

Relativity, Inequality and Optimal Nonlinear Income Taxation

Ravi Kanbur and Matti Tuomala

Preliminary Draft

May, 2009

Contents

1. Introduction
2. Basic Model
3. The Rawlsian Case
4. The Utilitarian Case
5. Inequality and Relativity
6. Discussion and Conclusion

Abstract

How does concern for consumption relative to others ("relativity") affect the progressivity of the optimal income tax structure? In this paper we revisit this literature and present a more detailed analysis of the solution to the non-linear income tax problem with consumption interdependence than is currently available, generalizing some results and developing other results for cases with special objective functions and special distributions, as well as numerical simulations. Of particular interest for us is the interplay between inequality and relativity in determining the optimal tax schedule. We find support for greater progressivity in the tax structure as relative concern increases. But our numerical calculations show that this incremental impact is less at higher levels of inequality.

1. Introduction

There is growing empirical evidence that the assumption that individual preferences are independent, in the sense that people do not want things because others want them, may not be entirely appropriate.¹ The major alternative to this assumption is that an individual's well-being depends on his or her relative consumption – how it compares to the consumption of others. This “relativity” idea is not new of course. More than one hundred years ago Thorsten Veblen² maintained that consumption is motivated by a desire for social standing as well as for enjoyment of the goods and services per se. This implies that people compare consumption not leisure.³

Relative consumption (income) concern or status seeking creates negative externalities because gains in one's status reduce someone else's. If these externalities are important as empirical research seems to suggest, taxing consumption externalities might be welfare enhancing just in the same way as any other Pigouvian tax. This simple intuition does not tell us anything about the detailed effects of relative income concern on the tax schedule. Do status considerations lead to a more progressive tax system or a less progressive tax system? Is income tax an effective tool for reducing inequalities and attenuating possible externalities arising from relative income concerns? How does inequality and relativity together determine the shape of the optimal tax schedule?

There are few papers asking these questions in an optimal nonlinear income tax framework inspired by Mirrlees (1971)⁴--see Oswald (1983), Tuomala (1990), Ireland (2001)⁵. In this paper we revisit these questions and extend the earlier work in the literature. We present a more detailed analysis of the solution to the non-linear income tax problem with consumption interdependence, including cases with special objective functions and special distributions, as well as numerical simulations. Of particular interest for us is the interplay between inequality and relativity in determining the optimal tax schedule.

¹ Clark, Frijters and Shields (2006) provide a good survey.

² Later on Duesenberry (1949), Galbraith (1958), Hirsch (1976) and Frank (1985, 1997) among others have written about the importance of relative position as a dominant spending motivation.

³ More recent empirical research findings show that relative consumption concerns have important effects on consumption but little, if any, on leisure (Clark and Oswald, 1996).

⁴ Aronsson-Johansson-Stenman (2008) address public good provision in this framework.

⁵ Boskin-Sheshinski (1978) and Blomquist (1993) consider linear income tax policy with relative consumption. Bowles and Park (2005) consider a simple two-class tax model. Their model takes each individual's reference consumption to be exogenous.

The plan of the paper is as follows. Section 2 presents the optimal income tax model with relative consumption concern. In Section 3 we consider implications of relative concern in the optimal nonlinear income tax model with Rawlsian social objectives and with special distributional assumptions. Section 4 presents numerical simulations in the Utilitarian case. Section 5 elaborates on the interplay between relativity and inequality in determining optimal tax rates. Section 6 discusses the results in the broader context of optimal taxation and behavioural public economics, and Section 7 concludes.

2. Optimal non linear taxation and relative consumption concern

Do people make comparisons between or among individuals of similar incomes? Or is the lifestyle of the upper middle class and the rich a more salient point of reference for people throughout the income distribution? A comparison consumption level can be constructed as follows. Let x denote consumption, and let

$$\mu = \int \omega(n)x(n)f(n)dn \quad (1)$$

where a distribution of wages (productivities), denoted by n , on the interval $(0, \infty)$ is represented by the density function $f(n)$. There are a number of alternative interpretations of the variable μ . The simplest one is obtained if each of the ω weight is equal to one. In this case the average consumption is the comparison consumption level. We can choose the weights ω so that μ is the consumption of the richest individual (this corresponds to Veblen's idea), of the median individual or something in between the richest and the median. It is difficult to say without empirical evidence which is the most plausible interpretation. Moreover, as Layard (1980) suggests, that people may have different μ values. In this paper, we restrict attention to the case where $\omega=1$ for all n so that μ is the average consumption of people in the economy.

We consider a one period model with labour as the only source of income. There is a continuum of individuals, each having the same preference ordering, which is represented by an additive utility function

$$u = U(x) + \psi(\mu) - V(y) \quad (2)$$

where x is a composite consumption good, μ is a comparison consumption level, and hours worked are y , with $U_x > 0$, $\psi_\mu < (>)0$ and $V_y < 0$ (subscripts indicating partial derivatives) and where $V(\cdot)$ is convex. As typical in optimal tax literature, we have to make simplifying assumption like this separability assumption to be able make progress in our understanding of the optimal schedules. Workers differ only in the pre-tax wage n they can earn. Gross income

$$z = ny \tag{3}$$

and consumption, x , is after-tax income.

Suppose that the aim of policy can be expressed as maximizing the following social welfare criterion

$$S = \int_0^{\infty} W(u(n))f(n)dn \tag{4}$$

where $W(\cdot)$ is an increasing and concave function of utility.

We should note before moving on that there are many difficult problems with formulation of the social welfare function. For example, we must decide whether the government ought to accept relative income concerns in social welfare. This is closely related to the awkward question of whether we should include antisocial preferences such as envy, malice etc. in social welfare function or not. If so, it would be important to consider the case where the government is “non-welfarist” (paternalistic). But it could be argued that to the extent relative concerns or Veblen effects are real, it should be respected when evaluating social welfare.⁶ In this paper we follow the latter, “welfarist”, route.

The government cannot observe individuals’ productivities and thus is restricted to setting taxes and transfers as a function only of earnings, $T(z(n))$. The government maximizes S subject to the revenue constraint

$$\int_0^{\infty} T(z(n))f(n)dn = R \tag{5}$$

where in the Mirrlees tradition R is interpreted as the required revenue for essential public goods. The more non-tax revenue a government receives from external sources, the lower is R .

⁶ Examples of the first include Kanbur, Keen and Tuomala (1994) and Pirttilä and Tuomala (2004), while O’Donoghue and Rabin (2003), Bernheim and Rangel (2007) and McCaffery and Slemrod (2006) are examples of the latter. See Seade (1980) for seminal work.

