A CONTRIBUTION TO THE THEORY OF TAXATION

THE problem I propose to tackle is this: a given revenue is to be raised by proportionate taxes on some or all uses of income, the taxes on different uses being possibly at different rates; how should these rates be adjusted in order that the decrement of utility may be a minimum? I propose to neglect altogether questions of distribution and considerations arising from the differences in the marginal utility of money to different people; and I shall deal only with a purely competitive system with no foreign trade. Further I shall suppose that, in Professor Pigou's terminology, private and social net products are always equal or have been made so by State interference not included in the taxation we are considering. I thus exclude the case discussed in Marshall's *Principles* in which a bounty on increasing-return commodities is advisable. Nevertheless we shall find that the obvious solution that there should be no differentiation is entirely erroneous.

The effect of taxation is to transfer income in the first place from individuals to the State and then, in part, back again to rentiers and pensioners. These transfers will slightly alter the demand schedules in a way depending on the incidence of the taxes and the manner of their expenditure. I neglect these alterations; [1] and I also suppose that "a given revenue" means a given money revenue, "money" being so adjusted that its marginal utility is constant.

This problem was suggested to me by Professor Pigou, to whom I am also indebted for help and encouragement in its solution.

In the first part I deal with the perfectly general utility function and establish a result which is valid for a sufficiently small revenue, and takes a peculiarly simple form if we can treat the revenue as an infinitesimal. I prove, in fact, 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.

In the second part I assume that the utility function is quadratic, which means roughly that the supply and demand

[1] The outline of a more general treatment is given in the Appendix.

curves are straight lines, but does not exclude the most general possibilities of joint supply and joint demand. With this assumption we can show that the rule given above for an infinitesimal revenue is valid for any revenue which can be raised at all.

In the third part I give certain important special cases of these general theorems; and in part four indicate certain practical applications.

Part I

(1) I suppose there to be altogether n commodities on which incomes are spent and denote the quantities of them which are produced in a unit of time by $x_1, x_2 \ldots x_n$. Some of these commodities may be identical, save for the place or manner of their production or consumption; *e.g.*, we can regard sugar used in tea as a different commodity from sugar used in coffee, and corn grown in Norfolk as different from that grown in Suffolk. In order to avoid double reckoning we suppose that these commodities are all either consumed or saved; *e.g.*, we include household coal, but not industrial coal except in so far as an increase in the stock of industrial coal is a form of saving, so that this rate of *increase* can form one of our quantities x. The quantities $x_1, x_2 \ldots$ can be measured in any convenient different units.

(2) We denote by $u = F(x_1 \ldots x_n)$ the *net* utility of producing and consuming (or saving) these quantities of commodities. This is usually regarded as the difference of two functions, one of which represents the utility of consuming, the other the disutility of producing. But so to regard it is to make an unnecessary assumption of independence between consumption and production; to assume, for instance, that the utility of a hot bath is the same whether one does or does not work in a coal mine. This assumption we do not require to make.

(3) If there is no taxation stable equilibrium will occur for values of the x's which make u a maximum. Let us call these values $\bar{x}_1, \bar{x}_2 \ldots \bar{x}_n$ or collectively the point P. Then at P we have

$$\frac{\partial u}{\partial x_r} = 0 \qquad\qquad r = 1, \ldots n.$$

$$d^2u = \Sigma\Sigma \frac{\partial^2 u}{\partial x_r \partial x_s} dx_r dx_s \text{ is a negative definite form.}$$

Suppose now taxes are levied on the different commodities at the rates $\lambda_1, \lambda_2 \ldots \lambda_n$ per unit in money whose marginal utility is unity. Then the new equilibrium is determined by

$$\frac{\partial u}{\partial x_r} = \lambda_r{}^{1} \qquad r = 1, \ldots n \quad . \quad . \quad . \quad . \quad (1)$$

In virtue of these equations we can regard the λ's as functions of the x's, which vanish at P, and satisfy identically

$$\frac{\partial \lambda_r}{\partial x_s} = \frac{\partial \lambda_s}{\partial x_r} \cdot \left(= \frac{\partial^2 u}{\partial x_r \partial x_s}\right) \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Also the revenue $R = \Sigma \lambda_r x_r$.

We shall always suppose R to be positive, but there is no *a priori* reason why some of the λ's should not be negative; they will then, of course, represent bounties.

(4) Our first problem is this: given R, how should the λ's be chosen in order that the values of the x's given by equations (1) shall make u a maximum.

