ECONOMIC RESEARCH REPORTS

INTERPRETING "THE RAMSEY EQUATIONS" OF OPTIMAL TAX THEORY

by

Charles A.M. de Bartolome

R.R. #89-21

August 1989

C. V. STARR CENTER
FOR APPLIED ECONOMICS

NEW YORK UNIVERSITY
FACULTY OF ARTS AND SCIENCE
DEPARTMENT OF ECONOMICS
WASHINGTON SQUARE
NEW YORK, N.Y. 10003

INTERPRETING "THE RAMSEY EQUATIONS" OF OPTIMAL TAX THEORY

by

Charles A.M. de Bartolome*

ABSTRACT

"The Ramsey Equations," which characterize the optimum indirect tax structure, are interpreted using the intuitive ideas of excess burden and willingness to pay.

* New York University. The technical assistance provided by the C.V. Starr Center is gratefully acknowledged.

1. INTRODUCTION

An individual is affected not only by the level of tax revenue transferred from him to the government, but also by the tax structure used to effect the transfer. The use of "distorting" taxes to collect tax revenue imposes a welfare cost on the individual in excess of the resources transfered to the government. The optimal indirect tax structure (to collect a given tax revenue from a representative individual) is specified by "The Ramsey Equations," which are a foundation of the optimal tax literature, and may be used directly to derive the results of Corlett and Hague (1953-54), Diamond (1975), and Bradford and Rosen (1976). Similar equations ("The Ramsey Pricing Formulae") also occur in the optimal pricing of public utilities. The focus of this paper is the intuitive interpretation of these important equations.

At least since Dupuit (1844), it has been realized that the utility loss associated with the use of a particular tax instrument, or the utility loss in excess of the loss associated with the transfer of resources, can be loosely measured by the triangular area between the demand and supply curves. This area was termed the excess burden (or the deadweight loss) of the tax system: the optimal tax structure should minimize the excess burden. This methodology was developed and extended by Jenkin (1871-2), Marshall (1890), Pigou (1928), Hotelling (1938), Boiteux (1951), Debreu (1951 and 1954) and Harberger (1964). The excess burden was reformalized in terms of the expenditure function by Mohring (1971), and Diamond and McFadden (1974).

An alternative methodology is utility maximization. Ramsey (1927) calculated how indirect taxes (proportional sales taxes) should be levied on commodities in order to maximize the utility of the representative household, subject to the tax requirement. He showed that products should be taxed differentially, with higher taxes being placed on goods with lower demand elasticities. The equation system characterizing the optimum ("The Ramsey Equations") appears in Dixit (1970) and is reformulated in terms of compensated demands in Diamond and Mirrlees (1971), Stiglitz and Dasgupta (1971), Atkinson and Stern (1974), Diamond (1975), Sandmo (1976), Atkinson and Stiglitz (1980), and Samuelson (1986).

Kay (1980), Stutzer (1982) and Auerbach (1985) combine the two approaches, showing that the Ramsey Equations, derived by the constrained maximization of utility, have the same *form* as the equations derived by the constrained minimization of the excess burden. This is illustrated graphically and this paper completes their analysis by showing the identity of the constraints in the two equation sets.

The focus of this paper is an intuitive interpretation of the Ramsey Equations. At the optimum, it is equally costly to raise an extra unit of tax by each instrument, and the cost for *each* instrument is equal to the cost when *all* instruments are adjusted optimally. The cost is measured by the intuitive concepts of the increase in the excess burden, and of the willingness to pay to avoid the use of indirect taxes.

Closely related to the Ramsey Equations of the optimal tax literature are the "Ramsey Pricing Formulae" of public utilities (e.g., Boiteux (1956) and Baumol and Bradford (1970)). In this context, the planner must choose consumer prices to maximize consumer welfare subject to the public utility recovering fixed costs or achieving a fair return to shareholders. The intuition provided in this paper may be directly used to explain the optimal mark-up over marginal cost in these regulated industries.

This paper is organized as follows. Section 2 derives and interprets The Ramsey Equations by minimization of the excess burden. Section 3 rederives the Equations by constrained utility maximization, and interprets them using "willingness to pay" arguments. Section 4 concludes.

2. MINIMIZATION OF EXCESS BURDEN

The transfer of resources from the household to the government implies an inevitable lowering of household utility. Figure 1 represents the economic problem of the representative household and planner, in the simple case of a good (consumption good) and resource (labor). Indifference curves are drawn and are labelled by their associated utility levels $U^2 < U^1 < U^0$. Labor is the numeraire, and p units of labor are required to produce one unit of good, or p is the resource cost of the good. OA (of slope 1/p) is the production possibility function of the economy and, if there were zero tax requirement, the household could achieve utility U^0. The planner must extract, through taxation, resources R from the household, or must move the household to an allocation on the planner's budget line BC, horizonatlly displaced from OA by distance R. The highest utility achievable on BC is U^1, and is achieved at E^1. If the planner were to directly choose the household allocation, he would allocate the good (by moving up BC from B to E^1) until the household's willingness to pay for a marginal unit of good (measured as the inverse of the slope of its indifference curve) is equal to the resource cost of the good (measured by the inverse of the slope of the planner's budget line).

