Content Articles in Economics

In this section, the *Journal of Economic Education* publishes articles concerned with substantive issues, new ideas, and research findings in economics that may influence or can be incorporated into the teaching of economics.

DAVID COLANDER, Section Editor

Self-Selection, Optimal Income Taxation, and Redistribution

J. Atsu Amegashie

Abstract: The author makes a pedagogical contribution to optimal income taxation. Using a very simple model adapted from George A. Akerlof (1978), he demonstrates a key result in the approach to public economics and welfare economics pioneered by Nobel laureate James Mirrlees. He shows how incomplete information, in addition to the need to preserve incentives, acts as a limit to a government's redistributive power. The model and technical analysis allow easy handling of three self-selection constraints in a manner that is accessible to students with knowledge of only intermediate microeconomics and elementary algebra. The diagrammatic exposition allows him to present interesting and insightful results.

Keywords: efficient, income taxation, incomplete information, redistribution, self-selection
JEL codes: D61, H21, H71

The idea that incomplete information is a limit to a government's ability to redistribute income is the cornerstone of what one might call the new public economics pioneered by Mirrlees (1971). Boadway (1997, 1998) thoroughly reviews and discusses this literature and the implications of the Mirrlees approach.[1] Boadway (1997) observed that this "approach to the theory of economic policy, which is grounded on the importance of imperfect information as a constraint on public policy, has changed the very nature of public economic research and its implications for economic policy and the role of government" (753–54).

***J. Atsu Amegashie** is an associate professor in the Department of Economics at the University of Guelph (e-mail: jamegash@uoguelph.ca).* *The author thanks two anonymous referees and an associate editor for extremely helpful comments.*

The technical and formal analyses of the results in this literature are typically accessible only to graduate students and occasionally to advanced undergraduates. The purpose of this article is to demonstrate, at a very elementary level, the result that incomplete information, in addition to the need to preserve incentives, is a limit to a government's ability to redistribute income. The model and technical analysis allow easy handling of three self-selection constraints in a manner that is accessible to students with knowledge of only intermediate microeconomics and elementary algebra. I also demonstrate other interesting results. In the following sections, I present a simple model of taxation and redistribution and then extend the model.

THE BASIC MODEL

In this section, I consider a very simple two-job–two-type model as a stepping stone to a three-job–three-type model. There are two types of jobs in a society: a low-skill job called job L and a high-skill job called H. People are free to choose either L or H but not both. There are high-ability people who can work in L or H. Low-ability people can only work in L. If a low-ability person works in a job L, her income is Y_{LL}. If a high-ability person works in job L, her income is Y_{HL}. If a high-ability person works in job H, her income is Y_{HH}. There are N_H high-ability people and N_L low-ability people. Assume that if a high-ability person is indifferent between working in L or H, then she will choose to work in H. Also, we assume that $Y_{HH} > Y_{HL} > Y_{LL} > 0$.

The government knows all the information indicated in the previous paragraph but does not know the identity of high-ability agents or low-ability agents. That is, the government knows the aggregate distribution of characteristics but does not know individual characteristics. It is in this sense that the government has incomplete information.

This basic model is similar to the model in section A of Akerlof (1978), where there are only two types of jobs—difficult and easy jobs—and two types of workers—skilled and unskilled workers. Skilled workers may work in either job, but unskilled workers can only work in the easy job. Indeed, my model could be seen as an argument for using the simple model in section A of Akerlof (1978) to illustrate the significance of the Mirrlees approach to undergraduate students with knowledge of only intermediate microeconomics and elementary algebra. However, my model is simpler than Akerlof's.[2]

I assume that the government wants to redistribute income from the high-ability types (i.e., the rich) to the low-ability types (i.e., the poor). If individuals have identical utility functions, which are increasing and strictly concave in income, then the pre-redistributive marginal utility of a dollar is higher for the low-ability agents than it is for the high-ability agents. Hence a government with a utilitarian social welfare function might want to redistribute income from the rich to the poor. I do not have such a justification in my model. Indeed, I do not specify a formal Bergson-Samuelson-type objective function for the government. However, one need not specify an explicit and formal social welfare function to justify redistribution. For example, a government may want to redistribute income because it has some degree of aversion to inequality (see Boadway and Keen 2000). Also, an incumbent government may engage in the redistribution of income from the rich to the poor to win the next election or strengthen its political power if the poor are in the majority.[3]

ANALYSIS OF THE BASIC MODEL

Clearly, the efficient allocation is for all low-ability people to work in L and all high-ability people to work in H. Note that this is efficient because if any high-ability person works in job L, the aggregate income of the economy will be lower than the aggregate income in the allocation where all high-ability people work in job H.

