

UDK 336.226.11.055.01  
336.226.11.055:303.725.3

# CRITICAL REVIEW OF THE (SECOND WAVE) OPTIMAL TAX THEORIES

Dushko Josheski,\*  
University Goce Delcev-Shtip , R.North Macedonia

Tatjana Boshkov \*  
University Goce Delcev-Shtip , R.North Macedonia

---

## Abstract

*James Mirrlees (1971) launched the second wave of optimal tax models by suggesting a way to formalize the planner's problem that deals explicitly with unobserved heterogeneity among taxpayers. So, in this paper optimal income taxation theories are subject of investigation following the classic paper in public finance by Mirrlees (1971). This provides analytical solutions for the second-best efficient tax system in presence of such an adverse selection. Until late 1990s, Mirrlees results were not closely connected to empirical tax studies and had little impact on tax policy recommendations. Next, the famous result Diamond-Mirrlees efficiency theorem Diamond-Mirrlees (1971a), Diamond-Mirrlees (1971b), has been reviewed. This theorem is important because it states that there should be no taxes on intermediate goods, and that private and public production should be based on same prices. Also, taxation should not violate efficiency of production. Solution to the Mankiw problem on the other hand states that small open economy, labor bears 100% of small capital income tax. The availability of the eight documents does not indicate outright budget process transparency nor do we suggest that in means absolute accountability of the authorities however it demonstrates a step towards increased citizens informed and active civic participation.*

**Keywords:** *Optimal taxation, asymmetric information, Diamond-Mirrlees efficiency theorem, Mankiw problem*

---

## INTRODUCTION

The subject of the optimal design of tax system has been topic that has long fascinated economic theorist. This paper explores some theories that policy makers must consider when designing optimal tax systems. Mainly the material in this paper draws material from the very important paper ,actually the foundation of modern tax theory by Mirrlees (1971). Some lessons from the optimal tax theory are presented in this paper

---

\* Assistant professor, UGD-Shtip, R.North Macedonia, dusko.josevski@ugd.edu.mk

\* Associate professor, UGD-Shtip, R.North Macedonia, tdzaleva@gmail.com

also namely: Optimal tax schedule depends on the shape of the distribution of the ability, the optimal marginal income tax schedule could incline at high incomes, a flat tax with a universal lump-sum transfer could be optimal, the extent of redistribution rises with wage inequality, taxes should depend on the ability (some personal characteristics such as IQ and the working hours), only final goods to be taxed (no taxes on intermediate goods), capital income ought to be untaxed, also Mankiw problem shows that that in a small open economy, labor bears 100% of small capital income tax. This paper is divided in following sections: at first it starts with the literature review, followed by the explanation of the Mirrlees (1971) model, which is better understood under Saez (2001). This section is briefly followed by the Diamond-Mirrlees theorem Diamond-Mirrlees (1971a), Diamond-Mirrlees (1971b). This section is followed by the Mankiw problem and its solution. Mankiw problem was posted online by Nicholas Gregory Mankiw, on his blog.<sup>1</sup> Later paper finishes by the conclusions part.

## Literature review

The central element of the theory of optimal taxation is information. Public policies apply to the individuals based on what the government knows about them. Second welfare theorem states, that where several convexity and continuity assumptions are satisfied, an optimum is a competitive equilibrium once initial endowments have been suitably distributed. Mirrlees (1986), elaborates that a good way of governing is to agree upon objectives, then to discover what is possible and to optimize. 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. 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).<sup>2</sup> The choice of the optimal redistributive tax involves tradeoffs between three kinds of effects: equity effect (it changes the distribution of income), the efficiency effect from reducing the incentives, the insurance effect from reducing the variance of individual income streams, Varian, H.R. (1980). Diamond, Helms and Mirrlees (1978), analyze the presence of uncertainty in the analysis of optimal taxation, with Cobb-Douglas utility function, with elasticity of substitution between labor and leisure  $< 1$  so that backward bending labor supply curve can be observed. Two period model with uncertainty showed how stochastic economies differ from the economies without uncertainty, since these second-best insurance/redistribution programs differ in the outcomes from the first best result economies without government intervention. In general, Varian (1980) finds for linear and non-linear optimal tax, that if the consumption values are bounded, the optimal tax will always exist and would be a continuous function of observed income. In early contribution Ramsey (1927), supposed that the planner must raise tax revenue only through imposition of tax on commodities only. In his model taxes should be imposed in inverse proportion to the representative customer's elasticity of demand for the good, so that commodities with more inelastic demand are taxed more heavily. Governments in real world however cannot observe individual ability. Mirrlees (1971), in the basic version of the model allowed individuals to differ in their innate ability. The planner 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. Saez (2001) argued that "unbounded distributions are of much more interest than bounded distributions to address high income optimal tax rate problem". In all the cases that Saez (2001) investigated (four cases)<sup>3</sup> 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%. Diamond and Mirrlees (1971a) and Diamond, Mirrlees (1971b), are proposing alternative in Ramsey

1) <http://gregmankiw.blogspot.com/2017/10/an-exercise-for-my-readers.html>

2) A key determinant of the optimal tax schedule (tax bracket) is the shape of the ability of the distribution.

3) Utilitarian criterion, utility type I and II and Rawlsian criterion, utility type I and II.

proposition by allowing the social planner to consider a numerous tax systems. In the first paper Diamond and Mirrlees (1971a), they prove how some market imperfections eg. capital market imperfections (consumers can lend but not borrow), the market situation will alter the optimal tax structure. Diamond and Mirrlees (1971b), are proposing tax rules for single good economy (changes on demand due to the tax structure differ from proportionality with larger than average percentage fall in the demand for goods with large income derivatives (elasticities) ,in three-good economy the tax rate is proportionately greater for a good with smaller cross elasticity of compensated demand with the price of labor, in many commodities economy, households with low social marginal income utility predominate among the purchasers of the commodity, that commodity should be taxed more heavily, and vice versa, this taxation increases total welfare. Diamond and Mirrlees (1971a), continue to point out that there should not be taxation on intermediate goods such as capital held by the producers, also see Judd (1999). The general result Judd (1999) finds is that optimal tax on capital should be zero except for the initial period. Judd (1985), also found a zero optimal long-run capital income tax rate for steady states of the general competitive equilibrium and heterogeneous infinitely-lived agents with non-separable preferences. But the famous Atkinson, Stiglitz (1976) results (result on the role of indirect taxation with an optimal nonlinear income tax) states that commodity taxes are not useful under these assumptions about the utility function: weak separability of function, and homogeneity across individuals in sub-utility of consumption. The Atkinson-Stiglitz result is obtained by embedding the Ramsey model within Mirrlees model. Also zero-top tax rate suggest important task for the policy makers to identify the shape of the high-end of the ability distribution (they cannot observe the effort and ability in direct way but they can observe income). Tuomala (1990), confirms that marginal tax rate decreases as income increases except at income levels within a bottom decile. Ordober, J., Phelps, E. (1979), provided that if consumption have weakly separable utility functions and government has instruments that allow it to fix the capital stock on the socially optimal level, then the optimal tax rate on capital is zero, Salanie (2003). Chamley (1986), results on zero capital income tax states: "When the consumption decisions in a given period have only a negligible effect on the structure of preferences for periods in the distant future, then the second-best tax rate on capital income tends to zero in the long run". But these are (Ramsey capital income tax) two period models if more periods are included than the optimal tax formula would be more complex, as in Auerbach, Kotlikoff (1987a), and Auerbach, Kotlikoff (1987b). Feldstein (1978), showed that elimination of tax on capital income is only optimal only when the structure of preferences satisfy certain separability condition. And for the capital taxation to be optimal it must be that uncompensated elasticity of savings (elasticity of the Marshallian demand for savings) is zero, even when the compensated elasticity of consumption of old population (Hicksian demand for consumption) is high (he reported result of -0.75). Now, if the labor and consumption are equivalent for the individuals, but savings pattern are different, results is that individuals will save more with consumption tax, than with labor tax. In OLG closed economy capital stock is due to life-time savings. The full neutrality result implies extra savings of young is equal to the consumption of old capital stock plus new government deficit (no change in capital stock).<sup>4</sup> In equilibrium where endowment is zero at equilibrium, and Hicksian demand for consumption is infinite i.e. compensated elasticity of consumption when old is infinite. But according to Saez, Stantcheva (2016a), because individuals derive utility from wealth, micro foundations for this wealth in the utility function are: bequest motives, entrepreneurship, or services from wealth it means that steady-state features finite supply elasticities of capital to capital tax rates. And because there is bi-heterogeneity of the agent's income and capital, Atkinson-Stiglitz zero-tax result does not apply herein. The optimal tax rate on inheritance (bequest in utility) case is zero, when the elasticity of bequest is infinite nesting the zero tax result. However, when in the model are imputed bequests, inequality is bi-dimensional and earnings are no longer the unique determinant of lifetime resources. That means that here A-S zero-tax result fails, see Piketty, T., Saez, E., (2013), Farhi and Werning (2010). Also, Stiglitz, J. (1982), showed that when leisure and goods are separable, differential taxation of commodities cannot be used as a basis of separation between the two and therefore is sub-optimal, Saez (2002). Commodity taxation is desirable when government is using social weights that are correlated with the con-

