An Important Theorem on Income Tax [1, 2]

JOHN BROOME

Birkbeck College

It is not generally known that the best marginal rate of income tax is 58·6 per cent. Were this fact better publicized, economists would be spared a lot of trouble.

I use the following assumptions. In case some of them do not look quite authentic, I have supplied a pedigree for each.

(1) Each person has an " income-earning ability " n. If he works for a fraction y of the day, his gross income is ny. This is the standard model in the theory of optimal income tax. It originates in Mirrlees [4].

(2) The government considers only linear tax functions, so that after-tax income is $x = any+b$. a and b are chosen by the government. The marginal tax rate is $(1-a)$, and b is the guaranteed minimum income. The assumption of linearity is made by Atkinson [1], by Feldstein [3] and by Sheshinski [6]. (Atkinson [2] also uses it, and compares the results with the more general case.)

(3) The government has no expenditure of its own. It taxes only to redistribute. The same thing is assumed by Atkinson [1] and [2] and by Sheshinski [6]. See the note under " Qualifications " below.

(4) The government's aim is to maximize the utility of the person with the least utility. Provided there are some people with an ability sufficiently small that they do not find it worthwhile to work, this comes to the same thing as maximizing b. The reason is that these are bound to be the people with the least utility, because anyone, whatever his ability, has the option of achieving the same utility as them simply by stopping working (see Phelps [5] for this argument in more detail). This " maximin " objective comes from Atkinson [2] and from Phelps. Its ancestry is well known to be of the highest breeding.

(5) Everyone has the same utility function $u = \log x + \log(1-y)$. This is the function adopted by Mirrlees in his concrete examples. See the note under " Qualifications " below.

(6) Only a few people have an ability less than 0·414 of the mean ability. Mirrlees uses a lognormal distribution with parameter $\sigma = 0{\cdot}39$ to " represent a realistic distribution of skills within the population ". Such a distribution has less than 2 per cent of the population below this critical level.

A person of ability n decides how much to work by maximizing his utility

$$\begin{aligned} u &= \log x + \log(1-y) \\ &= \log(any+b) + \log(1-y) \end{aligned}$$

with respect to y, subject to $y \geqq 0$

$$\frac{du}{dy} = \frac{an}{any+b} - \frac{1}{1-y}.$$

[1] *First version received March* 1974; *final version accepted June* 1974 (*Eds.*).

[2] When announcing a discovery such as this, it is comforting to have the encouragement of one's colleagues. Hugh Davies, Ben Fine, John Muellbauer, Richard Portes and John Weeks have all given me help, and some of them have given me encouragement.

Setting $du/dy = 0$, we find

$$y = \frac{1}{2} - \frac{b}{2an}.$$

$\frac{1}{2} - \frac{b}{2an}$ is non-negative provided $n \geqq b/a$, so in that case $y = \frac{1}{2} - \frac{b}{2an}$. If $n < b/a$, $y = 0$. For the moment we shall assume that everyone's ability is at least b/a.

The gross income of this person with ability n is

$$ny = \frac{n}{2} - \frac{b}{2a},$$

and his net income

$$any + b = \frac{an}{2} + \frac{b}{2}.$$

Let there be k people, with abilities $n_1, n_2, \ldots, n_k$. The average ability is $\Sigma n_i/k$, which we shall call m. The sum of the people's gross income must be equal to the sum of their net incomes. This is the government's constraint. That is:

$$\frac{1}{2}\Sigma n_i - \frac{kb}{2a} = \frac{a}{2}\Sigma n_i + \frac{kb}{2}$$

$$\Rightarrow \quad b = \frac{a(1-a)}{1+a} m.$$

The government wishes to maximize b with respect to a.

$$\frac{db}{da} = \frac{(-a^2-2a+1)}{(1+a)^2} m.$$

For the maximum $a^2 + 2a - 1 = 0$, whose positive root is

$$a = -1 + \sqrt{2}.$$

The marginal tax rate is then $(1-a) = 2 - \sqrt{2} = 0{\cdot}586$.

The minimum net income, b, may be calculated from the government's constraint.

$$b = \frac{a(1-a)}{1+a} m = (3 - 2\sqrt{2})m = 0{\cdot}172m.$$

This may be compared with the net income of the average man as follows. The average man has ability m, so his work time is

$$y = \frac{1}{2} - \frac{b}{2am} = \frac{1}{2} - \frac{(3-2\sqrt{2})}{2(-1+\sqrt{2})}\frac{m}{m} = (1 - \tfrac{1}{2}\sqrt{2}) = 0{\cdot}293.$$

0·293 of a 24-hour day is 7 hours and about $1\frac{1}{2}$ minutes, a reasonable day for the average man. His before tax income is $(1 - \frac{1}{2}\sqrt{2})m$, and it may easily be calculated that his after-tax income is the same.

He gets $(1 - \frac{1}{2}\sqrt{2})/(3 - 2\sqrt{2}) = (1 + \frac{1}{2}\sqrt{2}) = 1.71$ times the poorest man.

Lastly, we need to calculate b/a, since people with abilities smaller than this will not work.

$$\frac{b}{a} = \frac{(3-2\sqrt{2})}{(-1+\sqrt{2})} m = (-1+\sqrt{2})m = 0{\cdot}414m.$$

If there were a substantial number of people with abilities below this level, my calculations would not be accurate. The constraint on the government would be different. On the other hand, at least one person must have an ability at or below this level, or else maximizing b would not be same as maximizing the utility of the person with least utility.