Suppose the government wants to redistribute income from high-ability people (i.e., the rich) to low-ability people (i.e., the poor). If a worker chooses job H, the government can correctly infer that she is a high-ability person. Therefore, the government's problem is to set up a *self-selection* mechanism such that all high-ability agents have the incentive to work in job H (i.e., will self-select into job H). The government would then tax those in job H because they would have revealed their ability through their choice of job.

The government's goal is to maximize the transfer to a representative low-ability (poor) person while simultaneously preserving incentives (i.e., keeping high-ability people in job H). This is equivalent to maximizing aggregate tax revenue[4] by taxing high-ability people while preserving incentives. We can think of the solution to this problem in the following steps:

(1) The government announces a tax, T, per worker in job H.
(2) Based on T, high-ability agents decide to work in L or H.
(3) The government collects total tax revenue.
(4) The tax revenue is shared equally among *all* workers in job L.

To simplify the analysis, I assume that high-ability people cannot change their labor supply or hide their income in job H in response to the tax on their income. The only possible disincentive effect of government taxation is that it may cause a high-ability person to switch to job L and earn a lower income. These assumptions are consistent with the model in Akerlof (1978). In addition, there are no administrative or monitoring costs of redistribution.

Note that if a high-ability person chooses to work in job H, her post-tax income is $Y_{HH} - T$. Now consider an equilibrium in which all high-ability agents are in job H. If a high-ability agent switches to job L, her post tax-transfer income, *given that all other high-ability agents are in job H*, is

$$Y_{HL} + \frac{(N_H - 1)T}{N_L + 1}. \tag{1}$$

Note that the expression in equation (1) takes into account the fact that a high-ability person who switches to job L is also entitled to the transfers given to all workers in job L because her identity is *not* known. This is where the incompleteness of the government's information kicks in.[5]

Typically, N_H and N_L will be very large relative to a single person. Hence, given that each agent is atomistic in this economy, I can use the following approximations: $N_L + 1 \approx N_L$ and $N_H - 1 \approx N_H$. Therefore, a high-ability person will choose job H, if

$$Y_{HH} - T \geq Y_{HL} + \frac{N_H T}{N_L}. \tag{2}$$

The expression in (2) is called the *self-selection constraint* or the *incentive-compatibility constraint*.

The maximum tax that the government can collect from a person in job H must solve (2) with equality. This gives

$$T^{\max} = \frac{N_L(Y_{HH} - Y_{HL})}{N_L + N_H}. \tag{3}$$

If the government were to set a tax greater than $T^{\max}$, then all high-ability agents would switch to job L, leaving no income to be redistributed to the low-ability agents. Therefore, the need to preserve incentives restrains the government from setting a tax beyond $T^{\max}$. Notice also that $T^{\max}$ is decreasing in Y_{HL}. Hence the size of agents' outside options constrains a government's redistributive ability.

Let us compare this maximum tax with the maximum tax that the government can collect if it has complete information. Under complete information, the government *knows the identity* of high-ability and low-ability agents. Therefore, unlike the case of incomplete information, a high-ability person who works in job L will *not* get the transfer on the right-hand side of equation (2). Indeed, she will not get any transfer from the government, because her identity is known. Given complete information, a reasonable objective for the government is to equalize posttax incomes while preserving Pareto efficiency. It follows that[6]

$$Y_{HH} - T = Y_{LL} + \frac{N_H T}{N_L}. \tag{4}$$

This gives a maximum tax of

$$T^c = \frac{N_L(Y_{HH} - Y_{LL})}{N_L + N_H}. \tag{5}$$

Given $Y_{HL} > Y_{LL}$, it follows that $T^c > T^{\max}$. Hence the government can engage in more income redistribution under complete information. It is in this sense that incomplete information acts as a limit on the government's ability to redistribute income.

I have demonstrated two results. First, by comparing the maximum tax that the government can impose on high-ability agents when it has complete information with the maximum tax when it has incomplete information, I was able to show how incomplete information acts as a constraint on the government's redistributive ability. Second, I showed how the need to preserve incentives also limits the government's redistributive ability (i.e., the tax on a high-ability person cannot exceed $T^{\max}$).