4) Aggregate interest rate should equal to interest rate for the government debt.

sumption patterns and are conditional on income, or when the consumption patterns are related to the intrinsic earning ability and leisure choices. Saez, E., S. Stantcheva (2016b), define social marginal welfare weight as a function of agents consumption, earnings, and a set of characteristics that affect social marginal welfare weight and a set of characteristics that affect utility. Chari and Kehoe (1999), besides developing stronger zero-optimal capital income tax rate than Chamley (1986), are developing Barro's (1979) result on tax smoothing, where in deterministic concept, optimal tax rates are constant, while in stochastic economy with incomplete markets tax rates follow a random pattern generated by a martingale process. Auerbach, A. (2009), Kaplow (1994), propose equivalence of consumption taxes and labor taxes: a linear consumption at some inclusive rate, is equivalent to a labor tax income combined with the initial wealth. In this setting consumption tax is equal to labor tax if there is no initial wealth and differences in wealth arise only from wealth preferences.

**Optimal taxation models: Mirrlees (1971)**

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 of the agents in the economy Mirrlees (1971) model:

**Equation 1**

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

Where  $y = \theta l$  and  $\theta$  represents the level of skills of the worker. Now his social welfare function SWF is:  $SWF(v) = \log(v)$ . Now let's find the distribution of skills when  $T(y) = 0.3$  which is Pareto with  $h(y) = ky^{-k-1}y^k$ . Equation for the distribution of skills is

$$f(\theta) = h(y(\theta))y'(\theta), \text{ from the quasi-linear utility functions: } U(c, y, \theta) = c - \frac{1}{2}\left(\frac{y}{\theta}\right)^2. \text{ And}$$

the tax function  $T(y) = \tau y$ , individual with skill level  $\theta$  solves:

**Equation 2**

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

---

5) And if in the presence of optimal income taxation whether if a small commodity tax can be replicated by a small income change, and when this is not a case commodity taxation allows government to expand its own taxation power and therefore it is desirable.

6) Martingale is a sequence of random variables (i.e., a stochastic process) for which, at a particular time, the conditional expectation of the next value in the sequence, given all prior values, is equal to the present value.

FOC is given as :  $(1 - \tau) - \frac{y}{\theta^2} = 0$ , which implies that  $y = (1 - \tau)\theta^2$  and

$$f(\theta) = h(y(\theta))y'(\theta) = k(\theta)^{-k-1}y^k 2(1 - \tau)\theta = k((1 - \tau)\theta^2)^{-k-1}y^k 2(1 - \tau)\theta = 2k(1 - \tau)^{-k}\theta^{-2k-1}y^k = 2k\theta^{-2k-1}\theta_i^{2k}$$

.By integration one could get

$$F(\theta) = \int_{\theta_i}^{\theta} f(\theta) d\theta = \int_{\theta_i}^{\theta} 2k(1 - \tau)^{-k}\theta^{-2k-1}y^k d\theta = [-(1 - \tau)^{-k}\theta^{-2k}y^k]_{\theta_i}^{\theta} = (1 - \tau)^{-k}\theta_i^{-2k-1}\left((1 - \tau)\theta_i^2\right)^k - (1 - \tau)^{-k}\theta^{-2k}y^k = 1 - \theta^{-2k}\theta_i^{2k}$$

.Now we can solve for numerical optimum. Let's use  $y = 2$  and  $k = 4$  and truncate the distribution<sup>7</sup> at the top  $x$  percentile for some small  $x$ .

In this case :  $\max_{v(\theta), u(\theta)} \int_{\theta_i}^{\theta^h} W[v(\theta)]f(\theta) d\theta$ . Subject to :

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

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

Transversality conditions :  $\eta(\theta_i) = \eta(\theta^h) = 0$ . 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 3**

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

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 4**

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

<sup>7</sup> 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 .

In the previous expression  $\tau$  are taxes,  $\bar{g}$  is the ratio of social marginal utility for the top bracket taxpayers to the marginal value of public funds for the government, which depends on the social welfare function <sup>8</sup>. Utility in the social welfare function provides a guideline for the government for achieving optimal distribution of income, Tresch, R. W. (2008) . In the Saez (2001) optimal tax formula also:

**equation 5**

$$\bar{\epsilon}^u = \int_{\bar{w}}^{\infty} \epsilon_w^u w h(w) dw / w_m ,$$

where Marshallian demand for labor is given as  $w = w(1 - \tau, R)$  where  $R$  is the non-labour income, and  $w$  are earnings(wages)<sup>9</sup>. And compensated elasticity of earnings is :

**equation 6**

$$\bar{\epsilon}^c = \frac{1-\tau}{w} \left( \frac{\partial w}{\partial (1-\tau)} \right) \Big|_u .$$

Those two are related by the Slutsky equation :  $\epsilon^c = \epsilon^u - \eta$ , when there are no behavioral responses there is only mechanical effect denote by  $M$  and  $M = [w_m - \bar{w}] d\tau$ , where  $w_m - \bar{w}$  represents the earnings of the agent above medium population earnings. Behavioral responses are equal to :

$dw = -\frac{\partial w}{\partial (1-\tau)} d\tau + \frac{\partial w}{\partial R} dR = -\left( \epsilon^u w - \frac{(1-\tau)\partial w}{\partial R} \right) \left( \frac{d\tau}{1-\tau} \right)$ , or the total behavioral response  $\beta = -\left( \epsilon^u w \frac{(1-\tau)\partial w}{\partial R} \right) \left( \frac{\tau d\tau}{1-\tau} \right)$ . Saez(2001) result for high income earners is given as :

**equation 7**

$$\frac{\tau}{1-\tau} = \frac{(1-\bar{g})(w_m/(\bar{w}-1))}{\bar{w} - \int_{\bar{w}}^{\infty} \eta_w h(w) dw} .$$

