Manuscript received: 18.03.2022
Accepted:

International Journal of Economics, Management and Tourism
Vol. 2, No. 1, pp. 7-29
Online: ISSN **2671-3810**
UDC: 336.226.1.01
DOI https://doi.org/10.46763/IJEMT2221007j
Original scientific paper

PARETO EFFICIENT INCOME TAX SCHEDULES AND NUMERICAL SOLUTIONS TO MIRRLEES' OPTIMAL INCOME TAX MODEL: A CRITICAL SURVEY

Dushko Josheski [1*], Tatjana Boshkov [2] Mico Apostolov [3]

[1] *Associate professor Goce Delcev University of Stip, Faculty of tourism and business logistics, R.North Macedonia, dusko.josevski@ugd.edu.mk*

[2] *Associate professor, Goce Delcev University of Stip, Faculty of tourism and business logistics, R.North Macedonia, tatjana.dzaleva@ugd.edu.mk*

[3] *Professor, Goce Delcev University of Stip, Faculty of tourism and business logistics, R.North Macedonia, mico.apostolov@ugd.edu.mk*

* *Corresponding author: dusko.josevski@ugd.edu.mk*

Abstract

This paper aim is to study the Pareto efficient income taxation schedules in Mirrlees' optimal tax model. These Pareto efficient income tax schedules reveal the famous zero-at- the- top results. Namely that marginal tax rates should be low where probability density function of the distribution of income falls rapidly. Propositions for optimal Pareto income taxation are: There should be no distortionary taxation on the individual with the highest ability while the labor supply of the less able is distorted, there should be no commodity taxation on either high or low ability individuals if leisure and consumption are separable. Atkinson-Stiglitz theorem provides different result than Ramsey (1927) about the commodity taxation but also about optimal taxes on labor income change when in presence of commodity taxes.

Keywords: *Pareto efficient taxation, Mirrlees tax model, Optimal taxation*
JEL: H21

Introduction

Economist ever since 19th century were making first attempts to show that progressive taxation can be justified on more fundamental principles, see Stiglitz (1987). Edgeworth (1868) and Edgeworth (1897) tried to show that utilitarianism implied progressivity[1].New welfare economics from 1930's onwards limited economists in their role in identifying Pareto efficient allocations, and in finding Pareto inefficiencies and show they could be eliminated. But later in 1970's re-examined the tax structures by making use of the utilitarian social welfare functions, but also broadening their analysis to investigate the consequences of a wider class of social welfare functions (Rawlsian and with Pareto weights). The major advances on earlier work done by Edgeworth came from Mirrlees (1971). This work recognized the importance of incentives associated with taxation, namely: the presence of trade-offs between equity and efficiency considerations, that have been long recognized and there had been some formal modeling, see Fair (1971).Mirrlees calculations had provided support for the advocates of progressivity. Or as Mirrlees said :"I must confess that I had expected the rigorous analysis of income-taxation in the utilitarian manner to provide an argument for high tax rates," Professor Mirrlees wrote. "It has not done so." .The points made by Mirrlees include: Linear tax schedule is desirable, except supply of highly educated labor is much more inelastic from the utility function, and especially negative income tax is recommended for the workers that earn lower than some level, Income taxation is of no use when battling inequality, Some complementary taxes for the income tax will be of use here…such as taxes that depend on the time spent at

[1] Since all individuals have same utility of income function, and they exhibited diminishing marginal utility, and because social welfare is the sum of all individual utilities, it immediately followed that a decrease in utility from taking a dollar away from a rich person was less than the decrease in social welfare from taking away dollar from poor person.

Manuscript received: 18.03.2022
Accepted:
International Journal of Economics, Management and Tourism
Vol. 2, No. 1, pp. 7-29
Online: ISSN **2671-3810**
UDC: 336.226.1.01
DOI https://doi.org/10.46763/IJEMT2221007j
Original scientific paper

work and workers ability and the income from such labor. The problem lies here as Mirrlees wrote:" but if it is true, as our results suggest, that the income tax is not a very satisfactory alternative, this objection must be weighed against the great desirability of finding some effective method of offsetting the unmerited favors that some of us receive from our genes and family advantages". Mirrlees (1971) , in the basic version of the model allowed individuals to differ in their innate ability. The planer can observe income, but the planner cannot observe ability or effort. By recognizing unobserved heterogeneity, diminishing marginal utility of consumption, and incentive effects, the Mirrlees approach formalizes the classical tradeoff between efficiency and equity. In this framework the optimal tax problem is a problem of imperfect information between taxpayers and the social planner.So, the Pareto efficient income tax structures *i.e* the tax structures "which get the economy to the utilities possibilities schedule, given the limitations of government's information and other limitations of government's ability to impose taxes. The use of different social welfare functions can provide systematic way of thinking of the trade offs between efficiency and redistribution[2]. When individuals' income generating abilities are constant over time, Pareto efficient income taxation is the same as in static two-type (high and low ability) Mirrlees model. Efficient tax systems are stationary, low ability individuals face a positive marginal tax rate and high ability individuals face zero marginal tax rate, so their labor supply is undistorted. However, when ability types are not correlated over time efficient tax systems are non-stationary and the individuals whose labor supply is distorted are those who have been low ability, see Battaglini,Coate (2003).This paper also follows the Mirrlees (1971) mechanism design approach and models of information asymmetries that preclude non-distortionary taxation, see also Albanesi, Sleet (2003), Brito et al (1991), Diamond and Mirrlees (1978), Golosov, Kocherlakota and Tsyvinsky (2003), Golosov and Tsyvinski (2003) and Werning (2002). Mirrlees (1986), elaborates that a good way of governing is to agree upon objectives, then to discover what is possible and to optimize. The central element of the theory of optimal taxation is information. Public policies apply to the individuals on the basis of what the government knows about them. Second welfare theorem[3] states, that where a number of convexity and continuity assumptions are satisfied, an optimum is a competitive equilibrium once initial endowments have been suitably distributed. In general, complete information about the consumers for the transfers is required to make the distribution requires, so the question of feasible lump-sum transfers arises here. Saez (2001) argued that "unbounded distributions are of much more interest than bounded distributions to address high income optimal tax rate problem". In all of the cases that Saez (2001) investigated (four cases)[4] the optimal tax rates are clearly U-shaped. This paper by using the elasticity estimates from the literature, the formula for the asymptotic top rates suggests that the marginal rates for the labor income should not be lower than 50% and they could be as much as high as 80%. Usually the optimal tax systems combine flat marginal tax rate plus lump sum grants to all the individuals (so that the average tax rate rises with income even if the marginal does not), Mankiw NG, Weinzierl M, Yagan D.(2009). This paper is organized as follows: First it derives Pareto optimal income tax rates, second it reviews paper by Werning (2008), third it reviews results from Pareto efficient taxation by Stiglitz (2018),fourth it derives one version of Mirrlees 1971 model, and numerical solutions to Mirrlees original model are provided.

[2] "Most optimal tax models dealing with income redistribution assume that the government wants to redistribute from the well-off to the not so well-off, e.g., since low-income individuals are assumed to have higher marginal utility of consumption than high-income ones. We then often say that the government or the social planner is inequality averse "(see, Aronsson,T. Johansson-Stenman,O.(2015)).
[3] Second fundamental theorem is giving conditions under which a Pareto optimal allocation can be supported as a price equilibrium with lump-sum transfers, i.e. Pareto optimal allocation as a market equilibrium can be achieved by using appropriate scheme of wealth distribution (wealth transfers) scheme (Mas-Colell, Whinston et al. 1995)
[4] Utilitarian criterion, utility type I and II and Rawlsian criterion, utility type I and II.

Manuscript received: 18.03.2022
Accepted:

International Journal of Economics, Management and Tourism
Vol. 2, No. 1, pp. 7-29
Online: ISSN **2671-3810**
UDC: 336.226.1.01
DOI https://doi.org/10.46763/IJEMT2221007j
Original scientific paper

Pareto-optimal income taxation

Here we are going to assess the Pareto efficiency of a tax schedule. Here first assumption is that elasticity of labor supply is zero. Now, let ε_w^* represents the compensated elasticity of labor supply with respect to real wage. Let the distribution of income generated by the current tax system be Pareto:

equation 1

$$h(w) = k(w)^{-k-1}\underline{w}\,k \; for \; w \geq \underline{w} \; and \; k > 0$$

and now let's suppose that there is linear flat tax :

equation 2

$$t(w) = t + \tau(w)$$

Where τ represents marginal tax rate and intercept t.Here we assume that ε_w^* does not vary across individuals. This will be true in the case of this utility function[5]:

equation 3

$$u(c, w, \theta) = c - w\theta^{\alpha}$$

Now, starting from a general test for Pareto efficiency we will derive inequality for τ, ε_w^*, k. The starting point here is this inequality which states that marginal tax rate must be lower than 100% :

inequality 1

$$\frac{\tau(\theta)}{1-\tau(\theta)}\frac{\varepsilon_w^*}{\Phi}\left(-\frac{dlog\dfrac{\tau(\theta)}{1-\tau(\theta)}}{dlogw} - 1 - \frac{dlog\big(\varepsilon_w^*(w)\big)}{dlogw} - \frac{dlog(h^*(w)}{dlogw} - \frac{\partial MRS}{\partial c}w\right) \leq 1$$

The assumptions to use this inequality are as follows:

1. By quasi-linear utility preferences we have : $-\frac{\partial MRS}{\partial c} = 0$
2. A flat tax implies no convexity $t'' = 0$, a constant marginal tax rate $MTR = \tau(\theta) = \tau$ and also $\frac{dlog\frac{\tau(\theta)}{1-\tau(\theta)}}{dlogw} = 0$
3. Now, the logarithm of Pareto income density is given as:

equation 4

$$\log\big(h \cdot (w)\big) = \log k - (k+1)\log w + k\log \underline{w}$$

First of this log density with respect to income gives:

equation 5

$$\frac{d\log(h^*(w))}{d\log w} = \frac{d(\log k - (k+1)\log w + k\log \underline{w})}{d\log w} = \frac{-(k+1)d\log w}{d\log w} = -(k+1)$$

So the first inequality in this part $\frac{\tau(\theta)}{1-\tau(\theta)}\frac{\varepsilon_w^*}{\Phi}\left(-\frac{dlog\frac{\tau(\theta)}{1-\tau(\theta)}}{dlogw} - 1 - \frac{dlog(\varepsilon_w^*(w))}{dlogw} - \frac{dlog(h^*(w)}{dlogw} - \frac{\partial MRS}{\partial c}w\right) \leq 1$ would become:

[5] θ represents every individual's characteristics

Manuscript received: 18.03.2022
Accepted:
International Journal of Economics, Management and Tourism
Vol. 2, No. 1, pp. 7-29
Online: ISSN **2671-3810**
UDC: 336.226.1.01
DOI https://doi.org/10.46763/IJEMT2221007j
Original scientific paper

equation 6

$$\frac{\tau(\theta)}{1-\tau(\theta)}\varepsilon_w^* k \leq 1$$

The parameter k has been estimated by Saez (2001) to be of value 1.6[6]. The thicker the tail of the distribution, the smaller is α. Pareto distribution is given as PDF lower CDF and upper CDF [7].PDF (probability density function) :

equation 7

$$f(x, x_m, \alpha) = \frac{\alpha x_m^\alpha}{x^{\alpha+1}}$$

Lower cumulative distribution function (lower CDF):

equation 8

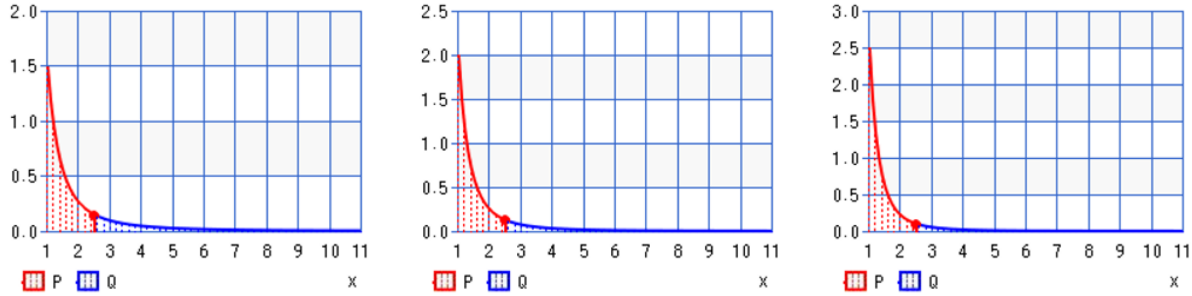
$$P(x, x_m, \alpha) = \int_{x_m}^{x} f(x, x_m, \alpha)dx = 1 - \left(\frac{x^m}{x}\right)^\alpha$$

Upper cumulative distribution function (upper CDF):

equation 9

$$Q(x, x_m, \alpha) = \int_{x}^{\infty} (x, x_m, \alpha)dx = \left(\frac{x_m}{x}\right)^\alpha$$

Figure 1 Pareto distribution function with shape parameter $\alpha \in (1.5,2,2.5)$



Source: Author's calculation

Table 1 Pareto distribution values

	$\alpha = 1.5$	$\alpha = 2$	$\alpha = 2.5$
Percentile x		2.5	
Scale parameter x_m		1	
Shape parameter α	1.5	2	2.5

Source: Author's calculation

Table 2 Pareto distribution probability density, lower CDF ,upper CDF

	$\alpha = 1.5$	$\alpha = 2$	$\alpha = 2.5$
pareto distribution		value	
probability density f	0.15	0.128	0.10
lower cumulative P	0.75	0.84	0.89
upper cumulative Q	0.26	0.16	0.10

Source: Author's calculation

[6] This value is approx..for US incomes above 0.3 m.

[7] This part is for readers that are not familiar with basic statistics

Manuscript received: 18.03.2022
Accepted:

International Journal of Economics, Management and Tourism
Vol. 2, No. 1, pp. 7-29
Online: ISSN **2671-3810**
UDC: 336.226.1.01
DOI https://doi.org/10.46763/IJEMT2221007j
Original scientific paper

Note that the Pareto distribution has unbounded variance for $a < 2$ and that several studies suggest that this parameter might be about 2 times higher in several European countries.

4. The compensated elasticity of labor supply with respect to real wage ε_w^* has been estimate approximately to be 0.5 see Gruber, Saez (2002).

5. So that $\frac{1}{\varepsilon_w^*} \in \left[\frac{1}{6}; \frac{10}{3}\right]$ or $\frac{1}{2*3} = \frac{1}{6}$ and $\frac{1}{0.2*1.5} = \frac{10}{3}$ which lies around a central value of $\frac{1}{0.5 \cdot 2.5} = 0.8$

And the second inequality from above now would become:

inequality 2

$$\frac{\tau(\theta)}{1-\tau(\theta)} \leq 0.8$$

Gruber, Saez (2002) estimate that for the US taxpayer with incomes above 100K$ have elasticity around 0.57. And those <100K$ have elasticity around 0.2 or even less. Then the inequality will be affected in two ways:

1. $\varepsilon_w^*(\theta)$ will be higher for higher incomes
2. $\frac{dlog(\varepsilon_w^*(w))}{dlogw} > 0$

The inequality then becomes:

Relative to the average-income constant elasticity benchmark case the upper bound on the marginal tax ratio $\frac{1}{\varepsilon_w^*\left(k - \frac{dlog(\varepsilon_w^*(w))}{dlogw}\right)}$ is affected as follows for high and low earners:

- For high earners :
  1. is directly negative affected by the factor $\frac{1}{\varepsilon_w^*}$
  2. is positively affected by the factor $\frac{dlog(\varepsilon_w^*(w))}{dlogw}$
- for low earners:
  1. is directly positively affected by the factor $\frac{1}{\varepsilon_w^*}$
  2. is positively affected by the factor $\frac{dlog(\varepsilon_w^*(w))}{dlogw}$

Thus, in order to pass the efficiency test:

- ✓ a higher maximal marginal tax rate for low-income earners is acceptable.
- ✓ the effect on the maximal tax rate for high-income earners is theoretically ambiguous even if I suspect the direct negative effect to dominate because locally the logarithm of elasticity is relative stable compared to the parameter k and hence a lower maximal marginal tax rate for high- income earners is acceptable. This is very intuitive: if low-income earners are less elastic, we can tax them relative more.

Now, let's see how progressivity would affects tax schedule in question here. Convexity implies that $\tau'' > 0$. To keep things simple, we continue to assume that there is:

- quasi-linearity of preferences: $-\frac{\partial MRS}{\partial c} w = 0$
- a constant compensated elasticity of labor supply with respect to the real wage: $\frac{dlog(\varepsilon_w^*(w))}{dlogw} = 0$.Then the inequality becomes:

Manuscript received: 18.03.2022
Accepted:

International Journal of Economics, Management and Tourism
Vol. 2, No. 1, pp. 7-29
Online: ISSN **2671-3810**
UDC: 336.226.1.01
DOI https://doi.org/10.46763/IJEMT2221007j
Original scientific paper

inequality 3

$$\frac{\tau(\theta)}{1-\tau(\theta)}\frac{\varepsilon_w^*}{\Phi}\left(-\frac{dlog\frac{\tau(\theta)}{1-\tau(\theta)}}{dlogw}-1-\frac{dlog(h^*(w)}{dlogw}\right)\leq 1$$

Given the convexity we also have because $\tau'(w) < 1$:

inequality 4

$$\Phi(w) = 1 + we_w^*(w)\frac{\tau''(w)}{1-\tau'(w)} > 1$$

Now we have that:
equation 10

$$\frac{dlog(h^*(w)}{dlogw} = \frac{dlog(h^*(w)\Phi(\mathrm{w})^{-1}}{dlogw}$$

We know from previously that $\Phi(\mathrm{w})$ increases in w or equivalently that $\Phi(\mathrm{w})^{-1}$ decreases in w and thus that the absolute value of the slope of the virtual density is higher than the real density $-\frac{dlog(h^*(w)}{dlogw} > -\frac{dlog(h(w)}{dlogw}$. Compared to the fat tax rate, the upper bound on the marginal tax ratio:
equation 11

$$\tau'(\theta) = \frac{1}{\frac{\varepsilon_w^*}{\Phi}\left(-\frac{dlog\frac{\tau(\theta)}{1-\tau(\theta)}}{dlogw}-1-\frac{dlog(h^*(w)}{dlogw}\right)}$$

And it is affected in three ways:

- positively by Φ
- positively by $-\frac{dlog\frac{\tau(\theta)}{1-\tau(\theta)}}{dlogw} < 0$
- negatively by the distinction between the virtual and the real density

We expect the positive effect to dominate and thus the upper bound on the marginal tax could then be higher. Werning (2008) proposed Pareto efficient income taxation with dual optimization problem in the original Mirrlees (1971) framework. Namely this model starts from the Mirrleesian framework with additively separable preferences like this:

equation 12

$$u(c, y, \theta) = u(c) - \theta h(y)$$

Where θ denotes heterogenous disutility from producing output y. Cardinality of preferences[8] is irrelevant and only ordinal preferences matter[9]. The expenditure function $e(v, y, \theta)$ is inverse from $u(\cdot, y, \theta)$,and $F(\theta)$ represents the distribution of θ in the population, and its PDF can be represented as $f(\theta)$. Some tax function is $t(y)$ and workers' utility $v(\theta)$ is maximized:
equation 13

$$v(\theta) \equiv \max_y u(y - t(y), y, \theta)$$

[8] In economics, a cardinal utility function or scale is a utility index that preserves preference orderings uniquely up to positive affine transformations, see Ellsberg (1954)
[9] In economics, an ordinal utility function is a function representing the preferences of an agent on an ordinal scale. Ordinal utility theory claims that it is only meaningful to ask which option is better than the other, but it is meaningless to ask how much better it is or how good

Manuscript received: 18.03.2022
Accepted:
International Journal of Economics, Management and Tourism
Vol. 2, No. 1, pp. 7-29
Online: ISSN **2671-3810**
UDC: 336.226.1.01
DOI https://doi.org/10.46763/IJEMT2221007j
Original scientific paper

$c(\theta) = e(v(\theta), y(\theta), \theta)$ is a consumption function dependent on workers' characteristics, $y(t) = y - t(y)$ and an allocation is resource feasible if :

inequality 5

$$\int \big(y(\theta) - c(\theta)\big) dF(\theta) + e \geq 0$$

Here e is an endowment. The allocation generated by some tax schedule is (constrained) Pareto efficient if there is no other tax schedule that induces a resource feasible allocation where nobody is worse off, and some workers are strictly better off. The marginal tax rate is :

equation 14

$$\tau(\theta) = t'\big(y(\theta)\big) = 1 + \frac{u_y(c(\theta), y(\theta), \theta)}{u_c(c(\theta), y(\theta), \theta)} = 1 - \frac{\theta h'\big(y(\theta)\big)}{u'^{(c(\theta))}} = 1 - e_y(v(\theta), y(\theta), \theta)$$

There is above mentioned dual problem for the social planer.Here is introduced Pareto planning problem and sufficient and necessary conditions for optimality of the solution to the planers' problem[10].

equation 15

$$\max_{\tilde{y},\tilde{v}} \int (\tilde{y}(\theta), -e(\tilde{v}(\theta), \theta) dF(\theta) \text{ s.t. } \tilde{v}(\theta) = \tilde{v}\big(\tilde{\theta}\ \big) - \int_{\theta}^{\bar{\theta}} u_\theta(e(\tilde{v}(z), \tilde{y}(z), z)\tilde{y}(z), z\ )dz$$

In previous $\tilde{y}(\theta)$ is non-increasing $\tilde{v}\big(\tilde{\theta}\ \big) \geq v(\theta)$. The objective is to maximize aggregate net resources, output minus consumption. FOC necessary to be verified in order allocation to be Pareto efficient. Lagrangian for the FOC's is:

equation 16

$$\mathcal{L} = \int (\tilde{y}(\theta), -e(\tilde{v}(\theta), \theta) dF(\theta) + \int \left(\tilde{v}(\theta) - \tilde{v}(\bar{\theta}) + \int_{\theta}^{\bar{\theta}} u_\theta(e(\tilde{v}(z), \tilde{y}(z), z)\tilde{y}(z), z\ )dz \right) d\mu(\theta)$$