Under competitive laissez-faire, the household buys the good (by selling labor) until its willingness to pay for an additional unit of the good is equal to the price it would have to pay for the additional unit (measured by the inverse of the slope of the household budget line). With a lump sum tax, the consumer price equals the resource cost, the houshold and planner budgets and marginal conditions coincide, and the household chooses the allocation at E^1. The utility loss necessitated by the resource transfer is U^0-U^1.

Household utility is derived from the quantity and *mix* of goods and leisure, and not from total resources per se. A planner, restricted to indirect taxation, must set the tax rate so that the household chooses an allocation on BC. This is achieved by raising the consumer price until the household budget is OD: faced with budget OD, the utility-maximizing household chooses E^2, which is on BC. Because the tax raises the consumer price above the resource cost, the household is behaving "as if" a consumption good has a higher resource cost than it really does, so that it buys "too few" goods, and consumes "too much" leisure. Put differently, with the tax rate at t, and if household labor supply were increased by one unit, 1/p extra goods would be produced but the household would need only 1/p(1+t) extra goods for compensation; therefore, household utility could be increased. The utility loss U^1-U^2 is associated with the use of indirect taxation to collect the tax revenue. It may be measured by the excess burden (or deadweight loss) of the tax structure.

Mohring (1971) formalized the concept of excess burden based on the equivalent variation.[1] The excess burden is the additional tax revenue (paid as lump-sum) which the household would be willing to pay in order to have the tax code replaced by a lump-sum levy of equal revenue R. In Figure 1, with the indirect tax structure, the household budget line is OD and it achieves utility U^2. It would achieve the same utility if it were to face budget line FG, or if it were to have a lump-sum levy equal to the horizontal distance of OA from FG. The excess burden, or the additional tax, is therefore the horizontal distance from BC to FG.

More generally, there are n commodities and labor. The consumption of the representative household of a typical commodity i is x_i and its labor supply is L. It is convenient to represent household consumption by the vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. Labor is the only input and is chosen as the numeraire; the wage is normalized to unity and is untaxed by assumption. The production of each commodity is assumed to show constant returns to scale: the production of one unit of good i requires p_i units of labor. The competitive producer price of good i is therefore p_i, independent of the quantity of production. Tax revenue R must be collected from the household and its collection is restricted to the indirect taxation of commodities; if the commodity i is taxed at rate t_i, its consumer price is $q_i = p_i(1+t_i)$. It is convenient to represent the producer and consumer prices by vectors, $\mathbf{p} = (p_1, p_2, \ldots, p_n)$ and $\mathbf{q} = (q_1, q_2, \ldots, q_n)$. Because $\mathbf{p}$ is fixed by the production technology, $\mathbf{q}$ is a representation of the tax structure.

The representative household is competitive and takes the after-tax price $\mathbf{q}$ as given when choosing its commodity purchase. The endowed non-labor income of the household is represented by M, (evaluation will be made at $M = 0$). Consumer demand is $\mathbf{x}(\mathbf{q}; M)$. The problem is about how the tax should be collected, and not about expenditure; the public project, on which the tax revenue is spent, is therefore ignored. The household receives utility $U(\mathbf{x}, L)$, for which the indirect utility function is $V(\mathbf{q}; M) \equiv U(\mathbf{x}(\mathbf{q}; M), L(\mathbf{q}; M))$. The excess burden for the tax code $\mathbf{q}$ is

$$EB(\mathbf{q}) = e(\mathbf{q}; V(\mathbf{q}; M)) - e(\mathbf{p}; V(\mathbf{q}; M)) - (\mathbf{q}-\mathbf{p})\cdot\mathbf{x}(\mathbf{q}; M),$$

where $e(\cdot;\cdot)$ is the expenditure function.

The planner's problem of this section is to choose the tax rates to minimize the excess burden subject to the revenue constraint. Remembering that $e(\mathbf{q}; V(\mathbf{q}; M)) \equiv M$, the planner's problem is

$$\min_{\mathbf{q}} \; M - e(\mathbf{p}; V(\mathbf{q}; M)) - (\mathbf{q}-\mathbf{p})\cdot\mathbf{x}(\mathbf{q}; M), \quad \text{s.t} \quad (\mathbf{q}-\mathbf{p})\cdot x(\mathbf{q}; M) = R.$$

