

# The Base for Direct Taxation: Commentary

Pierre Pestieau

Prepared for the Report of a Commission on  
Reforming the Tax System for the 21st Century,  
Chaired by Sir James Mirrlees

[www.ifs.org.uk/mirrleesreview](http://www.ifs.org.uk/mirrleesreview)

The Institute for Fiscal Studies

Full Report to be published by  
Oxford University Press



# The tax base for direct taxation

by

J. Banks, P. Diamond and J. Mirrlees

## *Comments<sup>1</sup>*

by Pierre Pestieau

February 21, 2008

The chapter of J. Banks, P. Diamond and J. Mirrlees (BDM) presents an excellent survey of the existing literature on optimal taxation and discusses a number of lessons that can be drawn from that literature. One of the main lessons on which their chapter focuses concerns the treatment of capital income. The authors argue that the finding that the optimal income tax schedule should not include tax on capital is based on too many restrictions, and is thus not robust for policy purposes. Another lesson is that taxation should vary with age. Not having any quarrel with these two points, I would like in this comment to discuss a number of points too quickly dealt with or deliberately neglected by BDM. They concern the issue of tagging, the problem raised by having more than one unobservable characteristic (besides ability), the issue of myopia and prodigality, the question of equal opportunities, the taxation of couples and the threat of tax competition.

### 1. Tagging

Supplementing optimal income taxation with tagging generally brings more welfare. Yet to qualify such a general statement, one has to distinguish different types of tagging according to whether or not it is costly, it brings stigmatisation and it conveys some particular information.

The tagging BDM have in mind when they recommend income taxation varying with age is both costless and neutral. In contrast, characteristics such as height, weight, language or colour, even though one can show that using them would be welfare improving, can have negative effects that more than offset those positive effects. This would be the case, e.g., of tagging Belgian citizens with different tax schedules for Dutch and French speaking along with intergroup transfers.<sup>2</sup>

Boadway and Pestieau (2006) have studied the issue of tagging with optimal income taxation in a two-group-two-skill-level setting. They show that tagging leads to horizontal inequity and more progressivity in the group comprising the higher proportion of unskilled workers.

---

<sup>1</sup> I am grateful to Robin Boadway and Helmuth Cremer for their helpful suggestions on an earlier draft of these comments.

<sup>2</sup> It is interesting to note that social attitudes towards the acceptability of some tagging (race, gender, ...) vary over time and across countries.

In general the characteristics considered are given and cannot be changed within a reasonable length of time. When they can be modified, we have an issue of moral hazard and the taxation problem becomes more difficult. One thinks of health status and family size that are important matters in the design of an optimal income tax structure.

When tagging is uncertain and costly, it is not clear that it is worth using. This issue has been extensively studied in work on disability insurance.<sup>3</sup>

Free tagging can convey different types of information. Not only a tagged group has a specific distribution of ability (in the two-type case, the relative number of skilled and unskilled), but also different needs (linked, e.g., to health or family size). Boadway and Pestieau (2003) have studied optimal income taxation when the tag reflects differences in needs, that is, differences in the resource required to achieve a given level of utility. In a two-ability setting, they show that the level of compensation given for needs exceeds the level of needs if a higher proportion of low ability households are needy, and vice-versa.

## **2. More than one characteristic<sup>4</sup>**

In the standard model of optimal income taxation, individuals only differ in productivity. This is clearly restrictive as we know that they also differ in other characteristics. When these characteristics are observable and convey some information on ability, we have the issue of tagging. When they are not observable, we face the analytical difficulty of dealing with a multidimensional principal agent problem. Let us assume that we have as unknown characteristics not only productivity, but also wealth, the rate of time preference or longevity. There is no doubt that these characteristics are not easily observable and also that they are partially correlated with productivity level.

As explained by BDM, with these and other characteristics, the Atkinson-Stiglitz proposition does not hold, but furthermore, we have a good case for taxing capital income. The intuition is simple: individuals with low discount rate, high wealth or high longevity tend to save more than those who are impatient, poor or have a low life expectancy.