Boadway (1997) describes these core results of the Mirrlees approach as follows:

> Reconciling the redistributive objective with the selection requirement constitutes the nub of the policy problem. The more one tries to redistribute to the less well off, the more attractive the consumption-income bundle of the less well off becomes to the better off. Policy must be designed to prevent the better off from "mimicking" the less well off. This is the so-called *incentive* or *self-selection* constraint and constitutes the *limit to redistribution*. (759)

Akerlof (1978) illustrates this basic tradeoff between redistribution and efficiency as follows:

> As taxes are raised and incomes are redistributed, there is a gain in welfare, because income is distributed to those who have greater need of it (higher marginal utility). But this gain must be balanced against a loss: as tax rates rise in relatively productive jobs and as subsidies rise in relatively unproductive jobs, workers are less willing to take the productive (and more willing to take the unproductive) jobs. Such switching, per se, results in a loss in U *(i.e., aggregate utility)* because each worker is choosing the amount of work, or the kind of job, which maximizes his private utility rather than the amount of work or kind of job which maximizes social utility. In general, the redistributive gains versus the *deadweight* losses caused by tax/transfer-induced *job* switching is the major tradeoff in the theory of optimal income taxes and welfare payments. . . . (11; italics added)

AN EXTENSION OF THE BASIC MODEL

I extend the basic model, maintaining all of the prior conditions, by adding a third job. Call this job M. Let this be a medium-ability job. Let there be a medium-ability group of people who can work in either job L or job M but not job H. In addition to the information above, low-ability people cannot work in job M. Assume that a medium-ability person who is indifferent between L and M will work in M.

Suppose that $N_k > 0$ is the number of k-ability people, where $k = L, M$, and H. Let Y_{kj} be the pre-tax-transfer income of a k-ability person who works in job j, where $j = L, M$, and H and $k = L, M$, and H. In addition to the restrictions in the basic model, the following restrictions are imposed: $Y_{HH} > Y_{HM} > Y_{HL} > Y_{LL}$, $Y_{MM} > Y_{ML} > Y_{LL}$, and $Y_{HM} > Y_{MM}$.

Continue to assume that the government does not know the identity of a worker (i.e., the government does not know whether a person has low, medium, or high ability). Again, a Pareto efficient allocation requires that a low-ability person works in job L, a medium-ability person works in job M, and a high-ability person works in job H. In this allocation, those in job H will be the high-income group, job M workers will be the middle-income group, and job L workers will be the low-income group.

By following a similar analysis as before, we shall find the optimal taxes T_M and T_H that the government can impose on workers in jobs M and job H, respectively, such that (1) the Pareto efficient allocation of workers to jobs is preserved, (2) the government maximizes its tax revenue, R, and (3) the total tax revenue is shared equally among all workers in job L.

High-Ability Agents Cannot Work in Job *M*

For now, suppose that high-ability people can choose to work in either job L or H but not job M. In this three-job–three-type case, we need to recognize that when a high-ability person decides to deviate from job H to L, she has to take into account the transfers that will accrue to those in job L from the taxes collected from those in jobs H *and* M. The same argument applies to a medium-ability person.

Then the self-selection constraints for medium-ability and high-ability people are respectively

$$Y_{MM} - T_M \geq Y_{ML} + \frac{N_H T_H + N_M T_M}{N_L} \tag{6}$$

and

$$Y_{HH} - T_H \geq Y_{HL} + \frac{N_M T_M + N_H T_H}{N_L}, \tag{7}$$

where $Y_{HH} > Y_{HL}$ and $Y_{MM} > Y_{ML}$.

Formally, the government solves the problem:

$$\underset{T_M, T_H}{\text{Max}}\ R = N_M T_M + N_H T_H,$$

subject to the self-selection constraints given by (6) and (7).

Assuming that (6) and (7) hold with strict equality, rewriting these equations gives

$$T_M = \frac{(Y_{MM} - Y_{ML})N_L}{N_L + N_M} - \frac{N_H}{N_L + N_M} T_H \tag{6a}$$

and

$$T_M = \frac{(Y_{HH} - Y_{HL})N_L}{N_M} - \frac{N_H + N_L}{N_M} T_H. \tag{7a}$$

In Figure 1, T_M is on the vertical axis and T_H is on the horizontal axis. The line *hH* represents the self-selection constraint for a high-ability person, *when this constraint holds with strict equality*. This is equation (7a). The line *mM* is similarly defined for a medium-ability person. The slope of the *hH* line is $-(N_H + N_L)/N_M$ and the slope of the *mM* line is $-N_H/(N_M + N_L)$. Because the origin (i.e., $T_M = T_H = 0$) satisfies both self-selection constraints, it follows that the set of points that satisfies *both* (3) and (4) is represented by the area *0HAm* in Figure 1.

