REFLECTIONS ON OPTIMAL TAX THEORY

by

Richard Arnott

Working Paper # 213

June 1992

Department of Economics
Boston College
Chestnut Hill, MA 02167

Reflections on Optimal Tax Theory*

Richard Arnott

This paper is a reaction to two papers in the volume, "Optimal Taxation and Government Finance" by James Mirrlees, and "Optimal Taxation, Progressivity, and Public Policy" by Robert Haveman. It addresses three questions raised in these papers: Is there a presumption that the second–best (when differential lump–sum taxation is impossible) level of a public good is lower than the first–best level? Can optimal tax theory be extended to address issues related to macroeconomic stabilization? And, how has optimal tax theory contributed to public economics?

The paper has two related themes. The first is that the primary contribution of optimal tax theory has been to provide a coherent theory of optimal economic policy. Following the method of optimal tax theory, optimal policy is solved for by maximizing a welfare function in which all desiderata are quantified subject to a complete description of the economy, which includes all relevant constraints—production possibilities, endowments, information, and politics. Contrary to the optimistic expectations early on in its development, optimal tax theory has provided little in the way of <u>concrete</u> guidance concerning the design of actual tax systems. Nevertheless, by providing a consistent and rigorous framework for normative analysis, it has contributed greatly not only to public economics but to applied economics generally.

The second theme concerns the methodological relationship between applied economic theory and public policy analysis. An applied theoretical model provides a way of conceptualizing and of focusing thought on <u>some</u> facets of a public policy issue. It does not, or should not, claim to capture all or even most of the important considerations relevant to a particular policy issue. The role of the policy economist is

* I would like to thank John Quigley for helpful comments on an earlier draft.

to draw together the insights derived from a variety of applied economic theoretical models and of econometric studies, as well as to consider practical issues that the academic literature has neglected, and then to recommend policy on the basis of this synthesis. It may be that some policy analysts have drawn too heavily on some simple optimal tax models in their policy prescriptions. Still, to criticize optimal tax theory for this reason is inappropriate.

Section I examines the first–and second–best levels of a public good, focusing on the contributions of the Mirrlees paper to this topic. Section II considers stabilization policy and optimal tax theory, taking Mirrlees' discussion as the point of departure. And Section III reacts to Haveman's paper by providing a defense of optimal tax theory.

1.1 First– and second–best levels of a public good

The first papers to address this issue were Stiglitz and Dasgupta (1971) and Atkinson and Stern (1974). Both considered economies with a single representative consumer. First–best finance was lump–sum taxation. Second–best finance was optimal differential commodity taxation with an untaxed good. With optimal differential commodity taxation, the tax rates are chosen to equalize across taxed goods the social cost of raising an extra dollar of revenue—one dollar plus the marginal excess burden. As was recognized at the time, the comparison of the first– and second–best levels of a public good with a single consumer is practically uninteresting since there is no good reason to rule out lump–sum taxation. The analysis was, however, considered interesting as a halfway house towards examination of the issue with heterogeneous consumers.

These papers demonstrated that the second–best level of the public good can exceed the first–best; for example, if goods whose optimal tax rates are low are complementary with the public good, the switch from lump–sum to differential commodity taxation may stimulate demand for the public good sufficiently that the optimal level increases, even though the marginal cost is higher. Nevertheless, the general insight derived from these papers was that the higher cost of the public good in the second–best economy "typically" causes its optimal level to be lower.

The optimal commodity tax literature started by examining a representative consumer economy with an untaxed commodity (Ramsey (1927), Corlett and Hague (1953/4), Diamond and Mirrlees (1971)). By the mid–seventies, the models were extended to treat heterogeneous individuals—the many–person Ramsey tax problem (Feldstein (1972), Diamond (1975) and Mirrlees (1975))—which permitted the integration of efficiency and equity. It was assumed that the government cannot identify individuals nor keep track of a specific individual's purchases; as a result, commodity taxation is anonymous and linear. It was also recognized that, under these assumptions, while the government cannot impose differential lump–sum taxation, it can impose a uniform poll tax. This is very important since it implies that the government then has a non–distortionary method of finance at the margin.

The many–person Ramsey tax problem was subsequently extended to incorporate public goods (e.g. Atkinson and Stiglitz (1980)). Remarkably, however, this paper by Mirrlees is the first[1] to examine the implications of having a non–distortionary method of finance at the margin in the second–best economy on the first–best (full lump–sum taxation) versus second–best level of public goods with heterogeneous consumers.

There are generally two different approaches to the comparison of the first– versus the second–best provision of public goods—the first–order approach and the levels approach. The first–order approach entails only a comparison of the first–order conditions in the first– and second–best problems. The levels approach attempts to compare the actual levels of the public goods. Since this entails a comparison of two general equilibria, the analysis is more difficult, but the results are also more informative.

[1] The paper by Wilson (1991) was written independently.

1.1.1. *The first–order condition approach*

Mirrlees analyzes the problem first using the first–order condition approach, and then using the levels approach. The basic procedure for the first–order condition approach is as follows. The Samuelson condition for the first–best level of the public good is $\Sigma MB = MC$. Since there is optimal lump–sum taxation, everyone has the same marginal utility of income and hence the same welfare weight. The first–order condition for the second–best level of the public good can be written as $\Sigma\tilde{M}B - \phi = MC$, where MC is the resource cost of the public good, $\Sigma\tilde{M}B$ is the welfare–weighted[2] sum of marginal benefits (since optimal lump–sum taxation is unavailable, individuals' marginal utilities of income differ at the second–best optimum) and ϕ is an additional term which shall be explained shortly. Then the two first–order conditions are compared. To simplify, assume that production possibilities are linear. Then the second–best level of the public good is greater then the first–best level if $(\phi + \Sigma\tilde{M}B)_2 > \Sigma MB_1$, where subscript 2 denotes evaluation with the second–best set of taxes and 1 evaluation with the first–best set of taxes. This inequality can be decomposed:

$$(\phi + \Sigma\tilde{M}B)_2 > \Sigma MB_1 \Leftrightarrow \phi + (\Sigma\tilde{M}B - \Sigma MB)_2 + (\Sigma MB_2 - \Sigma MB_1) > 0$$

The last term on the right–hand side of the second inequality is the difference in the welfare–unweighted sum of marginal benefits due to differences in the tax systems. The middle term is the difference between welfare–weighted and welfare–unweighted sum of marginal benefits under the second–best tax régime; it is positive if the covariance between welfare–weights and marginal benefits is positive. Mirrlees focuses on the first term, ϕ.

In Mirrlees' paper

---

[2]The welfare weight is $\frac{\partial\psi}{\partial v^h}\alpha^h \div \frac{1}{H}\sum_h \frac{\partial\psi}{\partial v^h}\alpha^h$, where ψ is the social welfare function, v^h is the utility of household h, and α^h the marginal utility of income of household h.

$$\phi = Ht(\bar{x}_b m^* - x_z),$$

where H is the number of individuals or households, t is the vector of optimal commodity (excise) tax rates, $\bar{x}$ is the vector of mean commodity demands, b is lump–sum income, z is the level of the public good, m^* is the mean welfare–weighted marginal benefit from the public good, and subscripts denote partial derivatives. $-\phi$ is the magnitude of the indirect revenue effect. Suppose that the government raises the poll tax by one dollar, spending the extra revenue raised on the public good. This extra revenue comes directly from the poll tax and indirectly from commodity taxes. The direct revenue increase is H. The perturbation changes households' lump–sum income, as well as the level of the public good. In response, households adjust their consumption bundles, which alters the revenue raised from commodity taxes. This indirect increase in revenue is $-H\phi$. The important point to note is that, by the envelope theorem, the welfare cost of the revenue raised is H. The indirect increase in revenue comes at no welfare cost. Thus, a positive indirect revenue effect can be viewed as either decreasing the marginal cost of the public good or as increasing the marginal benefit.