In the Mirrlees(1971) model government , maximizes<sup>10</sup> :

**equation 8**

$$SWF = \int_0^{\infty} G(u_w) f(w) dw .$$

In the previous expression  $G(u_w)$  represents the concave utility function<sup>11</sup>. The constraint here is given as:  $\int_0^{\infty} G(u_w) f(w) dw \leq \int_0^{\infty} w_i f(w) dw - E$ , where  $E$  are government expenditures. Now, about Pareto distributions it is well known fact that  $\frac{\text{ratio average}}{\text{threshold}} = \text{constant}$ . Now if we denote the average wage  $w^*(w) > w$ , and if  $w$  is a threshold, then  $w^*(w)$  can be expressed as :

<sup>8</sup> Social welfare function can be :  $SWF = \int U^i di$ -Utilitarian or Benthamite,  $SWF = \min_i U^i$ -

Rawlsian  $SWF = \int U^i di \rightarrow G(U) = \frac{U^{1-\gamma}}{1-\gamma}$  if  $\gamma = 0$  function is utilitarian , Rawlsian if  $\gamma = \infty$ . With

Pareto weights:  $SWF = \int \mu_i U^i di$  where  $\mu_i$  is exogenous.

<sup>9</sup> Income effects are captured through  $\eta = (1 - \tau)\partial w / \partial R$ , average income effects are

$\bar{\eta} = \int_{\bar{w}}^{\infty} \eta_w h(w) dw$

<sup>10</sup> Here we make assumption that wages =skill level

<sup>11</sup> Now, for a concave function  $f:(a, b) \rightarrow R$  is continuous in  $IntA$ . This function  $f:(a, b) \rightarrow R$  is concave in the interval  $(a, b)$ , if for every  $x_1, x_2 \in (a, b), a \in (0, 1)$ , it follows  $f(ax_1 + (1-a)x_2) < af(x_1) + (1-a)f(x_2)$ .

**equation 9**

$$w^*(w) = \int_{w_m > w} wf(w)dw$$

$$/ \int_{w_m > w} f(w)dw = \int_{w_m > w} dw/w^a / \int_{w_m > w} dw/w^{1+a} = \frac{aw}{a-1}$$

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

$$: 1 - F(w) = \left(\frac{k}{w}\right)^\alpha, \text{ and CDF of the function is given as } f(w) = \frac{\alpha k^\alpha}{w^{1+\alpha}},^{12} \text{ that is } \lim_{w \rightarrow \infty} \frac{\left(\frac{k}{w}\right)^\alpha}{\frac{\alpha k^\alpha}{w^{1+\alpha}}}$$

by

$$\lim_{x \rightarrow \alpha} [c \cdot f(x)] = c \cdot \lim_{x \rightarrow \alpha} f(x) \Rightarrow \frac{1}{\alpha k^\alpha} \cdot \lim_{w \rightarrow \infty} \frac{\left(\frac{k}{w}\right)^\alpha}{\frac{\alpha k^\alpha}{w^{1+\alpha}}} = \frac{1}{\alpha k^\alpha} \cdot \lim_{w \rightarrow \infty} (k^\alpha) = \frac{1}{\alpha k^\alpha} \cdot k^\alpha = \frac{1}{\alpha}$$

, hence the formula of marginal income for top earners  $\tau^* = \frac{1}{1+\alpha \cdot \varepsilon}$ . This is the tax rate that maximizes government revenues. The optimal tax government formula with Rawlsian government<sup>13</sup> would be :

**Equation 10**

$$\frac{T'(w(h))}{1-T'(w(h))} = \left(\frac{1+\varepsilon}{\varepsilon}\right) \frac{1-F(w)}{wf(w)} \text{ or } \frac{T'(w(h))}{1-T'(w(h))} = \left(\frac{1+\varepsilon}{\varepsilon}\right) \frac{\psi(w)-F(w)}{wf(w)}$$

Now if we divide and multiply by  $1 - F(w)$  we get :

**equation 11**

$$\frac{T'(w(h))}{1-T'(w(h))} = \left(\frac{1+\varepsilon}{\varepsilon}\right) \frac{\Psi(w)-F(w)}{1-F(w)} \frac{1-F(w)}{wf(w)}$$

In the previous formula  $\left(\frac{1+\varepsilon}{\varepsilon}\right) = A(w)$ , elasticity and efficiency argument,

$$\frac{\Psi(w)-F(w)}{1-F(w)} = B(w), \text{ measures the desire for redistribution : if the sum of weights } \psi(w)f(w)$$

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

$$\varepsilon = \frac{1-\tau}{z} \frac{dz}{d(1-\tau)}, \text{ where } z \text{ is taxable income and } z = y - x, \text{ where } y \text{ is the real income, and } x \text{ is sheltered income }^{14}, \text{ taxable income } s \text{ used in the calculation for Pareto parameter } \alpha = \frac{z}{z-\varepsilon}. \text{ Tax avoidance elasticity component is given as } \varepsilon_1 = \frac{1-\tau}{z} \frac{dx}{1-\tau}, \text{ and } \varepsilon_2 = \frac{1-\tau}{z} \frac{dy}{1-\tau} \text{ is}$$

the real labor supply elasticity. Now, when government raises slightly  $\tau \rightarrow d\tau$  there is:

$$1. \text{ mechanical effect from the increase in taxes i.e. } dM = (z - z^*)d\tau,$$

$$^{12} \frac{\left(\frac{k}{w}\right)^\alpha}{w \left(\frac{\alpha k^\alpha}{w^{1+\alpha}}\right)} = \frac{\frac{k^\alpha}{w^\alpha}}{w \left(\frac{\alpha k^\alpha}{w^{1+\alpha}}\right)} = \frac{\frac{k^\alpha}{w^\alpha}}{w \frac{\alpha k^\alpha}{w^{1+\alpha}}} = \frac{1}{\alpha}$$

<sup>13</sup> 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.

<sup>14</sup> Investments or investment accounts that provide favorable tax treatment, or activities and transactions that lower taxable income.

2. **welfare effect**  $dW = -\bar{g}dM = -\bar{g}(z - z^*)d\tau$ , where social marginal weight for individual is  $g_i = G'(u^i)u_c^i/\lambda$ , where  $\lambda$  is a multiplier of government constraint which is  $\int \tau(wl)f(w)dw \geq E$ , average income in economy is  $\bar{z} = \int zh(z)dz$ ,  $h(z)$  is a density  $\frac{\partial L}{\partial \tau(z)} = \lambda$ ,  $c = \bar{z} - E$ , while  $\bar{g}$  social marginal weight for top earners is given as:  $\bar{g} = \frac{\int g_i z_i}{z^* \cdot \int g_i}$ , where  $z - z^*$  is the mechanical redistribution effect.
3. And, the third effect is **behavioral response** of the top earners: 
$$dB = -\frac{\tau}{1-\tau} \cdot \frac{1-\tau}{z} \cdot \frac{dz}{d(1-\tau)} \cdot z \cdot d\tau = -\frac{\tau}{1-\tau} \cdot \varepsilon \cdot z \cdot d\tau.$$

From here it can be derived Diamond (1998) optimal tax formula :

**equation 12**

$$\frac{\tau(z)}{1-\tau(z)} = \frac{1}{\varepsilon(z)} \cdot \left[ \frac{1-F(z)}{z \cdot F(z)} \right] \cdot [1 - G(z)],$$

this is distribution shape parameter  $\frac{1-F(z)}{z \cdot F(z)}$ ,  $G(z)$  are social marginal welfare weights For numerical solutions of the Mirrless model (1971), one can look up to Brewer, M., E. Saez, and A. Shephard (2010),

**Equation 13**

$$\frac{\tau' z(h)}{1 - \tau' z(h)} = \left( 1 + \frac{1}{\varepsilon} \right) \frac{1}{h f(h)} \int_h^\infty \left( 1 - \frac{G' u(h)}{\lambda} \right) f(h) dh$$

Where  $\lambda = \int G'(u)dh$ . Few general conclusions about marginal tax rates in the literature appear:

1.  $\tau(z(h)) \geq 0$ , a in Mirrlees (1971),
2.  $\tau(z(\text{highest } h)) = 0$  ( $f(h)$  bounded above, Sadka, (1976),
3.  $\tau(z(\text{lowest } h)) = 0$  ( $h: y(h) > 0$ ), Seade, (1977).