For a given level of R, we can generate the locus of the combinations of T_M and T_H that give the same aggregate revenue. This gives an iso-revenue line similar to an indifference curve in consumer theory or an isoquant in producer theory. These iso-revenue lines, shown in Figure 1, are downward-sloping straight lines with slope $-N_H/N_M$. Clearly, $(N_H + N_L)/N_M > N_H/N_M > N_H/(N_M + N_L)$. Hence, any iso-revenue curve is steeper than the *mM* line but flatter than the *hH* line.

We want to attain the highest iso-revenue line, given that the set of feasible points is the area *0HAm*. Clearly, the maximum tax revenue occurs where the iso-revenue line $R*$ goes through the point A (a kink). At point A, the iso-revenue lines *hH* and *mM* intersect (i.e., $T_H = T_H^*$ and $T_M = T_M^*$). Hence the maximum tax revenue is $R^* = T_H^* + T_M^*$. Both self-selection constraints bind.

Note that because equations (6) and (7) hold with strict equality, it follows that $Y_{MM} - T_M^* - Y_{ML} = Y_{HH} - T_H^* - Y_{HL}$. This gives $T_H^* - T_M^* = (Y_{HH} - Y_{HL})-$

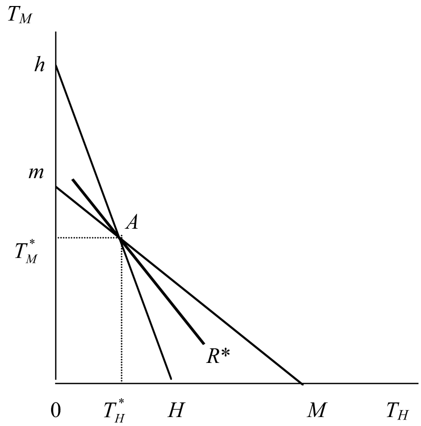

FIGURE 1. Optimal tax solution when high-ability types cannot work in job *M*.

$(Y_{MM} - Y_{ML})$. Therefore, $T_H^* - T_M^*$ has an ambiguous sign. High-ability types pay a higher tax than medium-ability types if the loss in gross income from switching to job L is larger for high-ability types than it is for medium-ability types, that is, if $(Y_{HH} - Y_{HL}) > (Y_{MM} - Y_{ML})$. This accords with intuition.

Suppose the government maintains the tax, T_M^*, but sets a tax for workers in job H that is higher than the tax, T_H^*. Are low-ability people (i.e., the poor) better off? The answer is *no*. To grasp this, note that if the government keeps T_M^* but sets a tax higher than T_H^*, then all workers in job H will now move to job L. Aggregate tax revenue will fall, and there will now be a smaller tax revenue to be shared among more workers in L. This will make the low-ability workers worse off. Hence, by not choosing a tax rate to keep the high-ability workers in job H, the government makes the low-ability workers worse off than would have been the case if it had respected incentives. The lesson is that a government that tries to redistribute too much through excessive taxation might paradoxically end up having a smaller national pie for redistribution. Hence the need to preserve incentives (efficiency) acts as a limit to the government's ability to redistribute income.

Suppose instead that the government knows the identity of high-ability, medium-ability, and low-ability workers, even if they do not reveal their identity through their choices of jobs. Again, the key here is to recognize that if the government knows the ability of everyone, then a person who deviates to job L but is not a low-ability person cannot fool the government by collecting the transfer intended for low-ability workers. Again, assuming that the government equalizes posttax income when it has complete information requires that

$$Y_{MM} - T_M = Y_{LL} + \frac{N_H T_H + N_M T_M}{N_L} \tag{8}$$

and

$$Y_{HH} - T_H = Y_{LL} + \frac{N_M T_M + N_H T_H}{N_L}. \tag{9}$$

The expression on the right-hand side of (8) and (9) is the post–tax transfer income of a low-ability person. We can rewrite (8) and (9) as

$$T_M = \frac{(Y_{MM} - Y_{LL})N_L}{N_L + N_M} - \frac{N_H}{N_L + N_M} T_H \tag{8a}$$

and

$$T_M = \frac{(Y_{HH} - Y_{LL})N_L}{N_M} - \frac{N_H + N_L}{N_M} T_H. \tag{9a}$$

Comparing equations (6a) to (8a) and (7a) to (9a), we see that equations (6a) and (8a) have the same slope and (7a) and (9a) also have the same slope. However, (8a) and (9a) have higher positive intercepts. Hence, if one were to draw the lines represented by (8a) and (9a) in Figure 1, they would be parallel *outward* shifts of the *hH* and *mM* lines. Therefore, under complete information, the government can attain a higher iso-revenue line than it can under incomplete information. This establishes the point that the redistributive power of the government is limited under incomplete information.[7]