Totally differentiating utility with respect to n , and making use of workers utility maximization condition, we obtain the incentive compatibility constraints,

$$\frac{du}{dn} = -\frac{yV_y}{n} = g. \quad (6)$$

Since $T = ny - x$, we can think of government as choosing schedules $x(n)$, $y(n)$ and μ . In fact it is easier to think of it choosing a pair of functions, $u(n)$, $y(n)$ and μ , which maximize welfare index (4) subject to the revenue requirement (5), the incentive compatibility condition (6) and the comparison condition (1). We focus on the case where $\omega=1$ for all n so that μ is the average consumption of people in the economy. Introducing Lagrange multipliers λ , $\alpha(n)$ and γ for the constraints (5), (6) and (1) and integrating by parts, the Lagrangean becomes

$$L = \int_0^{\infty} (W(u) + \lambda(ny - x))f(n) + \gamma(\mu - xf) - \alpha'u - \alpha g dn + \alpha(\infty)u(\infty) - \alpha(0)u(0) \quad (7)$$

Differentiating with respect to u , y and μ gives the first-order conditions⁸

$$L_u = [W' - h_u(\lambda + \gamma)]f(n) - \alpha'(n) = 0 \quad (8)$$

$$L_y = \lambda(n - h_y)f(n) - \gamma h_y f(n) + \alpha(n)(V_y + yV_{yy}) = 0 \quad (9)$$

$$L_\mu = \int -\lambda h_\mu f(n) dn + \gamma - \int \gamma h_\mu f(n) dn = 0 \quad (10)$$

⁷ The 1.order condition of individual's optimisation problem is only a necessary condition for the individual's choice to be optimal, but we assume here that it is sufficient as well. Assumptions that assure sufficiency are provided by Mirrlees (1976). Note also that while we here presume an internal solution for y , (6) remains valid even if individuals were bunched at $y=0$ since, for them, $du/dn = 0$.

⁸ Inverting utility we have $x = h(u, y, \mu)$ and calculating the derivatives

$$h_y = -V_y / U_x, h_x = 1 / U_x, h_\mu = -\psi_\mu / U_x$$

$$(10) \text{ implies } \frac{\gamma}{\lambda} = \frac{\int h_{\mu} f(n) dn}{1 - \int h_{\mu} f(n) dn} \quad (11)$$

(8) satisfies the transversality conditions

$$\frac{\partial L}{\partial u(0)} = \alpha(0) = 0; \frac{\partial L}{\partial u(\infty)} = \alpha(\infty) = 0$$

and

$$\mu(n) > 0, \text{ for } n \in (0, \infty),$$

Integrating in (8)⁹

$$\alpha(n) = \int_n^{\infty} [W' - \frac{(\lambda + \gamma)}{u_x}] f(p) dp \quad (12)$$

From the first order conditions of government's maximization, we obtain the following condition

for optimal marginal tax rate $t(z) = T'(z)$; [Note: $\frac{t}{1-t} = \frac{1}{1-t} - 1 = \frac{U_x n}{V_y} - 1$]

$$\frac{t}{1-t} = \frac{\gamma}{\lambda} + (1 + \frac{yV_{yy}}{V_y}) \frac{U_x}{\lambda n f(n)} \int_n^{\infty} [\frac{(\lambda + \gamma)}{U_x} - W'] f(p) dp \quad (13)$$

It is worth noting that the so called end-point results do not hold any more. From (13) and the transversality conditions $\alpha(0) = \alpha(\infty) = 0$ the marginal tax rates are positive at the both ends when $\psi_{\mu} < 0$ (see Oswald, 1983). This is also true with other comparators. Going beyond average consumption. As shown in Tuomala (1990) the separability assumption used in Oswald (1983) can be weakened so that μ affects individuals' choices. Unfortunately we are not able to say more on the shape of tax schedule with the weaker separability condition.

⁹ Integrating in (8)

$$\int_0^n \frac{d\alpha}{dn} dn = \alpha(n) - \alpha(0)$$

Multiplying and dividing (13) by $(1 - F(n))$ we can write the formula for marginal rates;

$$\frac{t}{1-t} = \frac{\gamma}{\lambda} + \underbrace{[1 + \zeta]}_{A_n} \underbrace{\left[\frac{[1 - F(n)]}{nf(n)} \right]}_{B_n} \underbrace{\left[\frac{\int_n^\infty \left[1 - \frac{W' U_x^{(p)}}{\lambda(1 + \gamma/\lambda)} \right] \frac{(1 + \gamma/\lambda) U_x}{U_x^{(p)}} f(p) dp}{(1 - F(n))} \right]}_{C_n}$$

(14)

$$\text{where } \zeta = \frac{yV_{yy}}{V_y}.$$

The first term on the right hand side of (14) for the marginal income tax rate is analogous to a Pigouvian tax correcting for an externality. It could also be called a first-best motive for taxation, as it corrects the individual activity to correspond to social preferences. From (14), there are in addition to the externality term three elements on the right hand side of (14) that determine optimum tax rates: elasticity and income effects (A&C), the shape of the skill distribution (B&C) and social marginal weights (C). On the basis of (14) we can notice that if the utility of individuals depends negatively on the comparison consumption the marginal tax rate of the highest income is positive.

It should however be clear from (14) that the variation of the optimal marginal tax rate with the level of income is a complex matter. It is clear that explicit solutions to the optimal income tax problem are difficult to obtain without simplifying assumptions. The terms in (14) simplify if we assume, as in Atkinson (1995) and Diamond (1998), quasi-linear preferences with $U_x = 1$. The marginal tax rate formula then reduces to:

$$\frac{t}{1-t} = \frac{\gamma}{\lambda} + \underbrace{[1 + \zeta]}_{A_n} \underbrace{\left[\frac{[1 - F(n)]}{nf(n)} \right]}_{B_n} \underbrace{\left[\frac{\int_n^\infty \left[1 - \frac{W'}{\lambda(1 + \gamma/\lambda)} \right] (1 + \gamma/\lambda) f(p) dp}{(1 - F(n))} \right]}_{C_n} \quad (15)$$

But this is still too complex, with a number of different influences in play, to allow useful interpretation. We turn therefore to further assumptions, on the government's objective function and on the distribution of n , to provide further insights.

3. The Rawlsian case

If we assume the Rawlsian social objective¹⁰ then the factor C_n in (14) is constant. Then the pattern of marginal tax rates depends only on B , that is, on the shape of the n -distribution:

$$\frac{t}{1-t} = \frac{\gamma}{\lambda} + \underbrace{\left[1 + \zeta\right]}_{A_n} \underbrace{\left[\frac{1-F(n)}{nf(n)}\right]}_{B_n} \underbrace{\left(1 + \frac{\gamma}{\lambda}\right)}_{C_n} \quad (16)$$

We specify further the case with a maximin criterion so that the upper part of the n -distribution is the unbounded Pareto distribution, $f(n) = \frac{1}{n^{1+a}}$ for $a > 0$, and the utility function is $u = x - \varphi \bar{x} - y^{1+\frac{1}{\varepsilon}}$

Then using (11) we have

$$\frac{t}{1-t} = \phi + \left[1 + \frac{1}{\varepsilon}\right] \frac{1}{a} [1 + \phi] \quad (17)$$

where $\phi = \frac{\varphi}{1-\varphi}$.

Hence using the Rawlsian social welfare function we do not obtain the rising part of the U-shaped marginal tax rates as in Diamond (1998).¹¹ The optimal top marginal tax rate depends negatively on a , which is a measure of the thinness of the tail of the n -distribution., and it is decreasing in ε , as expected. Finally, it is increasing in φ , which measures the importance of relativity in this framework.

We illustrate numerically top marginal rates in the following tables. The table 1 presents the top marginal tax rates for parameter value when $a=2$ and 3, $\varphi=0$ and 1/2 and $\varepsilon=1/3$, $1/2$ and 1.

¹⁰ Maximizing utility of worst off person in the society is not the original version of Rawls (1972). It is a kind of welfarist version of Rawlsian. "To interpret the difference principle as the principle of maximin utility (the principle to maximize the well-being of the least advantaged person) is a serious misunderstanding from a philosophical standpoint." Rawls, 1982)

¹¹ In the general additive case with maximin, C is $\int (f(p)/U_x) dp$. It is declining with n since $u(x)$ is concave and the intergral term declines in n . This might suggest declining marginal rates.