Integrating second term by parts we have:

equation 17

$$\mathcal{L} = \int (\tilde{y}(\theta), -e(\tilde{v}(\theta), \theta) dF(\theta) - \tilde{v}(\bar{\theta})\mu(\bar{\theta}) + \mu(\underline{\theta})\tilde{v}\big(\underline{\theta}\big) + \int \tilde{v}(\theta) d\mu + \int \mu(\theta) u_\theta\ (\tilde{v}(\theta), \tilde{y}(\theta), \theta) d\theta$$

About the efficiency conditions, the FOC for $\tilde{y}(\theta)$ evaluated at $\big(y(\theta), v(\theta)\big)$ gives:

equation 18

$$(1 - e_y(v(\theta), y(\theta), \theta) f(\theta) = -\mu(U_{\theta_c}(e(v(\theta), y(\theta), \theta) e_v(v(\theta), y(\theta), \theta) + u_{\theta_y}\big(e(v(\theta), y(\theta), \theta)\big)$$

Implying

$$\mu(\theta) = \tau(\theta) \frac{f(\theta)}{h'\big(y(\theta)\big)}$$

The FOC for $v(\bar{\theta})$ is $\mu(\bar{\theta}) \geq 0$, or if θ is bounded away from zero the FOC for $v\big(\underline{\theta}\big)$ gives $\mu(\bar{\theta}) \leq 0$. And so : $\tau(\bar{\theta}) \geq 0$ and $\bar{\tau}\big(\underline{\theta}\big) \leq 0$. For interior θ, the FOC with respect to $\tilde{v}(\theta)$ evaluated at $\big(y(\theta), v(\theta)\big)$ gives:

[10] A Pareto improvement would always be possible: if another allocation provided the same utility but increased net resources, then these resources can be used to construct another allocation that increases utility for some workers and is resource feasible.

Manuscript received: 18.03.2022
Accepted:

International Journal of Economics, Management and Tourism
Vol. 2, No. 1, pp. 7-29
Online: ISSN **2671-3810**
UDC: 336.226.1.01
DOI https://doi.org/10.46763/IJEMT2221007j
Original scientific paper

equation 19

$$\dot{\mu}(\theta) = \leq e_v(v(\theta), y(\theta), \theta) f(\theta)$$

By differentiation equation gives:
equation 20

$$\dot{\mu}(\theta) = \mu(\theta)\left(\frac{\tau'(\theta)}{\tau(\theta)} + \frac{f'(\theta)}{f(\theta)} - \frac{h''\big(y(\theta)\big)}{h'\big(y(\theta)\big)} y'(\theta)\right)$$

Substituting $\mu(\theta) = \tau(\theta)\frac{f(\theta)}{h'(y(\theta))}$ and $\dot{\mu}(\theta) = \mu(\theta)\left(\frac{\tau'(\theta)}{\tau(\theta)} + \frac{f'(\theta)}{f(\theta)} - \frac{h''(y(\theta))}{h'(y(\theta))} y'(\theta)\right)$ into the $\dot{\mu}(\theta) = \leq e_v(v(\theta), y(\theta), \theta) f(\theta)$ we get :
inequality 6

$$\tau(\theta)\left(\frac{d\log\tau(\theta)}{d\log(\theta)} + \frac{d\log f(\theta)}{d\log\theta} - \frac{d\log h'\big(y(\theta)\big)}{d\log\theta}\right) \leq 1 - \tau(\theta)$$

The integral form of this efficiency condition is given as:
equation 21

$$\frac{\tau'(\theta) f(\theta)}{h' y(\theta)} + \int_{\theta}^{\overline{\theta}} \frac{1}{u'\left(c(\tilde{\theta})\right)} f\big(\tilde{\theta}\big) d\tilde{\theta} \leq 0$$

Proposition 1 : Given the utility function $u(c, y, \theta)$ and a density of skills $f(\theta)$, a differentiable tax function $t(y)$ inducing an allocation .$\big(c(\theta), y(\theta)\big)$ is *Pareto efficient* if and only if condition$\frac{\tau'(\theta) f(\theta)}{h' y(\theta)} + \int_{\theta}^{\overline{\theta}} \frac{1}{u'\left(c(\tilde{\theta})\right)} f\big(\tilde{\theta}\big) d\tilde{\theta} \leq 0$ holds, where $\tau(\theta) = t'\big(y(\theta)\big)$.

Proof: Now, we define $\tilde{h}\,(\theta) = h\big(\tilde{y}(\theta)\big)$ and we will write the planning problem as:
equation 22

$$\max_{\underline{\tilde{v}}, \tilde{h}} \int \left(h^{-1}\left(\tilde{h}(\theta)\right) - u^{-1}\left(\underline{\tilde{v}} - \int_{\underline{\theta}}^{\theta} \tilde{h}(z) dz + \theta\tilde{h}(\theta)\right)\right) dF(\theta)$$

Subject to :

inequality 7

$$\underline{\tilde{v}} - \int_{\underline{\theta}}^{\theta} \tilde{h}(z) dz - v(\theta) \geq 0$$

And $\tilde{h}(\theta) \in ni(\Theta)$, where $ni(\Theta)$ is the set of non-increasing real-valued functions over Θ .This is a convex optimization problem the objective to be maximized is concave and the constraints are linear (convex). Now, $ni(\Theta)$ is a closed convex cone, closed under multiplication by positive scalars in the linear space of bounded functions $\mathcal{B}(\Theta)$ endowed with the supremum norm. Previous constraint $\underline{\tilde{v}} - \int_{\underline{\theta}}^{\theta} \tilde{h}(z) dz - v(\theta) \geq 0$ can be expressed as: $G(\tilde{h}) \in P$, where the mapping $G: ni(\Theta) \to c(\Theta)$ is convex, and P is the positive cone of the $c(\Theta)$.Previous constraint $\underline{\tilde{v}} - \int_{\underline{\theta}}^{\theta} \tilde{h}(z) dz - v(\theta) \geq 0$ allows for an interior point $\forall \underline{\tilde{v}} > v(\underline{\theta}); \tilde{h}(\theta) = h(\theta) = h\big(\tilde{y}(\theta)\big)$. All the conditions required in Luenberger (1969) are met and maximizing Lagrangian is sufficient and necessary for optimality. The Lagrangian here is:

Manuscript received: 18.03.2022
Accepted:

International Journal of Economics, Management and Tourism
Vol. 2, No. 1, pp. 7-29
Online: ISSN **2671-3810**
UDC: 336.226.1.01
DOI https://doi.org/10.46763/IJEMT2221007j
Original scientific paper

equation 23

$$\mathcal{L} = \int \left(h^{-1}(\tilde{h}(\theta) - u^{-1}\left(\underline{v} - \int_{\underline{\theta}}^{\theta} h(z)dz + \theta\tilde{h}(\theta)\right)\right) dF(\theta) + \int \left(\underline{v} - \int_{\underline{\theta}}^{\theta} \tilde{h}(z)dz - v(\theta)\right) d\lambda(\theta)$$

For some nondecreasing function $\lambda(\theta)$, the multiplier on the inequality $\underline{\tilde{v}} - \int_{\underline{\theta}}^{\theta} \tilde{h}(z)dz - v(\theta) \geq 0$, is normalized so that $\lambda(\bar{\theta}) = 0$. Fréchet derivative[11] is given by the following:

equation 24

$$\partial\mathcal{L}(h; \Delta_{\underline{v}}; \Delta_{\tilde{h}}) = \int \left((h^{-1})'\left(h(\theta)\right)\Delta_{\tilde{h}}(\theta) - (u^{-1})'\left(u(\theta)\right)\left(\Delta_{\underline{v}}(\theta) + \theta\Delta_{\tilde{h}}(\theta)\right)\right) + \int \Delta_{\tilde{v}}(\theta)d\lambda(\theta)$$

Where in previous:

equation 25

$$\Delta_{\tilde{v}}(\theta) = \Delta_{\underline{v}} - \int_{\underline{\theta}}^{\theta} \Delta_{\tilde{h}}(z)dz$$

Where z is the function of earnings. Now by substituting for $\Delta_{\tilde{v}}(\theta)$ and by integration by parts we get:

equation 26

$$\begin{aligned}\partial\mathcal{L}(h; \Delta_{\underline{v}}; \Delta_{\tilde{h}}) &= \int \left((h^{-1})'\left(h(\theta)\right) - (u^{-1})'(u(\theta)\theta)f(\theta)\Delta_{\tilde{h}}(\theta)d\theta\right) \\ &+ \int \left(\int_{\underline{\theta}}^{\theta} (u^{-1})'\left(u(z)\right)f(z)dz\right)\Delta_{\tilde{h}}(\theta)d\theta \\ &+ \int \lambda\,(\theta)\Delta_{\tilde{h}}(\theta)d\theta - \Delta_{\underline{v}}\left(\lambda(\underline{\theta})\right) + \int (u^{-1})'(u(\theta)f(\theta)d\theta)\end{aligned}$$

By collecting the terms we get :

equation 27

$$\partial\mathcal{L}(h; \Delta_{\underline{v}}; \Delta_{\tilde{h}}) = \int \mathcal{A}(\theta)\Delta_{\tilde{h}}(\theta)d\theta = \Delta_{\tilde{h}}(\underline{\theta}) \int_{\underline{\theta}}^{\theta} A(z)dz + \int \int_{\underline{\theta}}^{\theta} A(z)dz\, d\Delta_{\tilde{h}}(\theta)$$

Where:

equation 28

$$\mathcal{A}(\theta) = \left(\left((h^{-1})'\left(h(\theta)\right) - (u^{-1})'(u(\theta)\theta)f(\theta)\right)\right) + \int_{\underline{\theta}}^{\theta} (u^{-1})'\left(u(z)\right)f(\theta)dz + \lambda(\theta)$$

$\mathcal{L}(h; \Delta_{\underline{v}}; \Delta_{\tilde{h}})$ is convex, and the necessary and sufficient conditions for $\tilde{h}(\theta) \in ni(\Theta)$ to be maximized are :

inequality 8

$$\partial\mathcal{L}(h; \Delta_{\underline{v}}; \Delta_{\tilde{h}}) \geq 0\,; \forall\Delta_{\tilde{h}} \in ni(\Theta); \partial\mathcal{L}(h; \underline{v}; h) = 0$$

[11] It is commonly used to generalize the derivative of a real-valued function of a single real variable to the case of a vector-valued function of multiple real variables, and to define the functional derivative used widely in the calculus of variations.