Among the characteristics that can be introduced besides labor productivity, there is the risk of morbidity. As shown by Rochet (1989) (see also Boadway, Leite-Monteiro, Marchand, Pestieau (2003, 2006)), if morbidity risks are negatively related to income so that the poor face higher risks on average, then we have an obvious argument for social insurance. Social insurance combined with a standard distortionary income tax can redistribute more effectively. The reason is that redistributing through social insurance does not involve the same distortion as through income taxation. This result is shown to hold with moral hazard and adverse selection.

---

<sup>3</sup> See, e.g., Diamond and Sheshinski (1995).

<sup>4</sup> This section is further developed in the appendix. We make a difference between the tagging problem which arises when the characteristic just pertains to the distribution of skills and the problem of adding one characteristic which involves utility differences. On this issue see Kaplow (2008).

### 3. Preferences for leisure and responsibility

The problem of optimal taxation when people have different preferences for leisure raises difficult normative questions. A higher income may be due either to differences in innate productivity and skill levels or to differences in the degree of effort. Progressive taxes can therefore imply redistribution from those with a low preference for leisure to those with a high preference for leisure and from those with high productivity to those with low productivity. The ethical valuation of such redistributions depends on the interpretation given to the preference parameter. One may have ethical objections against redistributing from the hard working to the leisure-prone workers. Those objections would however disappear for redistribution from the skilled to unskilled. This distinction is linked to the notion of responsibility. One generally feels that people should be compensated for factors that are beyond their control. Innate skills are typical examples of such factors, which have led to the traditional literature on optimal income taxation. In contrast, it is felt that people should be held responsible for factors that are under their control. Typical examples of such factors are preferences for leisure.

Following Roemer (1996) [see also Fleurbaey and Maniquet (2006, 2007)], in a society consisting of individuals with the same productivity but different tastes for leisure, there should be no redistribution. Equality of opportunities would be achieved in such a society.<sup>5</sup>

A further difficulty arises when a low preference for work may mean two different things: taste for leisure and difficulty to work. As BDM observes "viewing a worker as lazy (liking leisure) is very different from viewing a worker as having difficulty working longer, perhaps for physical reasons." Cremer *et al.* (2007b) and Marchand *et al.* (2002) show that when a society consists of those two types of people, disabled and leisure prone, who have the same formal utilities the only way out in a second-best setting is either to resort to audits or to use indirect taxation given that they most likely don't have the same consumption needs.

### 4. Myopia and prodigality

There exists evidence that some households might undersave for two separate reasons: myopia and prodigality. Myopia comes from lacking self-control: individuals try to balance two objectives: instant gratification and retirement planning; quite often they err on the side of using too much of their resources for instant gratification and not enough to plan for retirement. Another source of insufficient saving occurs when individuals know that their government tends to bail out retirees without resources and thus they are tempted to consume all their resources during the active part of their life.

Myopia and prodigality<sup>6</sup> both provide reasons for individuals not to save for retirement, they make a strong case for the government to foster saving through subsidy on retirement saving or even mandatory pensions.

---

<sup>5</sup> See also Schokkaert *et al.* (2004).

<sup>6</sup> See Pestieau and Possen (2008).

In this particular case, we have an argument not to tax, but to subsidize some types of capital income.

## **5. Non welfarism and new paternalism**

We have just mentioned a number of instances where the social planner may be tempted not to follow individual preferences in assessing the social desirability of tax policy. There is the case of leisure prone or lazy individuals. The government might be tempted to induce them to work more than the disabled workers, from whom they cannot be distinguished. There is also the case of altruism where laundering out utilities has been advocated by, e.g., Hammond (1987).

There is also the whole range of situations where people make decisions against their own good intention. In these situations, individuals might want the government to intervene to induce behavior that is closer to what individuals wish they were doing. Procrastination, myopia, consumption of sin goods are examples of behavior that lead to what is called "new paternalism" and that implies non welfarist objective functions. (see Kanbur *et al.* (2006)).

Depending on the specific situations, what is called new paternalism is more or less accepted. There is little disagreement on the use of a paternalistic social welfare function for sin goods; for altruism whether or not individual preferences have to be laundered out is more disputed. Finally, the issue of distinguishing characteristics that comes from luck and those that come from responsible decisions made by individuals is controversial.