Atkinson and Stiglitz (1980, p. 497) clearly identify the indirect revenue effect in this context. Thus, Mirrlees' result is not new. He does, however, provide a more helpful discussion of the indirect revenue effect than do Atkinson and Stiglitz. He makes three interesting observations:

1. Suppose that commodity demands are independent of z; that, for whatever reason, goods' tax rates exceed factor tax rates in absolute value (since $q \equiv p + t$, a factor tax which causes the consumer factor price to be lower than the producer factor price corresponds to a negative t); and, that commodities are normal. Then the indirect revenue effect is negative. This indicates that the second–best level of the public good depends, inter alia, on the relative magnitude of factor and goods taxes (and, therefore, whatever determines these relative magnitudes).

2. Continue, for the purposes of illustration, to assume that the indirect revenue effect is negative. Return to Mirrlees' first–order condition for the second–best level of the public good, with MC = 1:

$$\Sigma \tilde{MB} - \phi = 1. \qquad \text{(iia)}$$

Atkinson and Stiglitz rewrite (iia) as

$$\begin{aligned} \Sigma \tilde{MB} &= 1 + \phi \\ &= \varphi, \text{ with } \varphi \equiv 1 + \phi, \end{aligned} \qquad \text{(iib)}$$

interpreting a negative indirect revenue effect as increasing the social cost of the public good. Mirrlees, in contrast, rewrites (iia) as

$$\Sigma \hat{MB} = 1, \text{ with } \hat{MB} \equiv \left(\frac{\Sigma \tilde{MB} - \phi}{\Sigma \tilde{MB}}\right)\tilde{MB}, \qquad \text{(iic)}$$

interpreting a negative indirect revenue effect as effectively causing the social valuation of the public good, $\hat{MB}$, to be lower than the "private valuation," $\tilde{MB}$. Since tax rates (and hence the consumer prices of commodities) are high, individuals overvalue the public good.

3. Mirrlees points out that, since poll taxation is possible, no distortion terms enter the formula for the second–best level of the public good. The distortions are inframarginal rather than marginal.

While these observations are insightful, they do not go very far towards identifying primitive conditions under which the second–best level of the public good is higher or lower than the first–best level. The first–order conditions are just too implicit.

1.1.2 *The levels approach*

Mirrlees' analysis here is characteristically elegant, clear, and insightful. He notes that with a poll tax, if everyone is identical the first– and second–best levels of the public good are the same. He then introduces a perturbation—a small amount of inequality—and compares the resulting change in the first– and second–best levels of the public good, under the simplifying assumption that utility is separable between public and private goods.

Individuals in the model are indexed by n (which can be interpreted as ability) and (with the assumed separability) have the utility functions $v(q,b,n) + g(z)$. Mirrlees demonstrates that the second–best level of the public good is higher than the first–best if $\frac{v_{bbn}}{v_{bn}} > \frac{v_{bbb}}{v_{bb}}$, and lower if the inequality is reversed.

Mirrlees provides a rigorous proof. Here I shall provide a heuristic derivation in order to elucidate the result. I shall simplify by assuming that there are only two individuals and that production possibilities are linear. In the initial situation, $q = p$ (no commodity taxes), b and n are the same for both individuals, and z is the optimal level of the public good. Then, <u>with z at its initial level,</u> a small mean–preserving spread in n is introduced, which is captured by assuming that one individual's n is $n + \Delta$ and the other's is $n - \Delta$. After the perturbation, the second–best level of the public good is higher than the first–best level if the sum of the marginal benefits is greater at the second–best optimum than that at the first–best optimum, i.e. $\Sigma MB_2 > \Sigma MB_1$.

Consider the <u>second–best</u> initially. It turns out that for this perturbation the effects from imposing differential commodity taxes are of higher order than the effects that are central; thus, we may assume that the public good is financed solely through a poll tax. Since, by assumption, the level of the public good and hence tax revenue are unaffected by the perturbation, the level of the poll tax is unaffected by the perturbation —label it $\bar{b}$. Thus,

$$\Sigma MB_2 = MB(p,\bar{b},n + \Delta) + MB(p,\bar{b},n - \Delta) \ , \qquad \text{(iiia)}$$

$$\left.\frac{\partial \Sigma MB_2}{\partial \Delta}\right|_{\Delta = 0} = (MB_n^+ - MB_n^-)_{\Delta = 0} = 0 \ , \text{ and} \qquad \text{(iiib)}$$

$$\left.\frac{\partial^2 \Sigma MB_2}{\partial \Delta^2}\right|_{\Delta = 0} = (MB_{nn}^+ + MB_{nn}^-)_{\Delta = 0} = (2MB_{nn})_{\Delta = 0} \ . \qquad \text{(iiic)}$$

The effect of the perturbation on ΣMB_2 is second–order and its magnitude is given by (iiic).

Consider next the <u>first–best</u> tax régime. The sum of the differential lump–sum taxes must raise the revenue to finance z, i.e. $b^+ + b^- = 2\bar{b}$. Hence

$$\Sigma MB_1 = MB(p,b^+, n + \Delta) + MB(p, b^-, n - \Delta) \ , \text{ and} \qquad \text{(iva)}$$

$$\left.\frac{\partial \Sigma MB_1}{\partial \Delta}\right|_{\Delta = 0} = \left(MB_b^+ \left(\frac{\partial b^+}{\partial \Delta}\right) + MB_n^+ + MB_b^- \left(\frac{\partial b^-}{\partial \Delta}\right) - MB_n^-\right)_{\Delta = 0} = 0 \qquad \text{(ivb)}$$

since $\frac{\partial b^+}{\partial \Delta} + \frac{\partial b^-}{\partial \Delta} = 0$. Now, the lump–sum taxes are chosen so that both individuals have the same marginal utility of income. Since marginal benefit equals g´(z) divided by the marginal utility of income, both individuals must have the same marginal benefit:

$$MB_b^+ \left(\frac{\partial b^+}{\partial \Delta}\right) + MB_n^+ = MB_b^- \left(\frac{\partial b^-}{\partial \Delta}\right) - MB_n^- \ . \qquad \text{(ivc)}$$

Combining (ivc) and (ivb) gives

$$\frac{\partial b^+}{\partial \Delta} = -\frac{MB_n^+}{MB_b^+} \quad \text{and} \quad \frac{\partial b^-}{\partial \Delta} = \frac{MB_n^-}{MB_b^-} \;. \tag{ivd}$$

Then

$$\left.\frac{\partial^2 \Sigma MB_1}{\partial \Delta^2}\right|_{\Delta=0} = (MB_{bb}^+ (\frac{\partial b^+}{\partial \Delta})^2 + MB_b^+ (\frac{\partial^2 b^+}{\partial \Delta^2}) + 2MB_{bn}^+ (\frac{\partial b^+}{\partial \Delta}) + MB_{nn}^+$$

$$+ MB_{bb}^- (\frac{\partial b^-}{\partial \Delta})^2 + MB_b^- (\frac{\partial^2 b^-}{\partial \Delta^2}) - 2MB_{bn}^- (\frac{\partial b^-}{\partial \Delta}) + MB_{nn}^-)_{\Delta=0}$$

$$= (2MB_{bb} (-\frac{MB_n}{MB_b})^2 + 4\, MB_{bn} (-\frac{MB_n}{MB_b}) + 2MB_{nn})_{\Delta=0}, \tag{ive}$$

using (ivd) and $\frac{\partial^2 b^+}{\partial \Delta^2} + \frac{\partial^2 b^-}{\partial \Delta^2} = 0$ from the revenue constraint.

Combining results gives

$$(z_2 - z_1)_{\Delta=0} > 0 \;\Leftrightarrow\; \left.\frac{\partial^2 \Sigma MB_2}{\partial \Delta^2}\right|_{\Delta=0} > \left.\frac{\partial^2 \Sigma MB_1}{\partial \Delta^2}\right|_{\Delta=0}$$

$$\Leftrightarrow\; (2MB_{bb} (-\frac{MB_n}{MB_b})^2 + 4\, MB_{bn} (-\frac{MB_n}{MB_b}))_{\Delta=0} < 0. \tag{v}$$

After substitution of $MB = \frac{g'}{v_b}$, (v) reduces to Mirrlees' result.

I have been unable to come up with a neat geometric characterization of the result. However, the following geometric characterization provides some insight.

