A Characterization for Marginal Income Tax Schedules

Oztek, Abdullah Selim

University College London

2 September 2019

A Characterization for Marginal Income Tax Schedules*

Abdullah Selim ÖZTEK[§]

University College London

Revised December 2019

Abstract

The paper studies the optimal income taxation with a finite number of types. It is shown that Rawlsian social welfare and maximax social welfare functions constitute upper and lower bounds for the second best optimal marginal tax schedules. Therefore any marginal tax schedule with a higher tax rate than Rawlsian bound or with a lower tax rate than maximax bound would be inefficient. Moreover, it is shown that reasonable marginal tax schedules between these two benchmarks could be supported as a second-best tax schedule with appropriate social weights. These results are also valid when bunching is optimal. Additionally, some characterization for the total tax rates at the top and bottom of the income distribution are given.

Keywords: Public Economics; Optimal Income Taxation

JEL Codes: H2; H21

1 Introduction

Efficiency and equity are the most important criteria that economists consider while assessing the outcomes of tax policy. Efficiency is about how resources are allocated in the society and it is not related to any normative judgements. However, equity is highly involved with the norms of society as it is about the distribution of the resources. The problem arises from the fact that it is not possible to fully achieve these two goals at the same time. Due to this trade-off, characterizing the properties of an efficient tax schedule becomes an important issue.

*I am very grateful to Guy Laroque and Ian Preston for their guidance. I would like to thank Attila Lindner, Konrad Mierendorff and Peter Spittal as well as many seminar participants at UCL for valuable comments. All errors are my own.

[§]Department of Economics, University College London. E-mail: abdullah.oztek.13@ucl.ac.uk

Following Mirrlees (1971), a huge literature emerged in optimal income taxation literature using models with a continuum of types. On the other hand, Guesnerie and Seade (1982), Stiglitz (1982), and Weymark (1986, 1987) analyse the non-linear income tax problem with a finite number of types. These two modelling techniques use different assumptions and arguments which makes it difficult to understand the underlying common principles. However, Hellwig (2007) forms a new perspective by developing a unified approach to optimal income taxation. He shows that by using the same assumptions in continuum and finite models, the theory of optimal income taxation could be regarded as "monolith" meaning that they are mathematically equivalent. Moreover, Bastani (2013) uses the discrete model to derive marginal tax rates and shows that continuum and discrete models give similar numerical results when the number of types is sufficiently high.

In both of the modelling techniques, the structure of the tax schedule is identified by a complex relation of several components such as ability distribution, redistributive tastes of government and the labour supply responses to a change in tax schedule. Since it is quite a complicated problem, a general utility specification leads to very few analytical results[1], and many studies have to rely on the numerical simulations (Tuomala 1990).

In order to deal with the complexity of the problem, and to have clear-cut results several studies use quasi-linear utility specifications. Lollivier and Rochet (1983), Weymark (1986,1987), Ebert (1992) papers conduct the analysis by using a quasi-linear utility which is linear in labour as it provides a closed-form solution. However, this type of utility seems rather restrictive as it leads to a tax schedule that is independent of income level. Moreover, Blundell and MaCurdy (1999) shows that substitution effects have a higher impact on the labour supply than income effects, therefore instead of using a quasi-linear in labour utility, adapting a utility function which is linear in consumption seems more relevant to the general case.

In this study, we analyse the optimal income tax schedule with a finite number of types by using a quasi-linear in consumption utility. Under this setup, we give a characterization for the efficient marginal tax schedules which is summarized in Figure 1[2]. We show that when we have a Rawlsian social welfare function, the resulting marginal tax schedule constitutes an upper bound (or benchmark) for the tax rates obtained by any weighted utilitarian social welfare function. There is a positive marginal tax rate along with the distribution except at the top as there is no distortion at the top hence we have zero marginal tax rate. In the weighted utilitarian social

[1]Such as, zero marginal tax rate at the top (if the skill distribution is bounded) and bottom (if the lowest skill is positive and no-bunching at the bottom), also a non-negative tax schedule between 0 and 1. (Mirrlees (1971), Sadka (1976), Seade (1977)).

[2]For the numerical simulations in Figure 1, we employ the utility function $u(c,l) = c - \frac{l^{1+1/\varepsilon}}{1+1/\varepsilon}$. The second term is a standard form commonly used in the optimal income tax literature and we assume that utility is quasi-linear in consumption. Following Mirrlees(1971) and Tuomala(1990,2010) we use a log-normal skill distribution with parameters (μ, σ) = (-1, 0.39). Frisch elasticity of labour supply ε is set equal to 0.5. For the Rawlsian social welfare function, we set all other social weights to zero except the least able agent. Conversely, we maximize the utility of the highest able agent for Maximax social welfare function. Second-Best schedule in Figure 1 is derived from a random social weight distribution.

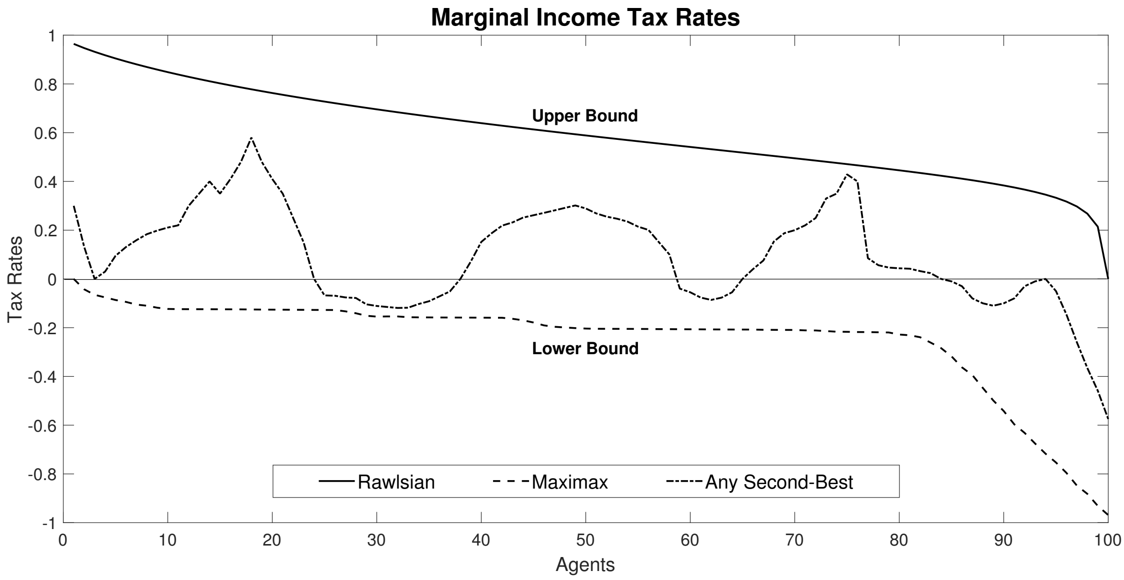

Figure 1: Marginal Income Tax Schedules

welfare function, social weights assigned by the government are generally aimed to redistribute from high-income earners to low-income earners, however other redistributive desires such as redistribution towards mean income or high-income earners are also possible. Since we allow for all social weight distributions, loosely speaking, one could say that the Rawlsian social welfare criterion gives an upper bound for all second-best marginal income tax schedules. Therefore any marginal tax schedule with a higher tax rate than Rawlsian bound would be inefficient. If the marginal tax rate is above this bound for any ability level, the government could reduce the tax rate and increase the tax revenue. Then redistributing this excess revenue would make the agents better-off. On the opposite, the maximax criterion that maximizes the utility of the highest able agent constitutes a lower bound for all second-best optimal tax schedules. In this case, there is no distortion at the bottom while there is a subsidy for other agents in the economy. Again it would be inefficient to set a lower marginal tax rate than this bound. So any efficient marginal tax schedule should be between these two benchmarks.

Moreover, we show that reasonable marginal tax schedules between these boundaries could be supported as a second-best efficient tax schedule via appropriate social weights.

Atkinson (1983) numerically shows that optimal linear tax rate is always higher under Rawlsian criterion than any other second-best case. A similar analysis is conducted by Saez (2001) for non-linear optimal taxation. Numerical simulations show that the Rawlsian criterion leads to higher marginal tax rates than the utilitarian social welfare function. In an optimal non-linear income tax model with extensive labour supply responses, Laroque (2005), and Choné and Laroque (2005) papers show that the Rawlsian social welfare criterion constitutes a benchmark for the tax schedules as well. Laroque (2005) shows that all utilitarian second best allocations are below the Laffer bound or the Rawlsian benchmark, and also proves that, under some mild conditions, any feasible allocation below Laffer bound corresponds to a second-best optimal allocation. The present study applies the same idea for marginal income tax schedules under intensive labour supply responses.

There is an ongoing discussion about the efficiency of marginal tax rates, especially for the top income earners. However, the effectiveness of the whole marginal tax schedule is not discussed very often. Bourguignon and Spadaro (2008) paper study the so-called "optimal inverse problem" which tries to recover the social welfare function (SWF) that would make the observed marginal tax rates optimal. They derive a necessary condition for the observed marginal tax rates that ensure the SWF to be Paretian[3]. They interpret this Paretian condition as a test on the relative position of the tax schedules with respect to the "Laffer Bound", defined as the revenue-maximizing tax system. They conclude that a tax system above the Laffer bound could only be optimal with non-Paretian social weights.

Lorenz and Sachs (2012) analyse the efficiency of the marginal tax rates in the phase-out region. They develop a similar test for marginal tax rates whether they are above the Laffer bound and thus second-best inefficient. The Laffer bound here is an extension of the Laffer argument to non-linear taxation, so the consideration is about whether the marginal tax rate at some specific income level is inefficiently high or not. They apply this test to Germany and find out that marginal tax rates are second best inefficient for the transfer phase-out region. The present study generalizes these kinds of tests to the entire distribution. Lorenz and Sachs (2012) use a quasi-linear in consumption utility. With this utility function maximizing the welfare by using Rawlsian social welfare function and maximizing the total tax revenue would generate exactly the same labour supply levels, since in the Rawlsian case government would collect the maximum possible revenue from other agents and give it to the least able individual as it only cares for the worst-off agent in the population. However, resulting allocations may not be the same due to different consumption levels. Since the marginal tax rates are independent of consumption in this utility specification, the resulting marginal tax rates will be the same.

To the best of our knowledge, Jacquet (2010) is the most relevant study for our study. By using a quasi-linear in consumption utility with an iso-elastic disutility of labour, Jacquet (2010) shows that a Rawlsian social welfare function always gives higher marginal tax rates than any second best redistributive utilitarian case. However, this result depends on the specific utility formulation and the redistribution desire of the government.

Moreover, by trusting the first order approach, Jacquet (2010) disregards the cases where bunching is optimal. The present paper shows that this result could be generalized to all labour supply functional forms, and to cases where bunching is optimal.

Due to the complexity of the problem, the characterization analysis is conducted under three separate parts.

First, we will analyse the Rawlsian case and show that Rawlsian is a benchmark for other social welfare functions where we have generally decreasing social weights ($GDSW$) for the agents. We show that under a Rawlsian SWF we have the highest downward distortion on

[3] A social welfare function is said to be Paretian if it assigns a positive social weight to each agent.

labour supply which leads to the result that we have the highest marginal tax rate.

Second, we apply the same analysis to the maximax case and any SWF with generally increasing social weights ($GISW$). Similarly, we show that with $GISW$, the upward distortion for labour supply could not exceed the upward distortion in maximax case. Therefore the marginal tax schedule under maximax SWF would be a lower bound for social welfare functions with $GISW$. Since there is a downward distortion in $GDSW$ cases, this result trivially holds for those possibilities as they imply positive marginal tax rates.

Third, although the $GDSW$ and $GISW$ cases form a significant part of all second-best cases, in general, there are many other possible second-best tax schedules. For this reason, to cover all second-best cases, we analyse all the possible second best allocations and show that in any case, the downward distortion for labour supply could not be greater than the downward distortion in Rawlsian case, and on the opposite, the upward distortion could not exceed the upward distortion in maximax case. Therefore, while Rawlsian gives an upper bound for the marginal tax rates, maximax constitutes a lower bound.

We show that these results are also valid when the monotonicity constraint is binding which corresponds to bunching. Results under bunching slightly differ from the cases where pooling is not optimal.

We conduct the same three-stage analysis on the converse of the result. We show that reasonable non-negative marginal tax schedules under Rawlsian benchmark could be supported as a second-best marginal tax schedule by choosing the appropriate social weights. Also at the other extreme, any reasonable non-positive tax schedule could be supported as a second-best schedule. For the general case, we need a complicated algorithm to show that any reasonable marginal tax schedule below Rawlsian and above Maximax would overlap with a second-best tax schedule with the appropriate social weights.

For the total tax rates, Rawlsian SWF gives a lower bound for the lowest able agent and an upper bound for the most productive agent. Total tax under Rawlsian SWF and any second-best tax schedule cross only once which means Rawlsian SWF constitutes a lower bound for the lower part of the population, and an upper bound for the remaining part of the ability distribution. Under reasonable distributions for ability and social weights, this intersection occurs near the median agent which is in line with Brett and Weymark (2015) and needs to be investigated further. On the opposite extreme, maximax SWF gives a lower bound for the highest able agent total taxes and an upper bound for least able individual's taxes.

The study is organized as follows: Section 2 presents the model. Sections 3 and 4 derives the results for Rawlsian and maximax social welfare functions respectively. Section 5 analyses the general case for all social weight distributions, and section 6 deals with bunching cases. Section 7 shows that reasonable tax schedules between these two benchmarks could be supported as a second-best efficient tax schedule. Section 8 presents the results for total tax rates, and section 9 concludes. Some of the proofs are left to the Appendix.

2 The Model

We study an optimal income taxation model with discrete ability types as in Guesnerie and Seade (1982), Stiglitz (1982) and Weymark (1986 and 1987). The only source for agent's heterogeneity is the labour productivity w, and in the economy, there are N productivity levels ranked in increasing order:

$$0 < w_1 < ... < w_N$$

The fraction of the population of ability w_i is π_i, with $\sum_{i=1}^{N} \pi_i = 1$. It is convenient to define the cumulative distribution function as $F_i = \sum_{j=1}^{i} \pi_j$, hence we have $F_N = \sum_{i=1}^{N} \pi_i = 1$.