The effect of small tax reform in Mirrless (1971) model is examined in Brewer, M., E. Saez, and A. Shephard (2010), where indirect utility function is given as  $U(1 - \tau, R) = \max_z ((1 - \tau)z + R, z)$ , where  $z$  represents the taxable income  $R$  is a virtual income intercept, and  $\tau$  is an imposed income tax. Marshallian labor supply is

$z = z(1 - \tau, R)$ , uncompensated elasticity of the supply is given as:  $\varepsilon^u = \frac{(1-\tau)}{z} \frac{\partial z}{\partial (1-\tau)}$ ,

income effect is  $\eta = (1 - \tau) \frac{\partial z}{\partial R} \leq 0$ . Hicksian supply of labor is given as:  $z^c((1 - \tau, u))$ , this

minimizes the cost in need to achieve slope  $1 - \tau$ , compensated elasticity now is  $\varepsilon^c = \frac{(1-\tau)}{z} \frac{\partial z^c}{\partial (1-\tau)} > 0$ , Slutsky equation now

becomes:  $\frac{\partial z}{\partial (1-\tau)} = \frac{\partial z^c}{\partial (1-\tau)} + z \frac{\partial z}{\partial R} \Rightarrow \varepsilon^u = \varepsilon^c + \eta$ , where  $\eta$  represents income effect

$\eta = (1 - \tau) \frac{\partial z}{\partial R} \leq 0$ . With small tax reform taxes and revenue change

i.e.:  $dU = u_c \cdot [-zdt + dR] + dz[(1 - \tau)u_c + u_z] = u_c \cdot [-zdt + dR]$ . Change of taxes and its impact on the society is given as:  $dU_i = -u_c dT(z_i)$ . Envelope theorem here says

$U(\theta) = \max_x F(x, \theta), s. t. c > G(x, \theta)$ , and the preliminary result is

$U'(\theta) = \frac{\partial F}{\partial \theta} (x^*(\theta), \theta - \lambda^*(\theta) \frac{\partial G}{\partial \theta} x^*(\theta), \theta)$ . Government is maximizing  
 $0 = \int G'(u^i) u_c^i \cdot \left[ (Z - z^i) - \frac{\tau}{d(1-\tau)} eZ \right]$ , mechanical effect is given as:  $dM = [z - z^*] d\tau$ ,  
welfare effect is:  $dW = -\bar{g} dM = -\bar{g} [z - z^*]$ , and at last the behavioral response is  
 $dB = -\frac{\tau}{1-\tau} \cdot e \cdot z d\tau$ . And lets denote that:

**Equation 14**

$$dM + dW + dB = d\tau \left[ 1 - \bar{g} [z - z^*] - e \frac{\tau}{1-\tau} \cdot z \right]$$

When the tax is optimal these three effects should equal zero i.e.  $dM + dW + dB = 0$  given that:  $\frac{\tau}{1-\tau} = \frac{(1-\bar{g})[z-z^*]}{e \cdot z}$ , and we got  $\tau = \frac{1-\bar{g}}{1-\bar{g}+\alpha \cdot e}$ ,  $\alpha = \frac{z}{z-z^*}$ , and  
 $dM = d\tau [z - z^*] \ll dB = d\tau \cdot e \frac{\tau}{1-\tau} \cdot z$ , where  $z^* > z^T$ , where  $z^T$  is a top earner income.  
Pareto distribution is given as:

**Equation 15**

$$1 - F(z) = \left( \frac{k}{z} \right)^\alpha, f(z) = \alpha \cdot \frac{k^\alpha}{z^{1+\alpha}}$$

$\alpha$  is a thickness parameter and top income distribution is measured as:

**Equation 16**

$$z(z^*) = \frac{\int_{z^*}^{\infty} s f(s) ds}{\int_{z^*}^{\infty} f(s) ds} = \frac{\int_{z^*}^{\infty} s^{-\alpha} ds}{\int_{z^*}^{\infty} s^{-\alpha-1} ds} = \frac{\alpha}{\alpha-1} \cdot z^*$$

Empirically  $\alpha \in [1.5, 3]$ ,  $\tau = \frac{1-\bar{g}}{1-\bar{g}+\alpha \cdot e}$ . General non-linear tax without income effects is given as:

**Equation 17**

$$\frac{T'(z_n)}{1 - T'(z_n)} = \frac{1}{e} \left( \frac{\int_n^{\infty} (1 - g_m) dF(m)}{z_n h(z_n)} \right) = \frac{1}{e} \left( \frac{1 - H(z_n)}{z_n h(z_n)} \right) \cdot (1 - G((z_n)))$$

Where  $G((z_n)) = \frac{\int_n^{\infty} g_m dF(m)}{1 - F(n)}$ , and  $g_m = G'(u_m)/p$  this is welfare weight of type  $m$ . But non-linear tax with income effect takes into account small tax reform where tax rates change from  $d\tau$  to  $[z^*, z^* + dz^*]$ . Every tax payer with income  $z$  above  $z^*$  pays additionally  $d\tau dz^*$  valued by  $(1 - g(z)) d\tau dz^*$ . Mechanical effect is:

**Equation 18**

$$M = d\tau dz^* \int_{z^*}^{\infty} (1 - g(z)) dz$$

Total income response is:  $I = d\tau dz^* \int_{z^*}^{\infty} \left( -\eta \frac{T'(z)}{1 - T'(z)} (z) \right) h(z) dz$ . Change at the taxpayers form the additional tax is:  $dz = -\varepsilon_{(z)}^{\tau} \frac{T'' dz}{1 - T'} - \eta \frac{d\tau dz^*}{1 - T'(z)} \Rightarrow -\eta \frac{d\tau dz^*}{1 - T'(z) + z \varepsilon_{(z)}^{\tau} T''(z)}$ , if one sums up all effects can be obtained:

**Equation 19**

$$\frac{T'(z)}{1 - T'(z)} = \frac{1}{\varepsilon_{(z)}^{\tau}} \left( \frac{1 - H(z^*)}{z^* h(z^*)} \right) \times \left[ \int_{z^*}^{\infty} (1 - g(z)) \frac{h(z)}{1 - H(z^*)} dz + \int_{z^*}^{\infty} -\eta \frac{T'(z)}{1 - T'(z)} \frac{h^*(z)}{1 - H(z^*)} dz \right]$$

With linear tax:  $\frac{\dot{z}_n}{z_n} = \frac{1 + \varepsilon_{(n)}^u}{n}$  and with non-linear tax:

**Equation 20**

$$\frac{\dot{z}_n}{z_n} = \frac{1 + \varepsilon_{(n)}^u}{n} - \dot{z}_n \frac{T''(z_n)}{1 - T''(z_n)} \varepsilon_{z(n)}^c$$

This model was later augmented with the migrations by Mirrless (1982). Migrations are of importance for the top incomes (brain drain). In the model earnings are fixed  $c$ , and  $p(c|z)$  represents the number of residents earning  $z$ , while  $c = z - T(z)$  represents the disposable income. Now, one small tax reform  $dT(z)$ , for those earning income  $z$ . Mechanical effect of net-welfare is:  $M + W = (1 - g(z))P(c|z)dT$ . Migration equal taxes average or total:  $M + W = \frac{\partial P(c|z) z - T(z)}{\partial c} \frac{z - T(z)}{P(c|z)}$ . Cost of imposing taxes are:  $B = -\frac{T(z)}{z - T(z)}$ . Optimal tax applies when:  $M + W + B = 0$ . And the formula for the optimal tax with migrations becomes:

**Equation 21**

$$\frac{T(z)}{z - T(z)} \frac{1}{\eta_m(z)} \cdot (1 - g(z))$$

$\eta_m(z)$  (elasticity of taxable top income) depends on the size of jurisdiction; it's large for the cities, and its small (zero) for the world, redistribution is easier in larger jurisdictions. Formula for maximizing the revenues from top incomes is:

**Equation 22**

$$\tau = \frac{1}{1 + a \cdot e + \bar{\eta}^m}$$

Where  $\bar{\eta}^m$  is elasticity of the top earners towards disposable income.

**Diamond-Mirrlees efficiency theorem Diamond-Mirrlees (1971a), Diamond-Mirrlees (1971b)**

Setup of the model here is:  $p$ -producer prices,  $q$ -consumer prices,  $t = q - p$  taxes

1.  $x^h(q)$ -net demand by consumer  $h$  (incomes are equal to zero)  $h = 1, 2, \dots, H$
2.  $u^h(x^h)$  utility function of consumer  $h$
3.  $v^h(q)$  indirect utility function  $v^h(q) = u^h(x^h(q))$
4.  $X(q) = \sum_h x^h(q)$  aggregate net demand
5.  $U(x^1, \dots, x^H)$  social welfare function
6.  $V(q) = U(x^1(q), \dots, x^H(q))$  indirect utility function
7.  $W(u^1(x^1), \dots, u^H(x^H))$  special case of an individualistic social welfare function

The aim here is to Maximize  $V(q)$

s.t.  $G(X(q)) \leq 0$

where  $G$  aggregate production constraint

**equation 23**

$$\frac{\partial V}{\partial q_k} = \lambda \sum p_i \frac{\partial X_i}{\partial q_k} = -\lambda \frac{\partial}{\partial t_k} (\sum t_i X_i) \quad k = 2, \dots, n$$

**equation 24**

$$\sum_{k=1}^n \left[ \frac{\partial V}{\partial q_k} - \lambda \sum p_i \frac{\partial X_i}{\partial q_k} \right] q_k = 0$$

Optimal tax structure for one consumer economy is presented as:

Slutsky equation is given as:

**equation 25**

$$\frac{\partial x_i}{\partial q_k} = s_{ik} - x_k \frac{\partial x_i}{\partial I}$$

$s_{ik}$  -derivative of compensated curve for  $i$  with respect to  $q_k$  and  $\frac{\partial x_i}{\partial I}$  is the derivative of the uncompensated demand curve

**equation 26**

$$-\alpha x_k = -\lambda \frac{\partial}{\partial t_k} (\sum t_i X_i) = -\lambda \left( x_k + \sum t_i \frac{\partial x_i}{\partial t_k} \right) = -\lambda x_k - \lambda \sum t_i s_{ik} + \lambda x_k + \sum t_i \frac{\partial x_i}{\partial I}$$

$$k = 1, 2, \dots, n$$

**equation 27**

$$\frac{\sum_i t_i s_{ik}}{x_i} = \frac{\alpha + \lambda - \lambda \sum t_i \frac{\partial x_i}{\partial I}}{\lambda} \quad ; \quad \frac{\sum_i t_i s_{ik}}{x_i} = -\theta$$

**equation 28**

$$\theta \sum_k t_k x_k = - \sum_{k,i} t_k s_{ki} t_i \geq 0$$

**equation 29**

$$\Delta x_k = \sum_i \int_0^{t_i} \frac{\partial x_k}{\partial t_i} dt_i = \sum_i \int_0^{t_i} s_{ki} dt_i = \sum_i s_{ki} \int_0^{t_i} dt_i = \sum_i s_{ki} t_i$$

**equation 30**

$$\sum_i \frac{\partial x_k}{\partial t_i} t_i + \frac{\partial x_i}{\partial I} \sum t_i x_i = -\theta x_k$$

**equation 31**

$$\frac{\sum_i \frac{\partial x_k}{\partial t_i} t_i}{x_k} = \theta - x_k^{-1} \frac{\partial x_i}{\partial I} \sum t_i x_i$$

Optimal tax formula is given as:

**equation 32**

$$q_k x_k = \sum_j [b_k \omega_j q_j - b_j \omega_k q_k]$$

**equation 33**

$$-\alpha q_k x_k = \lambda \left[ \sum_{j \neq k} b_j \frac{p_j}{q_j} \omega_k q_k - b_k \frac{p_k}{q_k} \sum_{j \neq k} \omega_j q_j \right] = \lambda \sum_j \left[ b_j \omega_k \frac{p_j q_k}{q_j} - b_k \omega_j \frac{p_k q_j}{q_k} \right]$$

$$\sum b_j = 1$$

**equation 34**

$$\sum_j \left[ b_j \omega_k q_k \left( \frac{p_j}{q_j} - \frac{\alpha}{\lambda} \right) - b_k \omega_j q_j \left( \frac{p_k}{q_k} - \frac{\alpha}{\lambda} \right) \right] = 0$$

Optimal taxation many consumer economy is given as:

**equation 35**

$$V(q) = W(v^1(q), \dots, v^H(q))$$

Differentiating with respect to  $q_k$  we have :

**equation 36**

$$V_k = \sum_h \frac{\partial W}{\partial u^h} v_k^h = - \sum_h \frac{\partial W}{\partial u^h} \alpha^h x_k^h$$

$\alpha^h$ -marginal utility of income

$\beta^h = \frac{\partial W}{\partial u^h} \alpha^h$  increase in social welfare form an increase in the income of consumer h

**equation 37**

$$-V_k = \sum_h \beta^h x_k^h$$

**equation 38**

$$\sum_h \beta^h x_k^h = \lambda \frac{\partial T}{\partial t_k}$$

$T = \sum_i t_i X_i$  -total revenue from taxes

**equation 39**

$$\sum_h \beta^h x_k^h = -\lambda \sum_i p_i \frac{\partial X_i}{\partial q_k}$$

In a Cobb-Douglas production function case we have :

**equation 40**

$$u^h = b_1^h \log(x_1^h + \omega^h) + \sum_2^n b_i^h \log x_i^h; \sum b_i = 1$$

Where  $x_i^h = q_1^{-1} b_i^h q_1 \omega^h; i = 1, 2, 3 \dots, n; x_1^h = -(1 - b_1^h) \omega^h$

**equation 41**

$$\frac{\partial x_i^h}{\partial q_k} = 0; k \neq i \neq 1 \quad \frac{\partial x_i^h}{\partial q_i} = -\frac{x_i^h}{q_i}$$