I.e., u is to be a maximum subject to $\Sigma \lambda_r x_r = R \left(\text{where } \lambda_r \text{ is } \frac{\partial u}{\partial x_r}\right)$.

We must have

$$0 = du = \Sigma \lambda_r dx_r \text{ for any values of } dx_r$$

subject to

$$0 = dR = \Sigma \lambda_r dx_r + \sum_r \sum_s x_s \frac{\partial \lambda_s}{\partial x_r} dx_r,$$

and so we have

$$\frac{\lambda_1}{\sum_s x_s \frac{\partial \lambda_s}{\partial x_1}} = \frac{\lambda_2}{\sum_s x_s \frac{\partial \lambda_s}{\partial x_2}} = \qquad = \qquad = \frac{\lambda_n}{\sum_s x_s \frac{\partial \lambda_s}{\partial x_n}} \quad . \quad . \quad (3)$$

$$= \frac{R}{\Sigma\Sigma \frac{\partial \lambda_s}{\partial x} x_r x_s} = -\theta \text{ (say)}.$$

(5) These equations determine values of the x's which are critical for u, and it remains to discuss the possibility of a plurality of solutions and to determine conditions under which they give a true maximum. We shall show that if R is small enough they will have a unique solution $x_1, x_2 \ldots x_n$, which tends to $\bar{x}_1, \bar{x}_2 \ldots \bar{x}_n$ as $R \to 0$, and that this solution will make u a true maximum.

[1] *E.g.*, if $u = u_1 - u_2$ (consumers' utility − producers' disutility) $\frac{\partial u}{\partial x_r} = \frac{\partial u_1}{\partial x_r} - \frac{\partial u_2}{\partial x_r}$ = demand price of rth commodity − supply price = tax.

For, since $d^2u = \Sigma\Sigma \frac{\partial \lambda_r}{\partial x_s} dx_r dx_s$ is negative definite at P, $(-)^n \frac{\partial(\lambda_1, \lambda_2 \ldots \lambda_n)}{\partial(x_1, x_2 \ldots x_n)}$ is positive at, and therefore near, P. Hence we can express the x's as functions of the λ's. The equations (3) then become

$$\lambda_r = R\psi_r(\lambda_1, \ldots \lambda_n) \qquad r = 1, 2, \ldots n.$$

For the denominator $\Sigma\Sigma \frac{\partial \lambda_s}{\partial x_r} x_r x_s$ is a negative definite form with d^2u and so cannot vanish near P (and therefore also $\theta > 0$). The Jacobian of these last equations with regard to the λ's will tend to 1 as R tends to 0, and they will therefore have a unique solution $\lambda_1, \ldots \lambda_n$ which tends to $0, 0 \ldots 0$ as R tends to 0. Hence the equations (3) have a unique solution tending to P as $R \to 0$.

We have now to consider the conditions for a maximum which are obtained most simply by Lagrange's multipliers.

If we consider $u + KR$

we should have $$\frac{\partial u}{\partial x_r} + K\frac{\partial R}{\partial x_r} = 0$$

or $1 + K - \frac{K}{\theta} = 0$ if θ has the meaning it has in equations (3).

or $$K = \frac{+\theta}{1-\theta}.$$

Then $$d^2u = d^2\left(u + \frac{\theta}{1-\theta}R\right)$$
$$= d^2u + \frac{\theta}{1-\theta}d^2R$$

(calculated as if the variables x were independent [1]), and in a sufficiently small neighbourhood of P we shall have $\theta <$ any assigned positive constant and so $d^2u + \frac{\theta}{1-\theta}d^2R$ negative definite with d^2u. This establishes the desired result.[2]

(6) Suppose now R and the λ's can be regarded as infinitesimals; then putting

$$\lambda_r = \sum_s \frac{\partial \lambda_r}{\partial x_s} dx_s$$

equations (3) give us, using (2),

[1] See, *e.g.*, de la Vallée Poussin, *Cours d'Analyse*, 4th ed., t. 1, p. 149.