The revenue constraint must hold with equality. The Lagrangian is $L_1 = M - e(\mathbf{p}; V(\mathbf{q}; M)) - (\mathbf{q}-\mathbf{p})\cdot x(\mathbf{q}; M) + \mu[R - (\mathbf{q}-\mathbf{p})\cdot x(\mathbf{q}; M)]$. The typical first-order condition is

$$\frac{\partial L_1}{\partial q_k} = - \frac{\partial e(\mathbf{p};V)}{\partial U} \frac{\partial V}{\partial q_k} - (1+\mu)[\sum_{i=1}^{n}(q_i-p_i)\frac{\partial x_i}{\partial q_k} + x_k] = 0.$$

Using Roy's Identity, $\partial V/\partial q_k = -\alpha x_k$, where $\alpha = \partial V(\mathbf{q}; M)/\partial M$ is the marginal utility of income, and rearranging,

$$\frac{x_k}{\sum_{i=1}^{n}(q_i-p_i)\frac{\partial x_i}{\partial q_k}+x_k} = \frac{1+\mu}{\frac{\partial e(\mathbf{p};V)}{\partial U}\cdot\alpha}, \qquad (k=1,\ldots,n).$$

This equation occurs in Dixit (1970), where the left-hand side is interpreted as the charge in consumer surplus if an additional unit of tax revenue were to be raised by adjustment of instrument q_k only. It can be rewritten as

$$\frac{\frac{\partial e(\mathbf{p};V)}{\partial U}\cdot\alpha x_k}{\sum_{i=1}^{n}(q_i-p_i)\frac{\partial x_i}{\partial q_k}+x_k} = (1+\mu), \qquad (k=1,\ldots,n). \tag{1}$$

Equation (1) is interpreted below. For later comparison, the equation can be rewritten, using the Slutsky Equation, as

$$-\frac{1}{x_k}\sum_{i=1}^{n}(q_i-p_i)\frac{\partial h_i}{\partial q_k} = 1-\frac{\alpha}{1+\mu}\frac{\partial e(\mathbf{p};V)}{\partial U}-\sum_{i=1}^{n}(q_i-p_i)\frac{\partial x_i}{\partial M}, \tag{2}$$

where $h_i(\mathbf{q};V)$ is the compensated demand of the ith good.

If a tax system $\mathbf{q}$ raises revenue $R(\mathbf{q})$ and has excess burden $EB(\mathbf{q})$, the representative household obtains utility $V(\mathbf{q};M) \equiv V(\mathbf{p};M-R(\mathbf{q})-EB(\mathbf{q}))$: $R(\mathbf{q})+EB(\mathbf{q})$ is the total purchasing power withdrawn from the household (evaluated at producer price $\mathbf{p}$).[2] If the consumer price of the kth commodity were to be raised from q_k to $q_k + dq_k$, causing a revenue increase of dR_k and an excess burden increase of dEB_k, the utility loss would be

$$dV = \frac{\partial V(\mathbf{q};M)}{\partial q_k} dq_k = - V_2(\mathbf{p};e(\mathbf{p};V(\mathbf{q};M))\cdot(dR_k+dEB_k). \tag{3}$$

where $M\text{-}R(\mathbf{q})\text{-}EB(\mathbf{q}) = e(\mathbf{p};\ V(q;\ M))$, and $V_2(\cdot;\cdot)$ is interpreted to mean the derivative of V with respect to the income position. The welfare cost of a small change in the tax code at prices $\mathbf{q}$ and consumer income M is evaluated by the income change at consumer prices $\mathbf{p}$ and consumer income $e(\mathbf{p};\ V(\mathbf{q};M))$ which would give an equal utility loss. Using Roy's Identity, $dV = -\alpha x_k dq_k$, or the change in utility is "as if" income $x_k dq_k$ is withdrawn. Hence, rearranging Equation (3),[3]

$$dR_k + dEB_k = \frac{\partial e(\mathbf{p};V)}{\partial U}\cdot\alpha x_k dq_k.$$

Noting $dR_k = d\sum_{i=1}^{n}(q_i-p_i)x_i = \sum_{i=1}^{n}(q_i-p_i)(\partial x_i/\partial q_k)dq_k + x_k dq_k$, the welfare loss per unit of marginal tax revenue raised through instrument q_k is

$$\frac{dR_k+dEB_k}{dR_k} = \frac{\dfrac{\partial e(\mathbf{p};V)}{\partial U}\cdot\alpha x_k}{\sum_{i=1}^{n}(q_i-p_i)\dfrac{\partial x_i}{dq_k} + x_k}.$$

The left hand side of Equation (1) is therefore the welfare cost of raising marginal tax revenue through instrument q_k, measured as the income change at prices $\mathbf{p}$ and income $e(\mathbf{p};V(\mathbf{q};M))$ which would give an equal utility loss. At the optimal tax structure, it is equally costly to raise marginal tax revenue through each instrument. Interpretating the Lagrangean multiplier as the change in the objective as the constraint is tightened, $1+\mu = d(R(\mathbf{q})+EB(\mathbf{q}))/dR$: the right-hand side is therefore the welfare cost of marginal tax revenue raised by optimal adjustment of *all* instruments. Equation (1) is now interpreted: at the optimum tax structure, it is equally costly to raise marginal tax revenue through adjustment of any instrument, as it is by optimal adjustment of all instruments.