## **6. The taxation of couples**

For a long time, the implicit picture of the two-person household was clearly one in which there is a complete division of labor between partners, with one specializing entirely in labor supply to the market, the other producing goods and services within the home. With such a picture, the analysis of optimal income taxation by Mirrlees (1971) that took the decision unit as a single individual dividing his time between market work and leisure didn't seem out of place. The spectacular growth in female labor force participation that took place in almost all developed countries called this picture into question and presents a new issue for tax policy. How to tax two-earner couples? The policy choice can be reduced to three alternatives:

- joint taxation in which the partners' income are added together and taxed at progressive marginal rates as if they had each earned one-half the income. This implies equality of marginal tax rates on partner income, or, that the tax rate on the last dollar of the husband's income was applied to the first dollar of the wife, as it was at times expressed;
- individual taxation in which each partner's income is taxed separately, but according to the same progressive tax schedule;
- selective taxation in which secondary earners are taxed on a separate, lower, progressive tax schedule than that of primary earners.

The paper by Boskin and Sheshinski (1983) is generally regarded as having established the conventional rule on this issue, namely that selective, and not joint or individual, taxation is optimal. That is, not only should women be taxed

separably from men, but they should be taxed on a lower rate schedule. This model suffers from the limitation that it rests on linear taxation.

Recently a number of papers<sup>7</sup> have been published dealing with this issue in a standard non-linear optimal tax framework without any a priori assumptions on the tax function of the household. Viewing the problem in this way gives arguments for spouses to face different marginal income tax rates, casting doubts on total family income as appropriate income tax base. It does not however vindicate the case for individual taxation. Both the productivity of spouses and the relative position of each have some bearing on the marginal rates faced by both of them.

What is also important is to take a dynamic view. It may happen that the tax structure deemed optimal in a static setting implies that one of the two spouses (traditionally the wife) stops working because of too high marginal tax. This outcome would not remain optimal if the possibility of divorce is taken into account, granted that stopping working strongly decreases the chance of getting a good and well paid job when needed.

## **7. Capital income taxation and tax competition**

The increased integration of the world capital market implies that the supply of capital becomes more elastic and therefore potentially a less efficient base for taxation. It is thus possible that taxing capital income is highly desirable in a closed economy and then becomes difficult in an open economy. This threat of a vanishing tax on capital income has led some countries to call for some type of cooperation, including the idea of a minimum withholding tax on capital income tax.

## **8. Implication for the tax treatment of capital income**

Some of the points just raised have admittedly little bearing on the gist of BDM's chapter, namely the taxation of capital income relative to that of earnings.

One clearly sees the implications of tax competition on capital income taxation. One also understands that introducing additional characteristics such as discount rates or longevity in the optimal tax model can lead to taxing capital income. Finally, the case of myopia and prodigality has clear implication as to the treatment of capital income.

What about the issue of equal opportunities? Let us use the two-period model of optimal income and consumption taxation with two unobservable characteristics: ability and wealth endowment. Assume that wealth endowment can result from either pure luck – unexpected bequest – or assistance to an ailing parent – exchange bequest. According to Roemer's view that there should be not redistribution for characteristics the individuals are responsible for, one can expect the case for capital income taxation to be stronger when bequests are accidental than when they result from family solidarity.

---

<sup>7</sup> Brett (2007), Cremer *et al.* (2007a), Kleven *et al.* (2007).

Concerning the taxation of couples, there exists no work addressing the issue of the relative effect of joint versus separate filing on capital income taxation.

