Петта Меѓународна Научна Конференција
ПРЕДИЗВИЦИТЕ ВО ТУРИЗМОТ И БИЗНИС ЛОГИСТИКАТА ВО 21 ВЕК »ISCTBL 2022«
Fifth International Scientific Conference
CHALLENGES OF TOURISM AND BUSINESS LOGISTICS IN THE 21ST CENTURY »ISCTBL 2022«

UDK 336.226.111:336.225]:303.725.3
DOI https://www.doi.org/10.46763/YFNTS2251016j

OPTIMAL LINEAR AND NON-LINEAR LABOR INCOME TAXATION: A CRITICAL SURVEY

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

[1] Assistant professor, UGD-Shtip, R.North Macedonia, dusko.josevski@ugd.edu.mk

[2] Associate professor, Goce Delcev University of Stip, email: tatjana.dzaleva@ugd.edu.mk

[3] Professor Goce Delcev University of Stip, email: mico.apostolov@ugd.edu.mk

Abstract

This paper considers optimal linear and non-linear labor income taxation, which is fair and efficient distribution of the tax incidence or tax burden across individuals with different earnings. There exists a large economic literature that casts light on the issue of optimal labor income taxation. Models in optimal tax theory typically posit that the tax system should maximize a social welfare function subject to a government budget constraint, considering how individuals respond to taxes and transfers. Social welfare is larger when resources are more equally distributed, but redistributive taxes and transfers can negatively affect incentives to work and earn income in the first place. This creates the classical trade-off between equity and efficiency which is at the core of the optimal labor income tax problem. This paper attempts critical survey on the main findings of this literature. This paper is finishing with the numerical solutions to optimal linear and non-linear taxation.

Keywords: *Optimal taxation, public finance, asymmetric information, non-linear tax rates*
JEL codes: H20, H21

Introduction

The modern tax theory has been heavily based on the paper by Mirrlees (1971). In the classical framework initiated by Mirrlees (1971), the theory studies the maximization of a utilitarian social welfare function by a benevolent planner who only observes the pretax labor income of agents whose wages differ, but whose preferences are identical. The other studies have relaxed the assumptions in order to take heterogeneity among agents into account. These studies include: Mirrlees (1976), Saez (2001), Choné and Laroque (2010), see Fleurbaey , Maniquet (2018). Mainly approach is based on asymmetric information. 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[1] 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. Usually the optimal tax systems combine flat marginal tax rate plus lump sum grants to all the individuals (so that the

[1] 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)

average tax rate rises with income even if the marginal does not), Mankiw NG, Weinzierl M, Yagan D.(2009). Previous is in spirit of the early contribution by Ramsey (1927) ,who 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. But form the standpoint of public economics, goal is to derive the best tax system. In perfect economy with absent of any market imperfection (externality), if the economy is described by the representative agent, that consumer is going to pay the entire bill of the government, so that the lump-sum tax is the optimal tax.Rigorous derivations of the optimal tax rates and formulas has been done in the literature namely: Atkinson,Stiglitz,(1980); Kaplow,(2008); Mirrlees(1976), Mirrlees(1986); Stiglitz,(1987); Tuomala,(1990). 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 form reducing the incentives, the insurance effect from reducing the variance of individual income streams, Varian,H.R.(1980).In his model Varian (1980) derives optimal linear and nonlinear tax schedule. He uses Von Neumann-Morgenstern utility function(VNM decision utility, or decision preferences) [2] , with declining absolute risk aversion, see Kreps (1988).Other important contribution in this are that it is necessary to be mentioned is the work by Diamond, Helms and Mirrlees (1978). They 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 s 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. The central goal of the optimal tax theory would be to cast light on the actual policy issues and to help to design better tax systems. Or as discussed in Diamond, Saez (2011),three conditions should be met in order theoretical analysis to be useful for policy makers (1) results must be based on economics mechanisms that are relevant empirically (2) results should be robust to modeling assumptions and particularly to the presence of heterogeneity of individual preferences (3) the tax policy prescription must be implementable and easy to explain and defend publicly and too complex to administer in practice[3].Other theoretical contribution in the theory of optimal labor income taxation includes Saez (2001) which 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)[4] the optimal tax rates are clearly U-shaped. Optimal linear income and linear capital tax are inversely related to the elasticity, the revenue maximizing tax rates are calculated when weights on capital and labor are zero. Saez, E. ,S. Stantcheva (2016),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.Related literature here. 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. The theory of optimal income taxation has reached maturity and excellent reviews of the field are available (Boadway (2012), Piketty and Saez (2013), Salanié (2011). Another important contribution to the theory of optimal labor income taxation is by Piketty, Saez,Stantcheva(2014). Their study derived optimal top tax rate formulas in a model where top earners respond to taxes through three channels: labor supply, tax avoidance, and compensation bargaining. This paper is organized as follows: First model foundations are outlined,

[2] This theorem serves as a basis of the expected utility theory. This theory actually represents maximizing the expected value of some function defined over the potential outcomes at some specified point in the future
[3] The set of possible tax systems evolves overtime with technological progress. If more complex tax innovations become feasible and can realistically generate large welfare gains, they are certainly worth considering.
[4] Utilitarian criterion, utility type I and II and Rawlsian criterion, utility type I and II.

then the link between commodity taxation as supplementary to income taxation is inspected, followed by the optimal linear and non-linear taxation and derivation of optimal non-linear taxes and optimal bottom tax rates in Mirrlees framework. And finally, numerical solutions and examples are presented.