3. MAXIMIZATION OF UTILITY

Auerbach (1985) notes that the excess burden can be written as

$$EB(\mathbf{q}) = M - e(\mathbf{p};\ V(\mathbf{q};\ M)) - R.$$

But $0 < \partial e/\partial U$, so that the minimization of the excess burden by choice of tax instrument $\mathbf{q}$, subject to the collection of given tax revenue R, is equivalent to the choice of $\mathbf{q}$ to maximize $V(\mathbf{q};\ M)$, subject to the same revenue constraint. The equivalence of the two problems is easily demonstrated using Figure 1. The revenue constraint implies an allocation on BC.[4] With convex indifference curves, movement along BC away from E^1 involves lower utility and (necessarily) higher excess burden.

"The Ramsey Problem" is to choose tax rates to maximize the utility achieved by the representative household subject to the revenue constraint. For expositional reasons, it is convenient to assume that tax revenue T may be collected as a lump-sum levy, (evaluation will be made at T = 0). The formal problem is therefore

$$\max_{\mathbf{q}} V(\mathbf{q};\ M\text{-}T), \qquad \text{s.t.} \quad R - T \leq (\mathbf{q}-\mathbf{p})\cdot x(\mathbf{q};\ M\text{-}T).$$

for which the Lagrangian is $L_2 = V(\mathbf{q};\ R\text{-}T) + \lambda[(\mathbf{q}-\mathbf{p})\cdot x(\mathbf{q};M\text{-}T) - R+T]$. The typical first order condition is

$$\frac{\partial L_2}{\partial q_k} = \frac{\partial V}{\partial q_k} + \lambda[\sum_{i=1}^{n}(q_i-p_i)\frac{\partial x_i}{\partial q_k} + x_k] = 0.$$

This compares to Equation (2.3.3) in Stiglitz and Dasgupta (1971): the change in utility per unit of extra tax revenue raised through instrument q_k must be equal across all commodities. Using Roy's Identity,

$$-\sum_{i=1}^{n}(q_i-p_i)\frac{\partial x_i}{\partial q_k} = \frac{\lambda-\alpha}{\lambda}\cdot x_k.$$

This corresponds to Equation (7) in Dixit (1970) and also in Sandmo (1976). Using Slutsky's Equation, and rearranging,

$$1 + \frac{1}{x_k}\sum_{i=1}^{n}(q_i-p_i)\frac{\partial h_i}{\partial q_k} = \frac{\alpha}{\lambda} + \sum_{i=1}^{n}(q_i-p_i)\frac{\partial x_i}{\partial M}, \qquad (k=1,\ldots n), \qquad (4)$$

or

$$-\frac{1}{x_k}\sum_{i=1}^{n}(q_i-p_i)\frac{\partial h_i}{\partial q_k} = \frac{\lambda-\alpha}{\lambda} - \sum_{i=1}^{n}(q_i-p_i)\frac{\partial x_i}{\partial M}, \qquad (k=1,\ldots n). \qquad (4')$$

This equation occurs in Diamond and Mirrlees (1971, Equation (37)), in Atkinson and Stern (1974, Equation (5)) and in Atkinson and Stiglitz (1980, Equation (12-16)). Equations (4) and (4') are the equations I will interpret.

It is normal to note that the substitution matrix is symmetric ($\partial h_i/\partial q_k = \partial h_k/\partial q_i$), to give "The Ramsey Equations,"

$$-\frac{1}{x_k}\sum_{i=1}^{n}(q_i-p_i)\frac{\partial h_k}{\partial q_i} = \frac{\lambda-\alpha}{\lambda} - \sum_{i=1}^{n}(q_i-p_i)\frac{\partial x_i}{\partial M}, \qquad (k=1,\ldots n). \qquad (5)$$

In addition to the authors cited above, this equation occurs in Stiglitz and Dasgupta (1971, Equation (2.3.7)), Sandmo (1976, Equation (9)), and Samuelson (1986, Equation (4)): the optimal tax design for small tax revenues ($(q_i-p_i) = \Delta q_i$) requires equal proportional reduction in compensated demands for all commodities. Ramsey's (1927) conclusion that "in raising an infinitesimal revenue by proportionate taxes on given commodities the taxes should be such as to diminish in the same proportion the production of each commodity taxed" is a special case, arising when all commodities except the numeraire have zero income elasticity, so that compensated and uncompensated demands are identical.

