary with the general level of taxes. To avoid the corresponding indeterminacy in optimal tax rates, we standardize taxes by setting a particular tax, t_n, equal to zero. This means that all taxes are relative to taxes on the n^{th} output. Hereafter, we use $t=(t_1,\ldots,t_{n-1})$ to denote the set of variable taxes. There exists a government that selects t so as to maximize a collective utility function,[2] $U[x(t)]$, subject to the social transformation function, $T[x(t)]=0$, and a tax feasibility constraint, $G(t)=0$.[3] A "first best" solution to this maximization problem exists when $G(t)=0$ is not a binding constraint, e.g., when $G(t)\equiv 0$ so that any set of taxes satisfies the tax feasibility constraint. When $G(t)=0$ is a binding constraint, then there is a "second-best" solution.[4]

When the constraint is not binding -- assuming now that $U[\cdot]$, $T[\cdot]$, $x(\cdot)$ are everywhere differentiable, that the Jacobian of $x(\cdot)$ does not vanish at the optimum, and that $\partial U/\partial x_n$ is always positive -- the traditional first-order conditions for a first-best solution, i.e.,

$$(1) \qquad \frac{\partial U/\partial x_j}{\partial U/\partial x_n} = \frac{\partial T/\partial x_j}{\partial T/\partial x_n}\ , \ j = 1,2,\ldots,n-1,$$

and easily derived. Adding that $U[\cdot]$ and $T[\cdot]$ are quasi-concave, these conditions are sufficient as well as necessary for an optimum.

II. The Source of the Policy Constraint

What is the source of the G(t)=0 constraint? While much of the applied literature appealing to "second best" arguments suggests all sorts of informational limitations on government decision makers, this is hardly reasonable given that the decision maker must know the U[t] and T[t] functions in order to ascertain the second best policy.[5] To be consistent with the information implicit in the above maximization problem, the constraint must instead come from externally imposed restrictions on the policies of informed decision makers. But why would anyone impose a restriction leading to Pareto inferior policies? Or, to cite the most common case considered in the literature, why wouldn't everybody vote for allowing an informed bureaucrat to subsidize the output of a single monopolist to induce an optimal output and simultaneously apply a lump-sum tax on him so as to make everybody better off? The answer must be that the efficient, per unit subsidy would not -- in fact -- be complemented by a lump-sum tax sufficient to compensate everybody. But then why wouldn't our informed bureaucrat compensate the potential losers, thereby preventing the restriction on his ability to achieve a first-best optimum and making everybody better off? The answer must be that transaction costs between the bureaucrat and the potential losers preclude it. A plausible reason for the inability of potential losers to inexpensively obtain an effective compensation commitment from a bureaucrat is that neither these voters nor their representatives can practically discover whether or not the voters have, in fact, been compensated. Under such conditions, it pays the potential losers to vote for a policy constraint to prevent the monopolist, and possibly other voters, from exercising their political power over the government decision maker. A contemporary example of such an effect is that our voter representatives in the U.S. have not allowed the prices of oil products to rise to levels that could make us all better off because they believe that the oil companies would not -- in fact -- pay sufficient lump-sums to compensate the great majority of their constituents for the price increases. Voters who do not themselves know the U[t] and T[t] functions are rationally suspicious of any policy that may harm them and help a special interest group. The voters' skepticism is based on their realization that, given their information disadvantage, a self-interested government official has an incentive to "sell out" to the special interest group and impose an overly small lump-sum tax on it. As a result, a rational electorate may impose binding policy constraints on its relatively informed bureaucrats. A very similar argument for the imposition of policy constraints has been developed by Brennan and Buchanan (1977).

For an overproduction distortion, such as arises with pollution-type external diseconomies, the argument would be that a per unit tax on the overproducers, together with a lump-sum subsidy to overproducers sufficient to prevent them from blocking the tax-subsidy bill, would not gain the support of voters at large because voters cannot measure the lump-sum compensation actually paid to producers and would not trust bureaucrats to make such compensation sufficiently small that the voters would gain from the tax-subsidy policy.