Model foundations

Fixed earnings: $F(w)$ is CDF of pre-tax earnings $F(w)$ is the fraction of the population with pre-tax earnings below and consumption $c = w - \tau(w)$ .The government chooses τ to maximize SWF (utilitarian):

equation 1

$SWF = \int_0^\infty u(w - \tau(w)dF(w)$ s.t. $\int_0^\infty \tau(w)dF(w) \geq E(\lambda)$

Where E represents exogenous revenue requirement and λ is an lagrangian multiplier, of the government budget constraint. FOC in $\tau(w)$is simply:

equation 2

$$u'\big(w - \tau(w)\big) = \lambda \Rightarrow w - \tau(w) => constant\ across\ w^5$$

Generalized SWF's like $\int G\big(u(c)\big)dF(w)$ are considered, there also $G(\cdot)$ increasing transformation of utilities[6] .The case where $G(\cdot)$ is ever concave is Rawlsian (maxi-min) SWF.For the heterogenous utility function $u(c)$ across individuals,the utilitarian optimum is such that $u_i'(c)$ is constant over population, see Piketty, Saez (2013). The concavity of $u(c)$ reflects society's value for redistribution rather than directly individual marginal utility of consumption[7].

Endogenous Earnings: The goal of the optimal income tax theory has been to extend the basic model to the case with endogenous earnings see Vickrey, (1945) and Mirrlees, (1971). Now, lets suppose general SWF of the following type :

equation 3

$$SWF = \int \omega_i G\left(u^i(c, w)\right) df(i)$$

In previous SWF ω_i are Pareto weights and $\omega_i \geq 0$ and they are independent of the individual choices on consumption and earnings $u(c, w)$ and $G(\cdot)$ increasing transformation of utilities, also $df(i)$ represents the distribution of individuals. Social marginal welfare weight[8] is given as:

equation 4

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

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).In Mirrless (1971) heterogeneity comes from the wages w_i only and utility function is given as : $u_i = \left(c_i, \frac{w}{w_i}\right)$ where $l = \frac{w}{w_i}$. Linear tax system considered here is augmented with demogrant[9] R

---

[5] The government imposes taxes on 100% of earnings, and funds its revenue requirement, then redistributes the remaining tax revenue equally across individuals.This result was first established by Edgeworth (1897).

[6] social welfare function

[7] If individuals have concave utility function they will prefer more redistribution policy by the government, Piketty, Saez (2013).

[8] 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.

[9] A grant awarded on purely demographic principles such as age and sex.

equation 5

$$c = (1-\tau)w + R$$

Now of the concepts of Labor supply one thing that comes across our mind is intensive margin[10].The maximization problem here is :

equation 6

$$\max u^i\big((1-\tau)w + R, w\big)$$

$$FOC: (1-\tau)\frac{\partial u^i}{\partial c} + \frac{\partial u^i}{\partial w} = 0$$

here Marshallian demand for labor is given as :$w = w(1-\tau, R)$ where R is the non-labour income, and w are earnings(wages). Income effects are captured through $\eta = (1-\tau)\partial w/\partial R$,average income effects are :$\bar{\eta} = \int_{\bar{w}}^{\infty} \eta_w h(w)dw$. And compensated elasticity of earnings is :

equation 7

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

Those two are related by the Slutsky equation : $\varepsilon^c = \varepsilon^u - \eta$, when there are no behavioral responses there is only meachnical 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 : $dw = -\frac{\partial w}{\partial(1-\tau)}d\tau + \frac{\partial w}{\partial R}dR = -\left(\varepsilon^u w - \frac{(1-\tau)\partial w}{\partial R}\right)\left(\frac{d\tau}{1-\tau}\right)$,or the total behavioral response:

equation 8

$\beta = -\left(\varepsilon^u w \frac{(1-\tau)\partial w}{\partial R}\right)\left(\frac{\tau d\tau}{1-\tau}\right)$

Saez(2001) result for high income earners is given as :

equation 9

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

Hicksian or compensated earnings supply f-ction is given as: $w_c^i(1-\tau, u)$, or formally $w_c^i(1-\tau, u)$ solves the following problem or inequality :

inequality 1

$\min_{w} c - (1-\tau)w$ s.t. $u(c,w) \geq u$

The Slutsky eq.relates $\varepsilon^c = \varepsilon^u - \eta$ so that we have :

inequality 2

$\bar{\varepsilon}^c = \frac{1-\tau}{w_c^i}\left(\frac{\partial w_c^i}{\partial(1-\tau)}\right)|_u > 0$; $\eta = \frac{(1-\tau)\partial w}{\partial R} \leq 0$; $\varepsilon^u = \frac{1-\tau}{w_u^i}\left(\frac{\partial w_u^i}{\partial(1-\tau)}\right) \gtreqless 0$

Virtual income formally is defined as "the non-labor income that the individual would get if her earnings were zero and she could stay on the virtual linearized budget" , $R = w - \tau(w) - (1-\tau'(w)) \times w$. Substitution effects: Hicksian labor supply: $w_c^i\ (1-\tau, u)$ minimizes cost needed to reach u given slope $1-\tau$.Slutsky equation is given as:

equation 10

$\frac{\partial w}{\partial(1-\tau)} = \frac{\partial w_c^i}{\partial(1-\tau)} + w\frac{\partial w}{\partial R} \Rightarrow \varepsilon^u = \varepsilon^c + \eta$

---

[10] Intensive margin refers to the degree (intensity) to which a resource is utilized or applied. For example, the effort put in by a worker or the number of hours the worker works