## References

- Apps, P. and R. Rees (2004), The taxation of couples, in A. Cigno, P. Pestieau and R. Rees (ed.), *Taxation and the Family*, MIT Press, forthcoming.
- Atkinson, A. and J. Stiglitz (1976), The design of tax structure: direct versus indirect taxation, *Journal of Public Economics*, 6, 55-75.
- Boadway, R. and P. Pestieau (2006), Optimal taxation with tagging, *Annales d'Economie et de Statistique*, 83-84, 123-50.
- Boadway, R., M. Marchand and P. Pestieau (2000), Redistribution with unobservable bequests. A case for taxing capital income, *Scandinavian Journal of Economics*, 102, 253-67.
- Boadway, R. and P. Pestieau (2003), Indirect Taxation and Redistribution: the Scope of the Atkinson-Stiglitz Theorem in *Imperfect Economics: Essays in Honor of Joseph Stiglitz*, R. Kanbur and R. Arnott (eds), MIT Press.
- Boadway, R., M. Leite-Monteiro, M. Marchand and P. Pestieau (2006), Social insurance and redistribution with moral hazard and adverse selection, *Scandinavian Journal of Economics*, 108, 279-98.
- Boadway, R., M. Leite-Monteiro, M. Marchand and P. Pestieau (2003), Social insurance and redistribution, in S. Cnossen (ed.), *Public Finance and Public Policy in the New Millenium*, Mc Millan, 333-358.
- Boadway, R., M. Marchand, P. Pestieau and M. Racionero (2002), Optimal redistribution with heterogeneous preference for leisure, *Journal of Public Economic Theory*, 4, 475-498.
- Boskin, M. and E. Sheshinski (1983), Optimal tax treatment of the family: married couples, *Journal of Public Economics*, 20, 281-97.
- Brett, C. (2007), Optimal non-linear taxes for families, *International Tax and Public finance*, 4, 225-263.
- Cremer, H., J-M. Lozachmeur and P. Pestieau (2007a), Income taxation of couples and the tax unit, CORE DP 2007/14.
- Cremer, H., J-M. Lozachmeur and P. Pestieau (2007b), Disability testing and retirement, *BE Journal of Economic Analysis and Policy*, vol. 7, issue 1.
- Cremer, H., P. Pestieau and J-Ch. Rochet (2001), Direct versus indirect taxation. The design of the tax structure revisited, *International Economic Review*, 42, 781-99.

- Diamond, P. and E. Sheshinski (1995), A model of social insurance with variable retirement, *Journal of Public Economics*, 10, 295-336.
- Fleurbaey, M. and F. Maniquet (2006), Fair income tax, *Review of Economic Studies*, 73, 55-83.
- Fleurbaey, M. and F. Maniquet (2007), Help the low-skilled or let the hardworking thrive? A study of fairness in optimal income taxation, *Journal of Public Economic Theory*, 9, 467-500.
- Hammond, P. (1987), Altruism. The New Palgrave. A Dictionary of Economics, Macmillan, London, 85-87.
- Kanbur, R., J. Pirttila and M. Tuomala (2006), Non welfarist optimal taxation and behavioural public economics, *Journal of Economic Survey*, 20, 849-868.
- Kaplow, L. (2008), Optimal policy and heterogeneous preferences, unpublished.
- Kleven, H.J., C.T. Kremer and E. Saez (2007), The optimal income taxation of couples, unpublished.
- Lozachmeur, J-M. (2006), Optimal age specific income taxation, *Journal of Public Economic Theory*, 8, 911-922.
- Marchand, M., P. Pestieau and M. Racionero (2003), Optimal redistribution programs when different workers are indistinguishable, *Canadian Journal of Economics*, 36, 911-922.
- Mirrlees, J. (1971), An exploration in the theory of optimal taxation, *Review of Economics Studies*, 38, 175-208.
- Pestieau, P. and U. Possen (2008), Prodigality and myopia: two rationales for social security, Manchester School (forthcoming).
- Roemer, J. (1996), *Theories of Distributive Justice*, Harvard University Press.
- Schokkaert, E., D. van de Gaer, F. Vandenbrouck and R. Ivan Luttens (2004), Responsibility sensitive egalitarianism and optimal linear income tax, *Mathematical Social Sciences*, 48, 151-182.

## Appendix

### Note on introducing additional features in the optimal income tax

---

To illustrate the issue at hand, we start with the following two period utility:

$$u(c, d, \ell) = u_1(c, \ell) + \beta u_2(d)$$

where  $c$  and  $d$  are first- and second-period consumption,  $\ell$  is labor in the first-period,  $\beta$  is a discount factor (reflecting either time preference or survival probability or both) and  $u$ ,  $u_1$  and  $u_2$  are utility functions with standard properties.

