Journal of Public Economics 27 (1985) 255–260. North-Holland

ACHIEVING THE FIRST BEST IN SMALL ECONOMIES

Franklin ALLEN*

Finance Department, University of Pennsylvania, Philadelphia, PA 19104, USA

Received December 1983, revised version received January 1985

It is shown in the context of a simple example with a small two-class population, that the first best utilitarian allocation can sometimes be achieved as a dominant strategy equilibrium, even though leisure is a normal good. This occurs because a person's optimal lump-sum tax depends both on his own announced ability and the distribution of announced abilities. It is also demonstrated that, in a wider class of cases, the first best can be achieved as a Bayes–Nash equilibrium.

1. Introduction

Mirrlees (1971, p. 205; 1974, p. 251) and Dasgupta and Hammond (1980, p. 150) have argued that if abilities cannot be directly observed and there is a continuum of individuals, then the lump-sum taxation necessary to achieve a first best utilitarian social optimum is not feasible, except in some unusual cases where leisure is not a normal good. The reason is that the first best allocation usually involves high ability people having less utility than those with low ability. To see this, consider the case where utility is separable in consumption and leisure. Optimality requires marginal utilities of consumption and hence levels of consumption to be equated across the population. Labour supplies must be such that the marginal disutility of labour for each unit of consumption produced is the same for everybody; this implies the high ability have to supply more labour than the low ability. They are therefore worse off and if offered the first best allocations will pretend to be low ability.

Hammond (1979, p. 271, and especially fn. 2, p. 281) has also argued, in the context of a model with a continuum of individuals but with two or more discrete classes of ability, that although the first best can sometimes be achieved, this will not be possible with a concave welfare integral if leisure is a normal good: in particular it will again not be possible with a utilitarian social welfare function if utility is separable in consumption and leisure [see

*I am grateful to P. Hammond, A. Postlewaite and two anonymous referees for helpful comments.

also Stiglitz (1982, p. 220)]. In this note it is shown how the first best can sometimes be achieved even though social welfare is utilitarian and utility is separable. Rather than assuming a continuum of individuals as Hammond does, there are taken to be a finite number of individuals as well as a discrete number of ability classes. An example is given for a small two-class economy where the first best allocation is achieved as a dominant strategy equilibrium. This is done by making the lump-sum tax imposed on each individual depend on both the announced distribution of abilities as well as the individual's announcement.

In the example given the first best can be achieved both in the standard case where only income can be observed and in the case considered by Dasgupta and Hammond (1980), where both income and labour supply are observable but people can choose their productivity to be less than or equal to their ability. In the incentive compatibility literature, it is usually argued that the ability to manipulate the terms of trade (here, the real efficiency wage) means the first best cannot be obtained in finite economies as a dominant strategy equilibrium [see, for example, Hurwicz (1972)]. The example shows that in economies where characteristics are known to belong to discrete classes and the population is finite it may be possible to obtain the first best in this way.

Finally it is shown that in a wider class of cases the first best can be achieved as a Bayes–Nash equilibrium using the same method.

2. The example

Consider an economy in which there are two people and two possible levels of ability: high (H) and low (L). There are then three possible distributions of ability:

State	1	2	3
Distribution of ability	2H; OL	1H; 1L	OH; 2L

Each person knows his own ability but this cannot be directly observed by anybody else. People with ability A_i have an income Y_i, which can be observed by the government, given by

$$Y_i = A_i l, \quad \text{for} \quad i = \mathrm{H}, \mathrm{L}, \tag{1}$$

where l is their labour supply, which is such that

$$0 \leqq l \leqq 1. \tag{2}$$

Initially it is assumed l cannot be observed by the government.

Utility is a concave, strictly increasing function of consumption C and leisure $1-l$:

$$U=U(C,1-l). \tag{3}$$

Leisure is taken to be a normal good.

In state 1 the utilitarian optimal allocation $(C_{\mathrm{H}}^1, l_{\mathrm{H}}^1)$ is given by the solution to

$$\max U(C_{\mathrm{H}}, 1-l_{\mathrm{H}}) \tag{4}$$

subject to

$$C_{\mathrm{H}}=A_{\mathrm{H}} l_{\mathrm{H}} \tag{5}$$

and (2).

Similarly for state 3.