Second thing that comes in our mind related to the concept of the labor supply is the thing called extensive margin[11]. Here we include fixed costs of searching job or discrete costs d^i_{costs}, so now we assume that the utility function is linear: $u_i = c_i - d^i_{costs} \cdot l_i$ .Here individual i works only and if only $w_i - \tau(w_i) - d^i_{costs} \geq \tau(0)$ where $l_i \in (0,1)$ is a work dummy variable. Or previous expression follows inequality:

inequality 3

$$d^i_{costs} \leq w_i - \tau(w_i) + \tau(0) = w_i \cdot (1 - \tau_p)$$

τ_p is the effective participation tax rate[12] (defined as the fraction of earnings taxed when the individual goes from not working and earning zero to working and earning w_i) which actually equals to following expression:

equation 11

$$\tau_p = \frac{(\tau(w_i) - \tau(0))}{w_i}$$

Ramsey (1927) model or Ramsey tax rule (commodity taxes and income taxation inverse elasticity rule)

In Ramsey (1927), utility function is given of type:$u = f(p_1, p_2, p_3, \ldots\ldots, w)$, $p_1, p_2, p_3, \ldots\ldots$ are prices and w is income. This result is known as Roy's identity , Roy (1947)[13], is :

equation 12

$\frac{\partial u}{\partial p_i} = -f_i \frac{\partial u}{\partial w}$.

With the horizontal demand curves, price of the producers is fixed, change in the goods price is only equal to the change in taxes. Than, $dp_1 = d\tau_1 > 0$, $dp_2 = d\tau_2 < 0$. Change in taxes must satisfy the following equation:

equation 13

$dU = \frac{\partial U}{\partial p_i} d\tau_1 + \frac{\partial U}{\partial p_2} d\tau_2 = 0$,and $\frac{d\tau_2}{d\tau_1} = -\frac{F_1}{F_2}$,

change in the revenues caused by the change in taxes is :$\frac{\partial(\tau_1 f_1)}{\partial t_1} = F_1 + \frac{\tau_1 df}{dp_1} = F\left(1 + \frac{\tau_1 dF_1 p_1}{p_1 dp_1 F_1}\right) = F_1\left(1 - \frac{\tau_1}{p_1}\varepsilon^1_u\right)$, where ε^1_u represents the compensated elasticity of the demand for good 1. Change of revenues as a result of change of taxes on good 2 is: $\frac{\partial(t_2 F_2)}{\partial t_2} = F_2\left(1 - \frac{\tau_2}{p_2}\varepsilon^2_u\right)$. With the optimal tax structure, this identity must holds:$\frac{t_2}{p_2}\varepsilon^2_u - \frac{t_1}{p_1}\varepsilon^1_u = 0$, for the linear demand curve results is :$\frac{t}{p} = \frac{kQ}{bp} = \frac{k}{\varepsilon^d_u}$. This conclusion is supported by the findings of Feldstein (1978), "when lump-sum taxation is not available (or, equivalently, when a tax on leisure is impossible), all other commodities should be taxed at differential rates (positive and negative) that depend on their relative demand elasticities and cross elasticities". Ramsey model was used in life cycle models, for best reference see Atkinson, A.B. and Stiglitz,J. (1976),Atkinson, A.B. and A. Sandmo (1980), Atkinson, A.B. and Stiglitz,J. (1980).

Commodity taxation supplementary to labor income taxation (Atkinson, Stiglitz theorem)

---

[11] Extensive margin refers to the range to which a resource is utilized or applied. For example, the number of people working is one measure that falls under the heading of extensive margin

[12] Participation tax rates are conceptually very similar, indicating the effective tax rate on the extensive margin, or the proportion of earnings paid as taxes and lost due to benefit withdrawal if a person moves from inactivity or unemployment to work

[13] The lemma relates the ordinary (Marshallian) demand function to the derivatives of the indirect utility function.

The government can implement differentiated commodity taxation in addition to non-linear income taxes. The usual hypothesis here is that commodity taxes have to be linear because of retrading , see Guesnerie,(1995) . Consider a model with k consumption goods $c = (c_1, \ldots, c_k)$ with pre-tax prices $p = (p_1, \ldots, p_k)$. Individual i derives utility from the k consumption goods and earnings supply according to a utility function $u_i(c_1, \ldots, c_K, w)$. The question here is whether government can increase social welfare by adding differentiated commodity taxation $\tau = (\tau_1, \ldots, \tau_k)$ in addition to nonlinear tax on earnings w . Atkinson and Stiglitz (1976) demonstrated the following theorem known as Atkinson, Stiglitz theorem:

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.,$u_i(c_1, \ldots, c_k, w) = u_i(v(c_1, \ldots, c_k), w)$ with the subutility function $v(c_1, \ldots, c_k)$ 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. $(p + t) \cdot c \le \gamma$ 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 $p \cdot c_i \ge w - \bar{\tau}(w)$ and we have that $\bar{\tau}(w) \ge \tau(w) + t \cdot c_i$ i.e. the government collects more taxes with $(\bar{\tau}(\cdot), t = 0)$ ■

Optimal linear taxation

First modern treatment of optimal linear tax was provided by Sheshinski (1972) following the nonlinear income tax analysis provided by Mirrlees (1971). In Sheshinski (1972) optimal linear tax formulae is given as:

equation 14

$$\int_0^{\infty} \tau(w) f(n) dn = \int_0^{\infty} (w - \alpha - \beta w) f(n) dn = 0$$