**equation 42**

$$\sum_h \beta^h x_k^h = \lambda p_k q_k^{-1} \sum_h x_k^h \quad k = 2, \dots, n$$

**equation 43**

$$\frac{q_k}{p_k} = \lambda \frac{\sum_h x_k^h}{\sum_h \beta^h x_k^h} = \lambda \frac{b_k^h q_1 \omega^h}{\beta^h b_k^h \omega^h}; k = 2, \dots, n$$

**equation 44**

$$-p_i \sum_h (1 - b_1^h) \omega^h + \sum_{i=2}^n \sum_h p_i q_i^{-1} b_i^h q_1 \omega^h = \gamma$$

$\gamma = \sum_{i=1}^n p_i z_i$  -maximized profit of production of net government needs

**equation 45**

$$\frac{q_1}{p_1} = \lambda \frac{\sum_h (1-b_1^h) \omega^h + \gamma p_1^{-1}}{\sum_{i=2}^n \sum_h \beta^h b_i^h \omega^h} = \lambda \frac{\sum_h (1-b_1^h) \omega^h + \gamma p_1^{-1}}{\sum_{i=2}^n \sum_h \beta^h (1-b_1^h) \omega^h}$$

$\gamma p_1^{-1}$  -depends on government expenditure decision and technology

Optimum taxes have to satisfy

**equation 46**

$$\frac{q_k}{p_k} \sum_h b_k^h (\omega^h)^{-\mu} \prod_{i=2}^n (b_i^h)^{\mu b_i^h} q^{\mu b_i^h} = \lambda \sum_h b_k^h \omega^h \quad (k = 2, 3, \dots, n)$$

**equation 47**

$$W = -\mu^{-1} \sum_h e^{-(\mu v^h)}$$

Optimal tax formulae here is :

**equation 48**

$$\sum_h \beta^h x_k^h = \lambda \sum_h \sum_i t_i \frac{\partial x_i^h}{\partial q_k} + \lambda \sum_h x_k^h$$

**equation 49**

$$\frac{\partial x_i}{\partial q_k} = s_{ik} - x_k \frac{\partial x_i}{\partial I} \quad \text{Slutsky equation}$$

**equation 50**

$$\frac{\partial x_i}{\partial q_k} = s_{ik} - x_k \frac{\partial x_i}{\partial I} = s_{ki} - x_k \frac{\partial x_i}{\partial I} = \frac{\partial x_k}{\partial q_i} - x_k \frac{\partial x_i}{\partial I} + x_i \frac{\partial x_k}{\partial q_i}$$

Next Diamond-Mirrlees theorem is put against efficiency. On the Consumption side

$z = z_1, z_2, \dots, z_n$  the vector of government supply Government

s.t.  $G(z) \leq 0$  ;  $z \leq g(z_2, z_3, \dots, z_n)$

$x(q) = z$  consumer demands  $x = (x_1, \dots, x_n)$  and prices faced by the consumers

$q = (q_1, \dots, q_n)$

$\max v(q)$  objective function s.t.  $G(x(q)) \leq 0$

$v(q) = u(x(q))$  ;  $v_k = \partial v / \partial q_k$  ;  $u_i = \partial u / \partial x_i$

$v_k = \sum u_i \frac{\partial x_i}{\partial q_k} = -\alpha x_k$  s.t. budget constraint  $\sum q_i x_i = 0$

$x_k + \sum q_i \frac{\partial x_i}{\partial q_k} = 0$  utility maximization implies  $u = \alpha q_i$

On the Production side:

$p = (p_1, \dots, p_n)$  producers prices they differ from prices faced by the consumers

$q_i = p_i + t$  ;  $y = y_1, \dots, y_n$  vector of privately supplied commodities

$y_i = f(y_2, \dots, y_n)$  equality on the production constraint

$p_i = -p_1 f_1(y_2, \dots, y_n)$

Constant returns to scale maximized profits are zero in equilibrium :  $\sum p_i y_i = 0$

$z = g(z_2, z_3, \dots, z_n)$  government constrain satisfied with equality

Walras' law implies that:

$x(q) = y + z$  all markets clear, Walras' law implies that if all agents satisfy their budget constraint and all markets but one are in equilibrium then the last market is also in equilibrium.

$\sum (q_i - p_i) x_i + \sum p_i z_i = 0 = \sum t_i x_i + \sum p_i z_i$  in normalization  $p_i = 1$  ;  $q_i = 1$  ;  $t = 0$

Thus levying a positive tax  $t > 0$  on consumer transactions is in fact subsidy and results in loss of revenues by the government.

Welfare maximization requires that :

$$y_2 = x_2 - z_2; \dots; y_n = x_n - z_n; \text{ s.t. } x_1(q) = y_1 + z_1 = f(x_2 - z_2, \dots, x_n - z_n)$$

$$\max v(q) \text{ s.t. } x_1(q) - f(x_2(q) - z_2, \dots, x_n(q) - z_n) - g(z_2, \dots, z_n) = 0$$

$$L = v(q) - \lambda [x_1(q) - f(x_2 - z_2, \dots, x_n - z_n) - g(z_2, \dots, z_n)] \text{ -Lagrangian}$$

Differentiate with respect to  $q_k$

$$v_k - \lambda \left( \frac{\partial x_1}{\partial q_k} - \sum_{i=2}^n f_i \frac{\partial x_i}{\partial q_k} \right) = 0 \quad k = 2, 3, \dots, n$$

Differentiating  $L$  to  $z_k$  we have

$$\lambda (f_k - g_k) = 0 \quad k = 2, 3, \dots, n$$

$\lambda \neq 0$  social cost associated with the marginal need for additional resources (i.e. equality of marginal rates of transformation):

$$MRT_{f_k} = MRT_{g_k}$$

Optimal tax structure is :

**equation 51**

$$\frac{\partial x_i}{\partial q_k} = \frac{\partial x_i}{\partial t_k}$$

$$v_k = \lambda \sum p_i \frac{\partial x_i}{\partial t_k} = \lambda \frac{\partial}{\partial t_k} \sum p_i x_i$$

$$\sum p_i x_i = \sum q_i x_i - \sum t_i x_i = -\sum t_i x_i$$

$$v_k = -\lambda \frac{\partial}{\partial t_k} \sum t_i x_i$$

$$x_k = \frac{\lambda}{\alpha} \frac{\partial (\sum t_i x_i)}{\partial t_k}$$

The ratio of marginal tax revenue from an increase in tax on that commodity is:

**equation 52**

$$\sum_h \beta^h x_k^h = \lambda \sum_h \sum_i t_i \frac{\partial x_k^h}{\partial q_i} + \lambda \sum_h \sum_i t_i \left( x_i^h \frac{\partial x_k^h}{\partial I} - x_k^h \frac{\partial x_i^h}{\partial I} \right) + \lambda \sum_h x_k^h$$

**equation 53**

$$\frac{\sum_h \sum_i t_i \frac{\partial x_i^h}{\partial q_i}}{\sum_h x_k^h} = \frac{1}{\lambda} \frac{\sum_h \beta^h x_k^h}{\sum_h x_k^h} - 1 + \frac{\sum_h (\sum_i t_i \frac{\partial x_i^h}{\partial I}) x_k^h}{\sum_h x_k^h} - \frac{\sum_h (\sum_i t_i x_i^h) \frac{\partial x_k^h}{\partial I}}{\sum_h x_k^h}$$