All agents have identical preferences over consumption c and labour supply l which are represented by a quasi-linear utility function $U : \mathbb{R} \times \mathbb{R}_+ \rightarrow \mathbb{R}$

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

where the function $v : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is assumed to be increasing and strictly convex with $v(0) = 0$, $\lim_{l \to \infty} v'(l) \rightarrow \infty$ and $v'''(.) > 0$[4]. While agents derive utility from private consumption working generates disutility i.e. $U_c' = 1 > 0$ and $U_l' = -v'(l) < 0$.

The economy is competitive, with constant returns to scale technology; therefore agent i's gross wage rate is equal to his productivity w_i. Agent i with productivity level w_i, earns a gross income $y_i = w_i l_i$ and pays an income tax from his gross income y_i. The government knows the functional form of the utility function and the skill distribution. However, it cannot observe the productivity of the agent nor the labour supply of the agent. Therefore the government is restricted to set a non-linear tax $T(y)$ as a function of gross income y_i.

Agents choose their optimal consumption and labour choice in the market by maximizing their utility subject to their budget constraint;

$$\max_{c,l} c - v(l)$$

$$s.t.$$

$$c = wl - T(wl)$$

or equivalently if we substitute out l as $l = \frac{y}{w}$;

$$\max_{c,y} c - v\left(\frac{y}{w}\right)$$

$$s.t.$$

$$c = y - T(y)$$

[4]This assumption is used in the literature involving risk and uncertainty, and is called "prudence" by Kimball (1990) which leads to precautionary savings. In the present setup it corresponds to convex marginal disutility of labour which says as the labour supply gets larger, the increase in additional labour supply that becomes unattractive becomes larger (Simula (2010)). This assumption provides a unique optimum. Several other studies (e.g. Hellwig (2007)) prove existence and uniqueness of the solution with different but weaker set of assumptions.

The first order optimality condition of the agent's problem would be

$$1 - T^{'}(y) \equiv \Omega(c, y, w) = v^{'}\left(\frac{y}{w}\right)\frac{1}{w}$$

where $\Omega(c, y, w)$ is the marginal rate of substitution of agent w which is independent of consumption c. This formulation allows to express the marginal tax rate as $T^{'}(y) = 1 - \Omega(c, y, w)$.

The single crossing property $\frac{\partial \Omega(c,y,w)}{\partial w} < 0$ is satisfied for this specific utility form. This condition states that at any point in the (y, c) space with y and c on the horizontal and vertical axes, respectively, the indifference curve of a more productive agent is flatter than the indifference curve of a less productive agent and these curves cross only once. The intuition is, in order to produce an additional unit of output, a high productive agent does not have to work as hard as a less able agent and hence needs less compensation. The single crossing property ensures that a more able agent will end up with a higher consumption-income allocation, so that second best taxation could separate types and guarantees incentive-compatibility. It can also be exploited to rule out the global incentive comparisons, meaning that it suffices to take into account the incentive compatibility constraints that compare adjacent individuals.

An allocation for this economy is a pair of consumption level and output for individuals with different skill levels, $a = (c_i, y_i)_{i=1}^{N} \in \mathbb{R} \times \mathbb{R}_+$

An allocation is feasible if

$$\sum_{i=1}^{N} \pi_i c_i \leq \sum_{i=1}^{N} \pi_i y_i \tag{1}$$

so total consumption does not exceed total output or income.

And the allocation is incentive-compatible if

$$c_i - v\left(\frac{y_i}{w_i}\right) \geq c_j - v\left(\frac{y_j}{w_i}\right) \quad \text{for all } i \text{ and } j \quad (IC_{ij}) \tag{2}$$

so nobody has an incentive to lie about his type. Henceforth we say that an allocation is incentive feasible if it is feasible and incentive compatible.

An incentive compatibility constraint is called adjacent or local when $i = j \pm 1$, and called non-local (global) if $i \neq j \pm 1$. Since the government cannot observe the private productivity parameter, incentive compatibility should be taken into account for implementing any desired allocation.

The aim of the government is to maximize the total social welfare, defined by a weighted utilitarian welfare function $W(a) : \mathbb{R}^N \times \mathbb{R}_+^N \rightarrow \mathbb{R}$

$$W(a) = \sum_{i=1}^{N} \pi_i \delta^i \left[c_i - v\left(\frac{y_i}{w_i}\right)\right] \tag{3}$$

where δ^i is the social weight of the type i agents. While the case $\delta^i = 1$ for all i gives the pure utilitarian social welfare function, the case $\delta^1 = 1$ and $\delta^i = 0$ for all $i \neq 1$ will generate the Rawlsian social welfare function where the government maximizes the utility of the lowest ability agents, whereas on the opposite extreme we have a maximax social welfare function when $\delta^N = 1$ and $\delta^i = 0$ for all $i \neq N$. We allow for all social weight distributions. Therefore redistribution does not necessarily take place from high able agents to low productive agents. Also, it is required that the function $W(.)$ be non-decreasing in each $U(c_i, l_i)$. Such welfare functions are called Paretian social welfare functions which ensure Pareto optimality of the solution. For later reference, it is practical to define the weighted cumulative social weight β_i that gives the summation of the social weight of the agents from agent 1 to agent i (i.e. $\beta_i = \sum_{j=1}^{i} \pi_j \delta^j$), and also we can normalize $\beta_N = 1$ since the objective function is homogeneous of degree one in δ.

By the taxation principle of Hammond (1979) and Guesnerie (1995), setting a non-linear tax schedule is identical with choosing a specific consumption-income bundle for each agent which satisfies the incentive compatibility constraints. The optimal income tax problem is to choose an allocation $a = (c_i, y_i)_{i=1}^N$ to maximize

$$\sum_{i=1}^{N} \pi_i \delta^i \left[c_i - v\left(\frac{y_i}{w_i}\right) \right]$$

subject to feasibility condition (1.1) which must be binding at the optimum since the utility function is increasing in consumption, and incentive compatibility constraints (1.2). The Lagrangian for this problem is;

$$\mathcal{L} = \sum_{i=1}^{N} \pi_i \delta^i \left[c_i - v\left(\frac{y_i}{w_i}\right) \right] + \lambda \sum_{i=1}^{N} \pi_i \left[y_i - c_i \right] + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ i \neq j}}^{N} \mu_{ij} \left[c_i - v\left(\frac{y_i}{w_i}\right) - c_j + v\left(\frac{y_j}{w_i}\right) \right]$$

where λ and μ_{ij} are non-negative multipliers. The maximization yields the following first-order conditions:

$$c_i : \pi_i \delta^i - \lambda \pi_i + \sum_{\substack{j=1 \\ i \neq j}}^{N} \mu_{ij} - \sum_{\substack{j=1 \\ i \neq j}}^{N} \mu_{ji} = 0$$

$$y_i : \pi_i \delta^i v'\left(\frac{y_i}{w_i}\right) \frac{1}{w_i} - \lambda \pi_i + \sum_{\substack{j=1 \\ i \neq j}}^{N} \mu_{ij} v'\left(\frac{y_i}{w_i}\right) \frac{1}{w_i} - \sum_{\substack{j=1 \\ i \neq j}}^{N} \mu_{ji} v'\left(\frac{y_i}{w_j}\right) \frac{1}{w_j} = 0$$

and complementary slackness conditions:

$$\lambda \sum_{i=1}^{N} \pi_i \left[y_i - c_i \right] = 0$$

$$\mu_{ij} \left[c_i - v\left(\frac{y_i}{w_i}\right) - c_j + v\left(\frac{y_j}{w_i}\right) \right] = 0 \text{ for all } i \text{ and } j.$$

However, this problem is complicated due to the number $(N(N-1))$ of the incentive compatibility constraints. It turns out that it is possible to relax the problem by reducing the number of incentive compatibility constraints with the following Lemmas.

Lemma 1. *For any incentive feasible allocation we have:* $y_i \geq y_{i-1}$ *and* $c_i \geq c_{i-1}$ *for all* $i \geq 2$. *Moreover we have* $c_i = c_{i-1}$ *if and only if* $y_i = y_{i-1}$.

Proof. See Appendix. □

Lemma 1 implies that two different types either differ in both income and consumption and they are monotonically increasing with ability, or have the same bundle, in which case they are said to be bunched. In order to reduce the number of IC constraints, the following Lemma shows that only local incentive compatibility constraints matter, therefore the focus could be solely on the local incentive compatibility.

Lemma 2. *A local incentive compatible allocation is incentive compatible.*

Proof. See Appendix. □

First, two local downward IC for adjacent agents i and $i-1$ ($IC_{i,i-1}$ and $IC_{i-1,i-2}$) imply the global downward IC between agents i and $i-2$ ($IC_{i,i-2}$). Second, two local upward IC constraints for agents i and $i+1$ imply the global upward IC between agents i and $i+2$. One can also show that $IC_{i,i-1}$, $IC_{i-1,i-2}$ and $IC_{i-2,i-3}$ imply $IC_{i,i-3}$, and etc. By starting from $i = N$ and proceeding inductively, it is possible to show that local downward incentive compatibility constraints imply all of the global downward incentive compatibility constraints. A similar argument applies to reverse direction that local upward incentive compatibility constraints imply all global upward incentive compatibility constraints. This feature is referred as the "transitivity property", which states that if the local IC constraints are satisfied then the allocation would be incentive compatible.

Then it is possible to set up the maximization problem by only using local IC constraints;

$$c_i - v\left(\frac{y_i}{w_i}\right) \geq c_{i-1} - v\left(\frac{y_{i-1}}{w_i}\right) \text{ for all } i \ (IC_{i,i-1})$$

$$c_i - v\left(\frac{y_i}{w_i}\right) \geq c_{i+1} - v\left(\frac{y_{i+1}}{w_i}\right) \text{ for all } i \ (IC_{i,i+1})$$

Hence, we have reduced the number of the necessary and sufficient IC constraints from $N(N-1)$ to $2(N-1)$, and the Lagrangian becomes;

$$\mathcal{L} = \sum_{i=1}^{N} \pi_i \delta^i \left[c_i - v\left(\frac{y_i}{w_i}\right)\right] + \lambda \sum_{i=1}^{N} \pi_i \left[y_i - c_i\right] + \sum_{i=2}^{N} \mu_{i,i-1} \left[c_i - v\left(\frac{y_i}{w_i}\right) - c_{i-1} + v\left(\frac{y_{i-1}}{w_i}\right)\right]$$
$$+ \sum_{i=1}^{N-1} \mu_{i,i+1} \left[c_i - v\left(\frac{y_i}{w_i}\right) - c_{i+1} + v\left(\frac{y_{i+1}}{w_i}\right)\right]$$

with $\mu_{1,0} = \mu_{0,1} = \mu_{N+1,N} = \mu_{N,N+1} = 0$. The first order conditions are:

$$c_i : \pi_i \delta^i - \lambda \pi_i + \mu_{i,i-1} - \mu_{i+1,i} - \mu_{i-1,i} + \mu_{i,i+1} = 0$$

$$y_i : \pi_i \delta^i v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} - \lambda \pi_i + \mu_{i,i-1} v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} - \mu_{i+1,i} v'\left(\frac{y_i}{w_{i+1}}\right)\frac{1}{w_{i+1}}$$
$$-\mu_{i-1,i} v'\left(\frac{y_i}{w_{i-1}}\right)\frac{1}{w_{i-1}} + \mu_{i,i+1} v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} = 0$$

By defining $\Psi_i = \mu_{i,i-1} - \mu_{i-1,i}$ for $i = 2, ..., N$ and setting $\Psi_1 = 0$, the first order conditions for consumption become;

$$\pi_i \delta^i - \lambda \pi_i + \Psi_i - \Psi_{i+1} = 0 \text{ for all } i$$

Summing up these conditions yield $\lambda = 1$. By starting with the condition for agent one, it is possible to solve for all Ψ_i and multipliers μ_i. For all i, we have;

$$\Psi_i = \beta_{i-1} - F_{i-1}$$

and the IC constraint multipliers μ's are given by;

$$\mu_{i,i-1} = \max(0, \Psi_i) \quad \mu_{i-1,i} = -\min(0, \Psi_i)$$

then; if $\Psi_i > 0$, downward $IC_{i,i-1}$ is binding,
if $\Psi_i < 0$, upward $IC_{i-1,i}$ is binding,
and if $\Psi_i = 0$, none of $IC_{i,i-1}$ and $IC_{i-1,i}$ are binding.

Hence, once we have the distribution of social weights δ and the population share parameter π, we can find which of the IC constraints are binding in the equilibrium.

The following matrix shows all the IC constraints in the very general problem, however as we showed in Lemma 2 the local IC constraints highlighted in the matrix are sufficient to have an incentive compatible solution. Hence corresponding multipliers for other IC constraints are zero.

$$\begin{bmatrix}
\boldsymbol{IC_{1,2}} & \boldsymbol{IC_{2,1}} & IC_{3,1} & IC_{4,1} & IC_{5,1} & IC_{6,1} & \cdots & IC_{N,1} \\
IC_{1,3} & \boldsymbol{IC_{2,3}} & \boldsymbol{IC_{3,2}} & IC_{4,2} & IC_{5,2} & IC_{6,2} & \cdots & IC_{N,2} \\
IC_{1,4} & IC_{2,4} & \boldsymbol{IC_{3,4}} & \boldsymbol{IC_{4,3}} & IC_{5,3} & IC_{6,3} & \cdots & IC_{N,3} \\
IC_{1,5} & IC_{2,5} & IC_{3,5} & \boldsymbol{IC_{4,5}} & \boldsymbol{IC_{5,4}} & IC_{6,4} & \cdots & IC_{N,4} \\
IC_{1,6} & IC_{2,6} & IC_{3,6} & IC_{4,6} & \boldsymbol{IC_{5,6}} & \boldsymbol{IC_{6,5}} & \cdots & IC_{N,5} \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
IC_{1,N} & IC_{2,N} & IC_{3,N} & IC_{4,N} & IC_{5,N} & \cdots & \boldsymbol{IC_{N-1,N}} & \boldsymbol{IC_{N,N-1}}
\end{bmatrix}$$

In the following part of the study, we analyse the optimal tax schedule under no bunching, however in a later section, we will consider the cases where bunching is optimal. In the no-bunching case, the following Lemma has to hold which is identical with the non-binding monotonicity constraints.

Lemma 3. *If there is no bunching at most one of the $IC_{i,i+1}$ and $IC_{i+1,i}$ binds.*

Proof. See Appendix. □

This is known as the "asymmetry property" in the literature (Homburg (2002)). If a downward IC constraint is binding then the corresponding upward IC constraint will be slack when the higher able agent has strictly more income *i.e.* no-bunching.