$f(n)$ is PDF of the individuals with ability n.Other symbols are defined as follows: α is a tax parameter and is a lump-sum tax if $\alpha < 0$ and tax-subsidy if $\alpha > 0$ given to an individual with no income.$1 - \beta$ is a marginal tax rate i.e. $0 \le \beta \le 1$ so that marginal tax rate is non negative in the linear tax function which is $\tau(w) = -\alpha + (1 - \beta)w$, after tax consumption is $c(w) = w - \tau(w) = \alpha + \beta w$.Optimal labor supply is given as: $\ell = \hat{\ell}(\beta n, \alpha)$.If λ is the lowest elasticity of labor supply function and it is equal to $\lambda = \lim_n \inf \left[\frac{\beta}{\hat{\ell}} \frac{\partial \hat{\ell}}{\partial \beta}\right]$ so that $\frac{\beta}{\hat{\ell}} \frac{\partial \hat{\ell}}{\partial \beta} \ge \lambda$. Revenue maximizing linear tax rate is given as: $\frac{\tau^*}{1 - \tau^*} = \frac{1}{e}$ or $\tau^* = \frac{1}{1+e}$.The government FOC given $SWF = \int \omega_i G\left(u^i(1 - \tau)w^i + \tau w(1 - \tau) - E, w^i\right)\big) df(i)$ is :

equation 15

$$0 = \frac{dSWF}{d\tau} = \int \omega_i G'(u_i) u_c^i \cdot \left((w - w^*) - \tau \frac{dw}{d(1 - \tau)}\right) df(i)$$

Social marginal welfare weight g_i is given as: $g_i = \frac{\omega_i G'(u_i)u_c^i}{\int \omega_j G'(u_j)u_c^j df(j)}$.So that optimal linear tax formula is:

equation 16

$$\tau = \frac{1-\bar{g}}{1-\bar{g}+e}$$

where $\bar{g} = \frac{\int g_i \cdot w_i df(i)}{w}$.

Optimal linear taxation and tax avoidance

Responses to tax rates can take form of tax avoidance[14] see Saez et al.,(2012).One approach that is being most recently employed is done by Piketty, Saez,Stantcheva(2014). If w are real wages and if we equal that to real income and x is the sheltered income[15] so that taxable income equals $w^\tau = w - x$, and the taxable income w^τ is taxed by a tax rate τ , while sheltered income is taxed by a constant rate t and $t < \tau$. Utility of an individual is given as:

equation 17

$$u_i(c, w, x) = c - h_i(w) - d_i(x)$$

Where $c = w - \tau w^\tau - tx + R = (1-\tau)w + (\tau - t)x + R$ is a disposable after tax income, and $h_i(w)$ are cost associated with earning income, and $d_i(x)$ are the costs associated with tax avoidance or costs of sheltering income. It is assumed that $h_i(\cdot), d_i(\cdot)$ are convex and increasing and $h_i'(0) = d'_i(0) = 0$and individual utility maximization implies that :

equation 18

$$h_i'(w_i) = 1 - \tau\, ; d_i'(x_i) = \tau - t$$

If we aggregate for all individuals we have $w = w(1-\tau) = \int w_i(1-\tau)$ and the income elasticity is given as: $\varepsilon_w = \frac{\left[\frac{(1-\tau)}{w}\right]dw}{d(1-\tau)} > 0$, and $x = x(t-\tau) = \int x_i(\tau - t)df(i)$, so $x(\tau - t = 0) = 0$ because there is tax avoidance (sheltering) only when $t < \tau$.

Aggregate taxable income $w^\tau = w^\tau(1-\tau, t) = w(1-\tau) - x(\tau - t)$ is increasing in $1-\tau$ and t. We denote by $e = [(1-\tau)/w]\partial w/\partial(1-\tau) > 0$ the total elasticity of taxable income w^τ with respect to $1-\tau$ when keeping t constant. Note that $e = (w/w^\tau)ew + ((1-\tau)/w^\tau)dx/d(\tau - t) > (w/w^\tau)ew$. Partial optimum of the linear tax in the case of taxa avoidance is given as:

equation 19

$$\tau = \frac{1 + t \cdot (e - \left(\frac{w}{w^t}\right)\varepsilon_w}{1+e}$$

General optimum of this tax in a case of tax avoidance and income sheltering is given as:

equation 20

$$t = \tau = \frac{1}{1+\varepsilon_w}$$

This tax rates optimiza global tax policy $\tau[w(1-\tau) - x(\tau - t)] + tx(\tau - t)$.

Labor tax reform and non-linear tax formula

---

[14] Tax avoidance opportunities typically arise when taxpayers can shift part of their taxable income into another form of income or another time period that receives a more favorable tax treatment, Piketty, Saez (2013)

[15] Sheltered Income means so called earned income, rebates, kick-backs, volume discounts, tier pricing, purchase commitment discounts, sales and service allowances, marketing allowances, advertising allowances, promotional allowances, label allowances, back-door income, etc.