Consider first the poll tax. The effect of the perturbation on the average marginal benefit is illustrated in Fig. 1.

Insert Figure 1

Next consider the first–best system of taxes.

Insert Figure 2

From Fig. 1, it is evident that the effect of the perturbation on the mean marginal benefit from the public good with the poll tax depends on the curvature of the MB– surface along $b = \bar{b}$. And Fig. 2 shows that the effect of the perturbation on the mean marginal benefit from the public good with lump–sum taxes depends on the curvature of marginal benefit contours.

In any event, the essential point made by Mirrlees is that even in an extremely simple situation which contains no obvious bias, the relative magnitude of the first– and second–best levels of the public good depends on higher–order (rather than concavity/convexity) curvature properties of the utility function, concerning which economic theory has little to say. With a finite amount of inequality, the comparison is further complicated by commodity taxes, the form of inequality aversion, and the form of the income distribution. Furthermore, the comparison then depends on global rather than local higher–order curvature properties of the utility functions. Thus, it is safe to say that there is no a priori presumption that the second–best (with the optimal poll tax) level of a public good is higher or lower than the first–best level.

It is commonly argued that the second–best level of a public good is typically lower than the first–best level on the basis that in the former case the use of differential commodity taxation raises the private cost of public funds. Mirrlees' paper demonstrates that this argument is fallacious. Differential commodity taxation is distortionary. However, when the set of policy instruments includes a poll tax, it is always possible to finance the marginal unit of the public good via a rise in the poll tax. Thus, when both differential commodity taxes and a poll tax can be employed, taxation is non–distortionary at the margin, and it is distortion at the margin that affects the private cost of public funds.

2. Stabilization Policy

Twenty years of students received their first exposure to graduate public finance through Richard Musgrave's The Theory of Public Finance. To these many generations of students it must be striking that the Stabilization Branch of Musgrave's Fiscal Department has all but disappeared from public economics. Its disappearance is, I think, a quirk of intellectual history. The "new public economics" (now not so new) insisted on rigorous microfoundations, which was hard, if not impossible, to reconcile with the Keynesian macroeconomics of that era (the early 1970's), which Musgrave espoused. Since then, of course, the microfoundations of macroeconomics have been made much stronger. But stabilization policy has not yet been brought back into public economics. This is regrettable since most fiscal policy analysis in macroeconomics provides a crude treatment of taxes and welfare economics.

In the first part of the paper, Mirrlees takes a step towards bringing stabilization policy back into the fold, through optimal tax theory. He initially assumes that nominal excise tax rates, nominal lump–sum transfers, and real levels of government purchases are decided on ex ante, in a static, uncertain environment. This leads to two related problems, on which Mirrlees focuses. First, at the time tax rates are set, the price level is unknown since it is state–contingent. Second, it may be impossible to finance the predetermined level of real government purchases––in some states, there may exist no set of prices such that the ex ante government purchases can be financed, given the nominal ex ante tax rates and lump–sum transfers.

Mirrlees provides the following example: Production possibilities are given by $y = c + z$, where y is aggregate output of a generic good which may be either privately consumed, c, or used by the government, z. Output is produced using labor, ℓ. The representative consumer's budget constraint is $(p + t)c = py(\ell) + b$, where b is the nominal lump–sum transfer and p is the endogenous price of y. The consumer decides on c and ℓ after the realization of the state. The consumer's maximization problem is therefore $\max_{c,\ell} u(c, \ell)$ s.t. $(p+t)c = py(\ell) + b$. The nominal revenue raised is $tc - b$, and the quantity of public goods purchased is $(tc - b)/p$. For purpose of illustration, suppose that the utility function is Cobb–Douglas, $u = c^{1/2}(1 - \ell)^{1/2}$ (the individual has one unit of time), and that $y(\ell) = w\ell$, where w is the realization of the random marginal (= average) product of labor. Then from the consumer's maximization problem, $1 - \ell = \frac{pw+b}{2pw}$ and $c = \frac{pw+b}{2(p+t)}$, which imply that real tax revenue is $\frac{tc - b}{p} = \frac{ptw - tb - 2bp}{2(p+t)\,p}$. The price level adjusts such that (if feasible) this equals z. Thus, p solves

$$p^2\,(2z) + p\,(2tz - tw - 2b) + tb = 0. \qquad \text{(vi)}$$

It is straightforward that if $tb > 0$ and $(2tz - tw + 2b)^2 < 8tbz$ or $2tz - tw + 2b > 0$, then there is no real, positive solution to (vi)––the situation shown in Mirrlees' Fig. 1.

Mirrlees then demonstrates that a potential non–existence problem may also arise if taxes and lump–sum transfers are indexed to prices. He then argues that: "An alternative view of [these potential non–existence problems] is that we have found here lurking in the microeconomic optimal–tax model some aspects of the macroeconomic problems of adjustment and stabilization; and that welfare maximization is not a satisfactory way of analyzing the policy issues in this context."

In the remainder of this subsection, I shall dispute these two statements, and discuss an alternative approach to stabilization policy from an optimal–tax perspective.

The potential non–existence problems that Mirrlees identifies are new and of some theoretical interest. But at the same time they appear similar to a non–existence problem that may arise in the standard optimal commodity tax problem. There, the government may commit itself to a technically feasible level of government expenditures whose cost exceeds the maximum amount that can be raised via taxation. In Mirrlees' model, nominal excise tax rates are fixed and there may be no price level for which enough revenue can be raised to finance a technically feasible level of government expenditures.

Mirrlees' model might be relevant to hyperinflated economies, but surely not to today's developed countries where the magnitude of fluctuations are "modest", the limits of fiscal capacity are not close to being exceeded, and the political process works at least moderately well. If there were a significant downturn, the political process would be sufficiently responsive that fiscal capacity would not be exceeded; government expenditures, tax rates, transfers, and monetary policy would be adjusted appropriately. I think Mirrlees is barking up the wrong tree.

From one perspective, there is no problem at a conceptual level in developing a theory of optimal stabilization policy based on welfare maximization. Consider the economy as a mechanism that transforms a set of stochastic dynamic input signals into a set of stochastic dynamic output signals (various economic time series). The object is to tinker with the mechanism so as to maximize the expected discounted welfare associated with the output signals. This is an exercise in optimal dynamic feedback control. It may be difficult to make this conceptualization operational for policy purposes. But we should be very hesitant to reject the welfare maximization approach since, like the assumption of rationality, it forces conceptual coherence.

The principal difficulty in developing a theory of optimal stabilization policy is to provide persuasive microfoundations not only for macroeconomics but also for the economic failures, market and non–market, which justify government intervention to stabilize the economy.

Macroeconomic fluctuations are not per se undesirable. They would occur in an Arrow–Debreu economy with uncertainty, which is well–known to be Pareto efficient, and with optimal lump–sum redistribution, which is optimal as well. There is a general perception, however, that macroeconomic fluctuations in developed economies are excessive — too large to be explained as the response of a well–behaved economy to the kinds of exogenous shocks that developed economies experience. Due to some failure, either the economy magnifies these exogenous shocks or else it creates the fluctuations internally. Support for this view is provided by the persistence, importance, and cyclicality of involuntary unemployment, which has no place in an Arrow–Debreu economy.

What is the source of economic failure? There is no shortage of candidate culprits: sticky or fixed prices and wages (Benassy (1982)), job search externalities (Diamond (1984)), coordination failure (Cooper and John (1988)), credit market imperfections (Greenwald and Stiglitz (1990)), staggered contracts (Taylor (1980)), implicit contracts (Arnott, Hosios, and Stiglitz (1988)), efficiency wages (Shapiro and Stiglitz (1984)), sunspots (Azariadis and Guesnerie (1986)), imperfect competition (Hart (1982)), etc. None has gained majority acceptance within the profession. Given this lack of consensus concerning appropriate microfoundations for macroeconomics, it would be misguided to search for the definitive theory of optimal stabilization policy. However, the more limited goal of developing a theory of optimal stabilization policy based on any of the above microfoundations seems realistic.