Notice that by shifting the *hH* and *mM* lines downward and varying their slopes accordingly, we can find a point of intersection between these lines where $T_M < 0$ but $T_H > 0$. In this case, medium-ability types are given a subsidy to stay in job M, and only high-ability types pay taxes. However, the subsidy to those in job M relaxes the self-selection constraint for high-ability types and allows the government to raise more tax revenue compared to the case of imposing a tax on job M workers. This case occurs when the intercept of the *mM* line (i.e., $Y_{MM} - Y_{ML}$) is sufficiently small and the intercept of the *hH* line (i.e., $Y_{HH} - Y_{HL}$) is sufficiently high. Clearly, if job L is sufficiently attractive to medium-ability types but not so attractive to high-ability types, then it is not surprising that to keep medium-ability types in job M, a subsidy may have to be paid to them.

High-Ability Agents Can Work in Job *M*

Now suppose that high-ability people can work in all three jobs and the government has incomplete information. This adds a third self-selection constraint to the analysis. That is, in addition to job L, we have to discourage the high-ability people from working in job M. This requires that $Y_{HH} - T_H \geq Y_{HM} - T_M$. Assuming that this constraint holds with equality gives

$$T_M = -(Y_{HH} - Y_{HM}) + T_H. \tag{10}$$

Given that high-ability types can now work in job M, are the taxes T_M^* and T_H^* still optimal taxes? Note that given these taxes, a high-ability person has no incentive to work in job L. If $T_M^* \geq T_H^*$, then a high-ability person has no incentive to work in job M, so T_M^* and T_H^* are optimal.

The more interesting case is $T_M^* < T_H^*$. Given $T_M^* < T_H^*$, a high-ability person has the incentive to switch to job M to pay a lower tax. The gain of this switch is $T_H^* - T_M^* \equiv \Delta_T > 0$. However, the cost of this switch is the loss of income equal to $Y_{HH} - Y_{HM}$. Therefore, a high-ability person will not switch to job M if $Y_{HH} - Y_{HM} \geq \Delta_T$. It follows that T_M^* and T_H^* are the optimal taxes if $Y_{HM} \leq Y_{HH} - \Delta_T$. Noting that $T_H^* - T_M^* = (Y_{HH} - Y_{HL}) - (Y_{MM} - Y_{ML})$, I can rewrite $Y_{HM} \leq Y_{HH} - \Delta_T$ as $(Y_{HM} - Y_{HL}) - (Y_{MM} - Y_{ML}) \leq 0$.

The result in the preceding paragraph can be demonstrated diagrammatically. To see this, refer to Figure 2, where the line *ZZ* represents equation (10). The *ZZ* line is an upward-sloping line with a negative intercept (i.e., – $[Y_{HH} - Y_{HM}]$). If $Y_{HH} - Y_{HM}$ is sufficiently high, then the *ZZ* line will be below the point (T_M^*, T_H^*). In this case, (T_M^*, T_H^*) are still the optimal taxes. The constraint given by the *ZZ* line does not bind. Interestingly, the condition $Y_{HH} - Y_{HM} \geq \Delta_T$ obtained in the previous paragraph is consistent with $Y_{HH} - Y_{HM}$ being sufficiently high.

Given $\Delta_T > 0$, suppose $Y_{HM} > Y_{HH} - \Delta_T$; then either T_M^* or T_H^* or both are no longer optimal. To prevent a high-ability person from switching to job M, job M must be less attractive or job H more attractive.[8] Note that $Y_{HM} > Y_{HH} - \Delta_T$ can be rewritten as $Y_{HH} - Y_{HM} < \Delta_T$. This gives $(Y_{HM} - Y_{HL}) - (Y_{MM} - Y_{ML}) > 0$, which requires that Y_{HM} be sufficiently high. This means that the negative intercept, $-(Y_{HH} - Y_{HM})$, of the *ZZ* line is sufficiently close to the origin. In this case, the *ZZ* line will intersect the *mM* line above the point (T_M^*, T_H^*), as shown in Figure 3. The optimal pair of taxes $(\tilde{T}_M, \tilde{T}_H)$ occurs at the point of intersection between the *mM* and *ZZ* lines, where $\tilde{T}_M > T_M^*$ and $\tilde{T}_H < T_H^*$. The revenue collected is represented