III. The Conventional Second Best Problem

Following the literature, suppose: (1) that n-1 of the industries, say industries 2,...,n, are standard competitive industries while the remaining industry contains a "distortion," i.e., an incentive system preventing the marginal equality in (1) from holding for j=1 at $t=\theta$, and (2) that it is not feasible to directly subsidize or tax the output of industry 1. The tax feasibility constraint therefore reads: $G(t)\neq 0$ whenever $t_1\neq 0$ and $G(t)=0$ for all $t_2,\ldots,t_{n-1}$ whenever $t_1=0$. Industry 1 is said to be "uncontrollable" while the others are "controllable."

To make the discussion more concrete, we shall make industry 1 a simple, nondiscriminating monopoly. To derive the monopolist's rational output rule, first consider the n-2 equilibrium conditions for the competitive industries relative to industry n for a given output of industry 1:

$$(2)\quad \frac{U_j}{U_n} = \frac{T_j(x_1;x_2^*,\ldots,x_n^*)}{T_n(x_1;x_2^*,\ldots,x_n^*)} + t_j \qquad j = 2,\ldots,n-1,$$

where x_2^* through x_n^* are the equilibrium outputs under competition, given x_1 and t. From (2) and $T[x]=0$, we obtain each competitive industry's output as a function of x_1 and t, or

$$(3)\quad x_j^*=x_j^*(x_1,t), \qquad j = 2,\ldots,n.$$

In contrast to the standard models, which assume that x_j^* is constant, our monopolist takes into account the n-1 equilibrium response functions expressed in (3) (Cf. Allingham and Archibald (1975) and Negishi (1972)). While the assumption that the monopolist knows all the general equilibrium effects of his output choice is unrealistic, it captures the idea that the monopolist is at least as aware of the interdependence among industries as is the government and cannot be "fooled" by indirect tax-subsidy policy imposed upon substitute or complementary commodities. Profit to the monopolist, in terms of Commodity n, is therefore written:

(4) $$\pi_1 = x_1 \frac{U_1}{U_n} [x_1, x_2^*(x_1,t), \ldots, x_n^*(x_1,t)] - C(x_1),$$

where $C(\cdot)$ is 1's total cost function in terms of Commodity n. We are following here the standard convention of disregarding the effect which the monopolist's output has on the relative prices of the goods which he consumes, and thus its effect on his optimal output choice. Implicitly, we are assuming that the monopolist consumes only the numeraire commodity. We further simplify the problem by assuming that the monopolist has no significant effect on input prices. Therefore, while his output choices significantly affect the outputs of others, the choices do not significantly affect his factor prices. In this case, of course, $C'(x_1) = T_1(x)/T_n(x)$.

Maximizing (4) by choice of x_1, given t, yields the first-order condition relative to the numeraire commodity produced by Industry n:

(5) $$\frac{U_1}{U_n} + x_1 \left[\sum_{j=2}^{n} \left(\frac{\partial(\frac{U_1}{U_n})}{\partial x_j^*} \frac{\partial x_j^*}{\partial x_1} \right) + \frac{\partial(\frac{U_1}{U_n})}{\partial x_1} \right] = \frac{T_1(x)}{T_n(x)}.$$

We assume that the term in large brackets is negative so that our monopolist is sufficiently conventional that he undervalues his output.

From (5) we can write 1's equilibrium output as a function of the tax rates, or

(6) $$x_1^* = f_1(t)$$

Inserting (6) into (3) yields

(7) $$x_j^* = f_j(t), \qquad j = 2,\ldots,n.$$

Note that the resulting equilibrium, x*, precludes a first best solution when $t_1 = 0$. A non-zero tax imposed on competitive industry i ($i \neq n$) violates (1); and, if zero taxes are applied everywhere, (5) violates the first best conditions in (1).