Manuscript received: 18.03.2022
Accepted:

International Journal of Economics, Management and Tourism
Vol. 2, No. 1, pp. 7-29
Online: ISSN **2671-3810**
UDC: 336.226.1.01
DOI https://doi.org/10.46763/IJEMT2221007j
Original scientific paper

Lemma 1. (optimality and FOC's to allow for Gateaux differentials[12] instead of Frechet derivatives Amador, Werning, and Angeletos (2006)) .Let f be a concave functional on P a convex cone in X.Take $x_0 \in P$ and define $h\,(x0) \equiv \{h : h = x - x0 \; and \; x \in P\}$.Then, $\exists \delta f\,(x_0, h)$ for $h \in h\,(x_0)$. Assume that $\exists \delta f\,(x_0, \alpha_1 h_1 + \alpha_2 h_2)$ for $h_1, h_2 \in h\,(x_0)$, and $\delta f\,(x_0, \alpha_1 h_1 + \alpha_2 h_2) = \alpha_1 \delta f\,(x_0, h_1) + \alpha_2 \delta f\,(x_0, h_2)$ for $\alpha_1, \alpha_2 \in R$. A necessary condition for $x_0 \in P$ to maximize f is that: $\delta f\,(x_0, x) \leq 0 \; \forall \, x \in P \; ; \delta f\,(x_0, x_0) = 0$. Thus, we obtain that a necessary and sufficient condition for the Lagrangian to be maximized at (u_0, w_0) over Φ and that is:

equation 29

$$\partial \mathcal{L}\left(\underline{w_0}; u_0; \underline{w_0}; u_0 \middle| \Lambda_0\right) = 0$$
$$\partial \mathcal{L}\left(\underline{w_0}; u_0; h_{\underline{w}}, h_u \middle| \Lambda_0\right) \leq 0 \; ; \forall \left(h_{\underline{w}}, h_u\right) \in \Phi$$

Since $\Delta_{\underline{v}} \lesseqgtr 0$ we obtain that:

equation 30

$$\lambda(\underline{\theta}) + \int (u^{-1})'(u(\theta)\theta) f(\theta) = 0$$

Because $\Delta_{\tilde{h}}(\theta) \lesseqgtr 0$ and $\Delta_{\tilde{h}} > 0$ it follows that we must have :

inequality 9

$\int_{\underline{\theta}}^{\theta} \mathcal{A}(z)dz = 0 \; ; \; \int_{\underline{\theta}}^{\theta} \mathcal{A}(z)dz \leq 0$

From $\partial \mathcal{L}(h; \underline{v}; h) = 0$, if the original $h(\theta) = h(y(\theta))$ strictly increasing near in neighborhood it follows that:

inequality 10

$$\int_{\underline{\theta}}^{\theta} \mathcal{A}(z)dz \leq 0 \Rightarrow \mathcal{A}(\theta) = 0$$

In addition we must have $\lambda(\theta)$, and by using the fact $h^{-1}(\tilde{h}(\theta) - (u^{-1})' \left(\underline{v} - u(\theta)\right)\theta = \tau(\theta)/h'(y(\theta))$ and that $(u^{-1})'(u(\theta)) = e_v(v(\theta), y(\theta), \theta)$ we obtain that :

equation 31

$$-\lambda(\theta) = \frac{\tau(\theta) f(\theta)}{h'(y(\theta))} + \int_{\theta}^{\overline{\theta}} e_v(v(z), y(z), z) f(z)dz$$

And previous expression is decreasing, by differentiation of this expressions and setting $-\lambda'(\theta) \leq 0$ gives $\tau(\theta)\left(\frac{d \log \tau(\theta)}{d \log(\theta)} + \frac{d \log f(\theta)}{d \log \theta} - \frac{d \log h'(y(\theta))}{d \log \theta}\right) \leq 1 - \tau(\theta)$■.

Now, if $F(\theta(y)) = 1 - G(y)$ which implies that $\frac{d \log \theta(y)}{d \log y} = -\varepsilon_{\theta,y}$ and $\theta'(y) < 0$. Where $\varepsilon_{\theta,y}$ is the elasticity of $\theta(y)$ with respect to y .

equation 32

$$\varepsilon_{\theta,y} \equiv \left|\frac{y\theta'(y)}{\theta(y)}\right| = -\frac{d \log(1 - t'(y))}{d \log y} - \frac{d \log u'(y - t'(y))}{d \log y} + \frac{d \log h'(y)}{d \log y}$$

[12] Gateaux differential or Gateaux derivative is a generalization of the concept of directional derivative in differential calculus. Like the Fréchet derivative on a Banach space, the Gateaux differential is often used to formalize the functional derivative

Manuscript received: 18.03.2022
Accepted:
International Journal of Economics, Management and Tourism
Vol. 2, No. 1, pp. 7-29
Online: ISSN **2671-3810**
UDC: 336.226.1.01
DOI https://doi.org/10.46763/IJEMT2221007j
Original scientific paper

And $f\big(\theta(y)\big) = -\frac{g(y)}{\theta'(y)}$

$$-\frac{d\log f\big(\theta(y)\big)}{d\log\theta}\varepsilon_{\theta,y} = \frac{d\log g(y)}{d\log y} - \frac{d\log -\theta'(y)}{d\log y} + 1 - \varepsilon_{\theta,y} - \frac{d\log \varepsilon_{\theta,y}}{d\log y}$$

And multiplying $\tau(\theta)\left(\frac{d\log\tau(\theta)}{d\log(\theta)} + \frac{d\log f(\theta)}{d\log\theta} - \frac{d\log h'\big(y(\theta)\big)}{d\log\theta}\right) \le 1-\tau(\theta)$ by $\varepsilon_{\theta,y}$ and by substituting this last expression:

inequality 11

$$-\frac{d\log\big(1-t'(y)\big)}{d\log y} - \frac{t'(y)}{\big(1-t'(y)\big)}\left(\frac{d\log g(y)}{d\log y} - \frac{d\log h'\big(y(\theta)\big)}{d\log\theta} + 1 - \varepsilon_{\theta,y} - \frac{d\log\varepsilon_{\theta,y}}{d\log y}\right) \le \varepsilon_{\theta,y}$$

By rearrangement this gives:

inequality 12

$$t'(y)\left(-\frac{d\log\big(1-t'(y)\big)}{d\log y} + \frac{d\log h'\big(y(\theta)\big)}{d\log\theta} - \frac{d\log g(y)}{d\log y} - 1 + \frac{d\log\varepsilon_{\theta,y}}{d\log y}\right) \le -2\frac{d\log\big(1-t'(y)\big)}{d\log y} + \frac{d\log h'\big(y(\theta)\big)}{d\log\theta} - \frac{d\log u'\big(y-t'(y)\big)}{d\log y}$$

Now one extension flat tax rate. We are assuming power utility function given as: $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$ and $h(y) = \alpha y^{\eta}$and we are supposing top tax rate:

equation 33

$$\bar{\tau} \equiv \lim_{\theta\to 0}\tau(\theta) = \lim_{y\to\infty} t'(y) < 1$$

equation 34

$\lim_{y\to\infty}\frac{d\log\big(1-t'(y)\big)}{d\log y} = 0\,;\ \frac{d\log\varepsilon_{\theta,y}}{d\log y} = 0$

For high income consumption becomes proportional to income:

equation 35

$\lim_{y\to\infty}\frac{d\log\big(1-t'(y)\big)}{d\log y} = -\sigma$ and $\lim_{y\to\infty}\frac{d\log h'\big(y(\theta)\big)}{d\log\theta} = \eta - 1$

Now by substituting these expressions in $t'(y)\left(-\frac{d\log\big(1-t'(y)\big)}{d\log y} + \frac{d\log h'\big(y(\theta)\big)}{d\log\theta} - \frac{d\log g(y)}{d\log y} - 1 + \frac{d\log\varepsilon_{\theta,y}}{d\log y}\right) \le -2\frac{d\log\big(1-t'(y)\big)}{d\log y} + \frac{d\log h'\big(y(\theta)\big)}{d\log\theta} - \frac{d\log u'\big(y-t'(y)\big)}{d\log y}$ gives:

inequality 13

$$\bar{\tau} \le \frac{\sigma+\eta-1}{\varphi+\eta-2}$$

Where $\varphi = -\lim_{y\to\infty}\frac{d\log g(y)}{d\log y}$, the value $\varphi - 1 > 0$ to ensure that income has finite mean, and it is called asymptotic Pareto distribution parameter. The Pareto distribution had a density that is a power function $g(y) = \mathcal{A}y^{-(\varphi)}$, so that these holds:

Manuscript received: 18.03.2022
Accepted:

International Journal of Economics, Management and Tourism
Vol. 2, No. 1, pp. 7-29
Online: ISSN **2671-3810**
UDC: 336.226.1.01
DOI https://doi.org/10.46763/IJEMT2221007j
Original scientific paper

equation 36

$$\frac{d\log g(y)}{d\log y} = -\varphi$$

In $\bar{\tau} \leq \frac{\sigma+\eta-1}{\varphi+\eta-2}$ if $\varphi \approx 3$ as per Saez (2001) , then $\sigma < 2$ and σ cannot be interpreted as risk aversion but as control variable[13] for controlling the income and substitution effects for labor.Now in a case of flat tax $t(y) = \bar{\tau}(y)$ for a flat tax rate following result is yielded:

equation 37

$$\tau(y) = \frac{\sigma+\eta-1}{-\frac{d\log g(y)}{d\log y} + +\eta - 2}$$

Now if we assume and transfers t_0 and that $t(y) = \bar{\tau}(y) - t_0$ where $t_0 > 0$ we get :

equation 38

$$-\frac{d\log u'\big(y-t(y)\big)}{d\log y} = -\sigma\frac{1-t'(y)}{1-\left(\frac{t(y)}{y}\right)} = \sigma\frac{1-\bar{\tau}}{1-\bar{\tau}+\frac{t_o}{y}} \leq \sigma$$

Which goes $-\frac{d\log u'(y-t(y))}{d\log y} \in (0,\sigma)$ for $y \in (0,\infty)$. So that $\frac{d\log \varepsilon_{\theta,y}}{d\log y} \geq 0$, aditionaly:

equation 39

$$\frac{dlog}{d\log y}\left(-\frac{d\log u'\big(y-t(y)\big)}{d\log y}\right) = \frac{dlog}{d\log y}\left(\frac{1-\bar{\tau}}{1-\bar{\tau}+\frac{t_0}{y}}\right) = \frac{\frac{t_0}{y}}{1-\bar{\tau}+t_0/y} \leq 1$$

Which implies that :

inequality 14

$$\frac{d\log \varepsilon_{\theta,y}}{d\log y} \leq \frac{\sigma}{\sigma+\eta-1}$$

And sufficient condition for $t'(y)\left(-\frac{d\log(1-t'(y))}{d\log y} + \frac{d\log h'(y(\theta))}{d\log\theta} - \frac{d\log g(y)}{d\log y} - 1 + \frac{d\log\varepsilon_{\theta,y}}{d\log y}\right) \leq -2\frac{d\log(1-t'(y))}{d\log y} + \frac{d\log h'(y(\theta))}{d\log\theta} - \frac{d\log u'(y-t'(y))}{d\log y}$ to hold is :

inequality 15

$$\bar{\tau} < \frac{\eta-1}{-\frac{d\log g(y)}{d\log y} + \eta - 2 + \frac{\sigma}{\sigma+\eta-1}} < \frac{\eta-1}{-\frac{d\log g(y)}{d\log y} + \eta - 1}$$

Pareto efficient taxation and expenditures: pre- and re-distribution

Now another model that we turn our attention not to models nut conclusion that are presented in Stiglitz (2018). This paper extends to some extent the findings in the original Atkinson-Stiglitz theorem.Atkinson and Stiglitz (1976) demonstrated the following theorem known as Atkinson, Stiglitz theorem[14]:

[13] A control variable (or scientific constant) in scientific experimentation is an experimental element which is constant and unchanged throughout the course of the investigation. Control variables could strongly influence experimental results, were they not held constant during the experiment in order to test the relative relationship of the dependent and independent variables. The control variables themselves are not of primary interest to the experimenter.