All compensated demands are evaluated at $V(\mathbf{q};M)$, so that Equations (2) and (4') are identical if[5]

$$\frac{1}{\lambda} = \frac{1}{1+\mu} \frac{\partial e(\mathbf{p};V)}{\partial U} .$$

Interpreting each Lagrangean multiplier as the change in the objective function as the constraint is tightened,

$$\lambda = - \frac{dV}{dR} \quad \text{and} \quad \mu = \frac{dEB(\mathbf{q})}{dR} = -\frac{\partial e(\mathbf{p};V)}{\partial U} \cdot \frac{dV}{dR} - 1.$$

Hence

$$\frac{1}{\lambda} = -\frac{1}{dV/dR} = \frac{1}{1+\mu} \frac{\partial e(\mathbf{p};V)}{\partial U} .$$[6]

The interpretation of Equation (4) and (4') is that it is equally costly to raise marginal tax revenue, by adjustment of each or all instruments. To interpret Equation (4), the marginal cost of instrument q_k is measured by the amount of extra tax, raised by adjusting instrument q_k, that a household is willing to pay in order to acquire a unit of income. If the household is willing to pay "much" tax revenue, then tax revenue from instrument q_k is of "low" cost. For each instrument q_k, the transaction can be thought of as two consecutive stages. In the first stage, the household acquires a unit of income; this raises its utility. In the second stage, the tax rate on the kth commodity is raised until the household's utility falls to its original value. The rise in utility in the first stage is the same for each instrument q_k. The rise in revenue at the second stage is therefore a measure of the relative sensitivity of utility to marginal revenue achieved by adjustment of instrument q_k. Little revenue implies the instrument is costly, or a "small" increase in revenue leads to a "large" fall in utility. In fact, the income transfer in the first stage increases the consumption of taxed goods and increases tax revenue. Our interpretation uses the total increase in tax revenue from the two stages as the measure of the marginal cost of instrument q_k: because the increase in the tax revenue in the first stage is the same for each instrument q_k, "little" tax revenue overall still implies "high cost".

When the household receives an income transfer of dM in the first stage, it is willing to allow the consumer price of the kth commodity to rise to q_k+dq_k in the second stage, or

$$V(q_1,\ldots,q_k,\ldots,q_n;\ M-T) \ = \ V(q_1,\ldots,q_k+dq_k,\ldots,q_n;\ M+dM\ -\ T).$$

I proceed intuitively. Expanding the left-hand side to first order terms gives

$$0 \equiv \frac{\partial V}{\partial q_k} dq_k + \frac{\partial V}{\partial M} dM,$$

or, using Roy's identity,

$$dM = -\frac{\partial V/\partial q_k}{\partial V/\partial M} dq_k \ = \ x_k dq_k.$$

The price rises by exactly the amount necessary for the increased price of the existing purchases to fully absorb the income transfer.

Remembering that utility is unchanged so that demand changes are changes in compensated demands,[7] the change in tax revenue is

$$dR_k' = d\sum_{i=1}^{n}(q_i-p_i)h_i = h_k dq_k + \sum_{i=1}^{n}(q_i-p_i)dh_i.$$

There are two effects: (1) "the price effect" being the extra tax collected on commodity h_k due to the higher tax rate, $h_k dq_k$, and (2) "the quantity effect" being the change in the tax revenue collected through the pre-existing tax structure due to the adjustment in quantities, $\sum_{i=1}^{n}(q_i-p_i)dh_i$.

The increase, by adjustment of instrument q_k, in tax revenue which the household is willing to pay to achieve a unit of income transfer is

$$\frac{dR_k'}{dM} = \frac{h_k dq_k + \sum_{i=1}^{n}(q_i-p_i)dh_i}{x_k dq_k} = 1 + \frac{1}{x_k}\sum_{i=1}^{n}(q_i-p_i)\frac{\partial h_i}{\partial q_k}.$$

This explains the left-hand side of Equation (4): it is the cost of raising extra tax revenue through instrument q_k, measured as the extra tax the household is willing to pay to acquire an extra unit of income. To acquire an extra unit of income, the household is willing to pay an additional unit of tax less the tax lost due to the adjustment of quantities in the household consumption bundle. At the optimum, it is equally costly to raise tax revenue through each instrument.

The right-hand side of Equation (4) is similarly interpreted as the extra tax the household is willing to pay in order to achieve an extra unit of income, with *all* tax rates being adjusted optimally. Again thinking of two consecutive stages, in the first stage the household receives an extra unit of income, tax revenue rises by $\sum_{i=1}^{n}(q_i - p_i)\partial x_i/\partial M$, and utility rises by α. In the second stage, tax rates on all commodities are changed, so that utility falls by α; the overall effect is to leave utility unchanged. By the interpretation of the Lagrangean multiplier, $\lambda = dV/dR$, so that in the second stage, additional revenue is raised of α/λ. The right-hand side of Equation (4) is therefore the total increase in tax revenue the household is willing to pay to acquire an extra unit of income; the first term corresponds to the extra tax raised by the adjustment in tax rates, and the second term to the extra tax raised by the extra income. Equation (4) is interpreted: at the optimum it is equally costly to raise marginal tax revenue by adjustment of each instrument as by the adjustment of all instruments, or the household is indifferent to small tax rate adjustments which leave total tax revenue unchanged.