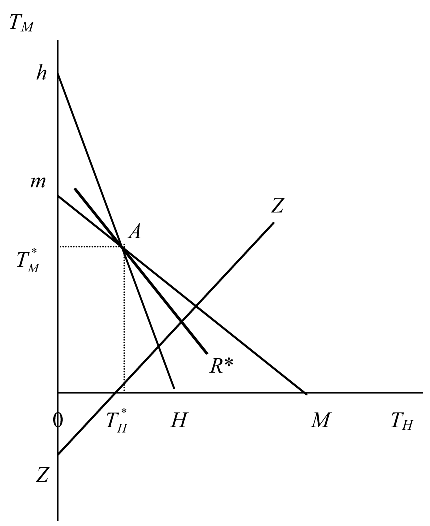

FIGURE 2. Optimal tax solution when high-ability types can work in job M and $Y_{HM} \leq Y_{HH} - \Delta_T$.

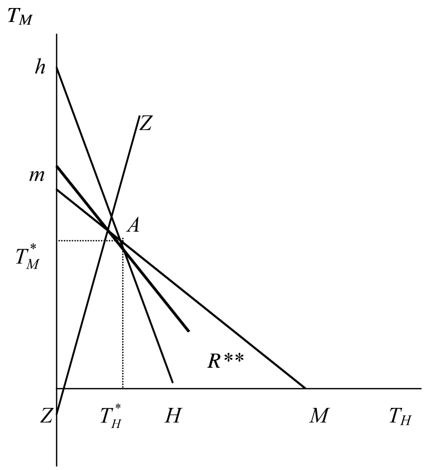

FIGURE 3. Optimal tax solution when high-ability types can work in job M and $Y_{HM} > Y_{HH} - \Delta_T$.

by the iso-revenue line R^{**}, which is lower than the revenue when high-ability types could not work in job M. In this case, the constraints given by *ZZ* and *mM* bind, but the *hH* constraint does not bind.[9]

The result—$\tilde{T}_H < T_H^*$—is interesting. It shows that when the high-ability people have more outside options (i.e., can work in job M in addition to job L) and these options are sufficiently attractive (i.e., $Y_{HM} > Y_{HH} - \Delta_T$), then the government must impose lower taxes on them compared to the scenario under which they had fewer options (i.e., can *only* work in job L). Therefore, government tax policy must take into account the options available to economic agents. For example, a tax on capital must consider whether agents can invest their capital outside the country's borders. Another example is that a tax by the U.S. government on basketball players, football players, doctors, actors, and so forth must take into account whether these professionals can sell their services in other countries.

There is also another possibility. Suppose $Y_{HH} - Y_{HM}$ is very close to zero and $(\tilde{T}_M, \tilde{T}_H)$ are the optimal taxes. Now if $Y_{HH} - Y_{HM}$ is close to the zero, then the *ZZ* line constraint (i.e., equation [10]) requires that $\tilde{T}_M - \tilde{T}_H$ be close to zero. The need to set T_H almost equal to T_M ties the government's hands in terms of how much it can extract from high-ability types. However, if $(Y_{HH} - Y_{HL})$ is sufficiently higher than $(Y_{MM} - Y_{ML})$, then the government may want to free itself from this restriction (i.e., $T_M = T_H$) and instead set an arbitrarily high T_M to make job M unattractive to both high-ability and medium-ability types but set a reasonably high T_H to keep the high-ability types in job H. The per-capita transfers to all workers in job L will then be higher than in the Pareto efficient case. Aggregate income is not maximized, because medium-ability types now work in job L.[10]

Formally, the government sets T_H to satisfy $Y_{HH} - T_H \geq Y_{HL} + N_H T_H/(N_L + N_M)$, given an arbitrarily high T_M. The optimal T_H satisfies this self-selection constraint with equality. This gives

$$\hat{T}_H = \frac{N_L + N_M}{N}\,(Y_{HH} - Y_{HL})\,, \tag{11}$$

where $N \equiv N_L + N_M + N_H$. Total revenue is $N_H \hat{T}_H$.