[2] Clearly also we shall get a maximum at any point for which d^2R is negative and $\theta < 1$; *i.e.*, if d^2R is everywhere negative (3) will give a maximum for all values of θ up to $\theta = 1$, which gives a maximum of R. This covers the case treated in Part II and so also any case approximating to that.

$$\frac{\sum_s \frac{\partial\lambda_r}{\partial x_s} dx_s}{\sum_s \frac{\partial\lambda_r}{\partial x_s} x_s} = \qquad = \qquad = -\theta = \frac{R}{\Sigma\Sigma x_r x_s \frac{\partial\lambda_r}{\partial x_s}} < 0,$$

and their solution is evidently given by

$$\frac{dx_1}{x_1} = \frac{dx_2}{x_2} = \qquad = \frac{dx_n}{x_n} = -\theta < 0 \quad . \quad . \quad . \quad (4)$$

i.e., the production of each commodity should be diminished in the same proportion.

(7) It is interesting to extend these results to the case of a given revenue to be raised by taxing certain commodities only. If the utility were the sum of two functions, one of the taxed and the other of the untaxed commodities, it is obvious that our conclusions would be the same as before. But in the general case the question is by no means so simple.

Let us denote the quantities of the commodities to be taxed by $x_1 \ . \ . \ . \ x_n$, and those not to be taxed by $y_1 \ . \ . \ . \ y_n$.

If $\quad \lambda_r = \frac{\partial u}{\partial x_r}$ then λ_r is the tax per unit on x_r,

and if $\quad \mu_r = \frac{\partial u}{\partial y_r}$, $\mu_r = 0$ (λ's and μ's functions of x's and y's), also as before

$$\frac{\partial\lambda_r}{\partial x_s} = \frac{\partial\lambda_s}{\partial x_r}, \ \frac{\partial\mu_r}{\partial y_s} = \frac{\partial\mu_s}{\partial y_r}, \text{ and } \frac{\partial\lambda_r}{\partial y_s} = \frac{\partial\mu_s}{\partial x_r} \quad . \quad . \quad . \quad (5)$$

and we have to maximise u subject to

$$\sum_{r=1}^{n} \lambda_r x_r = R, \ \mu_t = 0, \ t = 1, \ . \ . \ . \ m.$$

We have

$$0 = du = \sum_r \lambda_r dx_r$$

$$0 = dR = \sum_r \lambda_r dx_r + \sum_s\sum_r x_s \frac{\partial\lambda_s}{\partial x_r} dx_r + \sum_s\sum_t x_s \frac{\partial\lambda_s}{\partial y_t} dy_t$$

$$0 = d\mu_t = \sum_r \frac{\partial\mu_t}{\partial x_r} dx_r + \sum_u \frac{\partial\mu_t}{\partial y_u} dy_u, \qquad t = 1, \ . \ . \ . \ m.$$

Solving these last equations ($d\mu_t = 0$) for the dy's we obtain

$$dy_t = \sum_r \chi_{tr} dx_r \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

where $$\frac{\partial\mu_t}{\partial x_r} + \sum_{u=1}^{m} \frac{\partial\mu_t}{\partial y_u} \chi_{ur} = 0 \begin{Bmatrix} r = 1, \ . \ . \ . \ n \\ t = 1, \ . \ . \ . \ m \end{Bmatrix} \quad . \quad . \quad . \quad (7)$$

(The possibility of solution is guaranteed by the discriminants of d^2u not vanishing.)

Whence $\quad 0 = dR = \sum_r dx_r \left(\lambda_r + \sum_s x_s \frac{\partial\lambda_s}{\partial x_r} + \sum_s\sum_t x_s \frac{\partial\lambda_s}{\partial y_t} \chi_{tr}\right).$

$\therefore$ instead of equations (3) we have

$$\frac{\lambda_r}{\sum_{s=1}^n x_s\left(\frac{\partial\lambda_s}{\partial x_r}+\sum_{t=1}^m \chi_{tr}\frac{\partial\lambda_s}{\partial y_t}\right)} = \qquad = \quad . \quad . \quad (3')$$

It can be shown that these give a maximum of u with the same sort of limitations as equations (3) do.

(8) And if the λ's are infinitesimal

$$\begin{aligned}\lambda_r &= \sum_s \frac{\partial\lambda_r}{\partial x_s}dx_s + \sum_t \frac{\partial\lambda_r}{\partial y_t}dy_t \\ &= \sum_{s=1}^n dx_s\left(\frac{\partial\lambda_s}{\partial x_r}+\sum_{t=1}^m \chi_{ts}\frac{\partial\mu_t}{\partial x_r}\right) \quad \text{by (5), (6).}\end{aligned}$$