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 :
$U(1-\tau,R)=max_w\big((1-\tau)w+R,z\big)$, where w represents the taxable income R is a virtual income intercept, and τ is an imposed income tax. Marshalian labor supply is w= $w(1-\tau,R)$, uncompensated elasticity of the supply is given as:$\varepsilon^u=\frac{(1-\tau)}{w}\frac{\partial w}{\partial(1-\tau)}$, income effect is $\eta=(1-\tau)\frac{\partial w}{\partial R}\leq 0$.Hicksian supply of labor is given as:$w^c\big((1-\tau,u)\big)$, this minimizes the cost in need to achieve slope $1-\tau$, compensated elasticity now is : $\varepsilon^c=\frac{(1-\tau)}{w}\frac{\partial w^c}{\partial(1-\tau)}>0$, Slutsky equation now becomes: $\frac{\partial w}{\partial(1-\tau)}=\frac{\partial w^c}{\partial(1-\tau)}+z\frac{\partial z}{\partial R}\Rightarrow\varepsilon^u=\varepsilon^c+\eta$, where η represents income effect :$\eta=(1-\tau)\frac{\partial w}{\partial R}\leq 0$.
With small tax reform taxes and revenue change i.e.:$dU=u_c\cdot[-wdt+dR]+dw[(1-\tau)u_c+u_z]=u_c\cdot[-zdt+dR]$.Change of taxes and its impact on the society is given as:$dU_i=-u_c dT(w_i)$. Envelope theorem here says :$U(\theta)=\max_x F(x,\theta),s.t.c>G(x,\theta)$, and the preliminary result is :$U'(\theta)=\frac{\partial F}{\partial\theta}\Big(x^*(\theta),\theta-\lambda^*(\theta)\frac{\partial G}{\partial\theta}x^*(\theta),\theta\Big)$. Government is maximizing :

equation 21

$$0=\int G'(u^i)u_c^i\cdot\Big[(W-w^i)-\frac{\tau}{d(1-\tau)}eW\Big],$$

1. mechanical effect is given as:$dM=[w-w^*]d\tau$,
2. welfare effect is :$dW=-\bar{g}dM=-\bar{g}[w-w^*]$, and at last
3. the behavioral response is : $dB=-\frac{\tau}{1-\tau}\cdot e\cdot wd\tau$.

And let's denote that:
equation 22

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

When the tax is optimal these three effects should equal zero i.e. $dM+dW+dB=0$ given that:$\frac{\tau}{1-\tau}=\frac{(1-\bar{g})[w-w^*]}{e\cdot z}$, and we got $\tau=\frac{1-\bar{g}}{1-\bar{g}+a\cdot e}$, $a=\frac{w}{w-w^*}$,and $dM=d\tau[w-w^*]\ll dB=d\tau\cdot e\frac{\tau}{1-\tau}\cdot w, when\ w^*>w^T$, where w^T is a top earner income. Pareto distribution is given as:
equation 23

$$1-F(w)=\Big(\frac{k}{w}\Big)^a, f(w)=a\cdot\frac{k^a}{w^{1+a}}$$

a is a thickness parameter and top income distribution is measured as:

equation 24

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

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

equation 25

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

Where elasticity or efficiency $e=\Big[\frac{1-\tau}{w}\Big]\times\frac{dw}{d(1-\tau)}$. Where $G((w_n)=\frac{\int_n^{\infty}g_m dF(m)}{1-F(n)}$,and $g_m=G'(u_m)/\lambda$ this is welfare weight of type m.But non-linear tax witn income effect takes into account small tax reform where tax rates change from $d\tau$ to $[w^*,w^*+dw^*]$.Every tax payer with income $w>w^*$ pays additionaly $d\tau dw^*$ valued by $(1-g(w))d\tau dw^*$.Mechanical effect is :

equation 26

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

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

equation 27

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

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

equation 28

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

Optimal tax formula here if $dM + dW + dB = 0$ is given as : $\tau = \frac{1-\bar{g}}{1-\bar{g}+\alpha\cdot e}$; $\alpha = \frac{w}{w-w^*}$ where $\bar{g} = \frac{\int g_i \cdot w_i}{w \cdot \int g_i}$ and$g_i = G'(u')u_c^i$.

Formal derivation of optimal non-linear tax rates with no income effects

This point actually follows Mirrleees (1971) and Diamond (1998) , in deriving non-linear optimal tax rate with no-income effects.Utility function is quasi linear:

equation 29

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

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 30

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

When marginal tax rate τ is constant, the labor supply f-ction is given as: $l \rightarrow l\big(n(1 - \tau)\big)$ 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 31

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

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

equation 32

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

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:

equation 33

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

Now the Hamiltonian is given as:

equation 34

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

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

equation 35

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

FOC with respect to u is given as:

equation 36

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

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 37

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

We can rewrite FOC with respect to l_n as:

equation 38

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

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 39

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

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 $G(w_0 = 0) = 1$.

Derivation of the optimal bottom tax in Mirrlees model

Government maximizes following social welfare function :

equation 40

$$W = \int G(u_n) f(n) dn \ s.t. \int c_n f(n) dn \leq \int n l_n f(n) dn - E(p).$$

$F(n)$ is the distribution of skills with $f(n)$ its PDF and support $[0, \infty)$. Earnings equal $w = nl$ where n represents innate ability.Here 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:

equation 41

$$\frac{du_n}{dn} = -\frac{l_n u_l(c_n, l_n)}{n}$$

Now the Hamiltonian is given as:

equation 42

$$\mathcal{H} = [G(u_n) + \lambda \cdot (nl_n - c_n)] f(n) + \phi(n) \cdot \frac{-l_n u_l(c_n, l_n)}{n}$$

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

equation 43

$$\lambda \cdot \left(n + \frac{u_l}{u_c}\right) + \frac{\phi(n)}{n} \cdot \left[u_l + l_n u_{ll} + l_n u_{cl} \frac{u_l}{u_c}\right] = 0$$

At $n = n_0, l = 0, n_0 + \frac{u_l}{u_c} = n_0 \tau'(0)$, and this first order condition becomes:

equation 44

$$\lambda \cdot n_o f(n_0) \tau'(0) = \frac{\phi(n_0) u_l}{n_0}$$

Now ad $\frac{\partial c}{\partial u} = \frac{1}{u_c}$,the FOC with respect to u becomes:

equation 45

$$-\frac{d\phi(n)}{n} = \left[G'^{(u_n)} - \frac{\lambda}{u_c}\right] f(n) - \phi(n) \frac{l_n u_{cl}}{n u_c}$$