Indeed, much work has been done along these lines; but, none of it provides a rich treatment of fiscal policy, or develops a theory of optimal stabilization policy, which draws on optimal tax theory. My aim, in the model which follows, is to show that this can be done. The model's microfoundations of the macroeconomy are much too simple, but they are at least consistent. I hope that the model will encourage others to incorporate optimal tax theory into models with richer microfoundations.

The reasoning underlying the model specification is as follows. The model needs to incorporate uncertainty and a source of market failure, and to provide a description of the economy compatible with optimal tax theory. The theory of moral hazard satisfies these requirements. The model describes a farm economy. A farmer's output depends on his effort and the weather, which is imperfectly correlated over farms. Crop insurance is provided. This causes farmers to exert less effort, which in turn reduces average aggregate output and increases the variability of output. The efficiency loss associated with moral hazard can be mitigated through optimal taxation; for instance, by subsidizing agricultural inputs that reduce the probability of crop failure.

To set the stage, a preliminary model is presented first. The economy comprises identical individuals, each of whom farms a unit area of land. There is a single generic crop. Output per farm depends on both effort and the weather; specifically, there are two possible output levels—h (high) and ℓ (low). Effort is decided before the weather is known. In good years, the probability of low output on a particular farm is $p_g(e)$ and in bad years $p_b(e)$ ($p'_i < 0$, $p''_i > 0$, $p_i(0) = \bar{p}_i < .5$, i=g,b, and $p_g(e) > p_b(e)$ for all e.) Good and bad years occur with equal probability. Crop insurance is available. The insurer can observe whether a crop year was good or bad, and the total quantity of insurance purchased by a farmer, but neither the weather nor the effort expended on a particular farm. Consequently, an insurance contract in aggregate state i specifies a premium β_i and a net payout α_i. The insurance must break even in a particular crop year.

Let y be post–insurance income, and U(y,e) be the utility function. Then an individual's expected utility is

$$EU = \frac{1}{2}\left(\sum_i (1-p_i(e))\, U_i^0 + p_i(e) U_i^1\right), \qquad \text{(i)}$$

where $U_i^0 \equiv U^0(h-\beta_i, e)$ and $U_i^1 \equiv U^1(\ell+\alpha_i, e)$. The first–order condition for e (an interior solution is assumed) is

$$\left[\frac{1}{2}\sum_i p'_i\,(-U_i^0 + U_i^1)\right] + \left[\frac{1}{2}\sum_i \left((1-p_i)\frac{\partial U_i^0}{\partial e} + p_i\frac{\partial U_i^1}{\partial e}\right)\right] = 0, \qquad \text{(ii)}$$

which states that the expected marginal benefit from effort (the term in the first square bracket) equals the expected marginal cost (the term in the second square bracket). The solution to (ii) may be written compactly as $e = e(\alpha_g, \alpha_b, \beta_g, \beta_b)$. Normally[3], $\frac{\partial e}{\partial \beta_i} < 0$ and $\frac{\partial e}{\partial \alpha_i} < 0$; as more insurance is provided, the individual reduces effort.

Substitution of the effort equation into (i) gives expected utility as a function of the terms of the insurance contract, i.e. $V(\alpha_g, \alpha_b, \beta_g, \beta_b) \equiv \max_e EU(\alpha_g, \alpha_b, \beta_g, \beta_b)$.

The insurance is provided either by the government or by competitive insurance firms. In both cases, the insurance maximizes expected utility subject to the insurer at least breaking even <u>in each aggregate outcome</u>, i.e.

$$\max_{\alpha,\beta} V(\alpha,\beta) \quad \text{s.t.} \quad (1-p_i(e))\beta_i - p_i(e)\,\alpha_i \geq 0,\ i = g,b \qquad \text{(iii)}$$

(to simplify notation, we employ vectors where there is no ambiguity). The details of the solution need not concern us here; they are provided in Arnott and Stiglitz (1986). Three qualitative features of the solution are important, however. First, moral hazard causes individuals to exert too little effort relative to the first best, where the insurance

[3] And always with separable and event-independent utility: $EU = \frac{1}{2}\left(\sum_i (1-p_i(e))\, u(y_i^0) + p_i(e)\, u(y_i^1)\right) - e.$

contract specifies effort.[4] Second, the equilibrium is constrained efficient; even though it entails too little effort relative to the first best, it is efficient conditional on the unobservability of effort — there is no market failure. And third, effort is positively related to mean aggregate output and negatively related to the variability of aggregate output. Consequently, relative to both the no–insurance and first–best situations, the provision of insurance under moral hazard is destabilizing.

We now enrich the model in a natural way to include multiple goods. The resulting equilibrium is constrained inefficient, and the inefficiency is remedied by optimal commodity taxation. The intuition is straightforward. The equilibrium with insurance provided under moral hazard entails individuals expending too little effort relative to the first best. The efficiency loss associated with this distortion (relative to the first best) can be reduced by taxing substitutes to effort and subsidizing complements. The optimum balances off these efficiency gains against the efficiency losses from not pricing goods at marginal cost. Since the taxation stimulates effort, and since effort is negatively related to the variability of output, then in our economy optimal commodity taxation is stabilizing. A complete analysis of optimal commodity taxation with moral hazard is provided in Arnott and Stiglitz (1986). Here we simply illustrate the result.

We augment the model by assuming that a unit of the generic output can be costlessly transformed into either a unit of agricultural input or a unit of one of two consumer goods. Let f be the quantity of the agricultural input employed, and c_{ij}^1 and c_{ij}^2 be the quantity of goods 1 and 2 consumed with aggregate outcome i = g,b and output j = h,ℓ. The agricultural input increases the probability of high output (viz., $p_i = p_i(e,f)$ with $\frac{\partial p_i}{\partial f} < 0$, $\frac{\partial^2 p_i}{\partial f^2} > 0$, i = g,b). Consumption occurs after the outcome has been realized and insurance payments made. The government employs linear commodity

[4] The provision of insurance under moral hazard causes a substitution effect away from effort. With non-separable utility, there are also income effects which can operate in either direction, depending on the risk-aversion properties of the utility function.

taxes. Since the agricultural input is purchased before the realization of the outcome, its consumer price, q^f, is outcome–independent, and since the consumer goods are purchased after the realization of uncertainty, it is reasonable to assume that their consumer prices are outcome–contingent, q_i^1 and q_i^2. Since the demands for f and c are homogeneous of degree zero in post–insurance income and consumer prices, we need an additional normalization. There is no natural one; let it be $q_g^1 = 1$.

We shall skip the algebraic details. But it will be useful to explain the solution procedure. The consumer's maximization problem now has two stages. In the later stage, consumption decisions are made, taking e and f, consumer prices and the parameters of the insurance contract, as given. In the earlier stages the individual chooses e and f, taking into account the dependence of c on e and f, and taking consumer prices and the parameters of the insurance contract as given. It will simplify notation in what follows if we let x_{ij} be the transfer from the insurance company to the individual with aggregate output i and output j, and π_{ij} be the corresponding probability. Solving, one obtains c = e(y,q) etc. and hence

$$EU(y,q) = \sum_{ij} \pi_{ij}(y,q)\ V_{ij}(y,q), \tag{iv}$$

where V_{ij} is expected utility with i,j. The social planner's problem is then to choose the parameters of the insurance contract, as well as consumer prices, so as to maximize expected utility subject to budget balance in each aggregate outcome

$$B_i = [-\sum_j \pi_{ij}\, x_{ij} + \sum_{k=1}^{2}\sum_j \pi_{ij}\,(q_i^k - 1)\, c_{ij}^k + \sum_j \pi_{ij}\,(q^f-1)f] \div \sum_j \pi_{ij} = 0. \tag{v}$$