But $$\sum_t \chi_{ts}\frac{\partial\mu_t}{\partial x_r} = -\sum_t\sum_u \frac{\partial\mu_t}{\partial y_u}\chi_{ts}\chi_{ur} \quad \text{by (7)}$$

$$= \sum_t \chi_{tr}\frac{\partial\mu_t}{\partial x_s} \left(\text{by symmetry since } \frac{\partial\mu_t}{\partial y_u}=\frac{\partial\mu_u}{\partial y_t}\right).$$

So $$\lambda_r = \sum_s dx_s\left(\frac{\partial\lambda_s}{\partial x_r}+\sum_t \chi_{tr}\frac{\partial\lambda_s}{\partial y_t}\right), \text{ since } \frac{\partial\lambda_s}{\partial y_t}=\frac{\partial\mu_t}{\partial x_s},$$

and so equations (3′) are satisfied by

$$\frac{dx_1}{x_1} = \qquad = \frac{dx_n}{x_n},$$

i.e., as before the taxes should be such as to reduce in the same proportion the production of each *taxed* commodity.

(9) Further than this it is difficult to go without making some new assumption. The assumption I propose is perhaps unnecessarily restrictive, but it still allows scope for all possible first-order relations between commodities in respect of joint supply or joint demand, and it has the great merit of rendering the problem completely soluble.

I shall assume that the utility is a non-homogeneous quadratic function of the x's, or that the λ's are linear. This assumption simplifies the problem in precisely the same way as we have previously simplified it by supposing the taxes to be infinitesimal. We shall, however, make this new assumption the occasion for exhibiting a method of interpreting our formulae geometrically in a manner which makes their meaning and mutual relations considerably clearer.

It is not, of course, necessary, nor would it be sensible to suppose the utility function quadratic for all values of the variables; we need only suppose it so for a certain range of values round the point P, such that there is no question of imposing taxes large enough to move the production point (values of the

x's) outside this range. If we were concerned with independent commodities, this assumption would mean that the taxes were small enough for us to treat the supply and demand curves as straight lines.

PART II

(10) Let $u = \text{constant} + \Sigma a_r x_r + \Sigma\Sigma\beta_{rs}x_r x_s$, $(\beta_{rs} = \beta_{sr})$, and let us regard the x's as rectangular Cartesian co-ordinates of points in n-dimensional space.

The point $P(\bar{x}_1, \ldots \bar{x}_n)$ is given by $\dfrac{\partial u}{\partial x_r} = 0$,

and at that point

$$d^2u \doteq 2\Sigma\Sigma\beta_{rs}dx_r dx_s \text{ is a negative definite form.}$$

$$\therefore\ \Sigma\Sigma\beta_{rs}x_r x_s \text{ is a negative definite form,}$$

and the loci $u = \text{constant}$ are hyper-ellipsoids with the point P for centre.

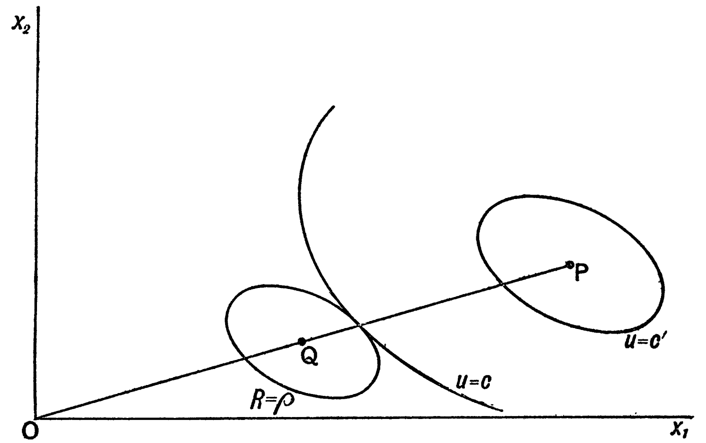
Since
$$\lambda_r = \frac{\partial u}{\partial x_r} = a_r + 2\sum_s \beta_{rs}x_s \quad \ldots \ldots \quad (8)$$

$$R = \Sigma\lambda_r x_r = \Sigma a_r x_r + 2\Sigma\Sigma\beta_{rs}x_r x_s \quad \ldots \quad (9)$$

and the loci $R = \text{constant}$ are hyper-ellipsoids with the point Q, whose co-ordinates are $\frac{1}{2}\bar{x}_1, \frac{1}{2}\bar{x}_2 \ldots \frac{1}{2}\bar{x}_n$, for centre.

(The equations for Q are those for P with their first degree terms doubled and their constant terms unaltered.)