Table 1 shows how the top marginal tax rate decreases when the elasticity of labour supply ϵ increases, the Pareto parameter a increases and the degree of relative consumption concern declines. The results in Table 1 depend on the chosen distribution of wages.

If the whole distribution of wages is an unbounded Pareto distribution, then optimal marginal tax rates are constant and positive. This implies that the optimal tax function is linear

$$T(z) = k + \tau z \quad (18)$$

The average tax rate is

$$\frac{T(z)}{z} = \frac{k}{z} + t \quad (19)$$

where $t = \frac{b}{1+b}$ is between zero and one and where $b = \phi + \left[1 + \frac{1}{\epsilon}\right] \frac{1}{a} [1 + \phi]$.

Equation (19) implies that average tax rates are increasing if and only if k is negative. If preferences are quasi-linear in consumption and the distribution of n is an unbounded Pareto distribution, a maxim criterion implies increasing average tax rates in income. When the elasticity of labour supply is not constant, the problem becomes more complicated. Then it is not possible without simulations to say anything about the shape of tax schedule.

We now consider two alternatives to the Pareto distribution: (i) the Champernowne (1952) distribution and (ii) the lognormal distribution (with parameters m and σ (see Aitchison and Brown, 1957)¹²). As is well known, the lognormal distribution fits reasonable well over a large part of income range but diverges markedly at the both tails. The Pareto distribution in turn fits well at the upper tail. Champernowne (1952) proposes a model in which individual incomes ARE assumed to follow a random walk in the logarithmic scale. Here we use the two parameter version of the Champernowne distribution. This distribution approaches asymptotically a form of Pareto distribution for large values of wages but it also has an interior maximum. As for the lognormal, the Champernowne distribution exhibits the following features: asymmetry, a left humpback and long right-hand tail; but it has a thicker upper tail than in the lognormal case.

The probability density function of the Champernowne distribution is

¹² $Ln(n; m, \sigma^2)$ with support $[0, \infty)$. The first parameter m is log of the median and the second parameter is the variance of log wage. The latter one is itself an inequality measure.

$$(18) \quad f(n) = \theta \left(\frac{m^\theta n^{\theta-1}}{(m^\theta + n^\theta)^2} \right)$$

in which θ is a shape parameter and m is a scale parameter. The cumulative distribution function is

$$(19) \quad F(n) = 1 - \frac{m^\theta}{(m^\theta + n^\theta)}$$

The inverse hazard rate is:

$$(20) \quad \lim_{n \rightarrow \infty} \frac{1 - F(n)}{nf(n)} = \lim_{n \rightarrow \infty} \frac{m^\theta + n^\theta}{\theta n^\theta} \rightarrow \frac{1}{\theta}.$$

Eq (20) confirms that the Champernowne distribution approaches asymptotically a form of Pareto distribution for large values of wages.

Calculating the inverse hazard ratio or Mills ratio $\frac{1 - F(n)}{nf(n)}$ for Champernowne distribution (with different parameter values of $\theta = 2$ and 3 and $m = e^{-1}$) and lognormal distribution (with different parameter values of $\sigma = 0.39$ and 0.7 and $m = e^{-1}$, see Figure 1a and b) we obtain from (16) the marginal tax rates with the Rawlsian case. The results are shown for different percentile points of the distribution.

Note that these (in Tables 1, 2 and 3) are marginal rates for all taxes that vary with income, and should be compared with the schedules for total of taxes on income and expenditures in real economies. From Tables 2 and 3 we can see that the marginal tax rates decrease with labor supply elasticities as expected. We also see that marginal tax rates are throughout much lower for lognormal case than for the Champernowne distribution. In other words the choice of the functional form of the n -distribution matters greatly. The results in Tables 2 and 3 again confirm that zero is a poor approximation even for the top 0.1 per cent. Finally and most importantly from our point of view, as the degree of relative consumption concern increases, (i) marginal tax rates increase throughout, (ii) they increase more at higher levels of income and (iii) with the result that the fall off of marginal tax rates is less steep, and in this sense the tax structure is more progressive.

5. The Utilitarian Case

The special cases considered in the previous section yield insights but within the framework of the assumptions made. How robust are these insights? For example, the Rawlsian objective embodies extreme inequality aversion. What happens at more moderate levels of inequality aversion? What happens when we move away from quasi-linearity? This section presents optimal tax schedules with alternative assumptions. We capture intermediate levels of inequality aversion through a utilitarian social welfare function with constant absolute utility-inequality aversion:

$$S(u) = -\frac{1}{\beta} e^{-\beta u} \quad (20)$$

where β measures the degree of inequality aversion (in the case of $\beta = 0$, we define $W = u$). We further move away from quasi-linearity and use the following utility function

$$U = \log x + \varphi \log \frac{x}{\mu} + \log(1 - y) \quad (21)$$

where μ is the comparison consumption level, φ is a degree of relative income concern. Of course, the form in (21) restricts the range of the elasticity of labor supply. It is important to note that (21) does not only affect directly individuals utility levels but it also has behavioural effects, namely, relativity concerns (φ) can change an individual's marginal rate of substitution between consumption and labour supply. This can be seen from individuals utility maximization condition; $x/(1 + \varphi)(1 - y) = 1 - t$. For distribution, we assume that $f(n)$ is lognormal density (m, σ) (mean, stand dev.)

The optimal tax schedules are calculated numerically. The results of the simulations are summarized below in Tables 4-8. In these Tables, R (or X/Z) is revenue requirement (R=0 means pure redistributive system), ATR is average tax rate and MTR is marginal tax rate. The Tables give labour supply, y , gross income, z , net income, x and optimal average (ATR) and marginal tax rates (MTR) at various percentiles of the ability distribution. The Tables also provide the decile ratio (P90/P10) for net income and gross income. Since marginal tax rates may be a poor indication of

the redistribution powers of an optimal tax structure we measure the extent of redistribution, denoted by RD, as the proportional reduction between the decile ratio for market income, z , and the decile ratio for disposable income, x . Tables 4-8 give comparisons as ϕ and σ vary. Figures 2 -6 show marginal tax rates for different parameters.

Several patterns emerge from the simulations presented here, focusing specifically on the impact of relativity on progressivity. As the parameter ϕ increases, (i) marginal tax rates increase at all levels of income, (ii) the drop off in marginal tax rates for higher income levels is mitigated, and (iii) our redistribution measure, RD, increases. The case for greater progressivity in the tax schedule, in these senses, comes through in the cases examined here—it is not just a property of the Rawlsian objective function, nor restricted to the Pareto or the Champernowne distributions.

To further examine how sensitive the shape of the tax schedule and working hours are to the choice of the parameter ϕ in the utility function and σ , inherent inequality we computed solutions for $\phi=1.0$ and 3.0 in the case of the utility function (21) and $\sigma=0.5$ and 0.7 . These results are shown in Figures 3 -6. We find that when ϕ and inherent inequality increase the marginal tax rates are higher and increasing with income up to around $F(n)=0.99$. These results reinforce the findings of Kanbur - Tuomala (1994) that when higher values of inherent inequality are used optimal marginal tax rates increase with the income over the majority of the population. It turns out that when we increase ϕ , individuals above the median work more in order to retain their relative position. When we increase simultaneously both ϕ and σ , then only those in the top decile increase working hours. To relate these results to empirical labour supply studies we give the values of the uncompensated elasticity, E^u and uncompensated elasticity E^u ¹³.