The optimal allocation in state 2, $(C_{\mathrm{H}}^2, l_{\mathrm{H}}^2)$ for H's and $(C_{\mathrm{L}}^2, l_{\mathrm{L}}^2)$ for L's, is found from

$$\max U(C_{\mathrm{H}}, 1-l_{\mathrm{H}})+U(C_{\mathrm{L}}, 1-l_{\mathrm{L}}) \tag{6}$$

subject to

$$C_{\mathrm{H}}+C_{\mathrm{L}}=A_{\mathrm{H}} l_{\mathrm{H}}+A_{\mathrm{L}} l_{\mathrm{L}} \tag{7}$$

and (2).

To implement the first best scheme, the government asks individuals to state their ability and then, depending on the announced configuration of abilities applies the appropriate first best lump-sum taxes. Thus, if both announce high, the optimal allocation for both is $(C_{\mathrm{H}}^1, l_{\mathrm{H}}^1)$ and no taxation is necessary. Similarly if they both announce low. If one announces high and the other low, the optimal allocations are $(C_{\mathrm{H}}^2, l_{\mathrm{H}}^2)$ and $(C_{\mathrm{L}}^2, l_{\mathrm{L}}^2)$. These are implemented by imposing a lump-sum tax of $A_{\mathrm{H}} l_{\mathrm{H}}^2 - C_{\mathrm{H}}^2$ on the person who announced high and $A_{\mathrm{L}} l_{\mathrm{L}}^2 - C_{\mathrm{L}}^2$ on the person who announced low. In addition, in all these cases the government must check to ensure the pre-tax incomes received correspond to those that are optimal for the announced ability level.

Given leisure is a normal good, it can be shown that a low ability person will never want to pretend to be a high ability.

Consider a high ability person's decision on what level to claim. Announcing high will be worthwhile if the other person announces high provided

$$U(C_{\mathrm{H}}^{1}, 1-l_{\mathrm{H}}^{1}) > U\left(C_{\mathrm{L}}^{2}, 1-\frac{A_{\mathrm{L}} l_{\mathrm{L}}^{2}}{A_{\mathrm{H}}}\right), \tag{8}$$

and if the other person announces low provided

$$U(C_{\mathrm{H}}^{2}, 1-l_{\mathrm{H}}^{2}) > U\left(C_{\mathrm{L}}^{3}, 1-\frac{A_{\mathrm{L}} l_{\mathrm{L}}^{3}}{A_{\mathrm{H}}}\right). \tag{9}$$

To see that (8) and (9) can both be simultaneously satisfied so that truth-telling is a dominant strategy, consider the case where

$$U = \alpha \log C + \log(1-l), \tag{10}$$

$$\alpha = 10; \qquad A_{\mathrm{H}} = 2; \qquad A_{\mathrm{L}} = 0. \tag{11}$$

The optimal utilitarian levels of consumption, labour supply and utility for the two types are then as shown in table 1.

Table 1

State	C_{H}	l_{H}	$U(C_{\mathrm{H}}, 1-l_{\mathrm{H}})$	C_{L}	l_{L}	$U(C_{\mathrm{L}}, 1)$
1	1.818	0.909	3.581	—	—	—
2	0.952	0.952	−3.532	0.952	0	−0.488
3	—	—	—	0	0	$-\infty$

It can be seen that (8) is satisfied since $U(C_{\mathrm{H}}^{1}, 1-l_{\mathrm{H}}^{1}) = 3.581$ whereas $U(C_{\mathrm{L}}^{2}, 1-A_{\mathrm{L}} l_{\mathrm{L}}^{2}/A_{\mathrm{H}}) = U(C_{\mathrm{L}}^{2}, 1) = -0.488$. Similarly (9) is also satisfied since $U(C_{\mathrm{H}}^{2}, 1-l_{\mathrm{H}}^{2}) = -3.532$ and $U(C_{\mathrm{L}}^{3}, 1-A_{\mathrm{L}} l_{\mathrm{L}}^{3}/A_{\mathrm{H}}) = U(C_{\mathrm{L}}^{3}, 1) = -\infty$. Thus truth-telling is a dominant strategy for high ability people. In state 2 the high ability person is worse off than the low ability person because he works harder. This is the basis for the argument in Hammond (1979) and Stiglitz (1982) that the first best cannot be achieved: the high ability envies the low ability's allocation. However, when making his decision on what level of ability to announce the scheme above is such that this is not the relevant comparison for a high ability person. Each person's announcement changes the announced distribution of abilities, so that (8) and (9) are the relevant comparisons and these can both be satisfied.

It can be seen that the example is such that the first best would still be attained if the social welfare function were more concave so that low utilities were weighted more heavily than high utilities. The first best will also be attained no matter what the actual distribution of abilities is, whether it be that corresponding to state 1, 2 or 3.