The solution (available on request) is

$$-\sum_i \lambda_i \sum_j \pi_{ij}[q^f - 1)\left(\frac{\partial f}{\partial q^f}\right)_\theta + \sum_{k=1}^{2}(q_i^k - 1)\left(\frac{\partial c_{ij}^k}{\partial q^f}\right)_\theta] + \sum_i \lambda_i \sum_j s_{ij}\left(\frac{\partial \pi_{ij}}{\partial q^f}\right)_\theta = 0 \tag{vi}$$

$$-\sum_i \lambda_i \sum_j \pi_{ij}[q^f - 1)\left(\frac{\partial f}{\partial q_{i'}^k}\right)_\theta + \sum_{k=1}^{2}(q_{i'}^k - 1)\left(\frac{\partial c_{ij}^k}{\partial q_i^k}\right)_\theta] + \sum_i \lambda_i \sum_j s_{ij}\left(\frac{\partial \pi_{ij}}{\partial q_{i'}^k}\right)_\theta = 0, \tag{vii}$$

where λ_i is the shadow price on the budget constraint with aggregate outcome i, $s_{ij} \equiv x_{ij} - (q^f - 1)f - \sum_{k=1}^{2} (q_i^k - 1)c_{ij}^k$ is the net transfer from the government to the individual for ij, and $(\frac{\partial m}{\partial q_n})_\theta$ denotes the change in m from a compensated change in q_n (the compensation occurs for all relevant i, j) and therefore captures the relevant substitution effects. Though they appear rather intimidating, eqs. (vi) and (vii) have an intuitive interpretation that has been alluded to earlier. The first term is the change in deadweight loss due to distorted consumer prices resulting from raising the consumer price of the good in question by one unit. The second term is the corresponding change in deadweight loss due to moral hazard. Thus, at the constrained optimum, the overall deadweight loss, relative to the first best, is minimized. Arnott and Stiglitz solve for the optimal tax rates in simple variants of their model with only one relative price. They find (subject to minor qualifications) that accident–prevention goods and goods complementary to effort should be subsidized. In terms of the model presented, these results suggest that the agricultural input should be subsidized,[5] as well as the consumer good that is more complementary to effort.[6] This accords with the intuition given earlier that commodity taxation/subsidization is desirable to stimulate effort which is inefficiently low due to moral hazard.

The above analysis has, I think, succeeded in its modest goal of demonstrating that optimal tax theory can be fruitfully employed in developing a theory of optimal stabilization policy.

However, the model is much too simple. It is non–dynamic.[7] It also ignores business cycles. Greenwald and Stiglitz (1991) cite four stylized facts that any satisfactory theory of business cycles, and hence of stabilization policy, should explain:

---

[5] Unless this causes the individual to reduce effort by so much that the probability of accident increases.

[6] Note that which goods will be taxed and which subsidized depends on the normalization employed for consumer prices.

[7] Stern (1992) provides an excellent discussion of the issues involved in extending optimal tax theory from a static to a dynamic environment.

(a) cyclical movements in real product wages; (b) cyclical patterns of output and investment, including inventories; (c) sensitivity of the economy to small perturbations; (d) persistence. To this list, I would add cyclical movements in unemployment. The model presented explains <u>none</u> of these stylized facts. To develop a satisfactory theory of optimal stabilization policy, it will be necessary to develop a satisfactory theory of business cycles. The books by Stokey and Lucas (1989) and by Greenwald and Stiglitz (in process) would appear to be good starting points.

Any theory of optimal policy has three essential ingredients: a rigorous model of general equilibrium, a fully–specified welfare function, and market failure. These three ingredients essentially define optimal tax theory, which I argue in the next section is, broadly interpreted, a theory of optimal economic policy. Thus, I contend that optimal tax theory is an <u>essential</u> component of any satisfactory theory of optimal stabilization policy.

I think the time is ripe to bring stabilization policy back into public finance. Building on progress made in other branches of economics over the last two decades, we should be able to extend the basic static, certain optimal tax model to stochastic dynamic settings in ways that permit insightful and conceptually sound analyses of optimal stabilization policy.

We should be grateful to Jim Mirrlees for advocating that the Stabilization Branch of Richard Musgrave's Fiscal Department be reopened.

3. In Defense of Optimal Tax Theory

Twenty years ago optimal tax theory was "in." Students packed the classrooms in graduate public finance courses, eager to learn about the revolution in the field. There was a sense of heady optimism. With a powerful new set of conceptual tools, there was the hope that significant advances would be made in solving for the optimal tax system.

The early models (1971–74) were by and large encouraging. They gave rise to interesting, intelligible, and non–trivial results: Tax goods according to their complementarity with leisure; tax goods such that the proportional reduction in consumption of all goods is the same; impose a marginal income tax rate of zero at the highest income level, etc. These first phase results (so I gather from Haveman's comments) have become part of the conventional wisdom and have formed the basis for some policy.

The next generation of models (1975–80) were less encouraging. Not only did they indicate that the qualitative results of the earlier models were non–robust,[8] but they also derived optimal tax formulae that were so implicit that they provided little insight into which goods should be taxed heavily and which lightly, or into the optimal progressivity of the income tax system.

The result was a decline of interest in the new public economics, as well as some disenchantment and rejection, and a shift within the field towards positive and more down–to–earth issues. As well, many of the contributors to optimal tax theory during its early incandescent phase moved on to other problems.

Optimal tax theory is now "out," or perhaps down but not out. Interest in graduate public finance is at a low ebb. Atkinson and Stiglitz' textbook, Lectures on Public Economics, the high water mark of optimal tax theory, is now out of print in North America.

The time has now come to take a sober, second look at optimal tax theory. I shall argue that optimal tax theory made several very important and lasting general conceptual contributions to economic theory, especially to the economic theory of policy, but that its specific contributions to public economics, over and above its contributions to economics generally, have been slim.

---

[8] Stiglitz (1982) provides an excellent discussion of the fragility of the results derived in the early optimal income tax literature.

3.1 The contributions of optimal tax theory

Optimal tax theory has had a profound and lasting effect on applied economic theory and practice. The theory has made five distinct contributions. First, by employing a cardinal welfare function, it permitted the integration of equity and efficiency in applied economics. Second, by formulating the optimal tax problem in a general equilibrium framework à la Arrow–Debreu, it catalyzed the use of general equilibrium analysis in applied economics. Third, by providing the first rigorous treatment of asymmetric information in applied economics, it stimulated application of the economics of information. Fourth, it gave rise to a richer and more reasonable view of the appropriate role of government in a market–oriented economy. And fifth, it provided rigorous foundations for the theory of the second best.

3.1.1 *The "new, new welfare economics"*

The new welfare economics, pioneered by Lionel Robbins (1932), declared inter–personal utility comparisons to be invalid. Economists were to restrict their attention to efficiency since efficiency analysis is scientific, and were to eschew equity concerns since they are metaphysical. This put normative economics in an intellectual straitjacket for over thirty years; the discipline still suffers from this misguided attempt to expel value judgements from economics, and to focus on efficiency. Policy makers could not, however, ignore equity. In policy analysis, equity was incorporated by tacking on an imprecise equity–efficiency tradeoff to the rigorous efficiency analysis.

The "new, new welfare economics" was developed primarily in the context of optimal tax theory, starting with Mirrlees' seminal paper on the optimal income tax. The basic idea of the new, new welfare economics is that value judgements should be admitted completely explicitly, via a cardinal social welfare function. That way efficiency and equity considerations can be integrated in policy analysis in a conceptually coherent manner. The equity–efficiency tradeoff can be formalized. Policy advisors may then offer the policy maker a range of policy options from which he may choose on the basis of his values. By introducing a rigorous formulation of equity and a precise language for discussing equity and efficiency together, the new, new welfare economics has led to a considerable improvement in the clarity and quality of policy debate.

These points can be illustrated with reference to cost–benefit analysis. Some assumptions concerning distributional weighting have to be made. In the new welfare economics, since inter–personal utility comparisons are deemed invalid, Occam's Razor or the principle of insufficient reason is used to justify the assumption that "a dollar is a dollar is a dollar". Furthermore, where equity considerations are evidently of importance, some discussion of equity effects is appended to the cost–benefit calculus. This procedure seems arbitrary and needlessly indirect at best.