The optimum is typically characterized by a certain fraction of individuals, at the bottom end, choosing not to work (where we have $dx/dn=dz/dn=0$, there is bunching of individuals of different n). When ϕ is zero there is very little bunching (eg, in the case of Table 4 it turns out to be

¹³ Differentiating the FOC of the individual maximization, $U_x n(1-t) + V_y = 0$, with respect to net wage, labour

supply and virtual income, b , we have after some manipulation elasticity formulas; $E^u = \frac{(V_y / y) - (V_y / U_x)^2 U_{xx}}{V_{yy} + (V_y / U_x)^2 U_{xx}}$,

(income effect parameter) $I = \frac{-(V_y / U_x)^2 U_{xx}}{V_{yy} + (V_y / U_x)^2 U_{xx}}$, and from the Slutsky equation $E^c = E^u - I$, then

$$E^c = \frac{(V_y / y)}{V_{yy} + (V_y / U_x)^2 U_{xx}}$$

practically zero ($F(n_0)=0.0003$). With greater relativity concern bunching increases slightly. In the case ($\sigma=0.5$ and $\varphi=1$ (table 4)) $F(n_0)=0.015$. Our numerical results reveal that the amount of bunching is quite sensitive to the greater inequality. When $\sigma=0.7$ and $\varphi=3$ (table 8) the amount of bunching is 11.7 %, i.e, $F(n_0)=0.117$. Hence the inherent inequality is very crucial in determining the amount of bunching.

5. Relativity and Inequality

In this section we look at the interaction of inequality and the strength of relativity in determining the optimal tax schedule. We know that progressivity increases with greater relativity concern. (Figure 1). The greater relativity concern increases the marginal tax rates throughout and they increase more at the higher level of income. We also know, from Kanbur and Tuomala (1994), that progressivity increases with inequality (confirmed in Figure 2). To see how the impact of greater relativity affects progressivity at successively higher levels of pre-tax inequality, we have computed solutions for different parameter values of relative consumption concern φ given pre-tax inequality, and then repeated the exercise at a higher level of inequality. From Figures 3 and 4 we see that the greater relativity increases progressivity but this impact is dampening with increasing inequality. Similarly, we can ask how the impact of greater inequality affects progressivity with greater strength of relativity. From Figures 5 and 6 we see that the greater inequality increases progressivity but this impact on progressivity is in turn declining in an increase of relativity. Thus it seems that, in these numerical simulations at least, relativity and inequality do not compound each other's incremental effect on progressivity. Further research is needed to understand the detailed nature of this result.

Given the inherent complexities of optimal non-linear income taxation, it is not straightforward to develop an intuition for this result. But we can take the first steps towards understanding as follows. Suppose first the case with a fixed cake. If average consumption increases by one Euro, individuals' utilities go down as average consumption goes up. Lowering taxes increase average consumption and consequently lower utility, other things equal. Hence marginal tax rates should be higher than otherwise. With the utility function (21) there is a simple relationship between λ and γ : $\gamma = \varphi\lambda$. Hence we see from (14) that an increase in φ increases marginal rates in the first best case. This is also true in the second best case at the endpoints of the distribution. Otherwise things are more

complicated in the second best world. Let us now focus on the C-term in (14). This term measures the social welfare gain from slightly increasing the marginal tax rate at n and distributing as a poll subsidy to those below n the revenue raised from consequent increase in average tax rates above n . The first term in the integrand tends to favour rising marginal rates. The higher n , the lower is W' , the lower would be the marginal utility of consumption, U_x . Hence at higher n the average value of $(W'U_x / \lambda(1 + \gamma / \lambda)) / (1 - F)$ is smaller and the whole term $1 - W'U_x / \lambda(1 + \gamma)$ is larger. With the utility function (21) this term becomes $1 - \frac{1}{\lambda x}$. So the direct effects of relativity concerns disappear.

Without simulations it is not possible say how greater relativity concern changes λx . On the other hand an increase in pre-tax inequality affects λ , the marginal cost of public funds. Kanbur-Tuomala (1993) shows the complications that can arise in signing the C-term as a function of mean preserving spreads in the distribution of n .

Finally, how are these relationships in turn affected by the elasticity of labor supply? As shown in Tables 4-8 both compensated and uncompensated labour supply elasticities are decreasing with income in all cases displayed in Tables. At the upper part of the distribution the labour supply elasticities are declining with greater relativity and inequality. Unfortunately there is little empirical evidence regarding the relationship between labour supply elasticities and wage rates¹⁴.

The income effects enter through the terms A and C. In the term A¹⁵ it affects how elasticities vary with skill. As shown by Chetty (2006) there is a relationship between risk aversion and the labour supply. Or to put it another way there is the connection between the curvature of the utility function

¹⁴ Røed and Ström (2002) (table 1 and 2) offer a review of the existing more recent evidence. They conclude that the limited evidence indicates that labour supply elasticities are declining with household income. Using Norwegian data Aaberge-Colombino (2006) provides support for declining elasticities. High labour supply elasticities among low-wage workers is also confirmed by empirical evaluations of various in-work benefit schemes operating in the US, UK and some other countries, see Blundell (2000) for a review. By contrast, there is empirical evidence on the elasticity of taxable income that higher elasticities are among high income individuals. See eg. Gruber-Saez (2002).

¹⁵ The marginal utility of consumption, U_x , and the term A in (14) has its origin in first order condition for labour supply (incentive compatibility condition). It shows the rate at which utility changes with n . The greater is U_x the more it changes as labour supply y is increased.

and the ratio of income and wage elasticities.¹⁶ We can see this link by differentiating of the FOC of individual's problem and using the Slutsky equation:

$$\frac{\partial y / \partial b}{\partial y^c / \partial n} = \frac{U_{xx} n}{U_x} \quad (22)$$

where y^c is a compensated labour supply and b is virtual income. As seen from (22) the curvature of the utility function with respect to consumption (the coefficient of relative risk aversion) is important because the labour supply response to an increase in income is related to how much the marginal utility of consumption changes as income changes. If U_{xx} is large, the marginal utility of consumption falls sharply as income rises, so that the taxpayer will reduce labour supply when his or her earnings rise. In fact this is the case here. We have assumed additively separable utility in comparison consumption (average consumption). This property rules out direct behavioural consequences in envy. But the utility function (21) does not abandon relativity effects on labour supply. The utility function (21) implies that U_{xx} is larger for low income people than high people. It also increases with ϕ . The labour supply is much smaller below the median income than without relativity concerns. In simulations those above median in turn seem to work harder.

6. Discussion and Conclusion

As noted in the introduction, there are a few papers in the literature that have attempted to analyze the structure of optimal income taxation in the presence of relative concerns. How does our paper compare with these exercises?

Boskin - Sheshinski (1978) construct an educational investment model in which the individual's income is determined by his income and demonstrated that increased concern for relative consumption in the optimal linear income tax leads to larger lump sum subsidies and higher tax rates. Oswald (1983) studies a more general optimal non-linear tax problem in a world in which there is altruism and envy. He takes the standard utility function as a function of consumption and leisure and adds a concern for consumption of others. He considers mainly the case where the comparison is average consumption (as we do here). Another simplifying assumption he makes is that envy (or altruism) has no effect on consumption decision and labour supply. With these

¹⁶ See the formula (7) in Chetty (2006)

assumptions and using a ceteris paribus argument he reaches the conclusion"optimal marginal tax rates are higher in a predominantly jealous world."