Substituting (6) and (7) into U[x], the government's second-best optimization problem is

$$(8) \qquad \max_{t} U[f_1(t), f_2(t), \ldots, f_n(t)], \quad \text{given } t_1 = 0.$$

Since the output functions in (7) already satisfy the transformation function, the latter is not included as an independent constraint on the maximization problem.

The necessary marginal conditions for (8) are:

$$(9) \qquad \sum_{k=1}^{n-1} \left(\frac{\frac{\partial U}{\partial f_k}}{\frac{\partial U}{\partial f_n}} - \frac{\frac{\partial T}{\partial f_k}}{\frac{\partial T}{\partial f_n}} \right) \frac{\partial f_k}{\partial t_i} = 0 \qquad i = 2, \ldots, n-1.$$

Equation (9) says that taxes are changed until the sum of the excesses of marginal social benefits over marginal social costs in each industry times the induced change in that industry's output is equal to zero. This condition must hold simultaneously for each variable tax rate. So, starting with zero taxes everywhere, while each sum has its last n-2 terms zero because of the competition in the corresponding industries, each of the first terms is generally non-zero because the bracketed term in (5) is non-zero (otherwise, no distortion could exist) while $\partial f_1/\partial t_i$ is generally non-zero. Hence, with $\partial f_1/\partial t_i$ non-zero for some i, a second best solution requires non-zero taxes on the controllable industries. Such taxes will create inequalities between marginal benefit and marginal cost in these industries.

Zero taxes on all controllable industries (i.e., laissez-faire policy) is indicated if and only if there is no ultimate effect of a tax change on the monopolist, i.e., if and only if $\partial f_1/\partial t_i=0$, all $i\geq 2$. This would occur, for example, if Industries 1 and n were separable from the rest of the economy in both the collective utility and commodity transformation function (see Faith-Thompson (1977)).

The above conclusions are similar to those found in the literature. Our model differs from previous analyses only in that (1) our monopolist accounts for the effects of changes in his output on the outputs of other industries and (2) our marginal conditions in (9) reflect the system's response to the available policy variables rather than assuming that the policy maker can directly select industry outputs. However, our marginal conditions expose a paradox in the conventional second best solution.

IV. The Paradox

Examining these marginal conditions, it is easily seen that any change in taxes moving the economy from a laissez faire position toward the above second best optimum requires an induced increase in the monopolist's output.[6] That is, conventional second best policy intervention works by inducing an expansion in the monopolist's output, reducing the degree of monopoly distortion by an amount that exceeds the distortions created between the competitive industries. But increasing a monopolist's output generally requires a relative increase in demand or a relative decrease in variable costs, both of which serve to increase the monopolist's profit.[7] Therefore, recalling that the reason for the infeasibility of first best policy is that voters fear a redistribution of wealth to the monopolist, the second best intervention should also be infeasible. It is, to say the least, paradoxical that the government cannot induce an increase in monopoly output by increasing monopoly profits via a direct subsidy but can induce the output increase by generating a like increase in monopoly profits via more costly, distortion-creating, indirect taxes and subsidies.

The paradox applies to any imperfection, not just monopoly. Suppose industry 1 is, rather than a monopoly, a competitive industry generating an external economy and that a per unit subsidy is precluded because the consumer-voters fear an insufficiently low lump-sum tax on industry 1 because of the latter's ability to form a powerful lobby. Then the second-best policy, using (9), would induce the industry to expand its output. But the only way to do this is to tax substitutes or subsidize complements in order to increase prices or reduce costs in industry 1. So there is still, in effect, a subsidy to industry 1. The only difference is that the feasible subsidy is more expensive than the infeasible one in that it generates new distortions.