In a *laissez-faire* economy, the individual faces the following budget constraint:

$$\omega + w\ell = c + d/R$$

where  $R$  is an interest factor,  $\omega$  is an initial endowment and  $w$ , the wage rate or the ability level.

Assume that individuals only differ in their ability. Assume further that the government just observes earnings  $y = w\ell$  and neither  $w$  or  $\ell$ . For the time being  $\omega = 0$ .

We want to know what would be the tax system that could lead to maximizing the sum of the above utilities. In the case of two types, we have:

$$\sum_{i=1,2} n_i \left[ u(c_i, d_i, \ell_i) - \mu(c_i + d_i/R - w_i \ell_i) \right]$$

where  $n_i$  is the relative number of individuals with ability  $i$ ,  $\mu$  is the Lagrangian multiplier and the optimal package  $(c_i, d_i, \ell_i)$  is subject to the self-selection constraint:

$$u\left(c_2, d_2, \frac{y_2}{w_2}\right) \leq u\left(c_1, d_1, \frac{y_1}{w_2}\right)$$

given that we assume  $w_2 > w_1$ .

The results of this standard problem are well-known: an optimal income taxation with zero marginal tax on  $y_2$ , and no indirect taxation, here a tax on savings, if there is weak separability between  $(c_i, d_i)$  and  $(\ell_i)$ . The first result is attributed to Mirrlees (1971) and the second to Atkinson and Stiglitz (1976).

We now modify this setting by introducing some additional features. We consider three ones that are mutually exclusive:

- (A) differential endowments  $\omega_i$ ,
- (B) differential discount factors  $\beta_i$ ,
- (C) the population can be divided according to a neutral and exogenous characteristic that provides information on the relative number of types. By neutral we mean that the characteristic has no negative connotation (e.g., skin color). By exogenous we mean that it cannot be changed (e.g., weight).

Another distinction one has to consider is the observability of these three features. They can be freely observable, observable at some cost, with or without errors and they are not at all observable. To keep things simple, we consider the two extreme cases: perfect observability or no observability. Table 1 summarizes our 6 cases.

**Additional features**

	Observable	Not observable
Endowment	A1	A2
Discount factor	B1	B2
Tag	C1	C2

The questions we want to address for each of these cases are, (i) how the introduction of this additional feature affects the progressivity of income taxation, (ii) whether it has an implication for the Atkinson-Stiglitz proposition, which in this particular setting means zero taxation on capital income.

A.1. If one observes those initial endowments, the obvious first step is to redistribute them equally. If there is any correlation between  $\omega$  and  $w$ , in other words if, e.g., individuals with higher endowment tend to be more skilled than those with lower endowment, we go to case C1 with observable tags.

A.2. If one does not observe  $\omega$  but only knows its distribution and its correlation with  $w$ , we have the problem dealt with by Cremer *et al.* (2001) and Boadway *et al.* (2000). The gist of these papers is that there is a good case for capital income taxation.

B.1. There is a big difference between differential endowments and different discount factors (longevity or time preference). In the latter case we deal with different utilities and summing them is questionable. In addition, if one judges that the feature at hand is endogenous – the individual is responsible for it – one might be reluctant to take it into account in the objective function.

Keeping aside these considerations, if the discount factor  $\beta$  is unrelated to  $w$ , there will be some redistribution from those with a lower discount factor to those with a higher discount factor. The Atkinson-Stiglitz proposition will hold. If there is any correlation between  $\beta$  and  $w$ , we have a tagging problem as in C1.

B.2. When  $\beta$  is not observable but is (e.g., positively) correlated with  $w$ , the AS proposition does not hold and one expects some taxation of saving on type 1's

individuals. Such a taxation relaxes the self-selection constraint preventing type 2's individuals to mimick type 1's individuals.

C.1. This is the problem studied by Boadway and Pestieau (2006). It appears clearly that using tags in optimal income taxation is always desirable, but generates horizontal inequity. Furthermore, there is more progressivity in the group with a higher proportion of skilled workers than in the group with a lower proportion of skilled workers. The AS proposition holds.

C.2. When the "tags" are not observable, it can't help. We then have a unique tax schedule and the AS proposition holds.