Ireland (2001) incorporates a social status-signalling mechanism into the Mirrlees model. In a model where individuals signal status with consumption, e.g. large houses, cars, boats etc., he finds status seeking leads to higher marginal tax rates, but not a more progressive rate structure. His results are based on quasi-linear preferences, unbounded Pareto distribution and utilitarian social welfare function. So he confirms Diamond's (1998) result – the U-shaped marginal tax rate structure - with the unbounded Pareto distribution. Our use of a maximin objective eliminates the rising part of the U-shaped marginal tax rate structure. This is also the case when we assume a truncated Pareto distribution. Ireland (2001) does not discuss any other distributions. In Ireland's model status does not affect the endpoint results. It seems to us that the negative externality (envy) story is more convincing especially at the top. Further, Ireland does not compute numerical solutions in the case with income effects.

Thus our paper supports the conclusion in the literature that relativity leads to higher marginal tax rates. It both generalizes some of the conditions under which this result is obtained in the literature, and fleshes out the detailed structure for optimal marginal tax rates for specific functional forms of distribution, utility function, and social welfare function. By and large, we find support for greater progressivity, as we define it, in the tax structure as relativity concern increases. And none of the papers in the literature, to our knowledge, highlights the interplay of relativity and inequality in determining the optimal structure of taxes.

More work is needed to further explore this interaction between relativity and inequality that our numerical simulations have uncovered. However, a major direction for further research is to explore what happens when the government does not accept the relative concerns of individuals and maximizes a non-welfarist objective function. The general approach proposed in Kanbur, Pirttilä and Tuomala (2006) is one way to go, but further research awaits.

References

- Aaberge, R. and Colombino, U. (2006), Designing optimal taxes with a microeconomic model of household labour supply, IZA, DP No.2468.
- Aronsson, T. and O. Johansson-Stenman (2008), When the Joneses' consumption hurts: Optimal public good provision and nonlinear income taxation, *Journal of Public Economics* 92,986–997.
- Atkinson, A.B. (1990), Public economics and economic public, *European Economic review* 34, 225-248.
- Atkinson, A.B., (1995), *Public economics in Action, The basic income/flat tax proposal, The Lindahl Lectures*, Oxford University Press.
- Blomquist, S. (1993), Interdependent behaviour and the Effect of taxes, *Journal of Public Economics* 51,211–18.
- Boskin, M.J. and E. Sheshinski (1978), Individual welfare depends upon relative income, *Quarterly Journal of Economics* 92,589–601.
- Bowles, S. and Y. Park (2005), Emulation, Inequality, and Work Hours: Was Thorsten Veblen Right? *Economic Journal* 115, F397–F412.
- Clark, P. Frijters and M. Shields (2007??), Relative Income, Happiness and Utility: An Explanation for Easterlin Paradox and Other Puzzles. IZA DP No. 2840.
- Clark, A.E. and A. Oswald (1996), Satisfaction and Comparison Income. *Journal of Public Economics* 61:3, 359–381.
- Champernowne, D.G. (1952), The graduation of income distributions, *Econometrica* 20, 591–615.
- Diamond, P. (1998), Optimal income taxation: an example with a U-shaped pattern of optimal marginal tax rates, *American Economic Review*, 88(1), pp.83–95.
- Duesenberry, J. S. (1949), "Income, Saving and the Theory of Consumer Behaviour". Harvard University Press, Cambridge, MA.
- Frank, R. (1985), *Choosing the Right Pond*. Oxford University Press, Oxford.
- Frank, R. (1997), The Frame of Reference as a Public Good. *Economic Journal* 107:445, 1832–1847.
- Galbraith, J.K. (1958), *The Affluent Society*. Houghton Mifflin Company, Boston.
- Gruber, J. and E. Saez, (2002), The Elasticity of Taxable Income: Evidence and Implications, *Journal of Public Economics*
- Hirsch, F. (1976), *Social Limits to Growth*. Harvard University Press, Cambridge Ma.

- Ireland, N. (2001), Optimal income tax in the presence of status effects, *Journal of Public Economics* 81, 193–212.
- Kanbur, R. and M. Tuomala (1993), Inherent inequality and the optimal graduation of marginal tax rates, *Tampere Economic Working Papers* 6/1993.
- Kanbur, R. and M. Tuomala (1994), Inherent inequality and the optimal graduation of marginal tax rates, *Scandinavian Journal of Economics* 96 (2), pp. 275–282.
- Kanbur, R., Keen, M. and Tuomala, M. (1994a) ‘Optimal non-linear income taxation for the alleviation of income-poverty’, *European Economic Review* 38, 1613-32.
- Kanbur R., Pirttilä J and M. Tuomala, 2006, Non-Welfarist Optimal taxation and Behavioural Public Economics, *Journal of Economic Surveys*, vol. 20, no.5, 849-868.
- Layard, R. (1980), Human satisfaction and public policy, *Economic Journal* 90: 360, 737–750.
- Mirrlees, J.A., (1971), An exploration in the theory of optimum income taxation, *Review of Economic Studies* 38, pp.175–208.
- O’Donoghue, T. and M. Rabin (2003) ‘Studying optimal paternalism, illustrated by a model of sin taxes’, *American Economic Review* 93, 186-191.
- Oswald, A. (1983), Altruism, jealousy and the theory of optimal non-linear taxation, *Journal of Public Economics* 20, 77–87.
- Rawls, J. (1971), *Theory of Justice*. Harvard University Press, Cambridge Ma.
- Rawls, J. (1982), Social unity and primary goods, in *Utilitarianism and Beyond*, eds by A.Sen and B.Williams, Cambridge University Press, 159-85.
- Rothschild, M and Stiglitz, J. (1971), Increasing risk I; A definition, *Journal of Economic Theory* 2, 225-243.
- Röed, K and S.Ström, 2002, Progressive taxes and the labour market: Is the trade-off between equality and efficiency inevitable? *Journal of Economic Surveys* vol 16 no1.
- Sandmo, A. (1983) ‘Ex post welfare economics and the theory of merit goods’, *Economica* 50, 19-33.
- Seade, J. (1980) ‘Optimal non-linear policies for non-utilitarian motives’, in D. Collard, R. Lecomber and M. Slater (eds.) *Income distribution: the limits to redistribution*, Bristol: Scientifica.
- Tuomala, M. (1990), *Optimal Income Tax and Redistribution*. Clarendon Press, Oxford.