[14] Atkinson and Stiglitz(1972) had shown that in the absence of an income tax, optimal commodity taxes could be described by a simple Ramsey-like formula incorporating distributional effects, which suggested that when distributional concerns were given sufficient weight (for instance, in a society with a high level of both inequality and inequality aversion) goods like food with a low price elasticity of demand but a low income elasticity of demand

Manuscript received: 18.03.2022
Accepted:

International Journal of Economics, Management and Tourism
Vol. 2, No. 1, pp. 7-29
Online: ISSN **2671-3810**
UDC: 336.226.1.01
DOI https://doi.org/10.46763/IJEMT2221007j
Original scientific paper

Theorem: Commodity taxes cannot increase social welfare if utility functions are weakly separable in consumption goods versus leisure and the subutility of consumption goods is the same across individuals, i.e.,$u_i(c_1,\dots,c_k,w) = u_i(v(c_1,\dots,c_k),w)$ with the subutility function $v(c_1,\dots,c_k)$ homogenous across individuals.
Laroque (2005) and Kaplow (2006) have provided intuitive prof of this theorem as follows:
Proof: A tax system $(\tau(\cdot),t)$that includes both nonlinear income tax and a vector of commodity taxes can be replaces by a pure income tax $(\bar{\tau}(\cdot),t=0)$ .This tax system keeps all individual utilities constant and raises at least as much tax revenue. Let $v(p+t,\gamma) = \max_c v(c_1,..c_k)$ s.t. $(p+t)\cdot c \leq \gamma$ be the indirect utility of consumption goods which is common to al individuals. Now if we consider replacing $(\tau(\cdot),t)$ this tax system with $(\bar{\tau}(\cdot),t=0)$ where $\bar{\tau}(w)$ is defined such that $v\big(p+t,w-\tau(w)\big) = v\big(p,w-\bar{\tau}(w)\big)$. Here $\bar{\tau}(w)$ naturally exists a $v(p,\gamma)$ is strictly increasing in γ .Which on turn implies that $u_i(v(p+t,w-\tau(w),w) = u_i(v(p+t,w-\bar{\tau}(w),w)$, $\forall w$ .So the utility and labor supply for $\forall i$ are unchanged .Attaining utility of consumption $v\big(p,w-\bar{\tau}(w)\big)$ at price p costs at least $w-\bar{\tau}(w)$.Now, let c_i be the consumer choice of individual i under the initial tax system $(\tau(\cdot),t)$.Individual i attains utility $v\big(p,w-\bar{\tau}(w)\big) = v\big(p,w-\bar{\tau}(w)\big)$ when choosing c_i. And, now $p\cdot c_i \geq w-\bar{\tau}(w)$ and we have that $\bar{\tau}(w) \geq \tau(w) + t\cdot c_i$ i.e. the government collects more taxes with $(\bar{\tau}(\cdot),t=0)$ ■
Paper Atkinson, Stiglitz (1976) showed that :" Even though there was a single "dimension" in which individuals differed (ability), in general, it seemed possible that one could extract information about that difference more efficiently by looking not just at the individual's labor supply, but also at his consumption patterns", see Stiglitz (2018).In the special case of separability[15] of utility function, Pareto efficient taxation required only an income tax; there was no benefit to be had by using information about consumption patterns. This paper Stiglitz (2018)sets the following proposition (link between commodity taxes and income taxes):

Proposition 2. There should be no distortionary taxation on the individual with the highest ability while the labor supply of the less able is distorted Stiglitz (2018)

for a further overview see Sadka (1976), Seade (1977), and (for an overview) Tuomala (1990) ,and see also Diamond (1998). The i[th] individual faces a before-tax wage (output per hour) of w_{ii}, and thus, in the absence of taxation, his budget constraint is simply
equation 40

$$\sum_j c_{ij} = w_i L_i$$

$c_{ij}$is the $i-th$ consumption of individual good j .The ith individual income is :$y_i = w_i L_i$ utility from consuming goods and disutility from work is:
equation 41
$u_i = u_i(c_i, L_i)$
Where $\frac{\partial u_i}{c_i} > 0$ and $\frac{\partial u_i}{\partial L_i} < 0$ is quasi concave. Vector of individual consumption function is :$c_i = \big(c_{i1},\dots,c_{ij},\dots\big)$. Individual maximization problem is:

would not be taxed at a high rate, but rather, that luxuries like perfume might face high rates, even though they have a higher price elasticity than food
[15] Function of 2 independent variables is said to be separable if it can be expressed as a product of 2 functions, each of them depending on only one variable.

Manuscript received: 18.03.2022
Accepted:

International Journal of Economics, Management and Tourism
Vol. 2, No. 1, pp. 7-29
Online: ISSN **2671-3810**
UDC: 336.226.1.01
DOI https://doi.org/10.46763/IJEMT2221007j
Original scientific paper

inequality 16

$$\max_{c_I,y_i} u_i(c_i,l_i)\ st \sum_j c_{ij} \le w_i l_I - \tau(w_i l_i)$$

FOC is:

equation 42

$$\frac{\frac{du_i}{dl_i}}{\frac{du_i}{dc_i}} = -w_i(1-\tau)$$

The left hand side is individual MRS, and right hand side is after-tax marginal return to working an extra hour. The problem of government is with Pareto efficiency.

Definition: *Pareto efficient tax structures are those (given the admissible set of taxes and the required public revenue) which are such that no one can be better off without making someone worse off.*

Here are identified properties of Pareto efficient tax structures which hold regardless of the social welfare function. See, e.g. Stiglitz (1982), Stiglitz(1987) and Brito et al (1990).Now, R is the government revenue and the required revenue is $\bar{R}$, n_i is the number of individuals of type i .If λ_i is the shadow price associated with the self-selection constraints and μ_i is the shadow price associated with the utility constraint.So, now the Lagrangian can be rewritten as:

equation 43

$$\mathcal{L} = n_2 v^2 + n_1\mu(v_1 - u_1) + n_2\lambda_2(v^2(c_2,y_2) - v^2(c_1,y_1)) + n_1\lambda_1(v^1(c_1,y_1) - v^1(c_2,y_2)) + \gamma((y_1 - c_1)n_1 + (y_2 - c_2)n_2 - \bar{R})$$

And the necessary FOC's :

1. $\frac{d\mathcal{L}}{dc_{1j}} = n_2\mu\frac{dv^1}{dc_{1j}} - n_2\lambda_2\frac{dv^2}{dc_{1j}} + n_1\lambda_1\frac{dv^1}{dc_{1j}} - \gamma n_1 = 0$
2. $\frac{d\mathcal{L}}{dc_{2j}} = n_2\mu\frac{dv^1}{dc_{2j}} + n_2\lambda_2\frac{dv^2}{dc_{2j}} - n_1\lambda_1\frac{dv^1}{dc_{2j}} - \gamma n_2 = 0$
3. $\frac{d\mathcal{L}}{dy_1} = n_1\mu\frac{dv^1}{dy_1} - n_2\lambda_2\frac{dv^2}{dy_1} + n_1\lambda_1\frac{dv^1}{dy_1} + \gamma n_1 = 0$
4. $\frac{d\mathcal{L}}{dy_2} = n_2\frac{dv^1}{dy_2} + n_2\lambda_2\frac{dv^2}{dy_2} - n_1\lambda_1\frac{dv^1}{dy_2} + \gamma n_2 = 0$

From this previous FOC's we derive that :

equation 44

$$\frac{\frac{dv^2}{dc_{2j}}}{\frac{dv^2}{dc_{1k}}} = 1\,;\frac{\frac{dv^2}{dy_2}}{\frac{dv^2}{dy_2}} = 1$$

equation 45

$$\frac{dv^1/dc_{1j}}{dv^1/dc_{1k}} = \frac{\gamma + \lambda_2 dv^2/dc_{1j}}{\gamma + \lambda_2 dv^2/dc_{1k}};\frac{dv^1/dc_{1j}}{dv^1/dy_1} = \frac{\gamma + \lambda_2 dv^2/dc_{1j}}{\gamma + \lambda_2 dv^2/dy_1}$$

The interpretation of previous expression is :

Manuscript received: 18.03.2022
Accepted:
International Journal of Economics, Management and Tourism
Vol. 2, No. 1, pp. 7-29
Online: ISSN **2671-3810**
UDC: 336.226.1.01
DOI https://doi.org/10.46763/IJEMT2221007j
Original scientific paper

equation 46

$$a_{jk}^{i} = MRS_{ijk}(c_1, y_1)$$

Marginal rate of substitution between j, k and the bundle (c_i, y_i).The two individuals differ in their abilities. Now let $\lambda_2 \frac{\partial v^2}{\partial c_{1k}} = b$.Then $\frac{\frac{dv^2}{dc_{2j}}}{\frac{dv^2}{dc_{1k}}} = 1$; $\frac{\frac{dv^2}{dy_2}}{\frac{dv^2}{dy_2}} = 1$ can be rewritten as: $a_{jk}^{1} = \frac{\gamma + b a_{jk}^{2}}{\gamma + b}$ and it follows that: $a_{jk}^{1} - a_{jk}^{2} = \{\gamma(1 - a_{jk}^{1})\}/\{\gamma + b\}$. And in the case of separability the MRS between j, k is unaffected by the amount of leisure , so $a_{jk}^{1} = a_{jk}^{2}$ or $a_{jk}^{1} = a_{jk}^{2} = 1$, the denominator of $a_{jk}^{1} - a_{jk}^{2} = \{\gamma(1 - a_{jk}^{1})\}/\{\gamma + b\}$ is positive:

Proposition 3. There should be no commodity taxation on either high or low ability individuals if leisure and consumption are separable.