V. Removing the Paradox

Summarizing the above, the conventional feasibility constraint creates a paradox in that, while the only plausible argument we have been able to find for the constraint is that voters do not trust a bureaucrat to lump-sum tax an underproducing sector sufficiently to compensate them for giving per unit subsidies to the sector, correct second-best policy intervention imposes a set of taxes and subsidies on controllable sectors that is beneficial only because it indirectly subsidizes the uncontrollable sector. If we have accurately captured the source of the standard constraint, then the paradox can be removed only by imposing sufficient additional tax constraints that the taxes on the controllable industries do not affect the profits, and hence the outputs, of the distorted one. So, for h, any member of H, the new controllable set, $\partial f_1/\partial x_h=0$. Since the exercise yielding (9) can be duplicated when $t_k=0$, $k\notin H$, rather than just $t_1=0$, our new conditions for optimal taxes are the same as (9) except that they apply only for $i\varepsilon H$ rather than for $i=2,\ldots,n-1$. Since $\partial f_1/\partial x_h=0$, it is immediately seen that having zero taxes on the controllable industries is necessary and sufficient for a second best optimum, i.e., laissez faire is the second best optimum.

But we have been assuming that the controllable sectors contain no distortions. Adding distorted, controllable sectors to our economy, our removal of the paradox tells us to apply classical economic policy to these sectors to arrive at our second best solution. For if the sector is controllable, it has no noticeable effect on the noncontrollable sectors, and, as we have seen, the standard optimality conditions apply. Thus, in developing policy toward any one sector, we can assume the rest of the economy is perfectly competitive (even though it certainly is not) and suggest a policy that would induce the otherwise distorted sector to behave as if it were perfectly competitive. In this way, our derived second-best-constraints serve to rationalize the classical, "piecemeal" approach to policy that Davis and Winston (1967) represent as having been destroyed by second-best theorists. This piecemeal approach is of great potential value to economics because it allows different economists to specialize on different sectors or problems but still come out with a collection of policy recommendations that would, if technically correct, achieve a Pareto optimum. Without this classical, piecemeal approach, each economist would have to theoretically model and empirically estimate the entire economy, with all of its imperfections and feasibility constraints, in order to logically derive a single policy proposal. It is little wonder that the many economists who have taken the conventional theory of second best seriously have been able to find few colleagues who concur with their models. What we are saying is that this distressing state of affairs is entirely unnecessary. Under a plausible model generating an uncontrollable, distorted sector, we have seen that such a sector makes related sectors uncontrollable, leaving us with a set of controllable sectors between which the standard optimality conditions apply. This enables us to employ classical, piecemeal policy to the controllable sectors, thereby permitting a potentially extremely valuable division of labor among economists.

Our formal argument for piecemeal policy despite the presence of rational political constraints does not apply when the constraints are not of the standard, controllability-uncontrollability form, i.e., when the underlying political transaction costs, rather than completely precluding all policy toward certain sectors, only restrict policy to special forms acceptable to the potential losers. In such cases, piecemeal policy will not generally achieve the second-best optimum because the optimal policy in one sector then depends on the specific forms of the policy constraints and imperfections in other sectors. Nevertheless, the economic advantage of piecemeal policy in permitting professional specialization within economics outweighs the cost of sacrificing a genuine second-best optimum if individual piecemeal policy suggestions are either totally ignored, as they are when a sector is uncontrollable, or implemented in a form that is only slightly distorted by political constraints, in which case the sector is close to controllable.