For agent i (where $i \neq 1, N$), there are four relevant IC constraints: $IC_{i-1,i}$, $IC_{i,i-1}$, $IC_{i,i+1}$, $IC_{i+1,i}$. If there is no bunching then we know that only one of ($IC_{i-1,i}$, $IC_{i,i-1}$) and ($IC_{i,i+1}$, $IC_{i+1,i}$) could be binding. Also if $IC_{i,i-1}$ and $IC_{i,i+1}$ bind at the same time, agent i will be undistorted. As noted above which of these constraints are binding at the equilibrium is identified by the magnitudes of β_{i-1}, β_i, F_{i-1}, and F_i. Table 1 gives these regions and binding IC constraints under each region. There will be nine possible cases for each agent.

Table 1: Binding IC Constraints by Model Parameters

Ψ_i/Ψ_{i+1}	$\beta_i > F_i$	$\beta_i < F_i$	$\beta_i = F_i$
$\beta_{i-1} > F_{i-1}$	$IC_{i,i-1} - IC_{i+1,i}$	$IC_{i,i-1} - IC_{i,i+1}$	$IC_{i,i-1}$
$\beta_{i-1} < F_{i-1}$	$IC_{i-1,i} - IC_{i+1,i}$	$IC_{i-1,i} - IC_{i,i+1}$	$IC_{i-1,i}$
$\beta_{i-1} = F_{i-1}$	$IC_{i+1,i}$	$IC_{i,i+1}$	–

Boadway et al. (2002) analyses the optimal income taxation with three ability levels. As they allow for all social weight distributions, there will be four different scenarios for binding IC constraints. By calling regimes these cases they characterize the optimal tax schedule and they also derive the conditions that make some specific IC constraints binding. Here the nine possible cases for each agent is just a generalization of this idea to a N-type model. There is a similar discussion in Stantcheva (2014) where she derives the conditions that make the downward local or upward local constraints binding. The optimality condition for agent i is as follows:

$$v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} = 1 - \frac{\mu_{i+1,i}}{[\pi_i\delta^i + \mu_{i,i+1} + \mu_{i,i-1}]}\left[1 - v'\left(\frac{y_i}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$$
$$-\frac{\mu_{i-1,i}}{[\pi_i\delta^i + \mu_{i,i+1} + \mu_{i,i-1}]}\left[1 - v'\left(\frac{y_i}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]$$

For ease of presentation, it would be better to briefly discuss the possible cases here, which would also make it easier to follow up the subsequent sections. From Table 1 we have the following nine possibilities for each agent i.

1-) If $\beta_{i-1} > F_{i-1}$ and $\beta_i > F_i$ then $IC_{i,i-1}$ and $IC_{i+1,i}$ bind. This is the usual case when the government has a redistributive desire from high income earners to low income earners (generally decreasing social weights). Only local downward incentive compatibility constraints are binding in the equilibrium and Rawlsian SWF is a special form of this case. Optimality condition:

$v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} = 1 - \frac{\mu_{i+1,i}}{\pi_i\delta^i+\mu_{i,i-1}}\left[1 - v'\left(\frac{y_i}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$ where the distortion depends on agent $i+1$.

2-) If$\beta_{i-1} > F_{i-1}$ and $\beta_i < F_i$ then $IC_{i,i-1}$ and $IC_{i,i+1}$ bind. Optimality condition:

$v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} = 1$. There is no distortion.

3-) If $\beta_{i-1} > F_{i-1}$ and$\beta_i = F_i$ then only $IC_{i,i-1}$ binds. Optimality condition:

$v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} = 1$. There is no distortion.

4-) If $\beta_{i-1} < F_{i-1}$ and $\beta_i > F_i$ then $IC_{i-1,i}$ and $IC_{i+1,i}$ binds. Optimality condition:

$v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} = 1 - \frac{\mu_{i-1,i}}{\pi_i\delta^i}\left[1 - v'\left(\frac{y_i}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right] - \frac{\mu_{i+1,i}}{\pi_i\delta^i}\left[1 - v'\left(\frac{y_i}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$ where the distortion depends on agents $i-1$ and $i+1$.

5-) If $\beta_{i-1} < F_{i-1}$ and $\beta_i < F_i$ then $IC_{i-1,i}$ and $IC_{i,i+1}$ bind. Here we have binding local upward IC constraints which is the case when we have a maximax SWF or generally increasing social weights . Optimality condition:

$v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} = 1 - \frac{\mu_{i-1,i}}{\pi_i\delta^i+\mu_{i,i+1}}\left[1 - v'\left(\frac{y_i}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]$ where the distortion depends on agent $i-1$.

6-) If $\beta_{i-1} < F_{i-1}$ and $\beta_i = F_i$ then only $IC_{i-1,i}$ binds. Optimality condition:

$v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} = 1 - \frac{\mu_{i-1,i}}{\pi_i\delta^i}\left[1 - v'\left(\frac{y_i}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]$ where the distortion depends on agent $i-1$.

7-) If $\beta_{i-1} = F_{i-1}$ and $\beta_i > F_i$ then only $IC_{i+1,i}$ binds. Optimality condition:

$v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} = 1 - \frac{\mu_{i+1,i}}{\pi_i\delta^i}\left[1 - v'\left(\frac{y_i}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$ where the distortion depends on agent $i+1$.

8-) If $\beta_{i-1} = F_{i-1}$ and $\beta_i < F_i$ then only $IC_{i,i+1}$ binds. Optimality condition:

$v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} = 1$. There is no distortion.

9-) If $\beta_{i-1} = F_{i-1}$ and $\beta_i = F_i$ then there is no binding IC constraints. Optimality condition:

$v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} = 1$. There is no distortion.

Under no bunching case there are three possibilities for the lowest able agent.

1-) If $\beta_1 > F_1$ then only IC_{21} binds.

$v'\left(\frac{y_1}{w_1}\right)\frac{1}{w_1} = 1 - \frac{\mu_{21}}{\pi_1\delta^1}\left[1 - v'\left(\frac{y_1}{w_2}\right)\frac{1}{w_2}\right]$ where the distortion depends on agent 2. Since IC_{21} is binding agent 2 can be either downward distorted or undistorted. So even the agent 2 is undistorted, agent 1 will be distorted downwards.

2-) If $\beta_1 < F_1$ then only IC_{12} binds. $v'\left(\frac{y_1}{w_1}\right)\frac{1}{w_1} = 1$ there is no distortion.

3-) If $\beta_1 = F_1$ then there is no binding IC constraints, so no distortion.

Similarly, there are three possibilities for the highest able agent.

1-) If $\beta_{N-1} > F_{N-1}$ then only $IC_{N,N-1}$ binds.

$v'\left(\frac{y_N}{w_N}\right)\frac{1}{w_N} = 1$. There is no distortion.

2-) If $\beta_{N-1} < F_{N-1}$ then only $IC_{N-1,N}$ binds.

$v'\left(\frac{y_N}{w_N}\right)\frac{1}{w_N} = 1 - \frac{\mu_{N-1,N}}{\pi_N \delta^N}\left[1 - v'\left(\frac{y_N}{w_{N-1}}\right)\frac{1}{w_{N-1}}\right]$ where the distortion depends on agent $N-1$. Since $IC_{N-1,N}$ is binding agent $N-1$ can be either upward distorted or undistorted. So even the agent $N-1$ is undistorted, agent N will be distorted upwards.

3-) If $\beta_{N-1} = F_{N-1}$ then there is no binding IC constraints, hence no distortion.

Since the population share parameter π is given for the economy, the only parameter that identifies the binding IC constraints is the social weight parameter δ. Hence depending on the redistribution taste of the government, the optimal solution will be characterized by these 9 possible cases for each agent. In the following two sections, first, we will analyse the Rawlsian case and show that Rawlsian is a benchmark for other social welfare functions where we have generally decreasing social weights ($GDSW$) for the agents. Second, we will show that maximax SWF is a benchmark for all other SWF with generally increasing social weights ($GISW$). Finally, we aim to show that all other possible solutions are between these two benchmarks.

3 The Rawlsian Benchmark

Most of the studies in the literature deal with the cases where the government has a redistributive desire from high-income earners to low-income earners (or from high able agents to low able agents). As it is a more interesting example we first analyse the Rawlsian social welfare function and any social welfare criterion with generally decreasing social weights, which is a special form of case 1 in our formulation. When the government has a redistributive desire from high able to low able (i.e. a decreasing social weight δ^i with the ability), Weymark (1986,1987), Hellwig (2007) and many other papers show that the allocation is a simple monotonic chain to the left, which means only the downward IC constraints are relevant and binding, therefore it is possible to relax the problem. In this environment, Rawlsian SWF yields the maximum downward distortion for the agent's labour supply except for the top agent (no distortion at the top). Since the marginal tax rate is independent of consumption, this downward distortion leads to the result that marginal tax rates are positive and always higher in the Rawlsian case. This result trivially holds when we compare the Rawlsian case and any other social welfare criterion with an increasing social weight pattern, because in that case except for the lowest able agent, there exists an upward distortion leading marginal tax rates to be negative.

Questions may arise about the existence and uniqueness of a solution for this kind of problem. Because $v(.)$ is strictly convex and $v'''(.) > 0$, the objective function is concave over the set $\mathbb{R}_+$. Therefore these first order conditions are both necessary and sufficient for an optimum[5]. The existence and uniqueness for this problem are discussed in several papers. Simula (2010) and Brett and Weymark (2015) mention that existence and uniqueness could be provided by

[5]Second order condition would be; $\mu_{i+1,i}v''\left(\frac{y_i}{w_{i+1}}\right)\frac{1}{w_{i+1}^2} - \left[\pi_i\delta^i + \mu_{i,i-1}\right]v''\left(\frac{y_i}{w_i}\right)\frac{1}{w_i^2} < 0$ which is satisfied as we have $v''' > 0$ and $\mu_{i+1,i} = [\beta_i - F_i] < \pi_i\delta^i + \mu_{i,i-1} = [\beta_i - F_{i-1}]$

using the same assumptions as we have here, however Hellwig (2010) paper shows that existence could be possible even with a weaker set of assumptions.

If we have a decreasing social weight distribution we know that only the downward IC constraints $IC_{i,i-1}$ and $IC_{i+1,i}$ are binding, and the optimality condition is:

$$v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} = 1 - \frac{\mu_{i+1,i}}{\pi_i\delta^i + \mu_{i,i-1}}\left[1 - v'\left(\frac{y_i}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$$

where $\mu_{i,i-1} = \beta_{i-1} - F_{i-1}$.

A sufficient condition for $\mu_{i,i-1} > 0$ is a decreasing social weight distribution (i.e. $\delta^i > \delta^j$ if $i < j$) as with these social weights β_{i-1} is always greater than F_{i-1}. Moreover for some social weight distributions, those are not always decreasing (increasing or constant for some weights), it is possible to fulfil the $\mu_{i,i-1} > 0$ condition[6]. This is also pointed out in Weymark (1987) as the social weights should not increase too rapidly with ability. Therefore a generally decreasing pattern is enough to have binding local downward IC constraints.

It is practical to rewrite the optimality conditions in the following form;

$$v'\left(l_i^{GD}\right)\frac{1}{w_i} = 1 - \frac{\beta_i - F_i}{\beta_i - F_{i-1}}\left[1 - v'\left(\frac{w_i l_i^{GD}}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$$

where the superscript GD stands for "generally decreasing" social weights, and the Rawlsian optimality condition can be found by setting all $\delta^i = 0$ for $i \neq 1$;

$$v'\left(l_i^{R}\right)\frac{1}{w_i} = 1 - \frac{1 - F_i}{1 - F_{i-1}}\left[1 - v'\left(\frac{w_i l_i^{R}}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$$

where the superscript R stands for Rawlsian SWF.

Since the optimal allocation is a simple monotonic chain to the left, the tax schedule is not differentiable, however it is possible to use the differentiability of the utility function to define implicit marginal tax rates as;

$$T_i'\left(y_i\right) = 1 - v'\left(\frac{l_i}{w_i}\right)$$

then optimal tax rates for $GDSW$ case will be as follows:

$T_N'\left(y_N^{GD}\right) = 0$ and

$T_i'\left(y_i^{GD}\right) = \frac{\beta_i - F_i}{\beta_i - F_{i-1}}\left[1 - v'\left(\frac{w_i l_i^{GD}}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$

and Rawlsian marginal tax rates are;

$T_N'\left(y_N^{R}\right) = 0$ and

$T_i'\left(y_i^{R}\right) = \frac{1 - F_i}{1 - F_{i-1}}\left[1 - v'\left(\frac{w_i l_i^{R}}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$

We know that both in the Rawlsian and $GDSW$ cases only downward IC constraints are

[6] Consider the following example; suppose we have 4 types, and $\pi_i = 0.25$ for all agents. A social weight distribution such as $\delta = [2, 0.8, 1, 0.2]$ is not decreasing in ability but the Lagrange multipliers for downward IC constraints are still positive.

binding. Therefore while there is no distortion at the top (zero marginal tax), it is optimal to impose a distortion on agent i to prevent agent $i+1$ from mimicking agent i by reducing his or her labour supply. For the rest of the population, there are two parts in the tax function. When we compare the first terms we can say that $\frac{1-F_i}{1-F_{i-1}}$ term is always greater than the $\frac{\beta_i-F_i}{\beta_i-F_{i-1}}$ term.

$T_i^{'}\left(y_i^{GD}\right)$ and $T_i^{'}\left(y_i^{R}\right)$ tax rates are for the agent i, however the income levels are different in these tax functions due to the different labour supplies l_i^{GD} and l_i^{R}. In a discrete model, it may not be possible to compare the tax rates for the same income levels, but it could be possible to make this comparison with a continuum of agents. In order to compare these two tax rates, one also needs to know the labour supply of the agent i under each social welfare function. If the Rawlsian labour supply is more downward distorted than any decreasing social weights, we can conclude that marginal tax schedule under Rawlsian is always greater than second best marginal tax schedules where there is a decreasing social weight distribution. Hence, a marginal tax schedule $T_i^{'}(y_i^{R})$ under Rawlsian SWF would be an upper bound for the possible marginal tax rates for the agent i. If the labour supply of the agent i increases with any social weight δ^j, we conclude that $l_i^{GD} \geq l_i^{R}$.

