# Pareto efficient taxation and expenditures: Pre- and re-distribution☆

## Joseph E. Stiglitz

*Columbia University, United States*

## ABSTRACT

This paper shows that there is a presumption that Pareto-efficient taxation entails a positive tax on capital. When tax and expenditure policies can affect the market distribution of income in ways that cannot be directly offset, those effects need to be taken into account, reducing the burden imposed on distortionary redistribution. The paper extends the 1976 Atkinson-Stiglitz results to a dynamic, overlapping generations model, correcting a misreading of the result on the desirability of a zero capital tax. That result required separability of consumption from labor and that the only unobservable differences among individuals were in (fixed) labor productivities. In a general equilibrium model, one needs to take into account the effects of policy changes on binding self-selection constraints. In a simple model with time separability but with non-separability between consumption and leisure, capital taxation depends on the complementarity/substitutability of leisure during work with retirement consumption.

The final section constructs a simple two class model, capitalists who maximize dynastic welfare and workers who save for retirement, whose productivity can be enhanced by (publicly provided) education. It derives a simple expression for the optimal capital tax, which is positive, so long as the social welfare function is sufficiently equalitarian and the productivity of educational expenditures is sufficiently high.

## 1. Introduction

Beginning with my supervision of Tony Atkinson in Cambridge in 1965–1966 while I was a junior research fellow at Gonville and Caius College, Tony and I enjoyed a close collaboration and friendship. One of our early results that received a great deal of attention was that when there was separability in the utility function between consumption and leisure,[1] if there existed an optimal income tax, it was optimal to have no commodity taxation (Atkinson and Stiglitz, 1976). An immediate corollary of that result was that, under the stipulated conditions, there should be no tax on interest income—treating consumption at different dates as different commodities. This and similar results from optimal tax theory were seized upon as a basis of policy by those critical of capital taxation.[2] Those who did so typically did not understand (or did

not want to understand) the limitations of the model. As always, one has to look carefully at the assumptions going into a model to judge whether they provide an appropriate basis for policy.

This paper argues that *even within the confines of a model in which differences in labor productivity are the only source of differences in income, the conclusion that there should be no capital taxation is in general wrong, even in the presence of separability, and that in a plausible model where individuals differ as well in the amount of inherited capital, there is a presumption for a possibly quite high tax on income from capital.*

This paper is divided into eight parts. The first presents some general reflections on our 1976 paper, our motivations in writing it, what we saw as some of the limitations, and some of the important extensions of the result in the subsequent literature. The second presents the basic model, based on Stiglitz (1982a), including some straightforward but important

[1] That is, utility U could be written U = U(u(**C**), L) where **C** is the vector of consumption goods and L is labor supply. We do not require additive separability.

[2] See, for instance, Mankiw et al. (2009) and Atkeson et al. (1999). Using a quite different framework, with individuals with infinite lives, Chamley (1986) has also concluded that the optimal tax rate on capital should be zero in the long run. In Section 3, we consider a social welfare maximization problem extending infinitely far into the future, but using an overlapping generations model, and show that his results obtains if and only if there is separability in the utility function between labor and consumption. Later sections present still more general frameworks, showing that even with separability, results on asymptotic zero capital taxation are not general.

extensions to Pigouvian taxation and the provision of public goods. In both cases, in the case of separability, results obtain that are much more akin to those that arise when government can impose out-of-lump-sum taxes than to those generated when government relies on distortionary Ramsey results. For instance, the sum of the marginal rate of substitution equals the marginal rate of transformation. In the standard Ramsey-Diamond-Mirrlees model with public goods with distortionary taxation, the Samuelson (1954, 1958) condition for optimal supply—that the marginal rate of transformation (MRT) equals the sum of the marginal rates of substitution (MRS)—is altered, and instead MRT < Σ MRS, the marginal rate of transformation is less than the sum of the sum of the marginal rate of substitution; and the level of consumption of public goods is reduced.[3] Section 3 shows that the Samuelsonian condition is restored with optimal non-linear taxation, provided there is separability between labor and consumption goods.[4] Similarly, first-best Pigouvian taxation is optimal with separability either between leisure and consumption, or between the externality on the one hand and consumption and labor on the other, though in the latter case, the Pigouvian tax has to be interpreted as the tax in excess of the optimal commodity tax (the shadow price).[5]

To analyze optimal interest taxation in general equilibrium, once has to construct a dynamic general equilibrium model. Section 4 does this, embedding optimal non-linear taxation in a general equilibrium overlapping generations model. While Atkinson and Sandmo (1980) had extended the Atkinson and Stiglitz (1972) model to a similar overlapping generations setting, this paper, a development of Stiglitz (1985a, 1985b), provides the corresponding extension of Atkinson and Stiglitz (1976). The conclusion of Atkinson and Sandmo (1980) was that "it is difficult to make a strong case *either* for the expenditure tax *or* for taxing interest income at the same rate as wage income." By contrast, we find that the basic results of our static model extend to this dynamic framework: (a) With separability, and fixed relative productivities, there should be no taxation of capital income *provided that relative wages are either observable or unaffected by taxation or investment (by the consumption/labor supplies which they induce)*; but (b) in the absence of separability between consumption and leisure, but with intertemporal separability, whether there should be an interest income tax or subsidy depends on whether first-period consumption is a complement or substitute for leisure.[6] But both separability and the italicized proviso are crucial. While the latter result is reminiscent of that of Corlett and Hague (1953) (in a model with no income taxation and just commodity taxation), we establish it in the presence of an optimal non-linear income tax. (This is important, because in many cases, the presence of one tax—in particular an income tax—markedly changes the taxes that should be imposed elsewhere in the system. For reasons we explain, with unconstrained non-linear taxes, results that hold only with lump-sum taxes are often restored.) As a prelude to the later discussion of public investment in the next section, the section also extends the analysis to show that the standard result on the desirability of productive efficiency holds.