Note that $(\tilde{T}_M, \tilde{T}_H)$ is obtained by solving equations (6a) and (10) simultaneously. This gives

$$\tilde{T}_H = \frac{N_L}{N}\,(Y_{MM} - Y_{ML}) + \frac{N_L + N_M}{N}\,(Y_{HH} - Y_{HM})\,. \tag{12}$$

As $(Y_{HH} - Y_{HM}) \to 0$, $\tilde{T}_M - \tilde{T}_H \to 0$, given equation (10). So total revenue is approximately equal to $\tilde{T}_H(N_H + N_M)$. Then the government will sacrifice efficiency if

$$N_H \hat{T}_H/(N_L + N_M) - (N_M + N_H)\tilde{T}_H/N_L > 0. \tag{13}$$

Note that $N_H/(N_L + N_M) < (N_M + N_H)/N_L$. However, $(Y_{HH} - Y_{HM}) \to 0$ and $\Delta_T = (Y_{HH} - Y_{HL}) - (Y_{MM} - Y_{ML}) > 0$ imply that $\hat{T}_H > \tilde{T}_H$. Hence, (13) will hold if $(Y_{HH} - Y_{HL}) - (Y_{MM} - Y_{ML})$ is sufficiently large.

In a first-best world (i.e., the government has complete information), the government would not sacrifice efficiency to improve the welfare of the low-ability people. Improving their welfare through transfers requires the maximization of aggregate output. However, in a second-best world (i.e., the government has incomplete information), it is possible that the government can improve the welfare of the low-ability people by sacrificing efficiency. Note that by setting a very high tax in the medium-ability job, the government effectively destroys this job. Hence, we now have only two jobs with three ability types. The bigger the difference between Y_{HH} and Y_{HL}, the bigger the tax revenue that the government can collect from the high-ability workers. On the other hand, the smaller the difference between $Y_{MM} - Y_{ML}$, the smaller the tax revenue that the government can collect from medium-ability people. Hence, it may be efficient for the government to collect tax revenue from only high-ability people if $Y_{HH} - Y_{HL}$ is sufficiently bigger than $Y_{MM} - Y_{ML}$. This is the intuition for the analysis in the preceding paragraph. And in this case, high-ability types are worse off relative to the Pareto efficient case.

CONCLUSION

I have used a simple model adapted from Akerlof (1978) to illustrate a key result in the approach to public economics and redistributive taxation pioneered by the Nobel laureate James Mirrlees. I demonstrated, at a very elementary level, that incomplete information and the need to preserve incentives limit a government's redistributive power. The analysis shows that if self-selection constraints are respected, then redistribution will have no effect on efficiency, and if they are violated, then redistribution will have an adverse effect on efficiency. However, there are situations in which redistribution can actually enhance efficiency or lead to an increase in aggregate output.[11] I also showed that the size of agents' outside options constrains a government's redistributive ability.

In the model, I assume that the government cannot force people to work in a particular job even if it has complete information. For example, the government cannot force high-ability people to work in job H and extract much more from them, nor can it ban them from working in job L. If the government had these powers, it could engage in even more redistribution to the extent of making the low-ability workers richer than the high-ability workers after taxes and transfers. But even so, caution is needed. Although I do not consider labor supply in our model, it is conceivable that the high-ability workers could reduce their supply of labor if they were forced to work in job H, resulting in earnings far below Y_{HH}. There are other efficiency costs of excessive taxation. For example, to force people to do things against their will, resources must be spent to monitor them. This will be another efficiency cost of forcing high-ability workers to work in job H. This explains why the efficiency costs of coercive systems, such as slavery, dictatorships, and authoritarian regimes, are high. These institutions are not only immoral but they are also grossly inefficient unless a group of people can be perpetually socialized to accept them.

My primary goal was to present a pedagogical optimal tax model and analyses that could be easily understood by undergraduate students with only knowledge of intermediate microeconomics and elementary algebra. I hope instructors and students in welfare economics, health economics, microeconomics, and, of course, public economics will find it useful.