Now, for $n \leq n_0$, $l_n = 0$, $u_n = u(c_0, 0)$, $u_c = u_c(c_0, 0)$ are constant with n so that this equation
becomes:

equation 46

$$-\frac{d\phi(n)}{n} = \left[G'^{(u_n)} - \frac{\lambda}{u_c}\right] f(n)$$

and can be integrated from $n = 0$ to $n = n_0$ and yields: $\phi(n_0) = \frac{p}{u_c}\left[1 - \frac{G'(u_0) u_c}{\lambda}\right] F(n_0)$.Now replacing the expression for $\phi(n_0)$ in the FOC for l at $n = n_0$ gives:

equation 47

$$n_0 f(n_0)\tau'(0) = \frac{u_l}{u_c n_0}\left[1 - \frac{G'\big(u_0(u_c)\big)}{\lambda}\right]F(n_0) = \big(1-\tau'(0)\big)\left[\frac{G'\big(u_0(u_c)\big)}{\lambda} - 1\right]F(n_0)$$

Previous expression can be written and simplified as:

equation 48

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

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 $n_0\big(1-\tau'(0)\big)u_c(c_0, 0) + u_l(c_0, 0) = 0$ 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:

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

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

equation 49

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

Numerical solutions and examples

Table 1 illustrates or proposes some illustrative calculations by using the optimal tax formula. The tax formula here is of the linear tax :
equation 50

$$\tau = \frac{1-\bar{g}}{1-\bar{g}+e}$$

The first column of the table follows realistic scenario with elasticity of range $e = 0.25$, as in Saez et al., (2012) and Chetty, (2012) , and Piketty, Saez (2013) .The second column is with estimates in range $e = 0.5$ which is high range elasticity scenario and a third scenario is $e = 1$ which is well above estimates in the current literature.

Table 1 Linear optimal tax rates per Piketty, Saez (2013)

	$e = 0.25$		$e = 0.5$		$e = 1$	
	$\bar{g}$	τ	$\bar{g}$	τ	$\bar{g}$	τ
Rawlsian revenue maximizing rate	0	0.8	0	0.67	0	0.50
Utilitarian CRRA=1 $u_c = \frac{1}{c}$	0.61	0.61	0.54	0.48	0.44	0.36
Median voter I $\frac{w_{median}}{w_{average}}$	0.7	0.55	0.7	0.38	0.7	0.23
Median voter II $\frac{w_{median}}{w_{average}}$	0.75	0.50	0.75	0.33	0.75	0.20
very low tax country 10%	0.97	0.1	0.94	0.1	0.88	0.1
low tax country 35%	0.87	0.35	0.807	0.35	0.46	0.35
high tax country 50%	0.75	0.5	0.5	0.5	0	0.5

Source: Author’s calculation

The first row of table 1 is Rawlsian criterion with $\bar{g} = 0$. The second row is utilitarian criterion with coefficient of risk aversion (CRRA) equal to one and social marginal welfare weights are proportional to $u_c = \frac{1}{c}$ where $c = (1 - \tau)w + R$ where R is disposable income. Chetty (2006) proved and showed that $CRRA = 1$ is consistent with empirical labor supply behavior and that is a reasonable benchmark. First scenario with $e = 0.25$ shows that revenue maximizing tax rate is 80% which is higher even for the countries with highest marginal tax rate which is around 50%. The optimal tax rate under Utilitarian criterion is 61%. The optimal tax rate for median earner is 55% or 38% under $e = 0.5$ and 36% under $e = 1$. In the examples with very low tax country one can see that a tax rate of 10% is optimal in a situation where $g = 0.97$ i.e. in a country with very low redistributive tastes. A tax rate of 50% would be optimal in a country with $\bar{g} = 0.75$. A high elasticity estimate $e = 0.5$ would generate tax rate of 67% above current rates in every country. The median voter tax rate in such a situation would be 38%, Utilitarian criterion generate tax rate of 48% in this situation. In the unrealistically high elasticity scenario $e = 1$ the revenue maximizing tax rate is 50% which is about the current rate in countries with highest $\frac{Tax}{GDP}$ ratios.

Example 2 Non-Linear taxes
In this table 2 non-linear taxes have been estimated by using this tax formulae:
equation 51

$$\tau = \frac{1 - \bar{g}}{1 - \bar{g} + \varepsilon_u + \varepsilon_c(\alpha - 1)}$$

Table consists of three global columns with supposed elasticities (uncompensated) $\varepsilon_u \in (0,0.2,0.5)$ and supposed compensated elasticities $\varepsilon_c \in (0.2,0.5,0.8)$.

Table 2 Non-linear income taxes under different uncompensated and compensated elasticities

	$\varepsilon_u = 0$			$\varepsilon_u = 0.2$			$\varepsilon_u = 0.5$		
$\varepsilon_c =$	0.2	0.5	0.8	0.2	0.5	0.8	0.2	0.5	0.8
				$\bar{g} = 0$					
a=1.5	0.91	0.80	0.71	0.77	0.69	0.63	0.63	0.57	0.53
a=2	0.83	0.67	0.56	0.71	0.59	0.50	0.59	0.50	0.43
a=2.5	0.77	0.57	0.45	0.67	0.51	0.42	0.56	0.44	0.37
				$\bar{g} = 0.25$					
a=1.5	0.88	0.75	0.65	0.71	0.63	0.56	0.56	0.50	0.45
a=2	0.79	0.60	0.48	0.65	0.52	0.96	0.52	0.43	0.37
a=2.5	0.71	0.50	0.38	0.60	0.44	0.35	0.48	0.38	0.31
				$\bar{g} = 0.5$					
a=1.5	0.83	0.67	0.56	0.63	0.53	0.45	0.45	0.40	0.36
a=2	0.71	0.50	0.38	0.56	0.42	0.33	0.42	0.33	0.28
a=2.5	0.63	0.40	0.29	0.50	0.34	0.26	0.38	0.29	0.23
				$\bar{g} = 0.75$					
a=1.5	0.71	0.50	0.38	0.45	0.36	0.29	0.29	0.25	0.22
a=2	0.56	0.33	0.24	0.38	0.26	0.20	0.26	0.20	0.16
a=2.5	0.45	0.25	0.17	0.33	0.21	0.15	0.24	0.17	0.13