Table 1**Rawlsian top marginal tax rates (%) when people care about relative consumption**

	$\epsilon=1/3$	$\epsilon=1/3$	$\epsilon=1/2$	$\epsilon=1/2$	$\epsilon=1$	$\epsilon=1$
Relative concern	$a=2$	$a=3$	$a=2$	$a=3$	$a=2$	$a=3$
$\varphi=0$	66.6	57	60	50	50	40
$\varphi=1/2$	83.3	78.6	80	75	75	70

Table 2**Rawlsian marginal tax rates (%) with the Champernowne distribution**

	$\epsilon=1/3$	$\epsilon=1/3$	$\epsilon=1/3$	$\epsilon=1/3$	$\epsilon=1$	$\epsilon=1$	$\epsilon=1$	$\epsilon=1$
	$\varphi=0$	$\varphi=1/2$	$\varphi=0$	$\varphi=1/2$	$\varphi=0$	$\varphi=1/2$	$\varphi=0$	$\varphi=1/2$
F(n)	$\theta=2$	$\theta=2$	$\theta=3$	$\theta=3$	$\theta=2$	$\theta=2$	$\theta=3$	$\theta=3$
0.10	94.7	97.2	93.3	96.4	89.2	94.6	86.7	93.4
0.20	91.5	95.6	85.9	92.5	83.2	91.6	73.8	86.9
0.50	81.2	90	74.1	86.3	66.6	83.3	56.9	78.4
0.75	79.8	86.3	66.1	82.1	56.9	78.4	47.4	73.7
0.90	70.4	84.3	61.4	79.8	52.3	76.2	42.3	71.2
0.95	69.2	83.7	60	79.1	50.9	75.5	41	70.5
0.99	67	82.6	58.2	78.1	48.4	74.2	39.5	69.5
0.999	58.7	78.4	52.2	75.1	39.9	69.2	33.5	66.8

Table 3**Rawlsian marginal tax rates (%) with the lognormal distribution**

	$\epsilon=1/3$	$\epsilon=1/3$	$\epsilon=1/3$	$\epsilon=1/3$	$\epsilon=1$	$\epsilon=1$	$\epsilon=1$	$\epsilon=1$
	$\varphi=0$	$\varphi=1/2$	$\varphi=0$	$\varphi=1/2$	$\varphi=0$	$\varphi=1/2$	$\varphi=0$	$\varphi=1/2$
F(n)	$\sigma=0.7$	$\sigma=0.7$	$\sigma=0.39$	$\sigma=0.39$	$\sigma=0.7$	$\sigma=0.7$	$\sigma=0.39$	$\sigma=0.39$
0.10	93.4	96.7	89.4	94.7	87.8	93.9	80.8	90.4
0.20	87.6	93.8	79.1	89.6	78	89	65.5	82.7
0.50	77.7	88.9	65.9	82.9	63.6	81.8	49.1	74.6
0.75	68.2	84.1	55	77.5	51.7	75.8	37.9	68.9
0.90	61.5	80.7	46.9	73.5	44.4	72.2	30.7	65.9
0.95	59.5	78.7	43.8	71.9	30.2	71.1	28.1	64
0.99	51.7	75.8	36.9	68.5	34.9	67.4	22.6	61.3
0.999	47.5	73.8	31.7	65.9	31.1	65.6	18.8	59.4

Table 4

$\beta=0$	$\sigma=0.5$	$R=0.0$	$\varphi=0$				
F(n)	y	z	x	ATR%	MTR%	E^u	E^c
0.10	0.32	0.06	0.09	-50	30	1.05	0.55
0.50	0.41	0.15	0.15	-4	29	0.72	0.22
0.90	0.46	0.32	0.28	13	26	0.58	0.08
0.99	0.48	0.57	0.47	18	23	0.56	0.06
P(90/10)		5.33	3.11				
RD%			41.7				

$F(n_0)=0.0003$

Table 5

$\beta=0$	$\sigma=0.5$	$R=0.0$	$\varphi=1$				
F(n)	y	z	x	ATR%	MTR%	E^u	E^c
0.10	0.22	0.04	0.13	-200	58	2.48	2.7
0.50	0.41	0.15	0.17	-14	60	0.85	1.09
0.90	0.53	0.37	0.26	30	61	0.41	0.66
0.99	0.59	0.68	0.38	44	60	0.27	0.52
P(90/10)		9.25	2.0				
RD%			78.4				

$F(n_0)=0.015$

Table 6

$\beta=0$	$\sigma=0.7$	$R=0.0$	$\varphi=1$				
F(n)	y	z	x	ATR%	MTR%	E^u	E^c
0.10	0.04	0.01	0.12	-	60	14.65	15
0.50	0.37	0.14	0.16	-21	65	0.82	1.15
0.90	0.53	0.48	0.28	42	67	0.26	0.59
0.99	0.61	1.06	0.47	55	66	0.11	0.44
P(90/10)		48	2.33				
RD%			95.1				

$F(n_0)=0.086$

Table 7

$\beta=0$	$\sigma=0.5$	$R=0.0$	$\phi=3$				
F(n)	y	z	x	ATR%	MTR%	E^u	E^c
0.10	0.17	0.03	0.15	-361	77	6.37	6.75
0.50	0.44	0.16	0.18	-9	78	0.84	1.03
0.90	0.60	0.42	0.23	44	79	0.63	0.65
0.99	0.69	0.79	0.31	61	80	0.42	0.45
P(90/10)		14	1.53				
RD%			89.1				

$F(n_0)=0.04$

Table 8

$\beta=0$	$\sigma=0.7$	$R=0.0$	$\phi=3$				
F(n)	y	z	x	ATR%	MTR%	E^u	E^c
0.12	0.01	0.06	0.14	-173	79	48.8	49
0.50	0.40	0.15	0.17	-16	80	0.57	0.94
0.90	0.61	0.55	0.25	55	82	0.13	0.43
0.99	0.72	1.26	0.37	69	81	0.12	0.39
P(90/10)		9.17	1.78				
RD%			80.6				