Moreover, the hyper-ellipsoids $u = \text{constant}$, $R = \text{constant}$ are all similar and similarly situated. The figure shows these relations for the case of two commodities only.



(11) If we are to raise a revenue ρ we must depress production to some point on the hyper-ellipsoid $R = \rho$.[1]

[1] We can depress production to any point we please because the connection between the x's and λ's is one-one.

To do this so as to make u a maximum we must choose a point on this hyper-ellipsoid at which it touches an ellipsoid of the family $u =$ constant. There will be two such points which will lie on the line PQ: one between Q and P making u a maximum, the other between O and Q making u a minimum. For the point of contact of two similar and similarly situated hyper-ellipsoids must lie on the line joining their centres. Since the maximum of u is given by a point on OP we have as before that

The taxes should be such as to diminish the production of all commodities in the same proportion.

And this result is now valid not merely for an infinitesimal revenue but for any revenue which it is possible to raise at all.

The maximum revenue will be obtained by diminishing the production of each commodity to one-half of its previous amount, *i.e.*, to the point Q.

(12) If in accordance with this rule we impose taxes reducing production from $\bar{x}_1, \bar{x}_2 \ldots \bar{x}_n$ to $(1-k)\bar{x}_1, (1-k)\bar{x}_2 \ldots (1-k)\bar{x}_n$.

We get from (8) $\lambda_r = a_r + 2(1-k)\sum_s \beta_{rs}\bar{x}_r$,

but at P $\quad \lambda_r = 0$, so that $0 = a_r + 2\sum_r \beta_{rs}\bar{x}_r$;

therefore $\quad \lambda_r = ka_r \quad \ldots \ldots \ldots \quad$ (10)

i.e., the taxes should be in the fixed proportions $\lambda_1 : \lambda_2 : \; : \lambda_n : : a_1 : a_2 \quad : a_n$ independent of the revenue to be raised.

Also $R = \Sigma\lambda_r x_r = k(1-k)\Sigma a_r \bar{x}_r$,

$= 4k(1-k) \times$ the maximum revenue (got by putting $k = \frac{1}{2}$).

(13) Since k is positive it follows from (10) that the sign of λ_r is the same as that of a_r, and unless the a_r are all positive some of the λ_r will be negative, and the most expedient way of raising a revenue will be by placing bounties on some commodities and taxes on others.

The sort of case in which this might occur is that of sugar and particularly sour fruits, *e.g.* damsons. A tax on sugar might reduce the consumption of damsons more than in proportion to the reduction in the total consumption of sugar and so require to be offset by a bounty on damsons.

(14) We can now consider the more general problem: a given revenue is to be raised by means of fixed taxes $\mu_1 \ldots \mu_m$ on m commodities and by taxes to be chosen at discretion on the remainder. How should they be chosen in order that utility may be a maximum?

We have $\lambda_1 = \mu_1, \ldots \lambda_m = \mu_m$, m hyperplanes ($n-1$ folds) whose intersection is a plane $n-m$ fold which we will call S. S will cut the hyper-ellipsoids $u =$ constant, $R =$ constant in hyper-ellipsoids which are similar and similarly situated and whose centres are the points P', and Q' in which S is met by the m-folds through P and Q conjugate to S in $u = c$ or $R = c$. As before the required maximum is given by the point of contact of two of these hyper-ellipsoids in S, which must lie upon the line $P'Q'$.

Now the hyperplane $\lambda_1 = \mu_1$ or $\dfrac{\partial u}{\partial x_1} = \mu_1$ is conjugate in $u = c$ to the diameter

$$x_2 = \bar{x}_2,\ x_3 = \bar{x}_3,\ \ldots\ x_n = \bar{x}_n.$$

Hence S is conjugate to the m-fold

$$x_{m+1} = \bar{x}_{m+},\ \ldots,\ x_n = \bar{x}_n,$$

and the co-ordinates of P' satisfy these equations, since they lie on this m-fold.

Similarly the co-ordinates of Q' satisfy

$$x_{m+1} = \tfrac{1}{2}\bar{x}_{m+1},\ \ldots\ x_n = \tfrac{1}{2}\bar{x}_n.$$

And so the desired production point lying on the line $P'Q'$ satisfies

$$\frac{x_{m+1}}{\bar{x}_{m+1}} = \frac{x_{m+2}}{\bar{x}_{m+2}} = \qquad = \frac{x_n}{\bar{x}_n},$$

i.e., the whole system of taxes must be such as to reduce in the same proportion the production of the commodities taxed at discretion.