Source: Author's calculation

Pareto distribution is given as PDF lower CDF and upper CDF [16]:

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

PDF (probability density function) :
equation 52

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

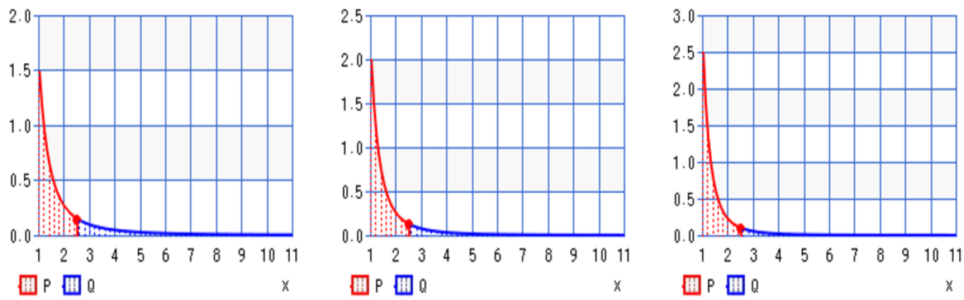
Lower cumulative distribution function (lower CDF):
equation 53

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

Upper cumulative distribution function (upper CDF):
equation 54

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

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



Source: Author's calculation

Table 3 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 4 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

From Table 2 highest on-linear income taxes are generated in the case with high redistributive tastes where $\bar{g} = 0$ and low uncompensated elasticity $\varepsilon_u = 0$ and a Pareto distribution parameter $\alpha = 1.5$. Generated tax rates are : $\tau \in (0.91, 0.8, 0.71)$.Lowest taxes are generated with high uncompensated elasticity $\varepsilon_u = 0.5$ and high $\varepsilon_c = 0.8$.Gnerated tax rates there are $\tau \in (0.22, 0.16, 0.13)$ with lowest tax rate generated under Pareto shape parameter $\alpha = 2.5$ and a very low taste for redistribution $\bar{g} = 0.75$ .Next, follows another example with fairly non-linear U-shaped taxes as per Diamond (1998).The formulae that we are using here is :

equation 55

$$\tau' = \frac{(e^{-1}+1)(1-g)}{[a+(e^{-1}+1)(1-g)]}$$

Table 5 Non-linear income tax rates as per Diamond (1998) and authors own calculations

	$g=0$			$g=0.25$			$g=0.5$			$g=0.975$		
a=	0.5	1.5	5	0.5	1.5	5	0.5	1.5	5	0.5	1.5	5
e												
0.2	0.92	0.8	0.55	0.90	0.75	0.47	0.86	0.67	0.38	0.23	0.09	0.03
0.5	0.86	0.67	0.38	0.82	0.60	0.31	0.75	0.50	0.23	0.13	0.05	0.01
0.75	0.82	0.61	0.32	0.78	0.54	0.26	0.70	0.44	0.19	0.10	0.03	0.01
1	0.80	0.57	0.29	0.75	0.50	0.23	0.67	0.40	0.17	0.09	0.03	0.01
1.5	0.77	0.53	0.25	0.71	0.45	0.20	0.63	0.36	0.32	0.08	0.03	0.01
2	0.75	0.50	0.23	0.69	0.43	0.18	0.60	0.33	0.13	0.07	0.02	0.01

Source: Author's calculation

Form previous table one can see that highest non-linear income taxes are generated with high tastes for redistribution where $g = 0$ and Pareto shape parameter $\alpha = 0.5$ and with labor elasticity $e = 0.2$.Gnerated tax rates are $\tau \in (0.92, 0.88, 0.55)$ for Pareto shape parameters $\alpha \in (0.5, 1.5, 5)$.For the same elasticities and Pareto shape parameters but with very low almost non-existent redistributive tastes generated low tax rates are: $\tau \in (0.23, 0.09, 0.03)$ respectively. On a very high (unrealistically high) labor elasticities generated are tending to zero $\tau \to 0$.

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 have improved the tax policies in the past. 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." .The points made by Mirrlees which are also support by the numerical results in this paper 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 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". So, in our opinion also as the analysis proved that not always implemented tax rates would be justified theoretically. Namely, optimal tax rates as this paper shows depend on redistributive tastes of the supposedly benevolent social planers. 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.As the numerical solutions in the non-linear optimal tax rates showed that high tax rates are obtained when there unrealistically low uncompensated and compensated elasticities, also the shape parameter of Pareto distribution must be lower. For high tax countries e.g. countries with highest tax burden around 50% the area that provides such high tax rates is where compensated elasticity is between 0.2 and 0.5 and uncompensated elasticity and unrealistically high compensated elasticity between 0.5 and 0.8 but medium redistributive tastes $\bar{g} = 0.5$. Or alternatively, if uncompensated elasticity is high $\varepsilon_u = 0.5$ than also the taste for redistribution must be high e.g. $\bar{g} \in (0,0.25)$.For low tax countries the area where those taxes are provided is in high Pareto distribution parameter and very low taste for redistribution. These are very loose results and are conditioned by themselves and their combinations. In turn there is not straightforward solution to the optimal linear or non-linear labor income tax problem.