Derivation of Mirrlees 1971 optimal taxation model and some numerical solutions

In the Mirrlees (1971) model, all individuals have same utility function which depends positively on consumption, and negatively on labor supply ,which can be denoted as $u(c, l)$.Let's suppose the utility function g the agents in the economy Mirrlees (1971) model:
equation 47

$$\widetilde{U}(c, l) = c - \frac{l^2}{2}$$

Where $y = \theta l$ и θ represents the level of skils of the worker. Now his social welfare function SWF is :$SWF(v) = \log(v)$.Now lets find the distribution of skills when $T(y) = 0.3$ which is Pareto with $h(y) = k y^{-k-1} y^{k}$[16].Equation for the distribution of skills is $f(\theta) = h(y(\theta))y'(\theta)$,from the quasi-linear utility functions : $U(c, y, \theta) = c - \frac{1}{2}\left(\frac{y}{\theta}\right)^2$.And the tax function $T(y) = \tau y$, individual with skill level θ solves :
equation 48

$$\max_{y}(1 - \tau)y - \frac{1}{2}\left(\frac{y}{\theta}\right)^2$$

FOC is given as :$(1 - \tau) - \frac{y}{\theta^2} = 0$, which implies that $y = (1 - \tau)\theta^2$ and $f(\theta) = h(y(\theta))y'(\theta) = k(\theta)^{-k-1} y^{k} 2(1 - \tau)\theta = k((1 - \tau)\theta^2)^{-k-1} y^{k} 2(1 - \tau)\theta = 2k(1 - \tau)^{-k}\theta^{-2k-1} y^{k} = 2k\theta^{-2k-1}\theta_l^{2k}$ .By integration one could get :$F(\theta) = \int_{\theta_l}^{\theta} f(\theta)d\theta = \int_{\theta_l}^{\theta} 2k(1 - \tau)^{-k}\theta^{-2k-1} y^{k} d\theta = [-(1 - \tau)^{-k}\theta^{-2k} y^{k}]_{\theta_l}^{\theta} = (1 - \tau)^{-k}\theta_l^{-2k-1}\left((1 - \tau)\theta_l^2\right)^{k} - (1 - \tau)^{-k}\theta^{-2k} y^{k} = 1 - \theta^{-2k}\theta_l^{2k}$.Now we can solve for numerical optimum. Let's use $y = 2$ and $k = 4$ and truncate the distribution[17] at the top x percentile for some small x.In this case : $\max_{v(\theta), u(\theta)} \int_{\theta_l}^{\theta^h} W[v(\theta)] f(\theta)d\theta$.Subject to :

$$\int_{\theta_l}^{\theta^h} (y(\theta) - e(v(\theta), y(\theta), \theta)) f(\theta)d\theta \geq 0; v'(\theta) = u_\theta\,[e(v(\theta), y(\theta), \theta)]$$

[16] This is a density of earnings function , dependent on v the skills of workers
[17] In statistics truncated distribution is a conditional distribution that comes as a result of the restriction of the domain of some other distribution or probability .

Manuscript received: 18.03.2022
Accepted:
International Journal of Economics, Management and Tourism
Vol. 2, No. 1, pp. 7-29
Online: ISSN **2671-3810**
UDC: 336.226.1.01
DOI https://doi.org/10.46763/IJEMT2221007j
Original scientific paper

$y(\theta)$ is non decreasing function. Hamiltonian is formed as $:H = W\,[v(\theta)]f(\theta) + \lambda\,(y(\theta) - e(v(\theta), y(\theta), \theta))f(\theta) + \eta\,(\theta)U_\theta\,[e(v(\theta), y(\theta), \theta), y(\theta), \theta]$.
Standard conditions are as:

1. $\frac{\partial H}{\partial y} = 0 \Rightarrow \lambda f(1 - e_y) + \eta[u_{\theta c}e_y + u_{\theta y}] = 0$
2. $\frac{\partial H}{\partial v} == \eta' \Rightarrow W'f - \lambda e_v f + \eta u_{\theta c}e_v = -\eta'$

Transfersality conditions : $\eta(\theta_l) = \eta(\theta^h) = 0$ .From $W = \log(v)$ и $u(c, y, \theta) = c - \left(\frac{1}{2}\right)\left(\frac{y}{\theta}\right)^2$ will get the following derivatives :$u_\theta = \frac{y^2}{\theta^3}; u_{\theta c} = 0;\ u\theta_y = \frac{2y}{3}$; $W' = \frac{1}{v}$.Let us remember that $v = u(e(v, y, \theta), y, \theta)$,we have $1 = u_c e_v$ и $0 = u_c e_y + u_y$,therefore :$e_v = \frac{1}{u_c}; e_y = -\frac{u_y}{u_c} = \frac{y}{\theta^2}$. If we substitute in the optimality and control equations about the state variables one can get :

equation 49

$$\lambda f\,\left(1 - \frac{y}{\theta^2}\right) + \eta\left[\frac{2y}{\theta^3}\right] = 0 \text{ и } \frac{f}{v} - \lambda f = -\eta'$$

If we solve in the first equation for $y(\theta)$ we get : $y(\theta) = \frac{\lambda f\,(\theta)\theta^3}{\lambda f\,(\theta)\theta - 2\eta(\theta)}$. With the equation $\eta'\eta'(\theta) = \left(\lambda - \frac{1}{v(\theta)}\right)f(\theta)$. If we substitute for $y(\theta)$ in the constraint $:v'(\theta) = u_\theta[e(v(\theta), y(\theta), \theta), y, \theta] = \frac{y^2}{\theta^3} = \left(\frac{\lambda f\,(\theta)}{\lambda f\,(\theta)\theta - 2\eta(\theta)}\right)^2\theta^3$.In Saez (2001) optimal tax formula is given as :

equation 50

$$\tau = \frac{1 - \bar{g}}{1 - \bar{g} + \bar{\varepsilon}^u + \bar{\varepsilon}^c(a - 1)}$$

In another example that follows Mirrleees (1971) and Diamond (1998) Utility function is quasi linear:

equation 51

$$u(c, l) = c - v(l)$$

c is disposable income and the utility of suuply of labor $v(l)$ is increasing and convex in l. Earnings equal $w = nl$ where n represents innate ability. CDF of skills distribution is $F(n)$, it's PDF is $f(n)$and support range is $[0, \infty)$. Government cannot observe abilities instead it can set taxes asa function of labor income $c = w - \tau(w)$.Individual n chooses l_n to maximize :

equation 52

$$\max(nl - \tau n(l) - v(l))$$

When marginal tax rate τ is constant, the labor supply f-ction is given as: $l\ \rightarrow\ l(n(1 - \tau))$ and it is implicitly defined by the $n(1 - \tau) =\ v'(l)$. And $\frac{dl}{d(n(1-\tau))}\ = \frac{1}{v''(l)}$, so the elasticity of the net-of-tax rate $1 - \tau$ is:

equation 53

$$e\ = \frac{\left(\frac{n(1-\tau)}{l}\right)dl}{d(n(1-\tau))} = \frac{v'(l)}{lv''(l)}$$

As there are no income effects this elasticity is both the compensated and the uncompensated elasticity.The government maximizes SWF :

equation 54

$$W\ = \int G(u_n)f\,(n)dn\ \ s.t. \int cnf\,(n)dn\ \ \leq\ \int nlnf\,(n)dn\ -\ E\,(\lambda)$$

u_n denotes utility, $w_n\ = nl_n$ denotes earnings, c_n denotes consumption or disposable income, and $c_n = u_n + v(l_n)$.By using the envelope theorem and the FOC for the individual, u_n satisfies following:

Manuscript received: 18.03.2022
Accepted:

International Journal of Economics, Management and Tourism
Vol. 2, No. 1, pp. 7-29
Online: ISSN **2671-3810**
UDC: 336.226.1.01
DOI https://doi.org/10.46763/IJEMT2221007j
Original scientific paper

equation 55

$$\frac{du_n}{dn} = \frac{lnv'(ln)}{n}$$

Now the Hamiltonian is given as:

equation 56

$$\mathcal{H} = \left[G(u_n) + \lambda \cdot \left(nl_n - u_n - v(l_n)\right)\right] f(n) + \phi(n) \cdot \frac{lnv'(ln)}{n}$$

In previous $\phi(n)$ is the multiplier of the state variable. The FOC with respect to l is given as:

equation 57

$$\lambda \cdot \left(n - v'(l_n)\right) + \frac{\phi(n)}{n} \cdot [v'(l_n) + l_n v''(l_n)] = 0$$

FOC with respect to u is given as:

equation 58

$$-\frac{d\phi(n)}{n} = [G'(u_n) - \lambda]$$

If integrated previous expression gives: $-\phi(n) = \int_n^{\infty}[\lambda - G'(u_m)]f(m)dm$ where the transversality condition $\phi(\infty) = 0$,and $\phi(0) = 0$, and $\lambda = \int_0^{\infty} G'(u_m)f(m)dm$ and social marginal welfare weights $\frac{G'(u_m)}{\lambda} = 1$. Using this equation for $\phi(n)$ and all previous $n - v'(ln) = n\tau'(w_n)$, and that

equation 59

$$\frac{[v'(l_n) + l_n v''(l_n)]}{n} = \left[\frac{v'(l_n)}{n}\right]\left[1 + \frac{1}{e}\right]$$

We can rewrite FOC with respect to l_n as:

equation 60

$$\frac{\tau'(w_n)}{1 - \tau'(w_n)} = \left(1 + \frac{1}{e}\right) \cdot \left(\frac{\int_n^{\infty}(1 - g_m)dF(m)}{nf(n)}\right)$$

In previous expression $g_m = \frac{G'(u_m)}{\lambda}$ which is the social welfare on individual m . The formula was derived in Diamond (1998) . If we denote $h(w_n)$ as density of earnings at w_n if the nonlinear tax system were replaced by linearized tax with marginal tax rate $\tau = \tau'(w_n)$ we would have that following equals $h(w_n)dw_n = f(n)dn$ and $f(n) = h(w_n)l_n(1 + e)$,henceforth $nf(n) = w_n h(w_n)(1 + e)$ and we can write previous equation as:

equation 61

$$\frac{\tau'(w_n)}{1 - \tau'(w_n)} = \frac{1}{e} \cdot \left(\frac{\int_n^{\infty}(1 - g_m)dF(m)}{w_n h(w_n)}\right) = \frac{1}{e} \cdot \left(\frac{1 - H(w_n)}{w_n h(w_n)}\right) \cdot \left(1 - G(w_n)\right)$$

In the previous expression $G(w_n) = \int_n^{\infty} \frac{dF(m)}{1-F(n)}$ is the average social welfare above w_n .If we change variables from $n \to w_n$, we have $G(w_n) = \int_{w_n}^{\infty} \frac{g_m dH(w_m)}{1-H(w_n)}$.The transversality condition implies $G(w_0 = 0) = 1$. In the Mirrlees(1971) model government , maximizes[18] :$SWF = \int_0^{\infty} G(u_w)f(w)dw$. In the previous expression $G(u_w)$ represents the concave utility function[19].The constraint here is given as: $\int_0^{\infty} G(u_w)f(w)dw \leqq \int_0^{\infty} w_l f(w)dw - E$,where E are government expenditures. Now, about Pareto distributions it is well known fact that

[18] Here we make assumption that wages =skill level

[19] Now, for a concave function $f: (a, b) \to R$ is continuous in $IntA$. This function $f: (a, b) \to R$ is concave in the interval (a, b) , if for every $x_1, x_2 \in (a, b), a \in (0,1)$, it follows $f(ax_1 + (1 - a)x_2) < af(x_1) + (1 - a)f(x_2)$.