Using a quasi-linear utility specification which is linear in labour, Weymark (1987) conducts a comparative statics for the welfare weights. A corresponding comparative analysis is done by Simula (2010) by using a utility function that is linear in consumption. Simula (2010) analyses the effects of increasing agent i's social weight δ^i while reducing the other agents social weights proportionately. In the absence of any normalization for the social weights, this is equivalent to increasing δ^i while holding the other social weights constant.

According to Simula (2010) when there is an increase in the social weight of agent i with all other social weights δ^j $i \neq j$ scaled down proportionately, there will be no change in the gross income y_N of the highest able agent. However, agents who have a lower ability level than agent i will have a higher gross income which means there will be a lower marginal tax rate for these agents. On the other hand, agent i's and more able agents' income levels will decrease and the marginal tax rates will increase. This analysis compares two different second best $GDSW$ income tax schedules. However our concern in this study is to compare the tax rates under Rawlsian and any second best tax schedule. So we need to check the effect of decreasing social weight of the least able agent while increasing the social weight of any agent.

Corresponding results for Theorem 2 in Weymark (1987) and Proposition 9 in Simula (2010) are as follows. When there is a change from Rawlsian to any $GDSW$ second-best we can investigate the situation as an increase in δ^j $j \neq 1$ with a decrease in δ^1. Otherwise comparison would be between any two second best cases. One can conclude that if δ^j increases, for the agents $i < j$ labour supply and income level will increase so there will be a lower tax for these agents. However for the agents $i \geq j$ there will be no change in labour supply and so in the income. Then the result is similar to Simula (2010) however in this case, if there is an increase in the δ^j agents $i \geq j$ labour supply and income will not be affected by this change. Under

$GDSW$, one could say that agents with a higher ability than agent j are in the same case as they are under the Rawlsian criterion. Theorem 2 in Weymark (1987) could be interpreted as the counterpart of this result with quasi-linear in labour utility. It is possible to adapt this theorem by analysing the effects of an increase in agent j's social weight δ^j with a corresponding decrease in δ^1. The result is the same but instead of labour, his paper compares the consumption levels. So while the agents $i < j$ would have a higher consumption, the consumption level for agents $i \geq j$ would not change. The following proposition shows that labour supply and income levels are weakly smaller under Rawlsian SWF than any utilitarian case with decreasing social weights.

Proposition 1. *For all i,* $l_i^{GD} \geq l_i^R$ *or equivalently* $y_i^{GD} \geq y_i^R$.

Proof. We need to show that if we switch from Rawlsian to any $GDSW$ utilitarian case labour supply of the agent i increases. From the optimality condition we have;

$v'(l_i)\frac{1}{w_i} = 1 - \frac{\beta_i - F_i}{\beta_i - F_{i-1}}\left[1 - v'\left(\frac{w_i l_i}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$

For agent i if any δ^j with $j \leq i$ increases, the term $\frac{\beta_i - F_i}{\beta_i - F_{i-1}}$ will be the same for both Rawlsian and $GDSW$ cases, because the increase in δ^j means a corresponding decrease in δ^1. Therefore the cumulative weight β_i will be the same in both cases since it contains the δ^j term. Then labour supply levels will be the same in this case. However if any δ^j with $j > i$ increases then the corresponding term will be lower in the $GDSW$ case. Here β_i does not contain the social weight δ^j hence β_i will be lower under $GDSW$ utilitarian case. From the optimality conditions we have;

$$\frac{\beta_i - F_i}{\beta_i - F_{i-1}} = \frac{\left[1 - v'\left(l_i^{GD}\right)\frac{1}{w_i}\right]}{\left[1 - v'\left(\frac{w_i l_i^{GD}}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} \text{ and } \frac{1 - F_i}{1 - F_{i-1}} = \frac{\left[1 - v'\left(l_i^R\right)\frac{1}{w_i}\right]}{\left[1 - v'\left(\frac{w_i l_i^R}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}$$

since $\frac{1-F_i}{1-F_{i-1}} > \frac{\beta_i - F_i}{\beta_i - F_{i-1}}$ we have

$$\frac{\left[1 - v'\left(l_i^R\right)\frac{1}{w_i}\right]}{\left[1 - v'\left(\frac{w_i l_i^R}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} > \frac{\left[1 - v'\left(l_i^{GD}\right)\frac{1}{w_i}\right]}{\left[1 - v'\left(\frac{w_i l_i^{GD}}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}$$

Note that the function $f(y) = \frac{\left[1 - v'\left(\frac{y}{w_i}\right)\frac{1}{w_i}\right]}{\left[1 - v'\left(\frac{y}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}$ is decreasing in y (or l) since the numerator of its derivative;

$v''\left(\frac{y}{w_{i+1}}\right)\frac{1}{w_{i+1}^2}\left[1 - v'\left(\frac{y}{w_i}\right)\frac{1}{w_i}\right] - v''\left(\frac{y}{w_i}\right)\frac{1}{w_i^2}\left[1 - v'\left(\frac{y}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$ is negative by convexity of $v'(.)$.

Hence, the inequality implies that $l_i^{GD} > l_i^R$ (or $y_i^{GD} > y_i^R$). □

Consequently, if a higher ability agent's social weight increases then the agent i labour supply increases, and there is no impact of an increase in social weight of a lower able agent.

With a Rawlsian SWF, the government collects the maximal amount of money from all agents, and transfers this money to the least well-off agent in the society. Intuitively, to increase the amount collected government will increase the marginal tax rates which creates a downward distortion for the labour supply. Since there is a positive social weight for other agents in the *GDSW* case, in order to increase the total welfare government should let the people work more and consume more than in the Rawlsian case. Hence, the labour supply level in Rawlsian SWF would be lower than any *GDSW* utilitarian criterion. After analysing the labour supply of agent i under two different social welfare functions, it is possible to compare the marginal tax rates under these two different cases.

Proposition 2. *For all* i, $T_i^{'}(y_i^R) \geq T_i^{'}(y_i^{GD})$.

Proof. *GDSW* case marginal tax function for agent i: $T_i^{'}(y_i^{GD}) = \frac{\beta_i - F_i}{\beta_i - F_{i-1}} \left[1 - v^{'}\left(\frac{w_i l_i^{GD}}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$

Rawlsian marginal tax function for agent i: $T_i^{'}(y_i^R) = \frac{1-F_i}{1-F_{i-1}} \left[1 - v^{'}\left(\frac{w_i l_i^R}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$

In Proposition 1 we showed that if there is an increase in δ^j for agents $j \leq i$, there will be no change in labour supply levels and $\beta_i = 1$. Hence, the marginal tax rate of the agent i is not affected by a change in social weight of himself or a lower able agent. However if δ^j for $j > i$ increases β_i will be lower than 1 and again from Proposition 1 we have $l_i^U \geq l_i^R$. Note that the function $f(y) = \left[1 - v^{'}\left(\frac{y}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$ is decreasing in y (or l) since its derivative $-v^{''}\left(\frac{y}{w_{i+1}}\right)\frac{1}{w_{i+1}^2}$ is negative. So the marginal tax rate for agent i will be higher in the Rawlsian case. □

Hence, the Rawlsian case where $\delta^i = 0$ for $i \neq 1$ constitutes an upper bound for any *GDSW* optimal marginal tax rates for agent i.

The inequality $\frac{1-F_i}{1-F_{i-1}} \geq \frac{\beta_i - F_i}{\beta_i - F_{i-1}}$ depends on the assumption that cumulative social weight $\beta_i \leq 1$. If we have the reverse then at least one of the social weights has to be negative (since we have $\beta_N = \sum_{j=1}^{N} \pi_j \delta^j = 1$), which means we do not have a Paretian SWF. Marginal tax rates above the Rawlsian might be achieved by a non-Paretian SWF. Hence, Rawlsian marginal tax rates constitute an efficiency bound. Above this bound we could not have a Pareto efficient tax schedule. This condition also says that if we have a Paretian SWF then the resulting marginal tax schedule has to be below the Rawlsian bound. However below Rawlsian some marginal tax schedules could still be inefficient due to the inefficient structure of the tax schedule itself. We will further discuss these issues in section 7.

There is a positive tax for all agents in the discrete ability setup, however Seade (1977) shows that the optimal tax rate at the bottom should be zero if there is no bunching at the bottom. This difference is due to the continuum setup he uses. In a continuum model, the mass of the worst-off agents is zero in the utilitarian objective. Therefore it is not possible to increase the social welfare by taxing these agents and redistributing the excess revenue. However, with a Rawlsian SWF, there will be a positive tax for the worst of agent even if we have a continuous ability distribution, as these agents are the only mass in the social welfare function. In a discrete model there is a positive mass of worst-off agents both in the Rawlsian and the utilitarian SWF. Therefore there would be equity gains from a positive marginal tax in either cases.

One of the general results in optimal income taxation is the zero marginal tax for the top agent (Sadka 1976, Seade 1977), as increasing the tax rate for the highest able agent will distort his labour supply without providing any additional revenue. However this result is only valid under a bounded ability distribution assumption since in an unbounded distribution there will always be a higher income earner than any income level. Since we have a finite setup, the zero marginal tax at the top result is valid for both Rawlsian and redistributive utilitarian social welfare functions.

The preceding analysis applies to cases where we have a proper redistribution from high income earners to low income earners. However for some distribution of social weights, upward IC constraints could be binding. Since the marginal tax rate is decreasing in income for the same agent, if there is an upward distortion it necessarily means a lower (negative) marginal tax rate for the agent. Hence, in an economy where all the agents are distorted upwards, the marginal tax schedule will always be below the Rawlsian tax schedule. The next section deals with the maximax benchmark and compare marginal tax rates with the social welfare functions where we have an increasing social weight distribution.

4 The Maximax Benchmark

On the opposite extreme, the maximax social welfare function gives a lower bound for the marginal tax schedule as the largest upward distortion occurs under this kind of social welfare function. When all upward IC constraints are binding one can say that all of the agents are distorted upwards. In this case an upward distortion is imposed on agent i to prevent agent $i-1$ pretending to be a high able individual. In the following, we only consider the second best cases with generally increasing social weights ($GISW$) and compare them to the results for maximax SWF.

In the maximax and $GISW$ cases we know that only upward IC constraints $IC_{i-1,i}$ and $IC_{i,i+1}$ are binding, and the optimality condition is:

$$v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} = 1 - \frac{\mu_{i-1,i}}{\pi_i\delta^i + \mu_{i,i+1}}\left[1 - v'\left(\frac{y_i}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]$$

where $\mu_{i,i+1} = F_i - \beta_i$.

So with increasing social weights where $\delta^i > \delta^j$ if $i > j$, F_i increases faster than β_i and multiplier $\mu_{i,i+1}$ will be positive. These multipliers could also be positive for some social weight distributions those are not entirely increasing (generally increasing social weights).

Brett and Weymark (2015) analyse the optimal tax rates identified by majority voting. In this setup while the lowest able agent proposes the Rawlsian tax schedule, the top agent votes for the maximax case. For the maximax case, second order conditions for an optimum could be problematic and they refer to this problem as "ill-behaved". However it is possible to have an optimum by imposing restrictions on the parameters[7].

By plugging the multipliers the first order condition could be written as;

$$v'\left(l_i^{GI}\right)\frac{1}{w_i} = 1 - \frac{F_{i-1}-\beta_{i-1}}{F_i-\beta_{i-1}}\left[1 - v'\left(\frac{w_i l_i^{GI}}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]$$

where superscript GI stands for "generally increasing" social weights. Maximax optimality condition could be found by setting all $\delta^i = 0$ for $i \neq N$.

$$v'\left(l_i^{M}\right)\frac{1}{w_i} = 1 - \frac{F_{i-1}}{F_i}\left[1 - v'\left(\frac{w_i l_i^{M}}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]$$

where superscript M states for Maximax SWF.

And the corresponding marginal tax functions;

$T_1'\left(y_1^{GI}\right) = 0$ and

$T_i'\left(y_i^{GI}\right) = \frac{F_{i-1}-\beta_{i-1}}{F_i-\beta_{i-1}}\left[1 - v'\left(\frac{w_i l_i^{GI}}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]$

and the tax rates for maximax case are;

$T_1'(y_1^M) = 0$ and

$T_i'(y_i^M) = \frac{F_{i-1}}{F_i}\left[1 - v'\left(\frac{w_i l_i^{M}}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]$

For both of these SWF while there is no distortion at the bottom, there is an upward distortion for all other agents. Again there are two parts in the tax function of the agents. When we compare the first terms we can say that $\frac{F_{i-1}}{F_i}$ is always greater than or equal to $\frac{F_{i-1}-\beta_{i-1}}{F_i-\beta_{i-1}}$. Also since there is an upward distortion, the second terms are both negative. Therefore, we can say that if the labour supply of the agent is more upward distorted in maximax, then maximax SWF gives lower tax rates than any $GISW$ marginal tax schedule. The analysis is very similar to the Rawlsian and $GDSW$ cases.

Proposition 3. *For all i, $l_i^M \geq l_i^{GI}$ or equivalently $y_i^M \geq y_i^{GI}$.*

Proof. We need to show that if we switch from Maximax to any $GISW$, the labour supply of agent i decreases. From the optimality condition we have;

$v'\left(l_i^{GI}\right)\frac{1}{w_i} = 1 - \frac{F_{i-1}-\beta_{i-1}}{F_i-\beta_{i-1}}\left[1 - v'\left(\frac{w_i l_i^{GI}}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]$

Since under maximax we have only δ^N this condition will be as follows;

$v'\left(l_i^{M}\right)\frac{1}{w_i} = 1 - \frac{F_{i-1}}{F_i}\left[1 - v'\left(\frac{w_i l_i^{M}}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]$

[7]Second order condition would be;
$\mu_{i-1,i}v''\left(\frac{y_i}{w_{i-1}}\right)\frac{1}{w_{i-1}^2} - [\pi_i\delta^i + \mu_{i,i+1}]v''\left(\frac{y_i}{w_i}\right)\frac{1}{w_i^2} < 0$ This would hold if
$\frac{F_{i-1}-\beta_{i-1}}{F_i-\beta_{i-1}} < \frac{v''\left(\frac{y_i}{w_i}\right)\frac{1}{w_i^2}}{v''\left(\frac{y_i}{w_{i-1}}\right)\frac{1}{w_{i-1}^2}}$. As discussed in Brett and Weymark (2015), this condition is harder to hold for the upper end of the ability distribution since the difference in adjacent agents abilities is higher, and more likely to hold at the bottom of the distribution. With a smooth small increment in ability, this condition could hold. Instead of making strong assumptions, we focus on the cases where we have an optimum.