The interpretation of Equation (4′) is similar. An alternative measure of the cost of raising marginal tax revenue through instrument q_k is the additional tax, raised with adjustment of instrument q_k, which the household is willing to pay for an additional unit of the tax revenue R to be collected by a lump-sum levy. As before, the process can be thought of as of two stages. Firstly, the household is levied an additional unit of lump-sum tax: this lowers utility but raises tax revenue. Secondly, the tax rate on the kth commodity is lowered until utility rises to its initial level. The fall in utility and rise in revenue in the first stage is the same across all commodities: the fall in revenue at the second stage is a measure of the relative sensitivity of utility to changes in tax revenue through adjustment of instrument q_k. A "small" fall in revenue in the second stage, giving a "large" increase in revenue overall, implies that a "small" change in tax revenue leads to a "large" utility change, or that the instrument is costly.

Formally, suppose when the lump-sum tax levy is increased by dT in the first stage, the household is willing to allow the price of the kth commodity to change by dq_k ($dq_k < 0$) in the second stage,

$$V(q_1,\ldots,q_k,\ldots,q_n;\ M-T) \;=\; V(q_1,\ldots,q_k+dq_k,\ldots,q_n;\ M-T\ -\ dT).$$

I proceed intuitively. Expanding to first order terms as before, and using Roy's Identity, $dT = -x_k dq_k$, the fall in disposable income in the first stage equals the fall in the price of the original allocation in the second stage. Remembering that utility is unchanged so that demand changes are changes in compensated demands, the change in total tax revenue

$$dR_k'' \;=\; d\left(\sum_{i=1}^{n}(q_i-p_i)h_i\right) + dT \;=\; h_k dq_k + \sum_{i=1}^{n}(q_i-p_i)dh_i\ +\ dT.$$

But $dT = -x_k dq_k = -h_k dq_k$: the increase in tax revenue in the first stage is exactly offset by the fall in tax revenues asssociated with the "price effect" in the second stage. Hence,

$$\frac{dR_k''}{dT} \;=\; -\,\frac{\sum_{i=1}^{n}(q_i-p_i)dh_i}{x_k dq_k}.$$

Thus the left-hand side of Equation (4′) is the cost of raising marginal tax revenue through instrument q_k, measured as the extra tax revenue which could be collected (by simultaneous adjustment of instrument q_k to leave utility unchanged) if the lump-sum levy was increased by a unit. The tax revenue increase is associated with the adjustment in quantities in the household consumption bundle. At the optimum, it is equally costly to raise tax revenue through each instrument.

The interpretation of the right-hand side of Equation (4') is similarly the cost of raising marginal tax revenue by the optimal adjustment of *all* tax instruments. Considering the two stage process, in the first stage the lump-sum levy is raised by dT. This increases tax revenue by

$$d\tilde{R}_1 = dT - \sum_{i=1}^{n}(q_i - p_i)\frac{\partial x_i}{\partial M}dT,$$

and lowers utility by αdT. In the second stage, tax rates are optimally adjusted to raise utility by αdT with minimum fall in tax revenue. By the interpretation of the Lagrangean multiplier, for each unit of increased revenue, utility falls by λ. Therefore, to achieve a utility increase of αdT, tax revenue must change by $d\tilde{R}_2 = -\alpha dT/\lambda$. Hence the household is willing to pay additional taxes, if one unit of the total tax revenue is levied as a lump-sum, of

$$\frac{d\tilde{R}_1 + d\tilde{R}_2}{dT} = 1 - \sum_{i=1}^{n}(q_i - p_i)\frac{\partial x_i}{\partial M} - \frac{\alpha}{\lambda}.$$

The right-hand side of Equation (4') therefore corresponds to the cost of marginal tax revenue, measured as the additional tax which could be collected if the lump-sum levy were increased by one unit. The first term corresponds to the lump-sum levy, the second term to the tax fall due to the fall in demand with the fall in disposable income, and the third term to the tax fall due to the necessary adjustment in tax rates to leave utility unchanged. The interpretation of Equation (4') is that at the optimum, it is equally costly to raise tax revenue by adjustment of each instrument or by the adjustment of all instruments, and the household is indifferent to small tax rate adjustments which leave tax revenue unchanged.

4. CONCLUSION

This paper provides the intuition for "The Ramsey Equations" of the optimal tax literature. When tax rates are set optimally, it is equally costly to raise marginal tax revenue through the adjustment of any instrument, or by the optimal adjustment of all instruments. The marginal cost of a tax instrument is measured by using the excess burden and the willingness to pay concepts.

REFERENCES

Atkinson, A.B., and N.H. Stern, (1974), "Pigon, Taxation and Public Goods," *Review of Economic Studies*, 41, 119-128.