Drèze and Stern (1987) provide an excellent description of the cost–benefit procedure employed in the new, new welfare economics. A project is accepted simply if it increases welfare. The general equilibrium of the economy is described as the solution to a welfare maximization problem subject to a complete listing of resource, informational, and political constraints. The Lagrange multipliers on the resource constraints (with an arbitrary normalization) are the appropriate shadow prices, and a small project should be accepted if it generates a profit evaluated using these shadow prices. The superiority of this approach compared to that of the new welfare economics is evident by contrasting the clarity of Drèze and Stern's treatment of the social rate of discount with the fuzziness of the earlier literature.

3.1.2 *Applied general equilibrium analysis*

International trade theory deserves the credit for being the first applied field in economics to employ general equilibrium analysis. But prior to optimal tax theory, almost all general equilibrium analysis in trade theory was in the context of the 2x2x2 Hecksher–Ohlin model. The first use of general equilibrium theory à la Arrow–Debreu in the applied economics literature was by Diamond and Mirrlees (1971) in their seminal papers on optimal commodity taxation. Since then, use of the Arrow–Debreu general equilibrium framework has become a quality standard in all branches of applied economics.

Prior to optimal tax theory, normative economic analysis not only separated equity and efficiency, but also (with the exception of international trade theory) relied almost exclusively on partial equilibrium analysis. The conditions under which single–market partial equilibrium welfare analysis is valid—equality of the marginal utility of income across consumers, and either no distortions elsewhere in the economy or zero income and cross–price (between the commodity under consideration and other commodities) effects—are extremely restrictive. And the extension of partial equilibrium analysis to multiple markets is fraught with pitfalls. A well–known example is from optimal commodity tax theory. Partial equilibrium analysis gives rise to the inverse elasticity rule. But if all goods are taxable, a uniform tax is optimal since it is equivalent to a lump–sum tax. General equilibrium is just much easier to do correctly. Nowadays, almost all normative analysis of quality is done employing general equilibrium models. Optimal tax theory led the way in this changeover.

A major contribution to applied economics over the last twenty years has been the development of increasingly sophisticated and detailed general equilibrium simulation models. This development was spurred not only by improvements in computer technology, but also by the switch to general equilibrium thinking by applied economists.

It is hard to overstate the importance of the general equilibrium revolution. The use of general equilibrium analysis not only enforces conceptual consistency—everything has to be accounted for—but also alters the focus of analysis from the single market to the whole economy. This change in perception has resulted in better economic analysis. And optimal tax theory was a major contributor to this revolution.

3.1.3 *Asymmetric information in applied economics*

It is now generally acknowledged that asymmetries of information are pervasive and are a major determinant of the structure of economic organization. Recognition of the importance of asymmetric information occurred only very recently, however. Insurers must have had at least an inchoate understanding of moral hazard and adverse selection for thousands of years. And the preference revelation problem for public goods has long been recognized. But Arrow's paper on the welfare economics of medical care (1963) is arguably the first to discern the broad importance of asymmetric information.

In the subsequent decade, there were a number of important contributions to the literature on asymmetric information. One was Mirrlees' paper on the optimal income tax (1971). The problem facing the government is to raise a given amount of revenue, at the least cost in terms of social welfare, from a population of heterogeneous individuals who differ in ability. If there were symmetric information, the government would tax on the basis of ability, equalizing everyone's marginal utility of income. But it is assumed that the government cannot observe ability, and instead bases its taxation on observable income. Then individuals have an incentive to reduce labor supplied so as to reduce their tax liability. Thus, redistribution has efficiency effects.

Mirrlees' paper was of historical importance in a number of respects. It was arguably the first paper on mechanism design (Vickrey's paper on auctions (1961) was not discovered until later), and can therefore be said to be one of the cornerstones of the burgeoning literature on the subject. Intensive study of the optimal income tax problem has also led to a greatly improved understanding of equilibrium with adverse selection with multiple types or a continuum of types.

The paper had a major impact on policy analysis as well. It was the first to impose informational constraints on the government's policy design problem. It was also the first to analyze formally the equity–efficiency tradeoff. And consequently, it was the first to model explicitly the role of informational constraints in limiting the scope for redistribution.

The optimal commodity tax literature has contributed less to the theoretical and applied literatures on asymmetric information, but not insubstantially. That asymmetric information generally causes shadow prices to deviate from market prices was first recognized in the context of optimal commodity taxation. Issues related to linear vs non–linear pricing were also scrutinized in the optimal commodity literature. With distortions, non–linear commodity taxation is generally desirable but requires that an individual's total consumption be observable; where only anonymous transactions can be taxed, linear pricing is implied.

3.1.4 *The role of government*

A generation ago, the judgement of most North American economists was that, apart from the enforcement of law and of property rights, governments should intervene only to correct demonstrable market failures—governments should provide public goods, regulate natural monopolies, and internalize production and consumption externalities. This belief was based primarily on the First and Second Theorems of Welfare Economics, derived from the Arrow–Debreu model of general equilibrium. Most economists also considered that some lump–sum redistribution should be undertaken on ethical grounds.

The insights derived from optimal tax theory have led to a fundamental reevaluation of the role of government. From the optimal income tax problem came the insight that lump–sum redistribution is infeasible, and consequently that efficiency cannot be separated from equity. Lump–sum redistribution must be based on exogenous characteristics of the individual, on the basis of which it is deemed fair to redistribute (ability in Mirrlees' model). Most such characteristics, however, are only imperfectly measurable. Redistribution is therefore based on measurable, endogenous characteristics of the individual (income in Mirrlees' model). But then individuals have an incentive to modify these characteristics so as to reduce their tax liability, which has efficiency effects.

The theory of optimal commodity taxation reinforced the insight that with an unalterable distortion, it is in general desirable to introduce offsetting distortions. For example, since redistribution via the optimal income tax distorts the labor–leisure tradeoff, it is in general desirable to impose a set of differential commodity taxes. Pursuit of this train of thought gave rise to the insight that the presence of asymmetric information upsets both the First and Second Welfare Theorems and causes shadow prices to deviate from market prices.

An implication of these results is that the potential scope for welfare–improving government intervention is considerable. However, whereas previously it was judged that where the market fails, the government should intervene, it is now recognized that those factors (transactions costs, asymmetric information, etc.) which cause the market to fail have implications for the effectiveness of government intervention. Thus, the modern view is that economists should search for the efficient structure of economic organization, which will generally include some mix of market, government, and private, non–market institutions.

The role of government is a central issue in economics. The issue will never be resolved, but optimal tax theory has raised the level of the debate.

3.1.5 *Theory of the second best*

The theory of the second–best is concerned with optimal policy response in the presence of unalterable distortions. Drawing together disparate second–best problems that had been treated, including Ramsey (1927) and Corlett and Hague (1953/4) from the literature on taxation, Lipsey and Lancaster (1956/7) provided a general statement of the problem. However, the formal development of the theory of the second–best occurred almost entirely in the context of optimal taxation.

The theory of the second–best is now widely employed in all branches of applied microeconomics. One common application is to public utility pricing and investment; there the distortion is a break–even or similar constraint in the presence of decreasing costs. Another common application is to urban transportation, where the distortion is that cars do not pay for the congestion they cause (e.g., Wilson,(1983)). The results of second–best pricing—Ramsey pricing—are especially well–known.

Insights derived from second–best theory—especially that policy rules derived from partial equilibrium analysis do not generally hold where there are distortions in other markets—are now found in almost all policy discussions. For example, twenty years ago most housing policy economists advocated minimal government intervention, whereas they now recognize that housing policy should be designed taking into account such distortions as the property tax, imperfect capital markets, and the infeasibility of lump–sum taxation.

Thus, optimal tax theory has profoundly altered the theory, philosophy, and practice of economic policy, and has led to significant improvement in both academic and popular discussion of economic policy.

3.2 Optimal income tax theory and income tax policy

In his paper in this volume, Haveman argues that optimal income tax theory has seriously misled policy makers into believing that income tax progressivity is undesirable. This has come about, he states, because the theory has neglected many important considerations, and because policy makers have incorrectly interpreted the optimal tax theory results concerning marginal progressivity to apply to average progressivity. Haveman appears to hold responsible optimal income tax theory and perhaps optimal income tax theorists as well, for this unfortunate state of affairs.