Qualifications

If the utility function is more general, $u = \alpha \log x + \log(1-y)$ (Mirrlees [4]), the marginal tax rate may be found in exactly the same way to be

$$\frac{(1+\alpha)-\sqrt{(1+\alpha)}}{\alpha}.$$

When $\alpha = 2$ this is 0·634, and when $\alpha = 1/2$ it is 0·551 (Atkinson [2]). A government which uses the rate of 0·586 will therefore not be far wrong.

If the government's expenditure is not zero but g (Feldstein [3]), then the work is more complicated. Take $\alpha = 1$ again. The government's constraint is now:

$$\frac{1}{2}\Sigma n_i - \frac{kb}{2a} = \frac{a}{2}\Sigma n_i + \frac{kb}{2} + g.$$

By the same process, we may find that

$$a = -1+\sqrt{2}\sqrt{\left(1-\frac{g}{km}\right)},$$

which makes the marginal tax rate $2-\sqrt{2}\sqrt{1-(g/(km))}$. Write $f = \sqrt{1-(g/(km))}$, so that $a = -1+f\sqrt{2}$ and $g/(km) = 1-f^2$. To assess the change, we need to express government expenditure as a fraction of national income. National income is

$$\frac{1}{2}\Sigma n_i - \frac{kb}{2a} = \frac{1}{2}km - \frac{kb}{2a}.$$

From the constraint:

$$\frac{kb}{2a} = \frac{km(1-a)-2g}{2(1+a)},$$

which makes national income

$$\frac{akm+g}{(1+a)}.$$

Suppose the government spends a fraction G of national income. Then

$$G = \frac{g(1+a)}{akm+g} = \frac{(g/km)(1+a)}{a+(g/km)} = \frac{(1-f^2)\sqrt{2}}{\sqrt{2}-f}.$$

So

$$f^2\sqrt{2} - fG - (1-G)\sqrt{2} = 0.$$

The positive root is

$$f = \frac{G+\sqrt{(G^2-8G+8)}}{2\sqrt{2}}.$$

Hence the marginal tax rate is

$$2 - \frac{G+\sqrt{(G^2-8G+8)}}{2}.$$

If $G = 1/3$ this is exactly $2/3 = 0{\cdot}667$. A large government expenditure, therefore, has a significant but not large effect on the best tax rate. (A tax rate of 2/3 is more convenient to administer, and probably more attractive to politicians,[1] than $2-\sqrt{2}$. These benefits

[1] I am indebted to John Weeks for pointing this out to me.

would come to a government which spent 1/3 of national income. However, I do not myself think that governments should decide their expenditure on such a basis.)

It has not been my intention to make fun of serious work in the field of optimal income tax. Indeed, for those who would like them, I have some serious conclusions to offer.

My model is strictly comparable with Mirrlees' numerical results [4]. The utility function is the same. The differences are: (*a*) Mirrlees' results assume a lognormal ability distribution, whilst I have only restricted the distribution to having a particular lower cut-off point, with just a few people below it. (*b*) I have limited the government to a linear tax function; Mirrlees' functions *turn out* to be roughly linear. (*c*) Most important, and the whole point of the comparison, is that Mirrlees works with a utilitarian objective, and mine is "maximin". His marginal tax rates are less than about 30 per cent. The more egalitarian objective puts them up to around 60 per cent.

Atkinson [2] studies in detail the effect of applying the maximin objective to Mirrlees' model. Readers of Atkinson will recognise 58·6 per cent as a special case of his conclusions. It is, however, the central case since the best tax rate varies from it only a little in response to reasonable changes in the assumptions, especially to reasonable changes in assumptions about the skill distribution. Atkinson's conclusions in the case of a Pareto distribution seem at first to contradict this, since in that case he finds that the best tax rate depends heavily on the parameter of the distribution. But this results from a misleading feature of the Pareto distribution. My method depends on there being a few, but not many, people with ability less than 0·414 of the mean ability. This is entirely reasonable (cf. Atkinson's conclusions for a lognormal distribution). A Pareto distribution is more or less bound to violate that condition, since it has many people with skills near its lower cut-off point, but absolutely none with skills below it. No one pretends that this is a sensible feature of that distribution, but it is the feature that determines Atkinson's results. A Pareto formula is useful for representing the upper part of a skill distribution, but the best linear income tax is determined by the lower end of the distribution, where its shape makes no sense. My method shows that the shape of the upper part of the distribution has no effect.

REFERENCES

[1] Atkinson, A. B. "How Progressive should Income Tax be?" in Michael Parkin (ed.), *Essays in Modern Economics* (Longman, 1973), 90-109.

[2] Atkinson, A. B. "'Maxi-min' and Optimal Income Taxation", *Review of Economic Studies* (forthcoming).

[3] Feldstein, M. "On the Optimal Progressivity of the Income Tax", *Journal of Public Economics*, **2**, 4 (November 1973), 357-376.

[4] Mirrlees, J. A. "An Exploration in the Theory of Optimum Income Taxation", *Review of Economic Studies*, **38**, 114 (April 1971), 175-208.

[5] Phelps, E. S., "Taxation of Wage Income for Economic Justice", *Quarterly Journal of Economics*, **87**, 3 (August 1973), 331-345.

[6] Sheshinski, E. "The Optimal Linear Income Tax", *Review of Economic Studies*, **39**, 119 (July 1972), 297-302.