Suppose there is an increase in δ^j where $j \geq i$, β_{i-1} will be zero. The optimality conditions would be exactly the same and there is no change in the labour supply. Now suppose δ^j increases where $j < i$, then β_{i-1} will be positive, and we have;

$\frac{F_{i-1}}{F_i} > \frac{F_{i-1}-\beta_{i-1}}{F_i-\beta_{i-1}}$ which implies $\frac{\left[1-v'\left(l_i^M\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{w_i l_i^M}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]} > \frac{\left[1-v'\left(l_i^{GI}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{w_i l_i^{GI}}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}$

Note that the function $f(y) = \frac{\left[1-v'\left(\frac{y}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}$ is increasing in y (or l) since the numerator of its derivative

$v''\left(\frac{y}{w_{i-1}}\right)\frac{1}{w_{i-1}^2}\left[1-v'\left(\frac{y}{w_i}\right)\frac{1}{w_i}\right] - v''\left(\frac{y}{w_i}\right)\frac{1}{w_i^2}\left[1-v'\left(\frac{y}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]$

is positive by convexity of $v^{'}(.)$. Otherwise, the second order condition would be violated as SOC implies;

$\frac{v''\left(\frac{y_i}{w_i}\right)\frac{1}{w_i^2}}{v''\left(\frac{y_i}{w_{i-1}}\right)\frac{1}{w_{i-1}^2}} > \frac{F_{i-1}-\beta_{i-1}}{F_i-\beta_{i-1}} = \frac{\left[1-v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_i}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}$

Hence, the inequality implies that $l_i^M > l_i^{GI}$ (or $y_i^M > y_i^{GI}$). □

We know that the marginal tax rate for agent i decreases with the upward distortion. Since we have a higher income level under maximax for everybody, we have a lower marginal tax rate. Similar with the corresponding proposition for $GDSW$ case, under maximax criterion marginal tax rates are always lower than in any $GISW$. So while Rawlsian is an upper bound for the second-best optimal marginal tax rates, maximax constitutes a lower bound for the tax rates.

Proposition 4. *For all i,* $T_i^{'}(y_i^{GI}) \geq T_i^{'}(y_i^M)$.

Proof. See Appendix. □

5 General Case

So far, we have shown that under generally decreasing social weights we have the Rawlsian benchmark. For the opposite case when we have generally increasing social weight distribution maximax is a lower bound and since these tax rates are negative they will always be below the Rawlsian marginal tax rates.

We previously assumed that social weights were generally decreasing or increasing with the ability. In this part, we will show that Rawlsian and Maximax SWF constitute benchmarks when there is no restriction on the distribution of social weights. It is possible that for some social weight distributions while some agents are distorted downwards, some agents could be distorted upwards. We may have different binding IC constraints and as we showed above there will be 9 different cases for each agent in the economy.

In cases 2,3,8, and 9 we have the first best so agents would be undistorted. Since we have a zero marginal tax in first best, Rawlsian marginal tax rates would be higher than these cases as we have a positive marginal tax for the agents. And, maximax marginal tax rates would be less than the first best as we have negative marginal tax rates (subsidies). The other 5 cases should be discussed separately.

Case 1 is the case we analysed above while discussing the Rawlsian benchmark. In Proposition 1 we showed that Rawlsian SWF leads to the highest downward distortion for the agents. In Proposition 2 we proved that Rawlsian gives the highest possible marginal tax rates when we have generally decreasing social weights. Clearly, since the maximax SWF gives a negative tax rate for all agents, it constitutes a lower bound for case 1.

We analysed case 5 when we discussed the maximax benchmark. While maximax constitutes a lower bound for the generally increasing social weights, the Rawlsian marginal tax schedule is an upper bound as in Rawlsian we have positive marginal tax rates whereas in case 5 we have negative rates for the agents.

Then cases 4,6, and 7 should be further investigated. Here, we will only discuss the possible magnitudes of the distortions. However, a formal proof will be provided in the following proposition with considering the changing IC constraint multipliers.

In case 6 we only have $IC_{i-1,i}$ binding. The optimality condition is;
$v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} = 1 - \frac{\mu_{i-1,i}}{\pi_i \delta^i}\left[1 - v'\left(\frac{y_i}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]$ where the direction of the distortion depends on agent $i-1$. Agent $i-1$ is undistorted if we have case 2 ($IC_{i-1,i-2}; IC_{i-1,i}$) or case 8 ($IC_{i-1,i}$). Otherwise if we have case 5 ($IC_{i-2,i-1}; IC_{i-1,i}$) agent $i-1$ is distorted upwards. So in any case agent i would be distorted upwards and this upward distortion could not be larger than the maximax case.

In case 7 only $IC_{i+1,i}$ binds. Optimality condition is;
$v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} = 1 - \frac{\mu_{i+1,i}}{\pi_i \delta^i}\left[1 - v'\left(\frac{y_i}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$ where the direction of the distortion depends on agent $i+1$. Agent $i+1$ is undistorted if we have case 2 ($IC_{i+1,i}; IC_{i+1,i+2}$) or case 3 ($IC_{i+1,i}$). Otherwise if we have case 1 ($IC_{i+1,i}; IC_{i+2,i+1}$) agent $i+1$ is distorted downwards. So in any case agent i would be distorted downward and this downward distortion could not be larger than Rawlsian case.

In case 4, $IC_{i-1,i}$ and $IC_{i+1,i}$ constraints are binding and optimality condition is; $v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} = 1 - \frac{\mu_{i-1,i}}{\pi_i \delta^i}\left[1 - v'\left(\frac{y_i}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right] - \frac{\mu_{i+1,i}}{\pi_i \delta^i}\left[1 - v'\left(\frac{y_i}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$ where the distortion depends on agents $i-1$ and $i+1$. As we showed in case 6, from agent $i-1$ we know for sure that there is an upward distortion effect for agent i. On the other side, in case 7 we showed that there will be a downward distortion effect from agent $i+1$. Direction of the final distortion is ambiguous but the idea is if we have a downward distortion effect it cannot exceed Rawlsian case and if we have a upward distortion it should be lower than the upward distortion under maximax SWF.

Consequently, suppose we have Rawlsian SWF with $\delta = [1, 0, 0, ..., 0]$, and we increase some δ^j which leads to a corresponding decrease in δ^1. After this increase if we still have a generally decreasing social weight distribution, as we have showed above, there will be no change for the agents where $i \geq j$. However, the labour supply of the agents below agent j will increase. If the increase in δ^j leads to a generally increasing social weight distribution, then this change will have an upward effect for all agents, and the labour supply levels will be higher than in the Rawlsian case.

For agent i, lower ability has an upward effect, and higher ability has a downward effect. If lower able agents' labour supply is distorted upwards, this leads to a higher upward distortion for agent i. On the other hand, any upward distortion for higher able agents means a lower downward effect on agent i. So in either case, when we move away from Rawls, there is an upward effect on agent i. This situation leads to the result that under Rawls we always have a higher marginal tax rate. On the other hand, upward distortion could not be greater than the maximax SWF which leads to the result that marginal tax rate is the lowest in maximax case. The example above analyses only the sign of the distortions. However, when we change the social weights, the multipliers for the IC constraints would change. In the proof for Proposition 5, we consider this change and prove the result.

Proposition 5. *For all i, $l_i^M \geq l_i \geq l_i^R$ or equivalently $y_i^M \geq y_i \geq y_i^R$.*

Proof. See Appendix. □

Proposition 6. *For all i, $T_i^{'}(y_i^R) \geq T_i^{'}(y_i) \geq T_i^{'}(y_i^M)$.*

Proof. Follows as Proposition 2. □

Hence, Rawlsian and maximax SWF appear to be the two extreme cases. Any second best Pareto efficient marginal tax schedule should be between these two benchmarks. Under this setup one could rationalize the negative marginal tax rates that are not possible under usual Mirrleesian setup. In order to have a negative tax rate, social weights should be increasing totally as in $GISW$ or partly increasing. For example, if we have a inverse-U shape for the social weight distribution where government gives the highest value to middle income earners, there will be negative rates for these agents.

The preceding sections assume that the monotonicity condition is satisfied, however in some cases it may not be the case and we could have pooling equilibria. The next section deals with the cases where bunching is optimal.

6 Optimal Allocation and Bunching

Bunching occurs when the income and consumption levels of two different agents are equal to each other. Similar with Boadway et al.(2002), we allow for all social weight distributions, hence both downward and upward IC constraints could be binding in the equilibrium. Lemma 3 shows that under no bunching at most one of the $IC_{i,i+1}$ and $IC_{i+1,i}$ binds. However, agents i and $i+1$ receive the same income-consumption bundle (i.e. bunching) if both constraints bind at the same time. It is convenient to analyse bunching under two different cases. First, there could be bunching due to violation of the non-negativity constraint for income level. This is called the $y = 0$ bunching, and only occurs at the bottom of the income distribution. Second, agents with different ability levels could be bunched along the income distribution which is due to the violation of the monotonicity.

Once we have a binding non-negativity condition, it is not possible to analyse the problem with first order conditions. In that case we have inequalities for the first order conditions and it is not possible to compare the magnitudes of the marginal tax rates. Therefore we will focus on the bunching that occurs at the interior of the ability distribution.

It is important to note that we need to have the same bunching sets under different social welfare considerations. Otherwise it is not possible to show the result without assuming specific distributions for social weights δ and for agent shares π.

For the sake of notational convenience we will focus on the bunching case of agent i and $i+1$. However it is possible to generalize the result for the cases where more than two agents are bunched. When agents i and $i+1$ are bunched it means that $IC_{i,i+1}$ and $IC_{i+1,i}$ are binding at the same time. The first order conditions for the other agents are exactly the same as for no-bunching. Denote y_b the income level of bunched agents, then first-order conditions for agents i and $i+1$ are:

$$(i):\ v'\left(\frac{y_b}{w_i}\right)\frac{1}{w_i}\left[\pi_i\delta^i + \mu_{i,i-1} + \mu_{i,i+1}\right] - \mu_{i+1,i}v'\left(\frac{y_b}{w_{i+1}}\right)\frac{1}{w_{i+1}} - \mu_{i-1,i}v'\left(\frac{y_b}{w_{i-1}}\right)\frac{1}{w_{i-1}}$$
$$=\pi_i\delta^i - \mu_{i-1,i} + \mu_{i,i-1} + \mu_{i,i+1} - \mu_{i+1,i}$$

$$(i+1):\ v'\left(\frac{y_b}{w_{i+1}}\right)\frac{1}{w_{i+1}}\left[\pi_{i+1}\delta^{i+1} + \mu_{i+1,i} + \mu_{i+1,i+2}\right] - \mu_{i,i+1}v'\left(\frac{y_b}{w_i}\right)\frac{1}{w_i} - \mu_{i+2,i+1}v'\left(\frac{y_b}{w_{i+2}}\right)\frac{1}{w_{i+2}}$$
$$= \pi_{i+1}\delta^{i+1} - \mu_{i,i+1} + \mu_{i+1,i} + \mu_{i+1,i+2} - \mu_{i+2,i+1}$$

by adding these conditions we have

$$\left[1 - v'\left(\frac{y_b}{w_i}\right)\frac{1}{w_i}\right] = \frac{\mu_{i-1,i}}{\pi_i\delta^i+\mu_{i,i-1}}\left[1 - v'\left(\frac{y_b}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right] - \frac{\pi_{i+1}\delta^{i+1}+\mu_{i+1,i+2}}{\pi_i\delta^i+\mu_{i,i-1}}\left[1 - v'\left(\frac{y_b}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$$
$$+\frac{\mu_{i+2,i+1}}{\pi_i\delta^i+\mu_{i,i-1}}\left[1 - v'\left(\frac{y_b}{w_{i+2}}\right)\frac{1}{w_{i+2}}\right]$$

By following similar steps as above, first we will analyse the Rawlsian and generally decreasing social weights cases. Second we will show that Maximax is a benchmark for all generally increasing social weights. Finally, we will show that it is possible to generalize this result for all other possible cases.

If we have generally decreasing social weights, we need to check whether labour supply or income level is higher in the Rawlsian case.

Proposition 7. *For all i those are bunched, $l_b^{GD} \geq l_b^R$ or equivalently $y_b^{GD} \geq y_b^R$.*

Proof. We need to show that if we switch from Rawlsian to any $GDSW$ case labour supply of the agent i increases. From the optimality condition we have;

$$\left[1 - v'\left(\frac{y_b^{GD}}{w_i}\right)\frac{1}{w_i}\right] = \frac{\mu_{i+2,i+1}}{\pi_i\delta^i+\mu_{i,i-1}}\left[1 - v'\left(\frac{y_b^{GD}}{w_{i+2}}\right)\frac{1}{w_{i+2}}\right] - \frac{\pi_{i+1}\delta^{i+1}}{\pi_i\delta^i+\mu_{i,i-1}}\left[1 - v'\left(\frac{y_b^{GD}}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$$

where $\mu_{i,i-1} = \beta_{i-1} - F_{i-1}$

By plugging $\mu_{i,i-1}$ we have;

$$\left[1 - v'\left(\frac{y_b^{GD}}{w_i}\right)\frac{1}{w_i}\right] = \frac{\beta_{i+1}-F_{i+1}}{\beta_i - F_{i-1}}\left[1 - v'\left(\frac{y_b^{GD}}{w_{i+2}}\right)\frac{1}{w_{i+2}}\right] - \frac{\pi_{i+1}\delta^{i+1}}{\beta_i - F_{i-1}}\left[1 - v'\left(\frac{y_b^{GD}}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$$

In Rawlsian case we have $\delta^i = 0$ except agent 1:

$$\left[1 - v'\left(\frac{y_b^{R}}{w_i}\right)\frac{1}{w_i}\right] = \frac{1-F_{i+1}}{1-F_{i-1}}\left[1 - v'\left(\frac{y_b^{R}}{w_{i+2}}\right)\frac{1}{w_{i+2}}\right]$$

First, suppose any δ^j where $j \leq i$ increases. Second term on the right hand side of the first equation is zero. For the first term, β_{i+1} will be equal to one, since it contains δ^j, and any increase in δ^j means a decrease in δ^1. Hence the conditions are exactly the same as for the labour supply levels.