So far it has been assumed that the government can observe only income Y. If in addition it can observe labor supply l but people can choose their productivity to be less than or equal to their ability, as in Dasgupta and Hammond (1980), then (8) and (9) are changed to:

$$U(C_{\mathrm{H}}^1, 1-l_{\mathrm{H}}^1) > U(C_{\mathrm{L}}^2, 1-l_{\mathrm{L}}^2), \tag{12}$$

$$U(C_{\mathrm{H}}^2, 1-l_{\mathrm{H}}^2) > U(C_{\mathrm{L}}^3, 1-l_{\mathrm{L}}^3). \tag{13}$$

Since in the example $l_{\mathrm{L}}^2 = l_{\mathrm{L}}^3 = 0$ it follows the first best can still be achieved in this case. Thus even though individuals can manipulate the terms of trade, which here is the real efficiency wage, this does not make the first best incentive incompatible as it usually does.

Truth-telling is not always a dominant strategy. For example if

$$\alpha = 1; \qquad A_{\mathrm{H}} = 2; \qquad A_{\mathrm{L}} = 0, \tag{14}$$

it can be shown that (8) is not satisfied. However, it is still the case that truth-telling can be achieved as a Bayes–Nash equilibrium, provided any high ability person's prior that state 2 is the actual state is strictly positive. To see this, suppose the high ability person's prior is

State	1	2	3
H's prior	π	$1-\pi$	0

(A high ability person knows state 3 is impossible since he himself is high ability.) No assumption is made as to what the actual distribution of abilities is and initially it is possible nobody, including the government, knows the actual distribution of abilities. Thus before any announcements are made π can take any value. In the case where only Y can be observed, given that the other person is telling the truth, a high ability person will correctly reveal his ability provided

$$\pi U(C_{\mathrm{H}}^1, 1-l_{\mathrm{H}}^1) + (1-\pi) U(C_{\mathrm{H}}^2, 1-l_{\mathrm{H}}^2),$$

$$> \pi U\left(C_{\mathrm{L}}^2, 1-\frac{A_{\mathrm{L}} l_{\mathrm{L}}^2}{A_{\mathrm{H}}}\right) + (1-\pi) U\left(C_{\mathrm{L}}^3, 1-\frac{A_{\mathrm{L}} l_{\mathrm{L}}^3}{A_{\mathrm{H}}}\right). \tag{15}$$

If the utility function is as in (10) and $A_{\mathrm{L}} = 0$, then provided only that $\pi < 1$, (15) will always be satisfied, since the left-hand side is finite and $U(C_{\mathrm{L}}^3,$

$1-A_{\mathrm{L}}l_{\mathrm{L}}^{3}/A_{\mathrm{H}})=U(C_{\mathrm{L}}^{3},1)=-\infty$. A similar result holds if l can be observed in addition to Y.

3. Concluding remarks

It has been shown in a simple example of a small two-class economy that the first best utilitarian solution can be achieved as a dominant strategy equilibrium. It has also been demonstrated that in a wider class of cases the first best optimum can be achieved as a Bayes–Nash equilibrium.

It should perhaps be pointed out that the example did not rely on the government using information on the distribution of abilities. It is therefore different from examples where everybody is punished if the announced distribution does not correspond to the distribution the government knows to be correct. Here, no assumption is made about the government's information on the distribution of abilities; all that is required is that the government knows people can be high or low ability and the form of their utility functions.

The model used was a very simple one, the crucial features being a finite population and the discrete number of classes. With greater numbers of people and more levels of ability, it should still be possible to construct similar examples. However it seems likely that, the greater the complexities of the model, the more special and unusual such examples will be.

References

Dasgupta, P. and P. Hammond, 1980, Fully progressive taxation, Journal of Public Economics 13, 141–154.

Hammond, P., 1979, Straightforward individual incentive compatibility in large economies, Review of Economic Studies 44, 263–282.

Hurwicz, L., 1972, On informationally decentralized systems, in: C.B. McGuire and R. Radner, eds., Decision and Organization (North-Holland, Amsterdam).

Mirrlees, J.A., 1971, An exploration in the theory of optimum income taxation, Review of Economic Studies 38, 175–208.

Mirrlees, J.A., 1974, Notes on welfare economics, information and uncertainty, in: M. Balch, D. McFadden and S. Wu, eds., Essays on economic behaviour under uncertainty (North-Holland).

Stiglitz, J.E., 1982, Self-selection and Pareto efficient taxation, Journal of Public Economics 17, 213–240.