Part III

(15) I propose now to explain what our results reduce to in certain special cases. First suppose that all the commodities are independent and have their own supply and demand equations, *i.e.*, we have for the rth commodity the demand price

$$p_r = \phi_r(x_r)$$

and the supply price $\qquad q_r = f_r(x_r).$

$$\therefore \qquad \lambda_r = p_r - q_r = \phi_r(x_r) - f_r(x_r),$$

and equations (3) become, since $\dfrac{\partial \lambda_r}{\partial x_s} = 0$, $r \neq s$,

$$\frac{\lambda_1}{x_1\{\phi_1'(x_1) - f_1'(x_1)\}} = \frac{\lambda_2}{x_2\{\phi_2'(x_2) - f_2'(x_2)\}} = \ldots = -\theta.$$

These equations we can express in terms of elasticities in the following way.

Suppose the tax *ad valorem* (reckoned on the price got by the producer) on the rth commodity is μ_r, then

$$\lambda_r = \mu_r q_r = \mu_r f_r(x_r),$$

and $$\phi_r(x_r) = f_r(x_r) + \lambda_r = (1 + \mu_r) f_r(x_r).$$

$$\therefore \qquad \theta = \frac{-\lambda_r}{x_r\{\phi_r'(x_r) - f_r'(x_r)\}} = \frac{+\mu_r}{x_r\dfrac{f_r'(x_r)}{f_r(x_r)} - (1+\mu_r)x_r\dfrac{\phi_r'(x_r)}{\phi_r(x_r)}},$$

now $x_r\dfrac{f_r'(x_r)}{f_r(x_r)}$ is the reciprocal of the elasticity of supply of the commodity reckoned positive for diminishing returns, and $-x_r\dfrac{\phi_r'(x_r)}{\phi_r(x_r)}$ is the reciprocal of the elasticity of demand, reckoned positive in the normal case.

Hence if we denote by ρ_r and ϵ_r the elasticities of demand and supply,

$$\mu_r = \theta\left(\frac{1}{\epsilon_r} + \frac{1+\mu_r}{\rho_r}\right),$$

or $$\mu_r = \frac{\left(\dfrac{1}{\epsilon_r} + \dfrac{1}{\rho_r}\right)\theta}{1 - \dfrac{\theta}{\rho_r}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

(valid provided the revenue is small enough, see § 5).

For infinitesimal taxes θ is infinitesimal and

$$\frac{\mu_1}{\dfrac{1}{\epsilon_1} + \dfrac{1}{\rho_1}} = \qquad = \qquad = \frac{\mu_n}{\dfrac{1}{\epsilon_n} + \dfrac{1}{\rho_n}} \quad . \quad . \quad (12)$$

i.e., the tax *ad valorem* on each commodity should be proportional to the sum of the reciprocals of its supply and demand elasticities.

(16) It is easy to see

(1) that the same rule (12) applies if the revenue is to be collected off certain commodities only, which have supply and demand schedules independent of each other and all other commodities, even when the other commodities are not independent of one another.

(2) The rule does not justify any bounties; for in stable equilibrium, although $\dfrac{1}{\epsilon_r}$ may be negative, $\dfrac{1}{\rho_r} + \dfrac{1}{\epsilon_r}$ must be positive.

(3) If any one commodity is absolutely inelastic, either for supply or for demand, the whole of the revenue should be

collected off it. This is independently obvious, for taxing such a commodity does not diminish utility at all. If there are several such commodities the whole revenue should be collected off them, it does not matter in what proportions.

(17) Let us next take the case in which all the commodities have independent demand schedules but are complete substitutes for supply; *i.e.*, with appropriate units the demand price

$$p_r = \phi_r(x_r),$$

the supply price $\qquad q_r = f(x_1 + \qquad + x_n).$

Let us put $\qquad z = x_1 + \qquad + x_n.$

We can imagine this case as that of a country in which all commodities are produced at constant returns by the application of one kind of labour only, the increase in the supply price arising solely from the increasing marginal disutility of labour, and the commodities satisfying independent needs. Then z will represent the amount of labour.