In this section, I shall take issue with two general points in Haveman's paper. The first is that theory is to blame for policies that are misguided because they are designed on the basis of unrealistically simple models. The second is that elaboration of the optimal income tax model to incorporate realistic complications will (for any social welfare function) most likely increase optimal progressivity.

3.2.1 *Applied theory and policy analysis*

I did not know that policy makers had interpreted the results of optimal income tax theory as implying that income tax progressivity is undesirable. A thorough reading of the optimal income tax literature suggests, instead, that quantitative and even qualitative results are non–robust. This is argued particularly forcefully in Stiglitz (1982): If income uncertainty is introduced, if there are two or more labor types (e.g., different kinds of ability) that are imperfect substitutes in production, or if income is measured with noise, then the qualitative properties of the optimal income tax model are altered. But, for the sake of argument, let us suppose that policy makers have indeed based policy on misinterpretations of simple optimal tax models, which seems quite plausible.

Good theorists are typically not good policy advisors. The mark of the good theorist is his ability to develop a useful conceptualization clearly. This requires a combination of sound intuition concerning the mechanism of the economy and an ability to simplify and abstract. Simplification and abstraction entail focussing on some important facets of a problem while ignoring others or paying them less attention than their empirical importance merits. Sound policy advising, meanwhile, requires good judgement concerning the relative importance of different aspects of a policy problem, sensitivity to political currents, and concern for practical detail. Thus, except for sound intuition, the qualities that make a good theorist are the antithesis of those that make a good policy advisor. A division of labor is therefore appropriate. Applied theorists should produce models and policy economists should use the insights derived from them (along with the results of econometric studies, and issues raised in public debate that have been ignored in the academic literature) as ingredients in their policy analysis.

According to this view, the applied theorist has done his job well if he has produced a model that sheds new light on some policy issues. He should not be held responsible for any policies that derive from overstressing or misinterpreting his model's results. The responsibility lies rather with the policy analyst whose job it is to weigh the policy significance of new models' results.

The issue is not that clear–cut, however, since many applied theorists make exaggerated claims for the policy relevance of their work. Many articles in applied theory present numerical examples and include some policy discussion. The intention of the numerical examples should be to show the sensitivity of results to certain parameters and to indicate the order of magnitude of various effects. But often applied theorists overstate the practical relevance of their numerical results, whether because they really believe that their models capture reality, or because they feel the need to oversell their product. Similarly, policy discussion in articles in applied theory should be presented as demonstrating how the model provides insights into a few features of a policy problem. Too often, however, the policy discussion is presented as being more comprehensive than it actually is, and caveats are often mumbled.

Policy analysts should have the common sense to discount inflated claims of policy relevance in articles in applied theory and to realize that famous theorists, while very smart, can be seriously deficient in policy sense. But at the same time, theorists need to be more circumspect and modest in their claims of policy relevance.

3.2.2 *Realistic complications and progressivity*

There are indeed many considerations relevant to the design of the income tax that optimal income tax theory has overlooked. Haveman lists a number of these neglected considerations which he feels should be incorporated into optimal income tax theory before it will provide reliable guidance concerning the optimal degree of progressivity. He argues that incorporation of these considerations would most likely support a more progressive income tax. While his discussion is insightful, sensitive, and well–informed, its balance is open to question.

A strong case can be made that the most important oversight of the basic income tax model is its neglect of administrative, "financial"[9], and compliance costs. According to this view, the income tax system should be radically simplified. Much of the complexity of the tax system stems from its marginal progressivity.[10] The more steeply the marginal tax rate rises with income, the greater the taxpayer's incentive to self–average income, which he can do by altering the timing of his expenses and receipts. If, indeed, the administrative, financial, and compliance costs associated with marginal progressivity are high, then a demogrant system is called for, which combines a poll subsidy with a flat–rate tax and hence exhibits average progressivity.

A strong case can also be made that the search for the optimal degree of income tax progressivity is a will–o'–the–wisp. Since the simple optimal income tax model is non–robust, a fortiori more complex models will be non–robust. Even the qualitative results will depend on the details of the model's specification and on parameter values. Then determining the optimal degree of progressivity will require precise parameter estimates and testing of competing models. But given the lack of agreement among empirical studies concerning the magnitude of such basic parameters as labor supply elasticities, it is open to question whether the simultaneous dynamic processes of improvements in data, theory, and econometric method will converge to a consensus concerning the optimal degree of progressivity.

---

[9]By "financial" costs, I mean the taxpayer's costs of altering his financial affairs without altering his real activities--relabelling income, changing the timing of receipts and expenditures, altering organizational form and financial structure, etc.

[10]This is argued forcefully in Blum and Kalven's book, An Uneasy Case for Progressive Taxation.

BIBLIOGRAPHY

Arnott, R., A. Hosios, and J. Stiglitz, "Implicit Contracts, Labor Mobility and Unemployment", American Economic Review 78 (1988), 1046–1066

Arrow, K., "Uncertainty and the Welfare Economics of Medical Care", American Economic Review 53, 1963, 941–973

Atkinson, A., and J. Stiglitz, Lectures in Public Economics (New York: McGraw–Hill, 1980)

Atkinson, A., and N. Stern, "Pigou, Taxation and Public Goods", Review of Economic Studies 41, 1974, 119–128

Auerbach, A., "The Theory of Excess Burden and Optimal Taxation", in A. Auerbach and M. Feldstein, eds., Handbook of Public Economics, v.1 (Amsterdam: North–Holland, 1985) 61–127

Azariadis, Costas, and Roger Guesnerie, "Sunspots and Cycles", Review of Economic Studies 53 (1986) 725–737

Benassy, J.P., The Economics of Market Disequilibrium (New York: Academic Press, 1982)

Blum, W. and H. Kalven, The Uneasy Case for Progressive Taxation (Chicago: University of Chicago Press, 1963)

Cooper, R., and A. John, "Coordinating Coordinations' Failures in Keynesian Models", Quarterly Journal of Economics 103, 1988, 441–463

Corlett, W.J., and D.C. Hague, "Complementarity and the Excess Burden of Taxation", Review of Economic Studies 21, 1953–4, 21–30

Diamond, P., "A Many–Person Ramsey Tax Rule", Journal of Public Economics 4, 1975, 335–342

Diamond, P.A., A Search–Equilibrium Approach to the Microfoundations of Macroeconomics, The Wicksell Lectures, (Cambridge, MA: M.I.T. Press, 1984)

Diamond, P., and J. Mirrlees, "Optimal Taxation and Public Production. I: Production Efficiency. II: Tax Rules", American Economic Review 61, 1971, 8–27 and 261–278

Drèze, J., and N. Stern, "The Theory of Cost–Benefit Analysis" in A. Auerbach and M. Feldstein, eds., Handbook of Public Economics, v.2 (Amsterdam: North–Holland, 1987) 909–989

Feldstein, M., "Distributional Equity and the Optimal Structure of Public Prices", American Economic Review 62, 1972, 32–36

Greenwald, B., and J. Stiglitz, "Macroeconomic Models with Equity and Credit Rationing", in R.G. Hubbard, ed., Asymmetric Information, Corporate Finance and Investment (Chicago: University of Chicago Press, 1990)

Hart, O., "A Model of Imperfect Competition with Keynesian Features", Quarterly Journal of Economics 97, 1982, 109–138

Lipsey, R., and K. Lancaster, "The General Theory of the Second Best", Review of Economic Studies 24, 1956–7, 11–32

Lucas, R., and N. Stokey, Recursive Methods in Economic Dynamics (Cambridge, MA: Harvard University Press, 1989)

Mirrlees, J., "An Exploration in the Theory of Optimal Income Taxation", Review of Economic Studies 38, 1971, 175–208

Mirrlees, J., "Optimal Commodity Taxation in a Two–Class Economy", Journal of Public Economics 4, 1975, 27–33

Musgrave, R., The Theory of Public Finance: A Study in Political Economy (New York: McGraw–Hill, 1959)

Ramsey, F., "A Contribution to the Theory of Taxation", Economic Journal 37, 1927, 47–61