NOTES

1. The survey in Boadway (1987, 1988) is nontechnical and mentions William Vickrey's important contribution to this literature. Stiglitz (1987) presents a technical analysis. Boadway (1987, 1998) refers to this literature as the *Mirrlees' approach* or the *information-based approach to public policy*, and Stiglitz (1987) called it the *new new welfare economics*. What one may call the *old welfare economics* or *old public economics* assumed complete information leading to the conclusion that one could separate issues of distribution from efficiency (i.e., the second welfare theorem). As is shown in a subsequent endnote, the second welfare theorem is significantly weakened in a world of incomplete information. See Boadway (1997) for a good discussion of this point.
2. Another simple model in this literature is that of Nichols and Zeckhauser (1982). However, the self-selection constraints are not explicitly written, because the bulk of the analysis is graphical. Furthermore, the behavior of the self-selection constraint as transfers change does not lend itself to a straightforward diagrammatic analysis. Last, the framework in Akerlof (1978) allows me to extend my model to three different groups with three self-selection constraints.
3. Harms and Zink (2003) presented a good survey of redistribution in a democracy and discussed several limits to redistribution. Boadway and Keen (2000) presented a more elaborate survey.
4. Maximization of tax revenue as an objective for governments is the basis of the famous Laffer curve and has been used by Stowhase and Traxler (2005), Keen and Kotsogiannis (2003), and Sanchez and Sobel (1993). However, governments have different reasons for maximizing tax revenue, and some of these reasons have nothing to do with the welfare of their citizens.
5. However, if the government can observe incomes, then it can identify a high-ability person in job L, given that $Y_{HL} \neq Y_{LL}$. One way of getting around this problem is to assume, as in, for example, Besley and Coate (1991), that the government cannot observe incomes. Another way is to define, as in Akerlof (1978), a nonpecuniary cost to high-ability types of working in job H stemming from the greater effort required in job H relative to job L. If $\theta > 0$ is the monetary equivalent of this cost, then this is a gain to a high-ability type who works in job L. So I can write

$Y_{HL} = Y_{LL} + \theta$, where θ is a component of income that is not observed by the government and is included in the self-selection constraint. I thank a referee for this point.

6. Notice that $Y_{HH} - T > Y_{HL} - T$ holds for any T. Hence, equalizing income under complete information is consistent with keeping high-ability workers in job H.
7. According to the second welfare theorem, a planner should be able to implement any Pareto efficient allocation as a competitive equilibrium (under certain conditions). Notice that this implies that, under certain conditions, any distribution of income that is consistent with Pareto efficiency can be attained. However, because the set of Pareto efficient distributions of income under incomplete information is only a subset of the set of Pareto efficient distributions of income under complete information, it follows that the second welfare theorem does not hold when the government has incomplete information.
8. Notice that increasing both taxes will violate the self-selection constraints in equations (6) and (7). Reducing both taxes is also not optimal, as our diagrammatic exposition shows.
9. I thank the referees for asking me to explore this case.
10. I thank a referee for asking me to explore this issue.
11. This is outside the article's scope. The reader may refer to Boadway and Keen (2000).

REFERENCES

Akerlof, G. A. 1978. The economics of "tagging" as applied to the optimal income tax, welfare programs, and manpower planning. *American Economic Review* 68 (1): 8–19.

Besley, T., and S. Coate. 1991. Public provision of private goods and the redistribution of income. *American Economic Review* 81 (4): 979–84.

Boadway, R. W. 1997. Public economics and the theory of public policy. *Canadian Journal of Economics* 30 (4): 753–72.

———. 1998. The Mirrlees approach to the theory of economic policy. *International Tax and Public Finance* 5 (1): 67–81.

Boadway, R., and M. Keen. 2000. Redistribution. In *Handbook of income distribution*, ed. A. B. Atkinson and F. Bourguignon, 677–789. Amsterdam: North Holland.

Harms, P., and S. Zink. 2003. Limits to redistribution in a democracy. *European Journal of Political Economy* 19 (4): 651–68.

Keen, M., and C. Kotsogiannis. 2003. Leviathan and capital tax competition in federations. *Journal of Public Economic Theory* 5 (2): 177–99.

Mirrlees, J. A. 1971. An exploration in the theory of optimum income taxation. *Review of Economic Studies* 38 (114): 175–208.

Nichols, A. L., and R. J. Zeckhauser. 1982. Targeting transfers through restrictions on recipients. *American Economic Review* 72 (2): 372–77.

Sanchez, I., and J. Sobel. 1993. Hierarchical design and enforcements of income tax policies. *Journal of Public Economics* 50 (3): 345–69.

Stiglitz, J. E. 1987. Pareto efficient and optimal taxation and the new new welfare economics. In *Handbook of public economics*, Vol. 2, ed. A. J. Auerbach and M. S. Feldstein, 991–1042. Amsterdam: North Holland.

Stowhase, S., and C. Traxler. 2005. Tax evasion and auditing in a federal economy. *International Tax and Public Finance* 12 (4): 515–31.