Equations (3) give us

$$-\theta = \frac{\lambda_r}{x_r\phi_r'(x_r) - zf'(z)}.$$

Or if μ_r represents the tax *ad valorem* and ρ_r the elasticity of demand for the rth commodity and ϵ the elasticity of supply of things in general, we get, by a similar process to that of § 15,

$$\mu_r = \frac{\left(\frac{1}{\rho_r} + \frac{1}{\epsilon}\right)\theta}{1 - \frac{\theta}{\rho_r}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

If the taxes are infinitesimal we have

$$\frac{\mu_r}{\frac{1}{\rho_r} + \frac{1}{\epsilon}} = \qquad = \qquad = \theta \quad . \quad . \quad . \quad . \quad (14)$$

In this case we see that if the supply of labour is fixed (absolutely inelastic, $\epsilon \to 0$) the taxes should be at the same *ad valorem* rate on all commodities.

(19) If some commodities only are to be taxed it is easier to work from the result proved in § 8 for an infinitesimal revenue, that the production of the commodities taxed should be diminished in the same ratio.

Suppose, then, $x_1, \ldots x_m$ are to be taxed, $x_{m+1} \ldots x_n$ untaxed.

Let $\qquad dx_1 = -kx_1, \ldots, dx_m = -kx_m.$

Let
$$z' = x_1 + x_2 + \ldots + x_m$$
$$z'' = x_{m+1} + \ldots + x_n.$$
$$\lambda_1 = \phi_1(x + dx_1) - zf(z + dz)$$
$$= \phi_1'(x_1)dx_1 - f'(z)dz.$$

$$\therefore \qquad \mu_1 = \frac{k}{\rho_1} - \frac{dz}{z\epsilon} \ldots, \mu_m = \frac{k}{\rho_m} - \frac{dz}{\epsilon z},$$

now
$$dz = dz' + dz'' = -kz' + dz'',$$

also
$$0 = -\frac{dx_{m+1}}{\rho_{m+1}x_{m+1}} - \frac{dz}{\epsilon z}.$$

$$\therefore \quad \frac{dx_{m+1}}{\rho_{m+1}x_{m+1}} = \frac{dx_{m+2}}{\rho_{m+2}x_{m+2}} = \frac{dx_n}{\rho_n x_n} = -\frac{dz}{\epsilon z} = \frac{dz''}{\sum\limits_{m+1}^{n}\rho_r x_r} = \frac{kz'}{\epsilon z + \sum\limits_{m+1}^{n}\rho_r x_r}.$$

$$\therefore \qquad \mu_1 = k\left(\frac{1}{\rho_1} + \frac{\sum\limits_{1}^{m} x_r}{\epsilon \sum\limits_{1}^{n} x_r + \sum\limits_{m+1}^{n}\rho_r x_r}\right), \text{ etc.}$$

As before we see that of two commodities that should be taxed most which has the least elasticity of demand, but that if the supply of labour is absolutely inelastic all the commodities should be taxed equally.

Part IV

(20) We come now to applications of our theory; these cannot be made at all exactly without data which I, at any rate, do not possess. The simplest result is the one which we have proved in the general case for an infinitesimal revenue (§ 8); this means that it is approximately true for small revenues, and that the approximation approaches perfection as the revenue approaches zero. It is thus logically similar to the theorem that the period of oscillation of a pendulum is independent of the amplitude. We have also extended the result to any revenue which does not take the production point outside a region in which the utility may be taken to be quadratic, *i.e.*, the supply and demand schedules linear.

The sort of cases in which our theory may be useful are the following:

(21) (*a*) If a commodity is produced by several different methods or in several different places between which there is no mobility of resources, it is shown that it will be advantageous to discriminate between them and tax most the source of supply which is least elastic. For this will be necessary if we are to maintain unchanged the proportion of production between the two sources (result analogous to § 19 with supply and demand interchanged).

(*b*) If several commodities which are independent for demand require precisely the same resources for their production, that should be taxed most for which the elasticity of demand is least (§ 19).

(*c*) In taxing commodities which are rivals for demand, like wine, beer and spirits, or complementary like tea and sugar, the rule to be observed is that the taxes should be such as to leave unaltered the proportions in which they are consumed (§ 14). Whether the present taxes satisfy this criterion I do not know.

(*d*) In the case of the motor taxes we must separate off so much of the taxation as is offset by damage to the roads. This part should be so far as possible equal to the damage done. The remainder is a genuine tax and should be distributed according to our theory; that is to say, it should be placed partly on petrol and partly on motor-cars, so as to preserve unchanged the proportion between their consumption, and should be distributed between Fords and Morrises, so as to reduce their output in the same ratio. The present system fails in both these respects.