$F(n_0)=0.117$

Figure 1 (1-F(n)/nf(n): Champernowne distribution

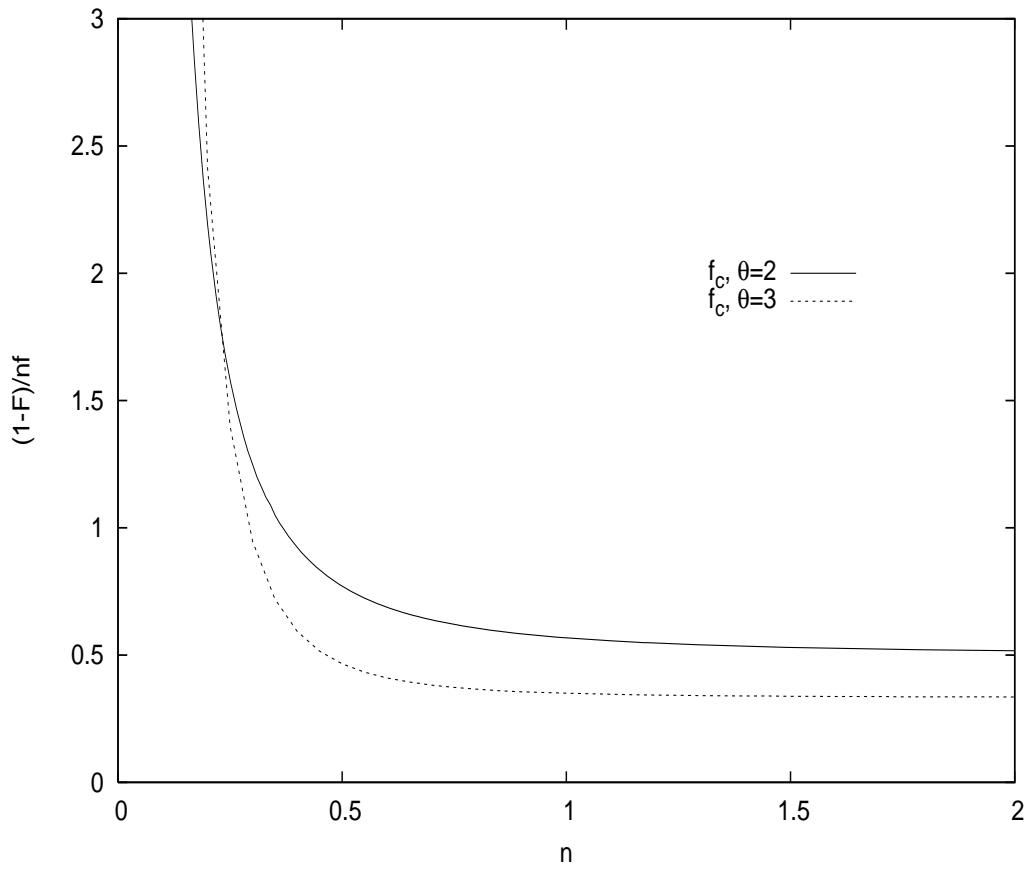


Figure 1 b (1-F(n)/nf(n): Lognormal distribution

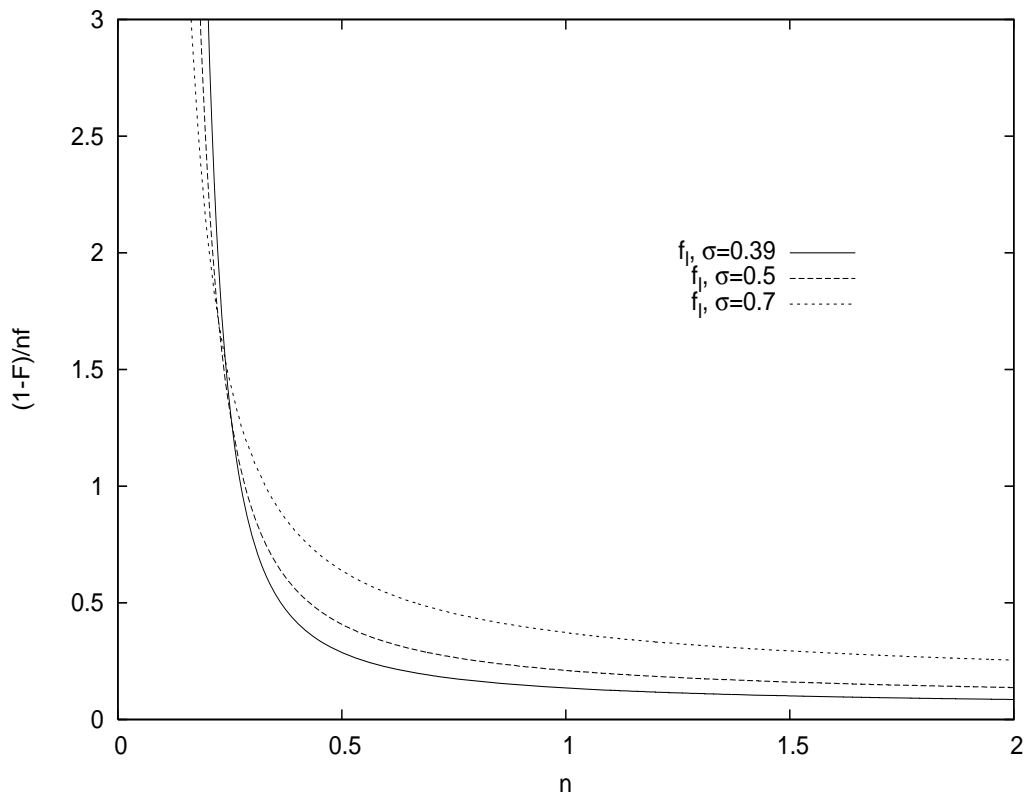


Figure 2 ($\sigma=0.5$)

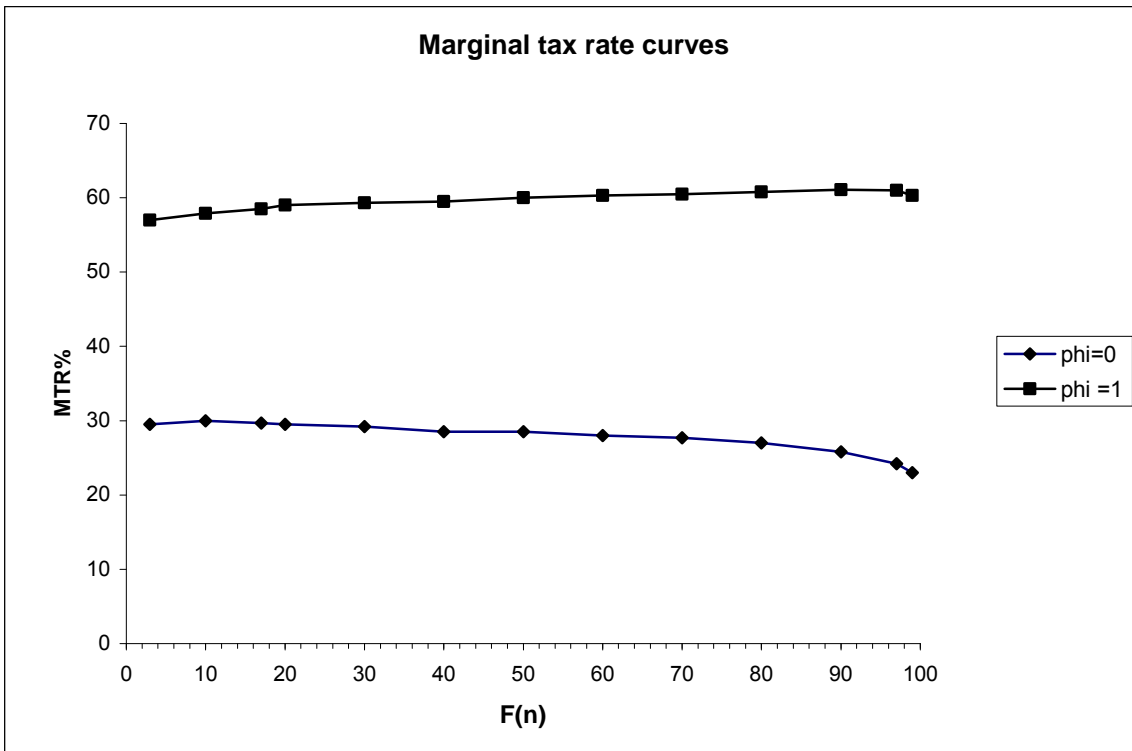


Figure 3 ($\varphi=1$)