References

1. Atkinson, A.B. and A. Sandmo (1980). Welfare Implications of the Taxation of Savings, Economic Journal, Vol. 90, 1980, 529-49. (web)
2. 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)
3. Atkinson, A.B. and Stiglitz,J. (1980).Lectures on Public Economics, Chap 14-4 New York: McGraw Hill, 1980. (web)
4. Auerbach, A. (2009). The choice between income and consumption taxes: A primer. Cambridge: Harvard University Press.
5. Boadway, Robin. (2012). From Optimal Tax Theory to Tax Policy. Cambridge and New York: MIT Press.
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. Chetty, R. (2006). A new method of estimating risk aversion. American Economic Review, 96(5), 1821–1834
8. Chetty,R. (2012). Bounds on elasticities with optimization frictions:A synthesis of micro and macro evidence on labor supply. Econometrica, 80(3), 969–1018
9. Choné, Philippe, and Guy Laroque. (2010). Negative Marginal Tax Rates and Heterogeneity. American Economic Review 100 (5): 2532–47.
10. Diamond, P. (1998). Optimal income taxation: An example with a U-shaped pattern of optimal marginal tax rates. American Economic Review, 88, 83–95.
11. Diamond,P., Saez,E. (2011).The case for a progressive tax:From basic research to policy recommendations. Journal of Economic Perspectives, 25(4), 165–190.
12. Diamond,P.,Helms,J. and Mirrlees (1978).Optimal taxation in a stochastic economy, A Cobb-Douglas example, M.I.T. Working Paper no. 217
13. Edgeworth, F.Y. (1897). The pure theory of taxation. Economic Journal, 7, 46–70, 226–238, and 550–571.
14. 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
15. Fleurbaey,Marc, Maniquet François (2018). Optimal Income Taxation Theory and Principles of Fairness. Journal of Economic Literature 2018, 56(3), 1029–1079
16. Guesnerie, R. (1995). A contribution to the pure theory of taxation. Cambridge, MA: Cambridge University Press
17. Kaplow, L. (1994). Taxation and risk taking: A general equilibrium perspective. National Tax Association 47 (4), 789–798.
18. Kaplow, L. (2008).The theory of taxation and public economics. Princeton: Princeton University Press.
19. 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.
20. Kreps, David M.(1988),Notes on the Theory of Choice. Westview Press ,chapters 2 and 5.
21. Laroque, G. R. (2005). Indirect taxation is superfluous under separability and taste homogeneity:A simple proof. Economics Letters, 87(1), 141–144.
22. Mankiw NG, Weinzierl M, Yagan D.(2009),Optimal Taxation in Theory and Practice. Journal of Economic Perspectives. 2009;23 (4) :147-174.
23. Mas-Colell, A., et al. (1995). Microeconomic Theory, Oxford University Press.
24. Mirrlees, J. A. (1971). An exploration in the theory of optimal income taxation. Review of Economic Studies,38, 175–208.
25. Mirrlees, J. A. (1976). Optimal tax theory:A synthesis. Journal of Public Economics, 6, 327–358.
26. 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.

27. Piketty, T., Saez,E. (2013), Chapter 7 - Optimal Labor Income Taxation, Elsevier,Volume 5, Pages 391-474, ISSN 1573-4420, ISBN 9780444537591
28. Piketty, Thomas, Emmanuel Saez, and Stefanie Stantcheva.(2014). Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities. American Economic Journal: Economic Policy 6, no. 1 (February 2014): 230–71.
29. Ramsey, F. (1927). A Contribution to the Theory of Taxation. Economic Journal. 37: 47 61. doi:10.2307/2222721.
30. Ramsey, F. (1927). A Contribution to the Theory of Taxation. Economic Journal. 37: 47 61.
31. Saez, E. ,S. Stantcheva (2016). Generalized social marginal welfare weights for optimal tax theory. The American Economic Review, 106(1), pp.24-45.
32. Saez, E., Slemrod, J., Giertz, S. (2012). The elasticity of taxable income with respect to marginal tax rates: A critical review. Journal of Economic Literature, 50(1), 3–50
33. Saez, E., Slemrod, J.,Giertz, S. (2012). The elasticity of taxable income with respect to marginal tax rates: A critical review. Journal of Economic Literature, 50(1), 3–50
34. Saez,E.(2001).Using elasticities to derive optimal income tax rates, The Review of Economic Studies, 68(1), pp.205- 229.
35. Salanie, Bernard. (2011). The Economics of Taxation, Second edition. Cambridge and London: MIT Press.
36. Sheshinski, E. (1972). The optimal linear income tax. Review of Economic Studies, 39(3),297–302.
37. Stiglitz,J. (1987). Pareto efficient and optimal taxation and the new new welfare economics. InA. J.Auerbach, & M. Feldstein (Eds.), Handbook of Public Economics.Vol. 2, (pp. 991–1042). Amsterdam: North-Holland.
38. Tuomala, M. (1990). Optimal income tax and redistribution. Oxford: Clarendon Press.
39. Varian,H.R.(1980).Redistributive taxation as social insurance. Journal of public Economics,14(1), pp.49-68.
40. Vickrey,W. (1945). Measuring marginal utility by reactions to risk. Econometrica, 13, 319–333.