Second, suppose we increase δ^j where $j > i+1$. Again the second term is zero, however in this case both β_{i+1} and β_i terms will reduce. By manipulating the terms we have;

$$\frac{\left[1-v'\left(\frac{y_b^{GD}}{w_{i+2}}\right)\frac{1}{w_{i+2}}\right]}{\left[1-v'\left(\frac{y_b^{GD}}{w_i}\right)\frac{1}{w_i}\right]} = \frac{\beta_i - F_{i-1}}{\beta_{i+1}-F_{i+1}} = \frac{\pi_1\delta^1 - F_{i-1}}{\pi_1\delta^1 - F_{i+1}} > \frac{1-F_{i-1}}{1-F_{i+1}} = \frac{\left[1-v'\left(\frac{y_b^{R}}{w_{i+2}}\right)\frac{1}{w_{i+2}}\right]}{\left[1-v'\left(\frac{y_b^{R}}{w_i}\right)\frac{1}{w_i}\right]}$$

by convexity of $v(.)$ function we have $y_b^{GD} > y_b^R$.

In the third case the difference between bunching and no-bunching appears. In no-bunching when we increase δ^{i+1} it does not have any effect on agent $i+1$. However in bunching, because it has an effect on agent i, it will change the income level of the agent $i+1$ as well. So suppose δ^j where $j = i+1$ increases, by manipulating the optimality condition we have;

$$\frac{\left[1-v'\left(\frac{y_b^{GD}}{w_{i+2}}\right)\frac{1}{w_{i+2}}\right]}{\left[1-v'\left(\frac{y_b^{GD}}{w_i}\right)\frac{1}{w_i}\right]} = \frac{\beta_i - F_{i-1}}{\beta_{i+1}-F_{i+1}} + \frac{\pi_{i+1}\delta^{i+1}}{\beta_{i+1}-F_{i+1}}\frac{\left[1-v'\left(\frac{y_b^{GD}}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^{GD}}{w_i}\right)\frac{1}{w_i}\right]} > \frac{1-F_{i-1}}{1-F_{i+1}} = \frac{\left[1-v'\left(\frac{y_b^{R}}{w_{i+2}}\right)\frac{1}{w_{i+2}}\right]}{\left[1-v'\left(\frac{y_b^{R}}{w_i}\right)\frac{1}{w_i}\right]}$$

Since $\frac{\left[1-v'\left(\frac{y_b^{GD}}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^{GD}}{w_i}\right)\frac{1}{w_i}\right]}$ term on the left hand side is greater than 1, we have the strict inequality which implies $y_b^{GD} > y_b^R$. □

Consequently; suppose the agents 3 and 4 bunched, then if δ^3 increases there will be no change for these agents. However when δ^4 increases both of the agents 3 and 4's labour supply levels and income levels will increase. This could be generalized to the cases where more than two agents are bunched. If the social weight of the first agent in the bunched group increases, there will be no change in the labour supply of the bunched group. However an increase in the social weight of the second or above agents affects the labour supply of all bunched agents.

After analysing the labour supply of agent i under two different social welfare functions, it is possible to compare the marginal tax rates under two different cases.

Proposition 8. *For all i,* $T_i^{'}(y_i^R) \geq T_i^{'}(y_i^{GD})$.

Proof. Follows as Proposition 2. □

A similar analysis could be conducted for Maximax SWF and any generally increasing social weights case.

Proposition 9. *For all i that are bunched,* $l_b^M \geq l_b^{GI}$ *or equivalently* $y_b^M \geq y_b^{GI}$.

Proof. Again suppose the agents i and $i+1$ are bunched at the optimum. First order optimality conditions for the other agents are exactly the same as with the no-bunching case however condition for agent i and $i+1$ are as follows;

$$\left[1-v'\left(\frac{y_b^{GI}}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]=\frac{\mu_{i-1,i}}{\pi_{i+1}\delta^{i+1}+\mu_{i+1,i+2}}\left[1-v'\left(\frac{y_b^{GI}}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]-\frac{\pi_i\delta^i}{\pi_{i+1}\delta^{i+1}+\mu_{i+1,i+2}}\left[1-v'\left(\frac{y_b^{GI}}{w_i}\right)\frac{1}{w_i}\right]$$

by plugging $\mu_{i-1,i}=F_{i-1}-\beta_{i-1}$ we have;

$$\left[1-v'\left(\frac{y_b^{GI}}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]=\frac{F_{i-1}-\beta_{i-1}}{F_{i+1}-\beta_i}\left[1-v'\left(\frac{y_b^{GI}}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]-\frac{\pi_i\delta^i}{F_{i+1}-\beta_i}\left[1-v'\left(\frac{y_b^{GI}}{w_i}\right)\frac{1}{w_i}\right]$$

The condition for Maximax case:

$$\left[1-v'\left(\frac{y_b^{M}}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]=\frac{F_{i-1}}{F_{i+1}}\left[1-v'\left(\frac{y_b^{M}}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]$$

First, suppose δ^j where $j\geq i+1$ increases, since we have $\beta_{i-1}=\beta_i=0$ the conditions are exactly the same as with the income levels.

Second, suppose δ^j where $j<i$ increases, then both β_{i-1} and β_i will increase and we have;

$$\frac{\left[1-v'\left(\frac{y_b^{GI}}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[1-v'\left(\frac{y_b^{GI}}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}=\frac{F_{i+1}-\beta_i}{F_{i-1}-\beta_{i-1}}>\frac{F_{i+1}}{F_{i-1}}=\frac{\left[1-v'\left(\frac{y_b^{M}}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[1-v'\left(\frac{y_b^{M}}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}$$

which implies that $y_b^M>y_b^{GI}$.

Third, again this is the difference from no-bunching. If δ^j where $j=i$ increases, in the no-bunching case we know that agents who have a higher ability level will have a lower labour supply level compared to maximax case. However when we have bunching, because δ^i affects agent $i+1$, it also affects agent i. We can manipulate the conditions as;

$$\frac{\left[1-v'\left(\frac{y_b^{GI}}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[1-v'\left(\frac{y_b^{GI}}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}=\frac{F_{i+1}-\beta_i}{F_{i-1}-\beta_{i-1}}+\frac{\pi_i\delta^i}{F_{i-1}-\beta_{i-1}}\frac{\left[1-v'\left(\frac{y_b^{GI}}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_b^{GI}}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}>\frac{F_{i+1}}{F_{i-1}}=\frac{\left[1-v'\left(\frac{y_b^{M}}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[1-v'\left(\frac{y_b^{M}}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}$$

The $\frac{\left[1-v'\left(\frac{y_b^{GI}}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_b^{GI}}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}$ term on the left hand side is always greater than one which leads to the strict inequality implying $y_b^M>y_b^{GI}$ again in this case. □

For the other social weight distributions, we can conduct the same analysis. However in this case the possible binding constraints for the agents are different than the separating equilibrium case. We left this analysis to the Appendix, however, in words, when we have a bunching between agent i and $i+1$ we have nine possible binding constraints, and under any of these possibilities, labour supply or income level cannot be lower than Rawlsian level and cannot exceed maximax labour supply level.

Proposition 10. *For all i those are bunched, $l_b^M\geq l_b\geq l_b^R$ or equivalently $y_b^M\geq y_b\geq y_b^R$.*

Proof. See Appendix. □

Proposition 11. *For all b, $T_b'(y_b^R)\geq T_b'(y_b)\geq T_b'(y_b^M)$.*

Proof. Follows as Proposition 2. □

The following section deals with the reverse of the relation. We try to characterize the efficient marginal tax schedules between Rawlsian and maximax benchmarks. We show that "reasonable" tax schedules could be supported as a second best efficient tax schedule with appropriate social weights.

7 Converse of the Relation

The preceding analysis aims to show that any second best marginal tax schedule lies below the Rawlsian SWF marginal rates and above the maximax SWF marginal rates. In this section, we will show that any reasonable marginal tax schedule between these two benchmarks can be supported as a second best optimal tax schedule by an appropriate distribution of social weights δ. Therefore we need to solve the optimal inverse problem. In the standard approach, by using a specific social welfare function and given population parameters optimal tax problem solves the marginal income tax schedule. However, here we seek for the social weights (or SWF) that are consistent with the actual marginal tax schedules. Bourguignon and Spadaro (2008) solves the inverse problem in a continuum model with a utility form that is linear in consumption and iso-elastic with respect to labour supply. The intuition is similar for discrete and continuum models. Once we observe the marginal tax schedule we could solve for the labour supply levels. The rest is just finding the appropriate social weights that make the labour supply levels second best efficient. Bourguignon and Spadaro (2008) is an empirical study as they infer the actual marginal tax schedule from income, tax and benefit data. Here our aim is to characterize the marginal tax schedules that could be supported as efficient tax schedules.

We will start with the usual redistributive (redistribute towards poor) marginal tax schedules. Given any non-negative marginal tax schedule $T_i^{'}(Y_i) < T_i^{'}(Y_i^R)$ for all i, one can find the allocation (l_i, c_i) for all i and the corresponding social weights δ^i as follow. From the agent market condition we know that $[1 - T_i^{'}(y_i)] = \frac{v^{'}(l_i)}{w_i}$. So for each agent labour supply level is $l_i = \frac{y_i}{w_i} = v^{'-1}[w_i[1 - T_i^{'}(y_i)]]$. Once we solve for labour supply levels, we can find the consumption from the incentive compatibility constraints and feasibility constraint. For each agent the consumption will be;

$$c_i = \left\{ \sum_{j=1}^{N} \pi_j y_j + \sum_{a=1}^{i-1} \pi_a \sum_{b=a}^{i-1} [R_{b+1}] - \sum_{a=i+1}^{N} \pi_a \sum_{b=i+1}^{a} [R_b] \right\}$$

where $R_i = v(\frac{y_i}{w_i}) - v(\frac{y_{i-1}}{w_i})$

This consumption function is the corresponding function in Weymark (1986, 1987) where he uses a quasi-linear in labour utility and derives the corresponding income function. Also this is the same term that appears in Simula (2010) when he studied optimal income taxation with a quasi-linear in consumption utility[8]. Both of these studies analyse the redistributive

[8]The corresponding consumption functions as in Weymark (1986, 1987) and Simula (2010) are as follows:

case and they solve a two step maximization problem which gives the same analytical results as we study here. The following proposition concludes the argument.

Proposition 12. *Any reasonable non-negative marginal tax schedule $T_i^{'}(y_i)$ is second best efficient with appropriate social weights.*

Proof. We solve for the social weights δ^i that lead to the marginal tax schedule $T_i^{'}(y_i)$. From the optimality condition of government problem we have $n-1$ equations $v^{'}\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} = 1 - \frac{\beta_i - F_i}{\beta_i - F_{i-1}}\left[1 - v^{'}\left(\frac{y_i}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$ for all $i \neq N$, and we also have the normalization $\sum_{i=1}^{N} \pi_i \delta^i = 1$. Since we have n equations and n unknowns we can solve for all δ^i as below:

$$\delta^i = \frac{\left[1 - v^{'}\left(\frac{y_i}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[v^{'}\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} - v^{'}\left(\frac{y_i}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} - \frac{\pi_{i-1}}{\pi_i}\left\{\frac{\left[1 - v^{'}\left(\frac{y_{i-1}}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[v^{'}\left(\frac{y_{i-1}}{w_{i-1}}\right)\frac{1}{w_{i-1}} - v^{'}\left(\frac{y_{i-1}}{w_i}\right)\frac{1}{w_i}\right]}\right\} \text{ for all } i \neq N.$$

and $\delta^N = \frac{1 - \sum_{i=1}^{N-1} \delta^i \pi_i}{\pi_N}$

Plugging the marginal tax terms is left to the Appendix. □

There should be a discussion for "reasonable marginal tax" schedules. It is important to note that this inversion procedure is inconsistent if the actual (or chosen) marginal tax schedule is not a solution to a problem that maximizes a social welfare function with respect to the budget constraint and IC constraints. Hence theory itself imposes restrictions on the chosen marginal tax rates. As an example if we have a marginal tax schedule which is positive for the first agent, negative for the second agent and positive for the third agent and follows like this, it is hard to find the social weights that support this tax schedule as a second best tax schedule. Even it is hard to believe that the government has such a strange redistribution desire, mathematically we could not disregard those possibilities, however it is not possible to characterize these cases. Therefore we have to restrict our attention to the "reasonable" marginal tax schedules.

First, any efficient tax schedule should satisfy the incentive compatibility constraints which also means that existing income levels should be strictly (if no bunching) monotonic. This will put conditions on the maximum and the minimum values for all marginal tax rates. From the market condition of the agent we know that $y_i = v^{'^{-1}}[w_i[1 - T_i^{'}(y_i)]]w_i$ holds for all agents in an efficient solution. So IC constraints imply that;

$$v^{'^{-1}}[w_{i+1}[1 - T_{i+1}^{'}(y_{i+1})]]w_{i+1} > v^{'^{-1}}[w_i[1 - T_i^{'}(y_i)]]w_i$$

which restricts the maximum and minimum values for marginal tax rates of agents i and $i+1$. Rewriting the inequality gives;

$$T_i^{'}(y_i) > 1 - v^{'}\left\{v^{'^{-1}}[w_{i+1}(1 - T_{i+1}^{'}(y_{i+1}))]\frac{w_{i+1}}{w_i}\right\}\frac{1}{w_i}$$

$$c_1 = \left\{\sum_{j=1}^{N} \pi_j y_j + \sum\left[\sum_{a=j+1}^{N} \pi_a\right]\left[v\left(\frac{w_j l_j}{w_{j+1}}\right) - v\left(\frac{w_{j+1} l_{j+1}}{w_{j+1}}\right)\right]\right\}$$

$$c_i = c_1 + \sum_{j=1}^{i-1}\left[v\left(\frac{w_{j+1} l_{j+1}}{w_{j+1}}\right) - v\left(\frac{w_j l_j}{w_{j+1}}\right)\right]$$

and

$$T_{i+1}'(y_{i+1}) < 1 - v'\left\{ v'^{-1}[w_i(1 - T_i'(y_i))] \frac{w_i}{w_{i+1}} \right\} \frac{1}{w_{i+1}}$$

So given the ability distribution, $T_{i+1}'(y_{i+1})$ sets the minimum value for $T_i'(y_i)$, while $T_i'(y_i)$ sets the maximum of $T_{i+1}'(y_{i+1})$. If the actual or desired tax schedule does not satisfy these conditions, it is not possible to reach that marginal tax schedule with usual optimal income tax problem.