The rest of the paper then provides a critique of the Atkinson and Stiglitz (1976) result. The fifth section focuses on the case where relative wages are not fixed. Even with separability between labor and consumption, (a) it is in general desirable to impose distortionary commodity taxation, even on high-ability individuals; (b) there should be a distortionary income tax, even on high-ability individuals; (c) there should be an interest income tax (or subsidy), even on high-ability individuals; (d) productive efficiency of the economy (where there are no production taxes and the marginal rate of transformation of all goods is the same among all firms) is not desirable[7]; (e) even asymptotically, the rate of return on public capital goods may not equal either the pure rate of time discount or the rate of return on private goods; the long-run return to some kinds of public capital may be lower than the social discount rate.

Recent policy discussions have suggested the desirability of focusing on policies which improve the market distribution of income, leaving less of a burden on *redistribution*.[8] Standard models which assume that relative wages are fixed cannot address this issue. Our analysis establishes that this intuition is correct. In each of the cases listed above, taxes and expenditures are adjusted at the margin to increase wages of the low skilled relative to the highly skilled.[9] Formally, the improvement in the distribution of income loosens the binding self-selection constraint, allowing for a Pareto improvement.

The optimal tax literature developed in the context of an economy of seemingly self-employed individuals, say making widgets. The only information problem was that of the government, which, seeing the output of widgets (or income from producing widgets) could not tell whether a high output was the result of hard work (long hours) or high ability. But most individuals in a modern economy work for others, and employers have an analogous information problem to that of government: they seek to discriminate among individuals, including by using self-selection mechanisms. But private screening is affected by public tax policy, and optimal taxation needs to take this into account. Section 6 does this, again showing that the Atkinson and Stiglitz (1976) result may not hold, even with separability.

Sections 7 and 8 consider optimal taxation with inheritances. Inheritances mean that individuals differ not just in their labor productivity, and optimal taxation ought to take that into account. This is so even in a world with dynastic families optimally sharing their wealth with their children. We show that the appropriate public policy response is taxation of inheritances and the return to capital. But the design of capital taxation is more nuanced than populist calls might suggest: careful attention has to be paid to incidence. A simplistic call for capital taxation with proceeds distributed to workers may actually leave workers worse off. In a simple two class model (workers and capitalists), we derive a simple formula for optimal taxation when the proceeds are invested in human capital or in productive capital investments. The real argument behind the taxation of capital is almost surely related to disparities in inherited capital and in the ability to obtain returns out of capital—and in luck—than with differences in productivities among workers, the subject of our 1976 paper.

Section 9 provides some concluding remarks, including putting the results presented here into some broader perspectives.

## 2. Reflections on the Atkinson-Stiglitz theorem

Our analysis was motivated in part as a critique of Ramsey taxation (Ramsey, 1927), which seemed to justify not just high taxes on basic necessities like food (which had a low elasticity of demand) but also patterns of pricing by monopolies like AT&T.[10] Ramsey had established that optimal commodity taxes should be inversely related to the

---

[3] See Pigou (1947), Dasgupta and Stiglitz (1971) and Atkinson and Stern (1974). Atkinson and Stern make the important point that one cannot infer from the fact that the sum of the MRT < Σ MRS the impact of distortionary taxation on the supply of public goods.

[4] Boadway and Keen (1993), using the Stiglitz (1982b) framework for the analysis of optimal taxation, had noted this result earlier. See also Kaplow (1996).

[5] Kaplow (2006a) had noted the former result. Sandmo (1975) had established that one could obtain results that were closely akin to first best Pigouvian taxes even in a Ramsey-Diamond-Mirrlees model.

[6] Similar in spirit to the Corlett and Hague (1953) analysis.

[7] One of the important results in Diamond-Mirrlees (1971) was establishing that optimal taxation entailed only consumption taxes, no production taxes, and that the economy was productively efficient. Dasgupta and Stiglitz (1971, 1972) had established that this was not so in the standard Diamond-Mirrlees model, but with restricted taxation (i.e. the government could not impose 100% taxes on pure rents or profits and could impose taxes on all goods and services at differential rates). The result here generalizes that of Naito (1999), whose analysis built on Stiglitz (1982a)

[8] See Hacker and Pierson (2011) and Stiglitz et al. (2015).

[9] In Stiglitz (1982a), I showed that this was true for taxes on labor income, but did not explore the full range of issues examined here.

[10] See, e.g. Baumol and Bradford (1970) and Boiteux (1956). Elsewhere, David Sappington and Stiglitz (1987) explained that while there were some similarities between the two problems, there were also some critical differences.