The ratio of marginal tax revenue from an increase in tax on that commodity to the quantity of the commodity is constant. Production efficiency in many consumer economy requires:

1.  $H$ -households  $u^h; x^h \quad h = 1, 2, \dots, H$
2.  $V(q) = U(x^1(q), x^2(q), \dots, x^n(q))$ -objective function
3.  $V(q) = W(u^1(x^1(q)), u^2(x^2(q)), \dots, u^n(x^n(q)))$ -Welfare function strictly increasing
4.  $\max V(q) \text{ s.t. } G(X(q)) \leq 0$
5.  $X(q) = \sum_h x^h(q) \quad G$  represents the production set

The efficiency argument requires that aggregate production functions  $X(q)$  are continuous then any small change in prices  $q$  will not change aggregate production requirements. Welfare can be written with a transfer  $\tau$   $V(q, \tau)$  s.t.  $X(q, \tau)$  being in  $G(q^*, \tau^*)$ . No commodity taxes are possible but poll tax is possible, A poll tax, also known as head tax or capitation, is a tax levied as a fixed sum on every liable individual. And part of the productions is privately controlled in such a way that it is uniquely determined by producer prices. Maximization here states that:

**equation 54**

$\max(p, \tau)$   $\tau$  -poll tax or subsidy

s.t.  $X(p, \tau) - y(p) = z$

where  $y(p)$  is private production vector and  $\tau < \tau^*$  where :

**equation 55**

$V(p^*, \tau) > V(p^*, \tau^*)$ ;  $X(p^*, \tau) - y(p^*)$  is not in  $G$ ,  $z^*$  is efficient in  $G$ .

Here  $H$ -households,  $x$ -consumption vector,  $C$ -consumption set s.t. budget constraint  $q \cdot x \leq 0$ ;  $C^h$  closed convex and bounded below by a vector  $\alpha^h$ , and contains vector with every component negative, the preference ordering is continuous, formally if  $x^2$  is preferred or indifferent to  $x^1$  and  $0 < t < 1$  then  $tx^2 + (1-t)x^1 > x^1$ , there is no satiation consumption in  $C^h$  (The economic principle of satiation is the effect whereby the more of a good one possesses the less one is willing to give up in order to get more of it).

If  $z \leq 0$ ;  $z \in G$ ;  $G$  is closed,  $\exists \bar{\alpha}$ ;  $z \leq \bar{\alpha} \forall z > 0$ ; in closure  $G$ ,  $G$  is convex.

$U(x^1, \dots, x^H)$  indirect utility function  $V(q) = U(x^1(q), \dots, x^H(q))$

Individual indirect utility function:  $v^h(q) = u^h(x^h(q))$

**equation 56**

$U(x^1, \dots, x^H) = W(u^1(x^1), \dots, u^H(x^H))$

Aggregate demand is :

**equation 57**

$$X(q) = \sum_h x^h(q)$$

The assumptions on the production possibility set are:

1.  $z \leq 0, z \in G$  every production plan in which nothing is produced in positive quantity
2.  $0 \in G$  i.e. complete inactivity
3.  $G$  is closed
4.  $\exists \bar{\alpha}; z \leq \bar{\alpha}$  in the convex closure of  $G$
5.  $G$  is convex

Indirect welfare function is given as:

**equation 58**

$V(q) = U(x^1(q), \dots, x^H(q))$

Individual indirect utility:

$U(x^1, \dots, x^H) = W(u^1(x^1), \dots, u^H(x^H))$

$U$  is a continuous function of  $(x^1, \dots, x^H)$

$\|x\| > M$  where all the attainable vectors in the set satisfy  $\|x\| < M$  we restrict attention to

**equation 59**

$q \geq 0 \sum_i q_i = 1$

$\{q | X(q) \in G\}$  this set is closed

**equation 60**

$$x' = X(q') ; (\|x'_n\|/\|x_n\|) \|\xi\| > \|\bar{\alpha}\|$$

**equation 61**

$$b = \sum_h \alpha_h ; x_n = \sum_h x_n^h$$

**equation 62**

$$(\|x'_n\|/\|x_n\|)x_n \cong (\|x'_n\|/\|x_n\|)b$$

$$\xi \cong 0$$

**equation 63**

$$V'(q) \cdot q = \sum \frac{\partial V}{\partial q_k} q_k = 0$$

$V$  is a homogenous of degree zero in  $q$ . And  $V' \cong 0$  if and only if  $\frac{\partial V}{\partial q_k} = 0$ .

**Mankiw problem**

An open economy has the production function  $y = f(k)$ , where  $y$  is output per worker and  $k$  is capital per worker. The capital stock adjusts so that the after-tax marginal product of capital equals the exogenously given world interest rate  $r$  the question here is how much will the tax cut increase wages? One solution here is Cobb-Douglas production function. Where:  $Y$ -the total output ;  $F(K,L)$ - production function ;  $K$ -total capital stock ;  $T$ -total taxes ;  $t$ -taxes per worker ;  $\tau$ -capital tax rate

**equation 64**

$$Y = F(K, L) = f(k)L, k = \frac{K}{L} ; Y = AK^\alpha L^{1-\alpha} = AK^\alpha L, \alpha \in [0,1]$$

With taxes, firms look to maximize:

**equation 65**

$$J = \max_{K,L} (1 - \tau)[F(K, L) - wL] - rK \quad J = \max_{K,L} (1 - \tau)[f(k) - wL] - rK$$

Where  $\tau$  is a capital tax rate,  $w$  is a wage per worker,  $r$  is exogenous tax rate. Then FOC then becomes:

**equation 66**

$$\frac{\partial J}{\partial K} = (1 - \tau)f'(k) - r ; \frac{\partial J}{\partial L} = (1 - \tau)[f'(k) - w]$$

Therefore, the optimality conditions imply that:

**equation 67**

$$r = (1 - \tau)f'(k) ; w = f(k) - f'(k)k$$

Total taxes are:  $T = \tau[F(K, L) - wL]$  ; taxes per worker are :

$$t = \tau[f(k) - w] = \tau f'(k)k$$

Then the change in taxes per worker, given the tax rate, is:

**equation 68**

$$\frac{dt}{d\tau} = f'(k)k + \tau[f'(k) + f''(k)k] \frac{dk}{d\tau} ;$$

$$\frac{dw}{d\tau} = [f'(k) - f'(k) - f''(k)k] \frac{dk}{d\tau} = -f''(k)k \frac{dk}{d\tau}$$

To find the change in wages, we differentiate the second optimality condition for the per worker wages to obtain:

**equation 69**

$$\frac{dr}{d\tau} = 0 \quad ; \quad \frac{dk}{d\tau} = \frac{f'(k)}{(1-\tau)f''(k)} \quad ; \quad \frac{dt}{d\tau} = f'(k)k \quad ; \quad \frac{dw}{d\tau} = \frac{f'(k)k}{1-\tau} = -\frac{1}{1-\tau} \frac{dt}{d\tau} \quad ; \quad \frac{dw}{d\tau} = -\frac{1}{1-\tau}$$

**equation 70**

$$\begin{aligned} \frac{dt}{d\tau} &= f'(k)k + \tau[f'(k) + f''(k)k] \quad ; \quad \frac{dk}{d\tau} = f'(k)k + \frac{\tau}{1-\tau} \frac{[f'(k)]^2 + f'(k)f''(k)k}{f''(k)} \\ &= \frac{\tau}{1-\tau} \frac{f'(k)^2}{f''(k)} + \frac{1}{1-\tau} f'(k)k = \frac{f'(k)}{1-\tau} \left[ \tau \frac{f'(k)}{f''(k)} + k \right] \end{aligned}$$