The second restriction is the Paretian condition. In order to have a Pareto efficient tax schedule we should have a Paretian social welfare function which says that each agent in the economy should have a non-negative social welfare weight. This will restrict the choice for marginal tax schedules. In order to have non-negative social weights the following condition has to be satisfied;

$$\frac{1-\Omega_i}{1-T_i'(y_i)-\Omega_i} \geq \frac{\pi_{i-1}}{\pi_i}\left[\frac{T_{i-1}'(y_{i-1})}{1-T_{i-1}'(y_{i-1})-\Omega_{i-1}}\right]$$

where

$$\Omega_i = v'[v'^{-1}[w_i(1 - T_i'(y_i))] \frac{w_i}{w_{i+1}}] \frac{1}{w_{i+1}}.$$ [9]

Therefore once we have the ability distribution, population share and marginal tax rate for T_{i-1}', the condition gives the minimum value for the T_i', similarly on the other side T_i' sets the maximum marginal tax rate for T_{i-1}'.

For the preceding analysis we assume that all downward IC constraints are binding. Indeed this is the case when all agents labour supplies are distorted downward while the top agent is undistorted. Since we know the ability distribution, we could find the first best labour supply of the each agent by $l_i = v'^{-1}[w_i]$, and could check whether the downward IC constraints are binding by comparing the labour supply in the first best and the labour supply level implied by the actual marginal tax schedule. If all agents are distorted downward then we could apply the preceding analysis. Marginal tax schedule should satisfy these two conditions, if not we could not support the existing tax schedule as a second best efficient tax schedule with a generally decreasing social weights. Therefore, the second condition also shows that whether it is possible to achieve the actual tax schedule by only binding downward IC constraints.

For the opposite case, if all of the agents (except the lowest) are distorted upwards, then we know that all upward IC constraints are binding. We could apply the same analysis for non-redistributive second best schedules and maximax, and find the appropriate social weight

[9] By using optimality condition we could derive;

$$v'\left(\frac{y_i}{w_{i+1}}\right)\frac{1}{w_{i+1}} = v'\left\{[v'^{-1}[w_i(1 - T_i'(y_i))]] \frac{w_i}{w_{i+1}}\right\} \frac{1}{w_{i+1}}$$

and

$$v'\left(\frac{y_{i-1}}{w_i}\right)\frac{1}{w_i} = v'\left\{[v'^{-1}[w_{i-1}(1 - T_{i-1}'(y_{i-1}))]] \frac{w_{i-1}}{w_i}\right\} \frac{1}{w_i}$$

By plugging these conditions to the equation for social weight δ^i, we have a condition for the relation of social weight δ^i with the marginal tax rates T_i' and T_{i-1}'.

$$\delta^i = \frac{\left[1-v'\left\{\left[v'^{-1}\left[w_i(1-T_i'(y_i))\right]\right]\frac{w_i}{w_{i+1}}\right\}\frac{1}{w_{i+1}}\right]}{\left[1-T_i'(y_i)-v'\left\{\left[v'^{-1}\left[w_i(1-T_i'(y_i))\right]\right]\frac{w_i}{w_{i+1}}\right\}\frac{1}{w_{i+1}}\right]} - \frac{\pi_{i-1}}{\pi_i}\left\{\frac{T_{i-1}'(y_{i-1})}{\left[1-T_{i-1}'(y_{i-1})-v'\left\{\left[v'^{-1}\left[w_{i-1}(1-T_{i-1}'(y_{i-1}))\right]\right]\frac{w_{i-1}}{w_i}\right\}\frac{1}{w_i}\right]}\right\}$$

distribution in a similar way with redistributive case. The corresponding consumption function would be;

$$c_i = \left\{ \sum_{j=1}^{N} \pi_j y_j + \sum_{a=1}^{i-1} \pi_a \sum_{b=a}^{i-1} [S_{b+1}] - \sum_{a=i+1}^{N} \pi_a \sum_{b=i+1}^{a} [S_b] \right\}$$

where $S_i = v\left(\frac{y_i}{w_{i-1}}\right) - v\left(\frac{y_{i-1}}{w_{i-1}}\right)$ and we can solve for all δ^i from optimality conditions

$$v^{'}\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} = 1 - \frac{F_{i-1} - \beta_{i-1}}{F_i - \beta_{i-1}}\left[1 - v^{'}\left(\frac{y_i}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right] \text{ for all } i \neq 1.$$

and the normalization rule. These conditions yield the following equations for social weights;

$$\delta^i = \frac{\left[1 - v'\left(\frac{y_i}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} - v'\left(\frac{y_i}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]} - \frac{\pi_{i+1}}{\pi_i}\left\{\frac{\left[1 - v'\left(\frac{y_{i+1}}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[v'\left(\frac{y_{i+1}}{w_{i+1}}\right)\frac{1}{w_{i+1}} - v'\left(\frac{y_{i+1}}{w_i}\right)\frac{1}{w_i}\right]}\right\} \text{ for all } i \neq N.$$

and

$$\delta^N = \frac{1 - \sum_{i=1}^{N-1} \delta^i \pi_i}{\pi_N}$$

The preceding two analysis are very robust and easy to implement. However, there are still marginal tax schedules between Rawlsian and Maximax that could not be covered by these two cases. In order to cover the other tax schedules between Rawlsian and maximax, we follow the following procedure. Table 2 shows the possible nine cases for each agent and the sign of the distortions for a 6 type example.

Table 2: Possible Binding IC Constraints and Sign of Distortions

Cases	***Distortion***	***Agent* 1**	***Agent* 2**	***Agent* 3**	***Agent* 4**	***Agent* 5**	***Agent* 6**
1	Downward	$IC_{2,1}$	$IC_{2,1}, IC_{3,2}$	$IC_{3,2}, IC_{4,3}$	$IC_{4,3}, IC_{5,4}$	$IC_{5,4}, IC_{6,5}$	–
7	Downward	–	$IC_{3,2}$	$IC_{4,3}$	$IC_{5,4}$	$IC_{6,5}$	–
5	Upward	–	$IC_{1,2}, IC_{2,3}$	$IC_{2,3}, IC_{3,4}$	$IC_{3,4}, IC_{4,5}$	$IC_{4,5}, IC_{5,6}$	$IC_{5,6}$
6	Upward	–	$IC_{1,2}$	$IC_{2,3}$	$IC_{3,4}$	$IC_{4,5}$	–
4	Ambiguous	–	$IC_{1,2}, IC_{3,2}$	$IC_{2,3}, IC_{4,3}$	$IC_{3,4}, IC_{5,4}$	$IC_{4,5}, IC_{6,5}$	–
2	Undistorted	$IC_{1,2}$	$IC_{2,1}, IC_{2,3}$	$IC_{3,2}, IC_{3,4}$	$IC_{4,3}, IC_{4,5}$	$IC_{5,4}, IC_{5,6}$	$IC_{6,5}$
3	Undistorted	–	$IC_{2,1}$	$IC_{3,2}$	$IC_{4,3}$	$IC_{5,4}$	–
8	Undistorted	–	$IC_{2,3}$	$IC_{3,4}$	$IC_{4,5}$	$IC_{5,6}$	–
9	Undistorted	*None*	*None*	*None*	*None*	*None*	*None*

As we mention above, when we pick a reasonable tax schedule between Rawlsian and Maximax, it is always possible to find the efficient labour supply levels from these tax rates as we have $l_i = v^{'^{-1}}[w_i[1 - T_i^{'}(y_i)]]$. Also we know the agents' labour supply level at the first-best case. After the comparison of these two labour supply levels, we could find the possible binding constraints for each agent. The rest is just finding the consistent binding IC constraints set. Once we found which IC constraints are binding we could solve for the consumption levels. Then we could find the social weight distribution that leads to the selected marginal tax schedule. This inversion procedure works if we do not cover cases where more than one agent have a zero marginal tax rate. Otherwise there will be less equations than the unknowns. For those cases the social weights could be solved with an optimization procedure subject to the first order conditions, inequalities for cumulative weighted social weights β and cumulative population share F. In a continuum setup since the IC constraint is just a differential equation, optimal inverse problem is easy to solve. However for the discrete setup, in order to solve the problem we need to find which IC constraints are binding in the optimum.

As long as the chosen marginal tax schedule (or actual tax schedule) satisfies the two efficiency condition stated above (monotonicity and Paretian), this inversion algorithm is applicable.

The next section considers the total taxes under different social welfare considerations.

8 Total Taxes

Rawlsian and Maximax social welfare functions constitute benchmarks for the total tax of the lowest and highest able individuals. Total tax paid for each agent is equal to the difference between his income and consumption;

$$T_i\left(y_i\right) = y_i - c_i = w_i l_i - c_i$$

For the redistributive cases we showed that consumption of agent i would be;

$$c_i = \left\{ \sum_{j=1}^{N} \pi_j y_j + \sum_{a=1}^{i-1} \pi_a \sum_{b=a}^{i-1} \left[R_{b+1}\right] - \sum_{a=i+1}^{N} \pi_a \sum_{b=i+1}^{a} \left[R_b\right] \right\}$$

where $R_i = v\left(\frac{w_i l_i}{w_i}\right) - v\left(\frac{w_{i-1} l_{i-1}}{w_i}\right)$

Total Tax at the top:

$$T_N(y_N) = w_N l_N - c_N = w_N l_N - \left\{ \sum_{j=1}^{N} \pi_j y_j + \sum_{a=1}^{N-1} \pi_a \sum_{b=a}^{N-1} \left[R_{b+1}\right] \right\}$$

and

$$\frac{\partial T_N(y_N)}{\partial \delta^k} = w_N \frac{\partial l_N}{\partial \delta^k} - \left\{ \sum_{j=1}^{k-1} \pi_j w_j \frac{\partial l_j}{\partial \delta^k} + \sum_{a=1}^{k-1} \pi_a \sum_{b=a}^{k-1} \left[\frac{\partial R_{b+1}}{\partial \delta^k}\right] \right\} < 0 \text{ for all } k.$$

Which means if we increase any social weight in the population this leads to a reduction in the total tax of the highest able agent. So whatever the total tax for the other agents, we can say that Rawlsian total tax level is an upper bound for all second best allocations. Total Tax at the bottom:

$$T_1(y_1) = w_1 l_1 - c_1 = w_1 l_1 - \left\{ \sum_{j=1}^{N} \pi_j y_j - \sum_{a=2}^{N} \pi_a \sum_{b=2}^{a} \left[R_b\right] \right\}$$

and

$$\frac{\partial T_1(y_1)}{\partial \delta^k} = w_1(\frac{\partial l_1}{\partial \delta^k}) - \left\{ \sum_{j=1}^{k-1} \pi_j w_j \frac{\partial l_j}{\partial \delta^k} - \sum_{a=2}^{N} \pi_a \sum_{b=2}^{a} \left[\frac{\partial R_b}{\partial \delta^k} \right] \right\} > 0 \text{ for all } k.$$

Then for the lowest able agent, Rawlsian total tax is a lower bound for the second best optimal tax rates. On the other hand maximax gives an lower bound for the top agent and an upper bound for the lowest able agent.

In maximax case we have the following total tax for the top agent;

$$T_N(y_N) = w_N l_N - \left\{ \sum_{j=1}^{N} \pi_j w_j \frac{\partial l_j}{\partial \delta^k} + \sum_{a=1}^{N-1} \pi_a \sum_{b=a}^{N-1} \left[\frac{\partial S_{b+1}}{\partial \delta^k} \right] \right\}$$

and

$$\frac{\partial T_N(y_N)}{\partial \delta^k} = w_N \frac{\partial l_N}{\partial \delta^k} - \left\{ \sum_{j=1}^{k-1} \pi_j w_j \frac{\partial l_j}{\partial \delta^k} + \sum_{a=1}^{k-1} \pi_a \sum_{b=a}^{k-1} \left[\frac{\partial R_{b+1}}{\partial \delta^k} \right] \right\} > 0 \text{ for all } k.$$

and for the lowest agent;

$$T_1(y_1) = w_1 l_1 - c_1 = w_1 l_1 - \left\{ \sum_{j=1}^{N} \pi_j w_j l_j - \sum_{a=2}^{N} \pi_a \sum_{b=2}^{a} [S_b] \right\}$$

and

$$\frac{\partial T_1(y_1)}{\partial \delta^k} = w_1 \frac{\partial l_1}{\partial \delta^k} - \left\{ \sum_{j=1}^{k-1} \pi_j w_j \frac{\partial l_j}{\partial \delta^k} - \sum_{a=2}^{N} \pi_a \sum_{b=2}^{a} \left[\frac{\partial S_b}{\partial \delta^k} \right] \right\} < 0 \text{ for all } k.$$

The following proposition summarizes these results:

Proposition 13. *For the highest able agent we have:* $T_N(y_N^M) \leq T_N(y_N) \leq T_N(y_N^R)$

For the lowest able agent we have: $T_1(y_1^R) \leq T_1(y_1) \leq T_1(y_1^M)$

The Rawlsian marginal tax schedule, any second best marginal tax schedule and maximax marginal tax schedule cross only once. So there is a critical point in the population below which the Rawlsian tax rate constitutes a lower bound and gives an upper bound above the threshold. It is the reverse for the maximax case, so Maximax constitutes an upper bound below the threshold ability level and gives a lower bound above the threshold. This threshold is identified by a complicated relation of social weight distribution δ, ability distribution w, and the share π of each ability in the distribution. One last point, with reasonable social weight distributions and population shares, numerical simulations show that Rawlsian and maximax SWF total tax rates are intersecting near median agent. This result is somehow similar with the optimal tax schedules those are identified by majority voting when the median voter theorem holds (Brett and Weymark (2015)).

9 Conclusion

In a discrete type setup assuming a quasi-linear in consumption utility, the present paper shows that Rawlsian social welfare and maximax social welfare functions constitute upper and lower bounds for the second best optimal marginal tax schedules. Also reasonable marginal tax schedules between these two benchmarks can be supported as a second best tax schedule with appropriate social weights. These results are also valid when the monotonicity constraint binds. Finally, we give some characterization for the total tax rates at the top and bottom of the income distribution.

The analysis for general setup is straightforward but a more unified approach would be better to present the results. The algorithm for finding the binding IC constraints is very robust for redistributive cases. Therefore it is possible to support any reasonable redistributive marginal tax schedule as a second best marginal tax schedule.

Under quasi-linear in consumption utility, marginal tax rates are independent of consumption. Hence social planner could change the consumption levels to satisfy constraints without any effect on the marginal rates. However, under the presence of income effect, marginal tax rates are affected by consumption. Our results in this study do not hold for the very general utility function. However, a set of weaker assumptions may be imposed to a more general utility function which would be a further contribution to the present study.