Manuscript received: 18.03.2022
Accepted:
International Journal of Economics, Management and Tourism
Vol. 2, No. 1, pp. 7-29
Online: ISSN **2671-3810**
UDC: 336.226.1.01
DOI https://doi.org/10.46763/IJEMT2221007j
Original scientific paper

:$\frac{ratio\ average}{threshold} = constant$ .Now if we denote the average wage $w^*(w) > w$, and if w is a threshold, then $w^*(w)$ can be expressed as :$w^*(w) = \int_{w_m > w} wf(w)dw / \int_{w_m > w} f(w)dw = \int_{w_m > w} dw/w^a \,/\, \int_{w_m > w} dw/w^{1+a} = \frac{aw}{a-1}$. In the previous expression a represents the shape parameter of the Pareto distribution. And $a = \frac{b}{b-1}$ i.e. $\frac{w^*(w)}{w} = b$ .About the Pareto distribution PDF of this distribution is given as :$1 - F(w) = \left(\frac{k}{w}\right)^a$,and CDF of the function is given as $f(w) = \frac{ak^a}{w^{1+a}}$,[20]that is $\lim_{w\to\infty} \frac{\left(\frac{k}{w}\right)^a}{w\cdot\left(\frac{ak^a}{w^{1+a}}\right)}$ by applying $\lim_{x\to a}[c \cdot f(x)] = c \cdot \lim_{x\to a} f(x) \Rightarrow \frac{1}{ak^a} \cdot \lim_{w\to\infty} \frac{\left(\frac{k}{w}\right)^a}{w\cdot\left(\frac{ak^a}{w^{1+a}}\right)} = \frac{1}{ak^a} \cdot \lim_{w\to\infty}(k^a) = \frac{1}{ak^a} \cdot k^a = \frac{1}{a}$ .hence the formula of marginal income for top earners

equation 62

$$\tau^* = \frac{1}{1+a\cdot\varepsilon}.$$

In Piketty, T., Saez.E., and Stantcheva,S.(2014), it is well defined aggregate elasticity of income as:

equation 63

$$\varepsilon = \frac{1-\tau}{z}\frac{dz}{d(1-\tau)} \ ,$$

where z is taxable income and $z = y - x$, where y is the real income, and x is sheltered income [21],taxable income s used in the calculation for Pareto parameter $a = \frac{z}{z-\bar{z}}$. Tax avoidance elasticity component is given as $\varepsilon_1 = \frac{1-\tau}{z}\frac{dx}{1-\tau}$, and $\varepsilon_2 = \frac{1-\tau}{z}\frac{dy}{1-\tau}$ is the real labor supply elasticity. The bottom tax formula in Mirrlees (1971) is derived in Pikety, Saez (2013).

equation 64

$$\frac{\tau'(0)}{1-\tau'(0)} = (g_0 - 1)\cdot\frac{F(n_0)}{n_0 f(n_0)} \Rightarrow \tau'(0) = \frac{g_0 - 1}{g_0 - 1 + \frac{n_0 f(n_0)}{F(n_0)}}$$

In previous expression $g_0 = \frac{G(u_0)u_c}{\lambda}$ is the social marginal weight of the non-worker. From previous we know that $n_0(1 - \tau'(0))u_c(c_0, 0) + u_l(c_0, 0) = 0$ which defines $n_0(1 - \tau'(0), c_0)$.The effect of $1 - \tau'(0)$ on n_0 is such that $\frac{\partial n_0}{\partial(1-\tau'(0))} = -\frac{n_0}{1-\tau'(0)}$. Hence,the elasticity of the fraction non-working $F(n_0)$ with respect to $1 - \tau'(0)$ is given as:

equation 65

$$e_0 \equiv -\frac{1-\tau'(0)}{F(n_0)}\frac{dF(n_0)}{d\big(1-\tau'(0)\big)}\Big|_{c_0} = --\frac{1-\tau'(0)}{F(n_0)}\cdot f(n_0)\cdot\frac{\partial n_0}{\partial\big(1-\tau'(0)\big)} = \frac{n_0 f(n_0)}{F(n_0)}$$

[20] $\frac{\left(\frac{k}{w}\right)^a}{w\cdot\left(\frac{ak^a}{w^{1+a}}\right)} = \frac{\frac{k^a}{w^a}}{w\cdot\left(\frac{ak^a}{w^{1+a}}\right)} = \frac{\frac{k^a}{w^a}}{w\cdot\frac{ak^a}{w\cdot w^a}} = \frac{1}{a}$

[21] Investments or investment accounts that provide favorable tax treatment , or activities and transactions that lower taxable income.

Manuscript received: 18.03.2022
Accepted:
International Journal of Economics, Management and Tourism
Vol. 2, No. 1, pp. 7-29
Online: ISSN **2671-3810**
UDC: 336.226.1.01
DOI https://doi.org/10.46763/IJEMT2221007j
Original scientific paper

So we can rewrite $\tau'(0)=\frac{g_0-1}{g_0-1+\frac{n_0 f(n_0)}{F(n_0)}}$ to :

equation 66

$$\tau'(0)=\frac{g_0-1}{g_0-1+e_0}$$

About social marginal weights: Social marginal welfare weight[22] is given as:
equation 67

$$g_i=\frac{\omega_i G'(u^i)u_c^i}{\lambda}$$

g_i measures the dollar/euro value (in terms of public funds) of increasing consumption of individual i by \$1 or €1. Under utilitarian criterion,$g_i=\frac{u_i^c}{\lambda}$ is directly proportional to the marginal utility of consumption. Under Rawlsian criterion all the $\forall g_i=0$ except for the most disadvantaged (poorest). Social welfare function can be :

1. $SWF=\int U^i\,di$-Utilitarian or Benthamite,
2. $SWF=min_i U^i$- Rawlsian $SWF=\int U^i di\rightarrow G(U)=\frac{U^{1-\gamma}}{1-\gamma}$ if $\gamma=0$ function is utilitarian , Rawlsian if $\gamma=\infty$.
3. With Pareto weights: $SWF=\int \mu_i U^i\,di$ where μ_i is exogenous.

The optimal tax government formula with Rawlsian government [23]would be :
equation 68
$\frac{T'(w(h))}{1-T'(w(h))}=\left(\frac{1+\varepsilon}{\varepsilon}\right)\frac{1-F(w)}{wf(w)}$ or $\frac{T'(w(h))}{1-T'(w(h))}=\left(\frac{1+\varepsilon}{\varepsilon}\right)\frac{\psi(w)-F(w)}{wf(w)}$

Now if we divide and multiply by $1-F(w)$ we get :$\frac{T'(w(h))}{1-T'(w(h))}=\left(\frac{1+\varepsilon}{\varepsilon}\right)\frac{\Psi(w)-F(w)}{1-F(w)}\frac{1-F(w)}{wf(w)}$ .In the previous formula $\left(\frac{1+\varepsilon}{\varepsilon}\right)=A(w)$, elasticity and efficiency argument, $\frac{\Psi(w)-F(w)}{1-F(w)}=B(w)$, measures the desire for redistribution :if the sum of weights $\psi(w)f(w)$ is below w is relative high to the weights above , the government will like to tax more, this part $\frac{1-F(w)}{wf(w)}=C(w)$ measures the density of the right tail of the distribution and higher density will be associated with higher taxes. This is ABC tax model by Diamond (1998) . Next, we will do numerical simulations on Mirrlees model and we can observe the marginal tax rate schedules.

Numerical solutions of Mirrlees optimal tax model

This simulation here captures section 8 (case I) of the original paper of MIrrlees 1971 paper and section 9. So the setup of the code due to Ben Lockwood (benlockwood.com) is as follows:
equation 69

$$\begin{cases} u=\alpha\log x+\log(1-y)\\ G(u)=-\frac{1}{\beta}e^{-\beta u}\\ f(n)=\frac{1}{n}\exp\left[-\frac{(\log n+1)^2}{2}\right]\end{cases}$$

[22] The marginal social welfare weight on a given individual measures the value that society puts on providing an additional dollar of consumption to this individual.
[23] The social welfare function that uses as its measure of social welfare the utility of the worst-off member of society. The following argument can be used to motivate the Rawlsian social welfare function.

Manuscript received: 18.03.2022
Accepted:

International Journal of Economics, Management and Tourism
Vol. 2, No. 1, pp. 7-29
Online: ISSN **2671-3810**
UDC: 336.226.1.01
DOI https://doi.org/10.46763/IJEMT2221007j
Original scientific paper

Skills are assumed to be lognormally distributed with the average $\bar{n} = \frac{1}{\sqrt{e}} = 0.607$. .So now, the equations :

equation 70

$$\begin{cases} \frac{dv}{dn} = -\frac{v}{n}\left(2 + \frac{nf'}{f}\right) - \frac{1}{n^2 u_1} + \frac{\lambda G'}{n^2} \\ \frac{du}{dn} = -\frac{yu_2}{n} \end{cases}$$

Would become:

equation 71

$$\begin{cases} \frac{dv}{dn} = -\frac{v \log n}{n} - \frac{x}{\alpha n^2} + \frac{\lambda}{n^2} e^{-\beta u} \\ \frac{du}{dn} = \frac{y}{n(1-y)} \end{cases}$$

Where : $v = \frac{\left[1+\frac{u_2}{nu_1}\right]}{\psi_y} = \frac{1-\frac{x}{\alpha n(1-y)}}{1/(1-y)^2} = (1-y)\left(1-y-\frac{x}{\alpha n}\right)$ and $e^u = x^\alpha(1-y)$ and when $\beta = 0$; $s = 1-y$ and $t = \log n$ .Now $r = s^{\frac{1-\alpha}{\alpha}}(s^2 - v)$,so that : $\alpha\frac{dr}{dt} = \frac{1-(1-\alpha)}{s}r$. The marginal tax rate is :

$$\tau = \frac{v}{s^2} ;$$

So as $\tau \to \infty$ $\theta \to 0$ in the original Mirrlees (1971) paper θ denotes marginal tax rate.And, it follows that :

equation 72

$\theta = \frac{v}{s^2} \sim \frac{1+\alpha}{\tau}$ or $\theta = \frac{v_t}{s_t^2} \to \frac{1+\alpha}{1+\alpha+\gamma}$

Assumed production function is linear : $x = z + a$, the average product of labor is x/z .In the full optimum it is maximized:

equation 73

$$\begin{cases} \int [\log x + \log(1-y)] f(n) dn \;\; s.t. \\ \int x f(n) dn = \int ny f(n) dn + a \end{cases}$$

where $x = x^0$ and $\log(1-y) + \frac{ny}{x^0}$ and maximization yields $y_n^0 = \left[1 - \frac{x^0}{n}\right]$.The value of x^0 is:

equation 74

$$x^0 = \int_{x^0}^{\infty} (n - x^0) f(n) dn + a$$

Where $\int_0^\infty f(n)dn = 1$.