Robbins, L., An Essay on the Nature and Significance of Economic Science, (London: MacMillan, 1932)

Shapiro, Carl, and Joseph E. Stiglitz, "Equilibrium Unemployment as a Worker Discipline Device", American Economic Review 74 (1984), 433–444

Stern, N., "From Static to Dynamic Taxation", Journal of Public Economics 47, 1992, 273–297

Stiglitz, J., and P. Dasgupta, "Differential Taxation, Public Goods, and Economic Efficiency", Review of Economic Studies 38, 1971, 151–174

Stiglitz, J., "Pareto Efficient and Optimal Taxation and the New New Welfare Economics", in A. Auerbach and M. Feldstein,eds., Handbook of Public Economics, v.2 (Amsterdam: North Holland, 1987) 991–1042

Stiglitz, J., "Self–Selection and Pareto Efficient Taxation", Journal of Public Economics 17, 1982, 213–240

Taylor, J., "Aggregate Dynamics and Staggered Contracts", Journal of Political Economy 88, 1980, 1–24

Vickrey, W., "Counterspeculation, Auctions, and Competitive Sealed Tenders," Journal of Finance 16, 1961, 8–37

Wilson, J. "Optimal Public Good Provision with Limited Lump–Sum Taxation", American Economic Review 81, 1991, 153–165

Wilson, J. "Optimal Road Capacity in the Presence of Unpriced Congestion", Journal of Urban Economics 13, 1983, 337–357

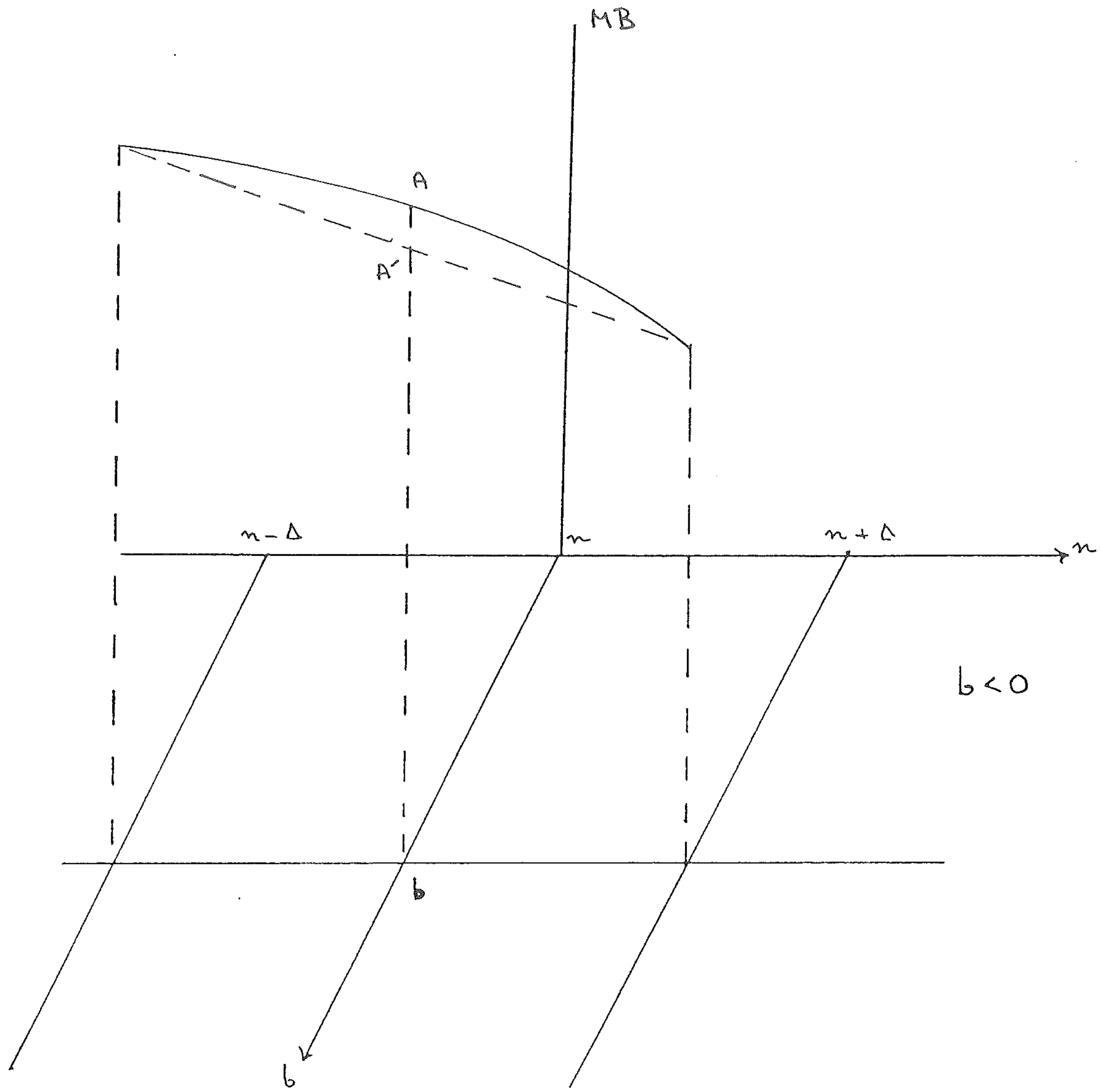

Figure 1: Effect of perturbation on average marginal benefit: poll tax

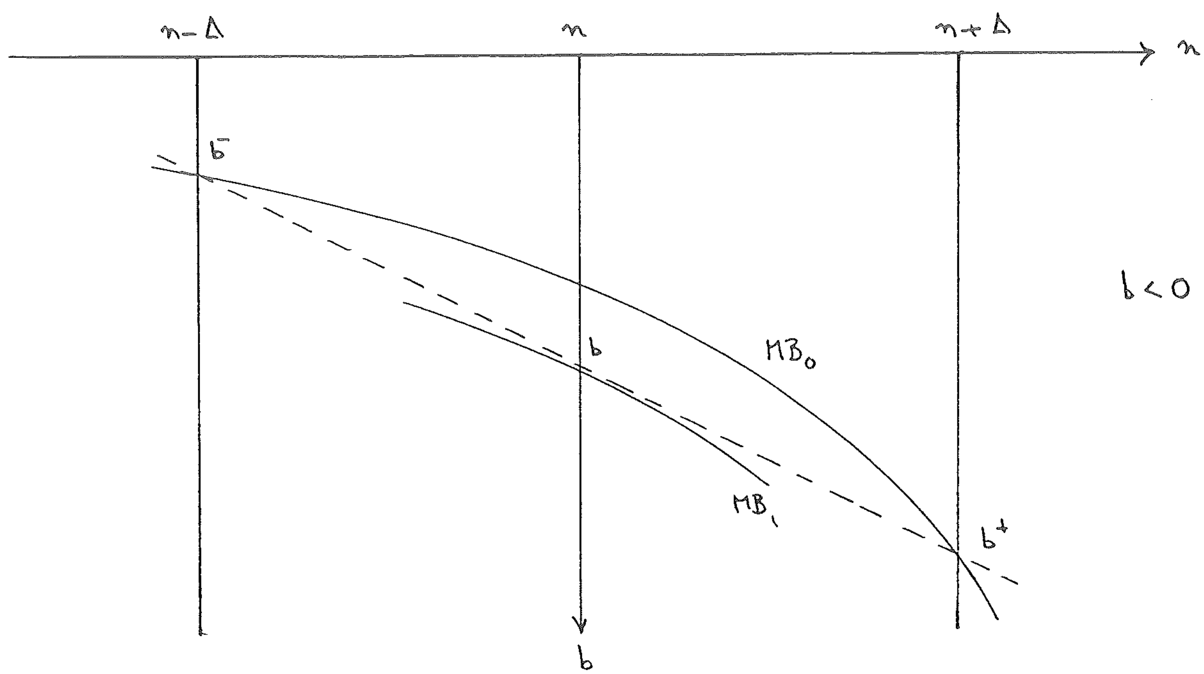

Figure 2: Effect of perturbation in average marginal benefit: first-best taxes