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

Auerbach, A.J., (1985), "The Theory of Excess Burden and Optimal Taxation," Chapter 2 of *The Handbook of Public Economics, Vol. 1*, edited by A.J. Auerbach and M. Feldstein. New York: Elsevier Science Publishers.

Baumol, W.J., and D.E. Bradford, (1970), "Optimal Departures from Marginal Cost Pricing," *American Economic Review*, 60, 265-283.

Boiteux, M., (1951), "Le "Revenue Distribuable" et Les Pertes Economiques," *Econometrica*, 19, 112-133.

———, (1956), "Sur la Gestion des Monopoles Publics Astreints a l'Equilibre Budgetaire," *Econometrica*, 24, 22-40. Translated and reprinted in *Journal of Economic Theory*, 3, 219-240.

Bradford, D. and H. Rosen, (1976), "The Optimal Taxation of Commodities and Income," *American Economic Association Papers and Proceedings*, 66, 94-101.

Corlett, W.J., and D.C. Hague, (1953-54), "Complementarity and the Excess Burden of Taxation," *Review of Economic Studies*, 21, 21-30.

Debreu, G., (1951), "The Coefficient of Resource Utilization," *Econometrica*, 19, 273-292.

———, (1954), "A Classical Tax-Subsidy Problem," *Econometerica*, 22, 14-22.

Diamond, P.A., (1975), "A Many-Person Ramsey Tax Rule," *Journal of Public Economics*, 4, 335-342.

Diamond, P.A., and D.L. McFadden, (1974), "Some Uses of the Expenditure Function in Public Finance," *Journal of Public Economics*, 3, 3-21.

Diamond, P.A., and J.A. Mirrlees, (1971), "Optimal Taxation and Public Production II: Tax Rules," *American Economic Review*, 61, 261-278.

Dixit, A.K., (1970), "On the Optimum Structure of Commodity Taxes," *American Economic Review*, 60, 295-301.

Dupuit, J., (1844), "De la Mesure de l'Utilite' de Traveaux Publics," *Annals des Ponts et Chaussees*, 8. Translated and reprinted in *Readings in Welfare Economics*, edited by K. Arrow and T. Scitovsky. Homewood, Ill.: Irwin.

Harberger, A.C., (1964), "Principles of Efficiency: The Measurement of Waste," *American Economic Association Papers and Proceedings*, 64, 58-76.

Hotelling, H., (1938), "The General Welfare in Relation to Problems of Taxation and of Railway Utility Rates," *Econometrica*, 6, 242-69.

Jenkin, F., (1871-72), "On the Principles which Regulate the Incidence of Taxes," *Proceedings of the Royal Society of Edinburgh*, Session 1871-72.

Kay, J.A., (1980), "The Deadweight Loss from a Tax System," *Journal of Public Economics*, 13, 111-115.

Marshall, A., (1890), *Principles of Economics*, London: Macmillan.

Mohring, H, (1971), "Alternative Welfare Gain and Loss Measures," *Western Economic Journal*, 9, 349-368.

Pigou, A.C., (1928), *A study in Public Finance*. London: Macmillan.

Ramsey, F.P., (1927), "A Contribution to the Theory of Taxation," *Economic Journal*, 37, 47-61.

Samuelson, P.A., (1986), "Theory of Optimal Taxation," *Journal of Public Economics*, 137-143.

Sandmo, A. (1976), "Optimal Taxation: An Introduction to the Literature," *Journal of Public Economics*, 6, 37-54.

Stiglitz, J.E., and P. Dasgupta, (1971), "Differential Taxation, Public Goods, and Economic Efficiency," *Review of Economic Studies*, 38, 151-174.

Stutzer, M.J., (1982), "Another Note on Deadweight Loss," *Journal of Public Economics*, 18, 277-284.

FIGURES

Figure 1 -- The Tax Problem of the Planner.