VI. Conclusion

We are left with the impression that standard second best theory is, rather than a general policy framework for economists wishing a more politically realistic view, just another abstract theory in search of an application. Along these lines, it may be useful to outline a quasi-realistic example of how conventional second best theory might currently be used and how our argument upsets the application. Conventional second best theory says that because it has been politically infeasible to lift the recently imposed U.S. price controls on oil products, a myriad of new taxes and subsidies on non-oil products has been required for a social optimum. But we have seen that the political feasibility constraint, being based on the rational voter fear that the government officials whose jobs would be to design and impose taxes on the oil companies sufficient to compensate the voters for the higher oil prices might in fact "sell out" to the oil companies, implies additional constraints on policy in other sectors. In particular, we have seen that any one of the myriad of conventional second-best interventions in the non-oil industries, being beneficial only in that it provides an indirect subsidy to producing oil, should be deemed infeasible for the same reason that the first-best policy is infeasible. The fact that oil price controls have not been accompanied by significant changes in tax and subsidy rates outside the oil industry indicates that political feasibility constraints do indeed extend beyond the distorted sectors to related, undistorted sectors. At the same time, it indicates that our amended theory of second best may well have some empirical relevance in explaining observed, government policy.

FOOTNOTES

[1]Armen Alchian, Geoffrey Brennan, Bob Clower, Jack Marshall, and Nick Tideman provided valuable comments on earlier drafts.

[2]For any x, aggregate output is allocated among all individuals while maintaining given levels of utility for all individuals but one and maximizing the utility of the one individual to obtain the utility index $U(x)$ with the conventional curvature. This is to be contrasted with Samuelson's (1956) attempts to construct social indifference contours that require lump-sum transfers to maintain "correct" distributions of utility. Here we are anticipating some final distributional solution when assigning utility levels.

[3]For notational simplicity, lump-sum taxes and subsidies and, correspondingly, a governmental budget balancing equation, are ignored, although they will continue to be included in our informal discussions on policy.

[4]Since we are dealing with a technological environment containing outputs only, the question of efficient use of inputs does not arise. Allingham and Archibald (1975) have shown that in a model with concave production functions, aggregate resource constraints, and an uncontrollable monopolist, second-best production takes place on the production frontier. Thus, we simply appeal to their results as a rationale for using a transformation constraint.

[5]Government decision makers could know these functions and thereby solve their maximization problem but still, conceivably, be unable to achieve the optimum because of an inability to observe actual outputs and thus enforce the requisite per unit taxes. However, in a world in which essentially all outputs are already monitored and taxed and relatively little is known about utility and transformation functions, it would be intolerably perverse to assume that government

decision makers know the utility and transformation functions but not the actual outputs. Therefore, while we are accepting the assumption that government decision makers know the utility, transformation, and supply functions, at least in the relevant regions, we take with it an assumption that actual outputs are also known.

The reader may believe that our informational assumptions imply that decentralized production is at best redundant because government decision makers know everything required for dictating a socially optimal set of outputs. However, while our assumptions do not preclude such redundancies -- indeed our results will apply to optimal policy in socialist as well as capitalist economies -- it is also possible that the decision makers' information on the nature of the U[x], T[x], and x(t) functions is based upon observations of parameters only present in decentralized systems. For example, free market prices in decentralized capitalist systems, in revealing private marginal rates of substitution and transformation, allow us to determine a first-best tax policy with knowledge of only equilibrium differences between private and social values. So no redundancy of free markets and capitalism is implied by our informational assumptions. Indeed, the assumption on the knowledge of U[t] and T[t] in relevant ranges may well be empirically useless in the absence of free markets, such markets serving to make an otherwise hopeless policy task just difficult.

[6] To see this, first multiply (9) by ∂t_i and note that at $t=\theta$ only the first terms on the left side of (9) may be non-zero. Then note that since the net social value of x_1 (i.e., the coefficient of ∂x_1) is positive, the net social value of the tax change (the entire left side of (9)) is positive if and only if ∂x_1 is positive.

[7] While it is conceivable that the second best taxes and subsidies could work to flatten the monopoly demand curve, inducing the monopolist to expand without increasing his profit, the same effect could be achieved by a non-linear, first-best, tax-subsidy schedule. Moreover, for externality imperfections, such an effect would not be present. The firms in a competitive industry-generating an external economy must <u>always</u> receive a higher price or pay a lower cost in order to expand output.