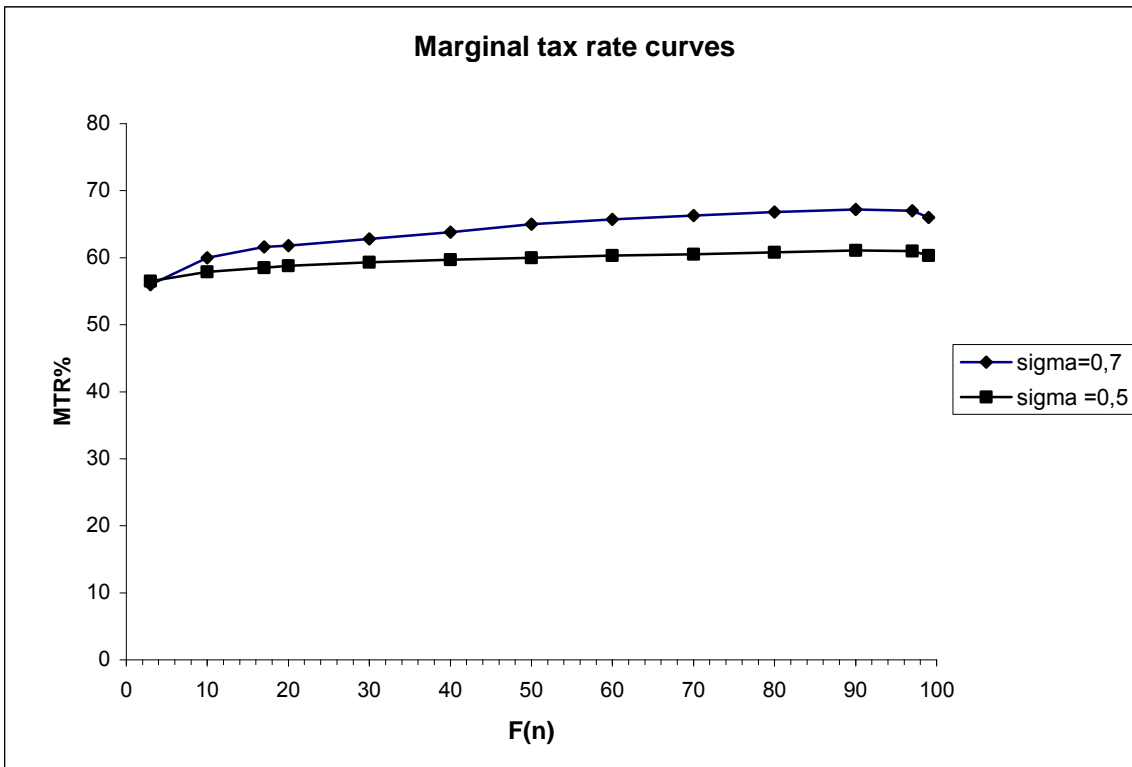


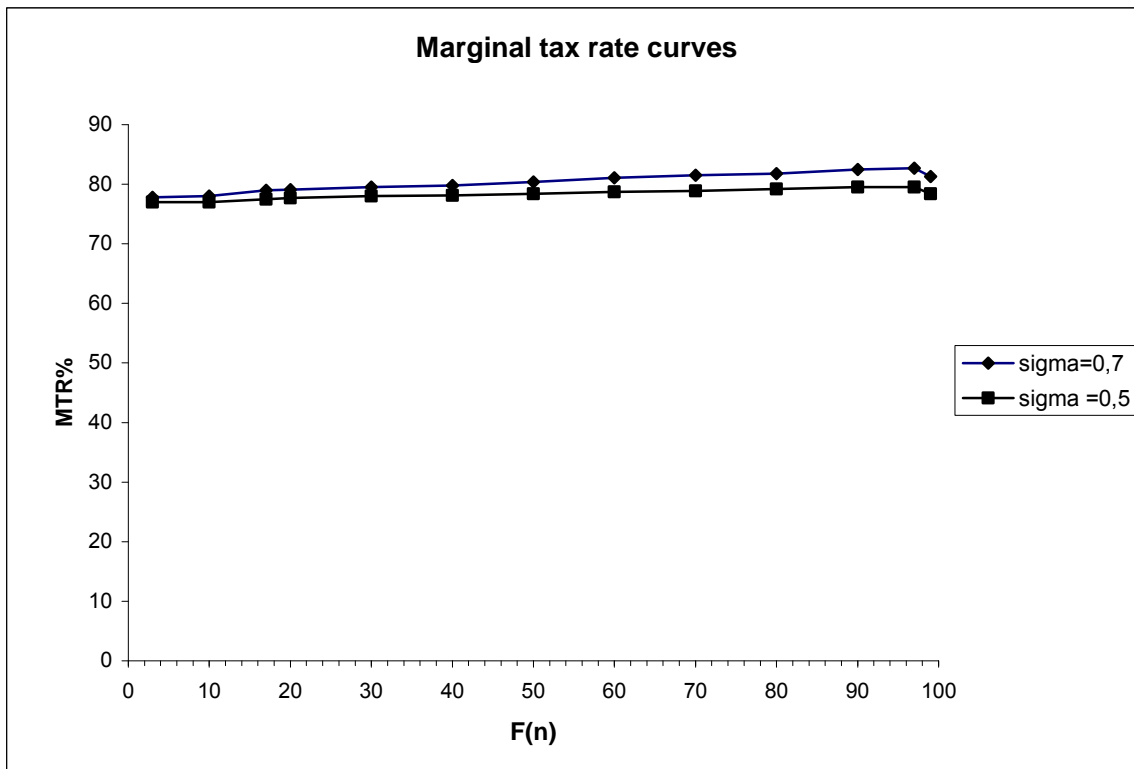
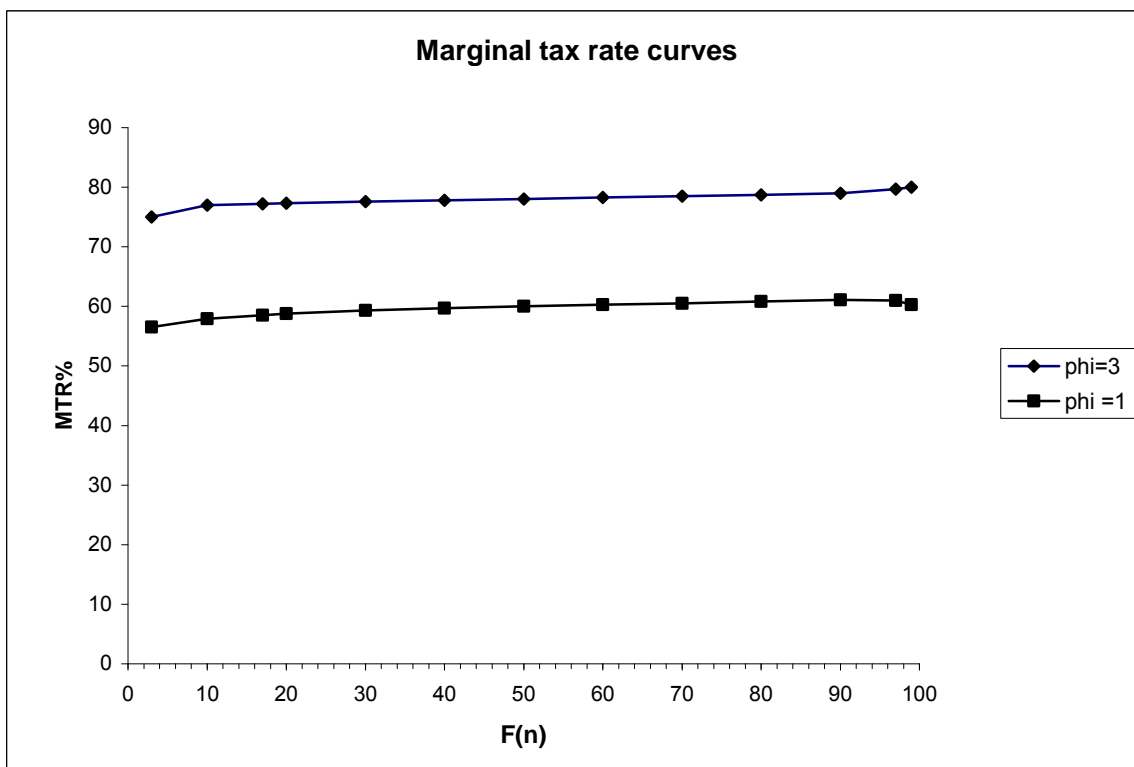
Figure 4 ($\varphi=3$)Figure 5 ($\sigma=0.5$)

Figure 6 ($\sigma=0.7$)