(22) (*e*) Another possible application of our theory is to the question of exempting savings from income-tax.[1] We may consider two uses of income only, saving and spending, and supposing them independent we may use the result (13) in § 17. We must suppose the taxes imposed only for a very short time [2] and that they raise no expectation of similar taxation in the future; since otherwise we require a mathematical theory considerably more difficult than anything in this paper.

On these assumptions, since the amount of saving in the very short time cannot be sufficient to alter appreciably the marginal utility of capital, the elasticity of demand for saving will be infinite, and we have

$$\mu_1 \text{ (tax on spending)} = \frac{\left(\frac{1}{\rho_1} + \frac{1}{\epsilon}\right)\theta}{1 - \frac{\theta}{\rho_1}},$$

$$\mu_2 \text{ (tax on saving)} = \frac{1}{\epsilon}\theta,$$

and we see that income-tax should be partially but not wholly remitted on savings. The case for remission would, however,

[1] No account is taken of graduation in this.

[2] Strictly, we consider the limit as this time tends to zero.

be strengthened enormously by taking into account the expectation of taxation in the future.

(23) It should be emphasized in conclusion that the results about " infinitesimal " taxes can only claim to be approximately true for small taxes, how small depending on data which are not obtainable. It is perfectly possible that a tax of 500% on whisky could for the present purpose be regarded as small. The unknown factors are the curvatures of the supply and demand curves; if these are zero our results will be true for any revenue whatever, but the greater the curvatures the narrower the range of "small" taxes.

On the other hand, the more complicated results contained in equations (3), (3′), (11), (13) may well be valid under still wider conditions. But these are, in the general case, too complicated to be worth setting down in the absence of practical data to compare with them.

APPENDIX

We can also say something about the more general problem in which the State wishes to raise a revenue for two purposes; first, as before, a fixed money revenue, R_1, which is transferred to rentiers or otherwise without effect on the demand schedules; and secondly, an additional revenue, R_2, sufficient to purchase fixed quantities, $a_1, a_2, \ldots a_n$ of each commodity.

Let us denote by p_r, q_r, as before, the demand and supply prices of the rth commodity, and the tax on it by λ_r. Then if x_r is the amount of the rth commodity consumed by the public (or by the State out of R_1), $x_r + a_r$ is the amount produced, and we have

$$\frac{\partial u}{\partial x_r} = \lambda_r = p_r(x_1, x_2, \ldots, x_n) - q_r(x_1 + a_1, x_2 + a_2, \ldots, x_n + a_n),$$

$$R_1 + R_2 = \Sigma\lambda_r x_r,\ R_2 = \Sigma a_r q_r;$$

so that u is to be a maximum subject to

$$\Sigma\lambda_r x_r - \Sigma a_r q_r = R_1 = \text{constant},$$

whence

$$\frac{\lambda_r}{\Sigma_s x_s \frac{\partial\lambda_s}{\partial x_r} - \Sigma_s a_s \frac{\partial q_s}{\partial x_r}} = \quad = \quad = -\theta$$

or

$$\frac{\lambda_r}{\Sigma_s (a_s + x_s)\frac{\partial q_s}{\partial x_r} - \Sigma_s x_s \frac{\partial p_s}{\partial x_r}} = \theta, \text{ which replace equations (3).}$$

Although these equations do not give such simple results as we previously obtained for an infinitesimal revenue or a quadratic utility function, in the cases considered in § 15 and § 17 they lead us again to the equations (11) and (13).

For, taking the case of § 15, in which the commodities are independent both for demand and supply, and, as before, denoting by μ_r the rate of tax *ad valorem* on the rth commodity and by ρ_r, ϵ_r its elasticities of demand and supply for the amounts x_r, $x_r + a_r$ respectively consumed and produced by the public, we have

$$\frac{\mu_r}{\dfrac{x_r + a_r}{q_r}\dfrac{dq_r}{d(x_r + a_r)} - \dfrac{x_r}{q_r}\dfrac{dp_r}{dx_r}} = \theta$$

or

$$\frac{\mu_r}{\dfrac{1}{\epsilon_r} + \dfrac{1 + \mu_r}{\rho_r}} = \theta$$

whence $\mu_r = \dfrac{\left(\dfrac{1}{\epsilon_r} + \dfrac{1}{\rho_r}\right)\theta}{1 - \dfrac{\theta}{\rho_r}}$, which is equation (11) again. And we can similarly derive equation (13) from the assumption of independence for demand and equivalence for supply.

F. P. Ramsey