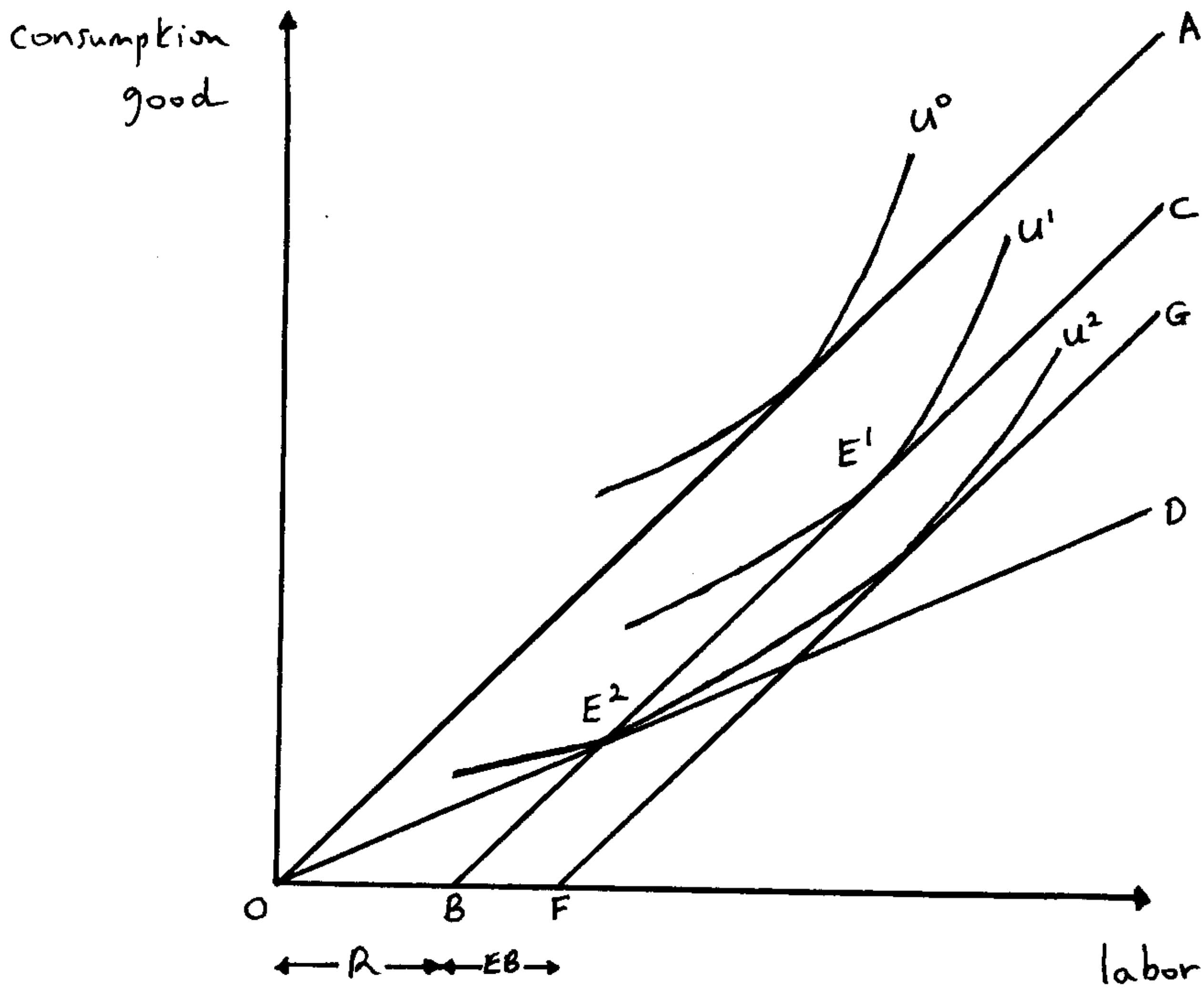

Figure 1: The Tax Problem of the Planner.

FOOTNOTES

[1] Diamond and McFadden (1974) propose an alternative measure of excess burden based on the compensating variation. However, minimization of this measure leads to an equation system which, although similar in form, is *not* identical to the Ramsey Equations: the relevant functions are evaluated at utility $U^0 = V(\mathbf{p}; M)$, whereas in the Ramsey Equations they are evaluated at $U^2 = V(\mathbf{q}; M)$.

[2] It should be noted that the total welfare cost of the tax is $R(\mathbf{q}) + EB(\mathbf{q})$. It is only correct to use the excess burden to compare the goodness of two tax codes *of equal revenue*. A representative household will achieve lower utility under tax code $\mathbf{q}_1$ than under tax code $\mathbf{q}_2$ iff $R(\mathbf{q}_2) + EB(\mathbf{q}_2) < R(\mathbf{q}_1) + EB(\mathbf{q}_1)$. Even when tax codes are set optimally for given revenue, raising tax revenue may be associated with falling excess burden. See also Footnote 6.

[3] Noting $V(\mathbf{p}; e(\mathbf{p}; U)) \equiv U$ or $V_2 \cdot (\partial e/\partial U) = 1$.

[4] In the special case of the figure with only one good being taxed, the planner has very limited choice of tax rates. In general, of course, many commodities and many different tax rate combinations yield revenue R.

[5] The formal comparison of the first-order conditions is informative because there has been some confusion in the literature. Equations which appear similar may have functions evaluated at different utility levels, see Footnote 1 and Diamond and McFadden (1974), Kay (1980) and Stutzer (1982).

[6] Note that $dV/dR < 0 \Rightarrow -1 < \mu$. Rising tax revenue will be associated with falling excess burden if $-1 < \mu < 0$. This is why the revenue constraint in the excess burden problem must hold with equality.

[7] The effect of the income change in the first stage, $\partial x_i/\partial M \cdot dM$, is exactly offset by the income effect of the price change in the second stage, $\partial x_i/\partial M \cdot x_k dq_k$.