Table 3 FOC's for the Mirrlees model

iteration	Func-count	f(x)	Norm of step	First-order optimality
0	3	1.37E-01		

Manuscript received: 18.03.2022
Accepted:
International Journal of Economics, Management and Tourism
Vol. 2, No. 1, pp. 7-29
Online: ISSN **2671-3810**
UDC: 336.226.1.01
DOI https://doi.org/10.46763/IJEMT2221007j
Original scientific paper

1	6	9.01E-04	0.000224	0.00276
2	9	2.13E-04	2.97E-01	0.000677
3	12	5.02E-08	4.86E-01	9.93E-06
4	15	2.87E-14	6.74E-03	7.50E-09

Table 4 skills, consumption and earnings for the Mirrlees model

F(n)-skills	x-cons.	y-income	x(1-y)	z-earnings
0	0.0424	0	0.0424	0
0.1	0.116	0.3894	0.0708	0.0869
0.5	0.18	0.4382	0.1011	0.1612
0.9	0.2888	0.4686	0.1535	0.2842
0.99	0.4315	0.4841	0.2226	0.4412

Table 5 average and marginal tax rates for Mirrlees model

z-earnings	x-consumption	average tax rate	marginal tax rate
0	0.0424	-Inf	0.2147
0.05	0.0847	-0.54	0.2336
0.1	0.1271	-0.1558	0.2223
0.2	0.214	0.0273	0.1993
0.3	0.3031	0.0817	0.1824
0.4	0.3937	0.1052	0.1698
0.5	0.4856	0.1171	0.1599

Figure 2 Earnings and taxes paid by Mirrlees schedule

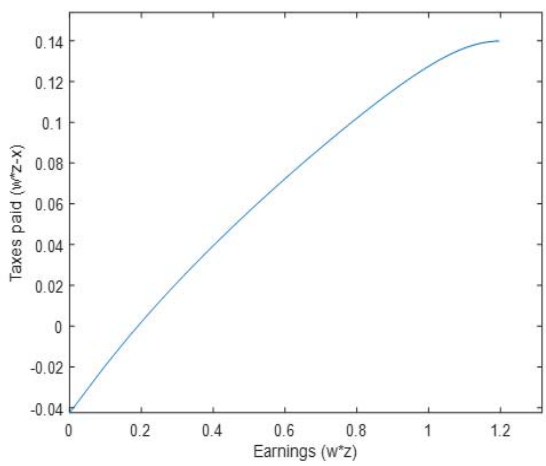
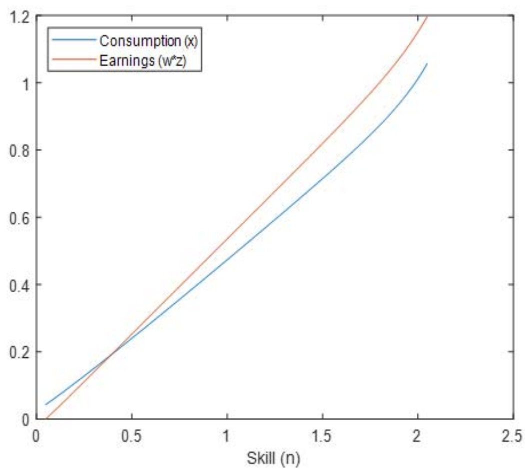


Figure 3 distribution of skills and earnings in Mirrlees model



Manuscript received: 18.03.2022
Accepted:

International Journal of Economics, Management and Tourism
Vol. 2, No. 1, pp. 7-29
Online: ISSN **2671-3810**
UDC: 336.226.1.01
DOI https://doi.org/10.46763/IJEMT2221007j
Original scientific paper

Conclusion

This paper has characterized the set of Pareto efficient income schedules in Mirrlees (1971) model. Namely the models obtained by Werning (2008) provides versions of the Pareto optimality condition that may be useful of testing this condition, or that may provide framework for quantity analysis. These models were avoiding the specification by normative welfare criterion, and the analysis was more able to focus on the elements of positive economy. The optimality conditions shed new light in on the importance of the skill distribution and other parameters in shaping efficient tax schedules. Atkinson-Stgilitz theorem (1976) on the other hand was a reminder that in presence of optimal non-linear income tax, the role of commodity taxation was limited see Stiglitz (2018). Further propositions for optimal Pareto income taxation are: There should be no distortionary taxation on the individual with the highest ability while the labor supply of the less able is distorted, there should be no commodity taxation on either high or low ability individuals if leisure and consumption are separable. These models of Pareto efficient taxation provide most useful insight into what economic theory has to say about the design of tax structures i.e. Pareto efficient tax structures maximize the utility of one individual (group) given the utility of others and given the budget balance and informational constraints on the government. The optimal income tax problem is in the middle or the asymmetric information with adverse selection. Individuals it is assumed that they differ in ability and productivity when they have same endowments and utility functions. The Mirrlees optimal tax problem could be thought of extracting information about those differences. The information about individuals did not just end up with different abilities, one could extract information about the differences by looking into consumption patterns also. Though in the case of separability of the utility function Pareto efficient taxation required only income taxation. Pareto efficient taxation takes positive and normative characterization in: redistribution vs efficiency and Pareto efficiency, unlike Mirrleesian optimal taxation such as: Mirrlees (1971), Diamond (1998), Saez (2001) which characterizes redistribution vs efficiency as positive criterion but utilitarian social welfare function is normative criterion. This paper provides similarities between Mirrleesian and Pareto efficient taxation embedded into Mirrleesian framework.

References

1. Albanesi,S.Sleet,C. (2003).Dynamic Optimal Taxation with Private Information.mimeo, Federal Reserve Bank of Minneapolis.
2. Amador, Manuel, Iván Werning, and George-Marios Angeletos, (2006). Commitment vs. Flexibility. Econometrica. 74 (2): 365–396
3. Aronsson,T. Johansson-Stenman,O.(2015). Pareto Efficient Income Taxation when People are Inequality Averse.Working paper, Research grants from the Bank of Sweden Tercentenary Foundation, the Swedish Council for Working Life and Social Research, and the Swedish Tax Agency
4. Atkinson, A. and J. E. Stiglitz(1972). The Structure of Indirect Taxation and Economic Efficiency, Journal of Public Economics, 1(1): 97-119
5. Atkinson, A., Stiglitz.J.E. (1976). The design of Tax Structure: Direct Versus Indirect Taxation, Journal of Public Economics, 6: 55-75.
6. Battaglini, M. Coate, S. (2003). Pareto Efficient Income Taxation with Stochastic Abilities. Journal of Public Economics. 92. 844-868. 10.1016/j.jpubeco.2007.06.006.
7. Brito , D.L., J.H. Hamilton S.M. Slutsky and J. E. Stiglitz, (1990).Pareto Efficient Tax Structures,Oxford Economic Papers, 42, pp. 61-77
8. Brito, D. Hamilton,J. Slutsky,S.Stiglitz,J. (1991). Dynamic Optimal Taxation with Commitment," Journal of Public Economics, 44,15-35.
9. Diamond, P. (1998). Optimal income taxation: An example with a U-shaped pattern of optimal marginal tax rates. American Economic Review, 88, 83–95.

Manuscript received: 18.03.2022
Accepted:
International Journal of Economics, Management and Tourism
Vol. 2, No. 1, pp. 7-29
Online: ISSN **2671-3810**
UDC: 336.226.1.01
DOI https://doi.org/10.46763/IJEMT2221007j
Original scientific paper

10. Diamond,P. Mirrlees,J. (1978).A Model of Social Insurance with Variable Retirement. Journal of Public Economics, 10, 295-336.
11. Edgeworth, F.Y.(1868).Mathematical physics: an essay on the application of mathematics to the moral sciences, C. KEGAN PAUL & CO., 1 paternoster square
12. Edgeworth, F.Y.(1887).The Pure Theory of Taxation. The Economic Journal Vol. 7, No. 28 (Dec., 1897), pp. 550-571
13. Ellsberg, Daniel (1954). Classic and Current Notions of 'Measurable Utility. Economic Journal. 64 (255): 528–556.
14. Fair,R.C. (1971).The Optimal Distribution of Income.Quart. J. Econ., Nov. 1971, 85, 55779.
15. Golosov, M., Tsyvinsky,A, (2003).Designing Optimal Disability Insurance. mimeo, University of Minnesota
16. Golosov, M.Kocherlakota,N. Tsyvinsky,A. (2003).Optimal Indirect and Capital Taxation. Review of Economic Studies, 70, 569-588.
17. Gruber, J.Saez,E.(2002). The Elasticity of Taxable Income: Evidence and Implications, Journal of Public Economics 84(1):1-32
18. Kaplow,L. (2006). On the undesirability of commodity taxation even when income taxation is not optimal. Journal of Public Economics, 90(6–7), 1235–50.
19. Laroque, G. R. (2005). Indirect taxation is superfluous under separability and taste homogeneity:A simple proof. Economics Letters, 87(1), 141–144.
20. Luenberger, D. G.(1969). Optimization by Vector Space Methods, John Wiley & Sons, Inc
21. Mankiw NG, Weinzierl M, Yagan D.(2009).Optimal Taxation in Theory and Practice. Journal of Economic Perspectives. 2009;23 (4) :147-174.
22. Mas-Colell, A., et al. (1995). Microeconomic Theory, Oxford University Press.
23. Mirrlees, J. A. (1971). An exploration in the theory of optimal income taxation. Review of Economic Studies,38, 175–208.
24. Mirrlees, J. A. (1986). The theory of optimal taxation. In K. J. Arrow, M. D. Intriligator (Eds.), Handbook of mathematical economics.Vol. 3 (pp. 1197–1249). Amsterdam: North-Holland.
25. Piketty, T., Saez,E. (2013), Chapter 7 - Optimal Labor Income Taxation, Elsevier,Volume 5, Pages 391-474, ISSN 1573-4420, ISBN 9780444537591
26. Piketty,T., Saez,E., Stantcheva,S.(2014). Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities,American Economic Journal: Economic Policy, American Economic
27. Sadka,E.(1976).On Income Distribution, Incentive Effects and Optimal Income Taxation,The Review of Economic Studies, Vol. 43, No. 2, (Jun., 1976), pp. 261-267
28. Saez,E.(2001).Using elasticities to derive optimal income tax rates, The Review of Economic Studies, 68(1), pp.205- 229.
29. Seade, J. K. (1977).On the shape of optimal tax schedules, Journal of Public Economics, 7, issue 2, p. 203-235.
30. Stiglitz, J.E.(1982). Self-Selection and Pareto Efficient Taxation, Journal of Public Economics, 17, 213-240.
31. Stiglitz, J.E.(1987). Pareto Efficient and Optimal Taxation and the New New Welfare Economics," in Handbook on Public Economics, A. Auerbach and M. Feldstein (eds.), North Holland: Elsevier Science Publishers, 991-1042
32. Stiglitz, J.E.(2018). Pareto efficient taxation and expenditures: pre- and re-distribution. NBER WORKING PAPER SERIES
33. Tuomala, M.(1990). Optimal Income Tax and Redistribution. New York: Oxford University Press.

Manuscript received: 18.03.2022
Accepted:
International Journal of Economics, Management and Tourism
Vol. 2, No. 1, pp. 7-29
Online: ISSN **2671-3810**
UDC: 336.226.1.01
DOI https://doi.org/10.46763/IJEMT2221007j
Original scientific paper

34. Werning, Ivan, (2002).Optimal Dynamic Taxation and Social Insurance. Doctoral Dissertation, University of Chicago.
35. Werning,I.(2007).Pareto efficient income taxation. NBER Public Economics meeting
36. Ramsey, F. (1927). A Contribution to the Theory of Taxation. Economic Journal. 37: 47 61. doi:10.2307/2222721.