The static effect of capital tax cut is :

$$\frac{dw}{d\tau} = -\frac{1}{1-\tau} \frac{dt}{d\tau} \quad ; \quad \frac{dw}{d\tau} = \frac{\frac{f'(k)k}{1-\tau}}{\frac{f'(k)[\tau \frac{f'(k)}{f''(k)} + k]}{1-\tau}} = -\frac{1}{\tau \frac{f'(k)}{k f''(k)} + 1}$$

Using the Cobb-Douglas production function, we have that:

**equation 71**

$$\frac{f'(k)}{k f''(k)} = -\frac{1}{1-\alpha} \quad ; \quad \frac{dw}{d\tau} = -\frac{1-\alpha}{1-\tau-\alpha}$$

Therefore, the **dynamic** effect of a capital tax cut on wages is:

**equation 72**

$$\frac{dw}{dt} = -\frac{1-\alpha}{1-\tau-\alpha}$$

This result shows that in small open economy, labor bears 100% of small capital income tax, The fact that, starting at a positive tax rate, the burden of a tax increase exceeds revenue collection due to the first-order deadweight loss

**Conclusion**

This paper made attempt to review the past and the current literature on the optimal tax theory, empirical and theoretical. The developments of the tax theory are improving the tax policies around the world. One of the most important improvements is the worldwide trend towards reduction of taxation of capital income. Also, the worldwide tendency toward flatter income tax rates. The motivation of the original Mirrlees (1971) paper was to provide a framework for which to derive an optimal structure of tax rates, which turned out to be flat for a broad range. 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." He was surprised by his conservative conclusions. Those conclusions are: 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 work and workers ability and the income from such labor. The problem lies here as prof.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". This is contradictory to the requirement of the optimal tax,i.e. that optimal tax must not have effect on the wealth of the society where it has being levied. The sum of mechanical effect, welfare effect, and behavioral effect should equal to zero.

**References:**

1. Atkinson, A.B. and Stiglitz, J. (1976) "The design of tax structure: Direct versus indirect taxation", *Journal of Public Economics*, Vol. 6, 1976, 55-75. (web)
2. Auerbach, A. (2009). "The choice between income and consumption taxes: A primer". Cambridge: Harvard University Press.
3. Auerbach, A., Kotlikoff, L. (1987a). "Evaluating Fiscal Policy with a Dynamic Simulation Model", *American Economic Review*, May 1987, 49-55
4. Auerbach, A., J., Kotlikoff, L. J., (1987b), "Dynamic fiscal policy", Cambridge University Press
5. Barro, R., J. (1979), "On the determination of public debt", *Journal of political economy*, 87:940-971
6. Brewer, M., E. Saez, and A. Shephard (2010) "Means Testing and Tax Rates on Earnings", in *The Mirrlees Review: Reforming the Tax System for the 21st Century*, Oxford, University Press, 2010. (web)
7. Chamley, C. (1986). "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives". *Econometrica*. 54 (3): 607-622. doi:10.2307/1911310.
8. Chari, V.V., Kehoe, P.J., (1999). "Optimal fiscal and monetary policy". *Handbook of macroeconomics*, 1, pp.1671-1745.
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.
10. Diamond, P., Helms, J. and Mirrlees (1978) "Optimal taxation in a stochastic economy, A Cobb-Douglas example", M.I.T. Working Paper no. 217
11. Farhi, E., Werning, I. (2010): "Progressive Estate Taxation," *Quarterly Journal of Economics*, 125 (2), 635-673. [1852, 1853, 1864-1866]
12. Feldstein, M. (1978), "The Welfare Cost of Capital Income Taxation", *Journal of Political Economy*, Vol. 86, No. 2, Part 2: Research in Taxation (Apr., 1978), pp. S29-S51
13. Judd, K.L. (1985) "Redistributive taxation in a simple perfect foresight model". *Journal of Public Economics* 28, 59-83
14. Judd, K.L. (1999) "Optimal Taxation and Spending in General Competitive Growth Models" *Journal of Public Economics* 71: 1-26
15. Kaplow, L. (1994). "Taxation and risk taking: A general equilibrium perspective". *National Tax Association* 47 (4), 789-798.
16. Mankiw NG, Weinzierl M, Yagan D. (2009), "Optimal Taxation in Theory and Practice". *Journal of Economic Perspectives*. 2009;23 (4) :147-174.
17. Mirrlees, J. A. (1971). "An Exploration in the Theory of Optimum Income Taxation". *The Review of Economic Studies*. 38 (2): 175-208. doi:10.2307/2296779. JSTOR 2296779
18. Mirrlees, J.A. (1982), "Migration and Optimal Income Taxes", *Journal of Public Economics*, Vol. 18, 1982, 319-341. (web)
19. Mirrlees, J.A., Diamond, P. (1971b). "Optimal Taxation and Public Production II: Production Efficiency". *American Economic Review*. 61: 261-278.
20. Mirrlees, J.A., (1986), "The theory of optimal taxation", Chapter 24 *Handbook of Mathematical Economics*, edited by K. J. Arrow and M. Intriligator, eds. (Amsterdam, North-Holland, 1986)
21. Mirrlees, J.A., Diamond, P. (1971a). "Optimal Taxation and Public Production I: Production Efficiency". *American Economic Review*. 61: 8-27.
22. Ordovery, J., Phelps, E. (1979). "The concept of optimal taxation in the overlapping generations model of capital and wealth". *Journal of Public Economics* 12: 1-26.
23. Piketty, T., Saez, E., (2013) "A Theory of Optimal Inheritance Taxation", *Econometrica*, 81(5), 2013, 1851-1886.
24. 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 Association, vol. 6(1), pages 230-71
25. Ramsey, F. (1927). "A Contribution to the Theory of Taxation". *Economic Journal*. 37: 47-61. doi:10.2307/2222721.
  26. 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
  27. Saez, E. (2001). "Using Elasticities to Derive Optimal Income Tax Rates". *Review of Economic Studies*. 68: 105–229. doi:10.1111/1467-937x.00166.
  28. Saez, E., S. Stantcheva (2016a). "A Simpler Theory of Optimal Capital Taxation", NBER Working Paper 22664, 2016.
  29. Saez, E., S. Stantcheva (2016b). "Generalized social marginal welfare weights for optimal tax theory". *The American Economic Review*, 106(1), pp.24-45.
  30. Saez, E. (2001). "Using elasticities to derive optimal income tax rates", *The Review of Economic Studies*, 68(1), pp.205-229.
  31. Salanie, B. (2003). "The Economics of Taxation". The MIT Press Cambridge, Massachusetts London, England
  32. Seade, J. K., (1977), "On the shape of optimal tax schedules", *Journal of Public Economics*, 7, issue 2, p. 203-235.
  33. Stiglitz, J. (1982) "Self-Selection and Pareto Efficient Taxation", *Journal of Public Economics*, 1982, 17, 213–240.
  34. Tresch, R. W. (2008). "Public Sector Economics". 175 Fifth Avenue, New York, NY 10010: PALGRAVE MACMILLAN. p. 67. ISBN 978-0-230-52223-7.
  35. Tuomala, M. (1990). "Optimal Income Tax and Redistribution". New York: Oxford University Press.
  36. Varian, H. R. (1980). "Redistributive taxation as social insurance". *Journal of Public Economics*, 14(1), pp.49-68.