The optimal income taxation and non-linear pricing have identical setups with minor differences. We believe that the idea in this study can be extended to non-linear pricing to give a characterization for the efficient price levels.

References

[1] Atkinson, A.B., (1983), "How Progressive Should Income Tax Be?", Chapter 15, Social Justice and Public Policy, MIT Press.

[2] Bastani, S., (2013), "Using the Discrete Model to Derive Optimal Income Tax Rates", Working Paper 2013:11, Uppsala Center for Fiscal Studies.

[3] Blundell, R. and MaCurdy T. (1999), "Labor Supply: A Review of Alternative Approaches", in O. Ashenfelter and D. Card (eds.), Handbook of Labor Economics, Vol. 3, North-Holland, Amsterdam, pp. 1559-1695.

[4] Boadway, R., Marchand, M., Pestieau, P., Del Mar Racionero, M. (2002), "Optimal Redistribution with Heterogeneous Preferences for Leisure", Journal of Public Economic Theory, 4-4, pp. 475-498.

[5] Bourguignon, F. and Spadaro, A., (2008), "Tax-benefit revealed social preferences", EUROMOD Working Papers

[6] Brett, C. and Weymark, J. (2015), "Voting Over Selfishly Optimal Nonlinear Income Tax Schedules", Working Paper

[7] Choné, P. and Laroque, G. (2005) "Optimal Incentives for Labor Force Participation", J. Public Econ. 89(2--3) (2005) 395-425.

[8] Diamond, P. (1980), "Income taxation with fixed hours of work", Journal of Public Economics 13(1), 101-110.

[9] Diamond, P. (1998), "Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates", The American Economic Review Vol. 88, No. 1, pp. 83-95

[10] Ebert, U. (1992), "A reexamination of the optimal nonlinear income tax". Journal of Public Economics, 49(1): 47-73

[11] Guesnerie, R., (1995). "A Contribution to the Pure Theory of Taxation", Econometric Society Monograph. Cambridge University Press

[12] Guesnerie R., Seade J. (1982), "Nonlinear pricing in a finite economy". Journal of Public Economics, 57:157-179

[13] Hammond, P., (1979), "Straightforward individual incentive compatibility in large economies". Review of Economic Studies 46, 263-282.

[14] Hellwig, M. F. (2007). "A contribution to the theory of optimal utilitarian income taxation". Journal of Public Economics, 91:1449-1477.

[15] Homburg, S. (2002). “The Optimal Income Tax: Restatement and Extensions”. Hannover Economic Papers (HEP) dp-252, Leibniz Universität Hannover, Wirtschaftswissenschaftliche Fakultät.

[16] Jacquet, L. (2010), “Optimal Labor Income Taxation under Maximin: An Upper Bound”, Economics Bulletin, 30(4): 3152-3160.

[17] Kimball, M., (1990), “Precautionary saving in the small and in the large”. Econometrica, 58:53-73

[18] Laroque, G. (2005), “Income Maintenance and Labor Force Participation”, Econometrica, 73(2): 341-376 American Economic Review, 88(1), 83-95.

[19] Lollivier S., Rochet J-C, (1983), “Bunching and second-order conditions: a note on optimal tax theory”. Journal of Economic Theory, 31:392-400

[20] Lorenz, N. and Sachs, D., (2012), “Optimal Participation Taxes and Efficient Transfer Phase-Out”. No 2012-37, Working Paper Series of the Department of Economics, University of Konstanz.

[21] Mirrlees, J. (1971), “An Exploration in the Theory of Optimum Income Taxation”, Review of Economic Studies, 38(114), 175-208.

[22] Piketty, T. and Saez, E. (2012). “Optimal labor income taxation”. NBER Working Papers 18521, National Bureau of Economic Research, Inc.

[23] Sadka, E. (1976), “On income distribution, incentive effects and optimal income taxation”, Review of Economic Studies, 43(2), 261-8.

[24] Sadka, E. (1976), “On Progressive Income Taxation”, The American Economic Review, Vol. 66, No. 5, pp. 931-935.

[25] Saez, E. (2001), “Using Elasticities to Derive Optimal Income Tax Rates”, Review of Economics Studies, 68, 205-229.

[26] Seade, J. (1977), “On the Shape of Optimal Tax Schedules”, Journal of Public Economics, 7(2), 203--235.

[27] Simula, L. (2010). “Optimal nonlinear income tax and nonlinear pricing: optimality conditions and comparative static properties”. Social Choice and Welfare, 35(2):199-220.

[28] Stantcheva, S., (2014), “Optimal Income Taxation with Adverse Selection in the Labor Market”, Review of Economic Studies 81: 1296-1329.

[29] Stiglitz, J. E. (1982), “Self-Selection and Pareto Efficient Taxation”, Journal of Public Economics, 17(2): 213--40.

[30] Tuomala, M. (1990), "Optimal Income Tax and Redistribution", Oxford University Press, USA.

[31] Tuomala, M. (2010), "On optimal non-linear income taxation: numerical results revisited", International Tax and Public Finance, vol. 17-3, pages 259-270

[32] Weymark, J.A., (1986), "A reduced-form optimal nonlinear income tax problem". Journal of Public Economics, 30:199--217

[33] Weymark, J.A., (1987), "Comparative static properties of optimal nonlinear income taxes". Econometrica, 55:1165--1185

Appendix

Proof of Lemma 1:

From $IC_{i,i-1}$ and $IC_{i-1,i}$ we have;

$$\left[v\left(\frac{y_i}{w_{i-1}}\right) - v\left(\frac{y_i}{w_i}\right)\right] - \left[v\left(\frac{y_{i-1}}{w_{i-1}}\right) - v\left(\frac{y_{i-1}}{w_i}\right)\right] \geq 0$$

The function $f(y) = v\left(\frac{y}{w_{i-1}}\right) - v\left(\frac{y}{w_i}\right)$ is increasing in y since its derivative;

$v'\left(\frac{y}{w_{i-1}}\right)\frac{1}{w_{i-1}} - v'\left(\frac{y}{w_i}\right)\frac{1}{w_i}$ is positive by the convexity of $v(.)$ and sorting condition $w_i > w_{i-1}$. Hence, $IC_{i,i-1}$ and $IC_{i-1,i}$ imply $y_i \geq y_{i-1}$. Rewrite $IC_{i,i-1}$ as $c_i - c_{i-1} \geq v\left(\frac{y_i}{w_i}\right) - v\left(\frac{y_{i-1}}{w_i}\right)$ so if $y_i \geq y_{i-1}$ then we have $c_i \geq c_{i-1}$. Suppose $c_i = c_{i-1}$ but $y_i \neq y_{i-1}$, then from $IC_{i,i-1}$ and $IC_{i-1,i}$ we have: $v\left(\frac{y_{i-1}}{w_i}\right) \geq v\left(\frac{y_i}{w_i}\right)$ and $v\left(\frac{y_i}{w_{i-1}}\right) \geq v\left(\frac{y_{i-1}}{w_{i-1}}\right)$, since these inequalities cannot hold at the same time, we should have $y_i = y_{i-1}$, otherwise one of the IC constraints would be violated.

Proof of Lemma 2:

First, local downward incentive compatibility constraints $IC_{i,i-1}$ and $IC_{i-1,i-2}$ implies global downward incentive compatibility constraint $IC_{i,i-2}$. From $IC_{i,i-1}$ and $IC_{i-1,i-2}$ we have;

$$v\left(\frac{y_{i-1}}{w_i}\right) - v\left(\frac{y_i}{w_i}\right) \geq c_{i-1} - c_i$$

$$v\left(\frac{y_{i-2}}{w_{i-1}}\right) - v\left(\frac{y_{i-1}}{w_{i-1}}\right) \geq c_{i-2} - c_{i-1}$$

Adding the conditions imply

$$v\left(\frac{y_{i-1}}{w_i}\right) - v\left(\frac{y_i}{w_i}\right) + v\left(\frac{y_{i-2}}{w_{i-1}}\right) - v\left(\frac{y_{i-1}}{w_{i-1}}\right) \geq c_{i-2} - c_i \ (\bigstar)$$

Note that the function $f(w) = v\left(\frac{y_{i-2}}{w}\right) - v\left(\frac{y_{i-1}}{w}\right)$ is increasing in w since its derivative $v'\left(\frac{y_{i-1}}{w}\right)\frac{y_{i-1}}{w^2} - v'\left(\frac{y_{i-2}}{w}\right)\frac{y_{i-2}}{w^2}$ is positive by the convexity of $v(.)$.

The LHS of $(\bigstar)$ is smaller than $v\left(\frac{y_{i-1}}{w_i}\right) - v\left(\frac{y_i}{w_i}\right) + v\left(\frac{y_{i-2}}{w_i}\right) - v\left(\frac{y_{i-1}}{w_i}\right) = v\left(\frac{y_{i-2}}{w_i}\right) - v\left(\frac{y_i}{w_i}\right)$ hence $IC_{i,i-2}$ is satisfied: $v\left(\frac{y_{i-2}}{w_i}\right) - v\left(\frac{y_i}{w_i}\right) \geq c_{i-2} - c_i$.

Second, local upward incentive compatibility constraints $IC_{i-1,i}$ and $IC_{i-2,i-1}$ implies global upward incentive compatibility constraint $IC_{i-2,i}$. From $IC_{i-1,i}$ and $IC_{i-2,i-1}$ we have

$$v\left(\frac{y_i}{w_{i-1}}\right) - v\left(\frac{y_{i-1}}{w_{i-1}}\right) \geq c_i - c_{i-1}$$

$$v\left(\frac{y_{i-1}}{w_{i-2}}\right) - v\left(\frac{y_{i-2}}{w_{i-2}}\right) \geq c_{i-1} - c_{i-2}$$

Adding the conditions imply

$$v\left(\frac{y_i}{w_{i-1}}\right) - v\left(\frac{y_{i-1}}{w_{i-1}}\right) + v\left(\frac{y_{i-1}}{w_{i-2}}\right) - v\left(\frac{y_{i-2}}{w_{i-2}}\right) \geq c_i - c_{i-2}$$

Similarly, the LHS is smaller than

$$v\left(\frac{y_i}{w_{i-2}}\right) - v\left(\frac{y_{i-1}}{w_{i-2}}\right) + v\left(\frac{y_{i-1}}{w_{i-2}}\right) - v\left(\frac{y_{i-2}}{w_{i-2}}\right) = v\left(\frac{y_i}{w_{i-2}}\right) - v\left(\frac{y_{i-2}}{w_{i-2}}\right)$$

hence $IC_{i-2,i}$ is satisfied:

$v\left(\frac{y_i}{w_{i-2}}\right) - v\left(\frac{y_{i-2}}{w_{i-2}}\right) \geq c_i - c_{i-2}$

Proof of Lemma 3:

If both $IC_{i,i+1}$ and $IC_{i+1,i}$ bind we have;

$c_{i+1} - v\left(\frac{y_{i+1}}{w_{i+1}}\right) = c_i - v\left(\frac{y_i}{w_{i+1}}\right)$ and $c_i - v\left(\frac{y_i}{w_i}\right) = c_{i+1} - v\left(\frac{y_{i+1}}{w_i}\right)$

Adding the conditions imply;

$v\left(\frac{y_{i+1}}{w_{i+1}}\right) - v\left(\frac{y_i}{w_{i+1}}\right) = v\left(\frac{y_{i+1}}{w_i}\right) - v\left(\frac{y_i}{w_i}\right)$

Note that the function $f(w) = v\left(\frac{y_{i+1}}{w}\right) - v\left(\frac{y_i}{w}\right)$ is decreasing in w. Since we have $w_{i+1} > w_i$, this can happen only in bunching case *i.e.* $y_{i+1} = y_i$ and $c_{i+1} = c_i$

Proof of Proposition 4:

Marginal tax function for agent i with $GISW$:

$T_i'(y_i^{GI}) = \frac{F_{i-1}-\beta_{i-1}}{F_i-\beta_{i-1}}\left[1 - v'\left(\frac{w_i l_i^{GI}}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]$

Since under maximax we have only δ^N this condition will be as follows;

$T_i'(y_i^M) = \frac{F_{i-1}}{F_i}\left[1 - v'\left(\frac{w_i l_i^M}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]$

In *Proposition* 3 we show if δ^j where $j \geq i$ increases, β_{i-1} term will be zero. So the optimality conditions would be exactly the same and there is no change in the labour supply. So marginal tax rates are same. However if δ^j where $j < i$ increases, then β_{i-1} will be positive, and we have higher labour supply in maximax which leads to the result.

Proof of Proposition 5:

The agent condition is:

$$v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} = 1 - \frac{\mu_{i+1,i}}{[\pi_i\delta^i+\mu_{i,i-1}+\mu_{i,i+1}]}\left[1 - v'\left(\frac{y_i}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$$
$$-\frac{\mu_{i-1,i}}{[\pi_i\delta^i+\mu_{i,i+1}+\mu_{i,i-1}]}\left[1 - v'\left(\frac{y_i}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]$$

There is no need to discuss the cases 2, 3, 8 and 9 since in these cases agent i would be undistorted. So labour supply will be between the Rawlsian and Maximax cases so the marginal tax rates. For the remaining cases we need to consider each case separately;

In Rawlsian case the condition is: $v'\left(\frac{y_i^R}{w_i}\right)\frac{1}{w_i} = 1 - \frac{1-F_i}{1-F_{i-1}}\left[1 - v'\left(\frac{y_i^R}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$

Similarly in maximax case: $v'\left(\frac{y_i^M}{w_i}\right)\frac{1}{w_i} = 1 - \frac{F_{i-1}}{F_i}\left[1 - v'\left(\frac{y_i^M}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]$

1-) If $\beta_{i-1} > F_{i-1}$ and $\beta_i > F_i$ then $IC_{i,i-1}$ and $IC_{i+1,i}$ bind. Optimality condition: $v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} = 1 - \frac{\mu_{i+1,i}}{\pi_i\delta^i+\mu_{i,i-1}}\left[1 - v'\left(\frac{y_i}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$ where the distortion depends on agent $i+1$. $\frac{\mu_{i+1,i}}{\pi_i\delta^i+\mu_{i,i-1}} = \frac{\beta_i-F_i}{\beta_i-F_{i-1}}$ which is less than or equal to $\frac{1-F_i}{1-F_{i-1}}$. Result follow by the convexity of $v(.)$. Since in maximax there is an upward distortion, result trivially holds for maximax case.

4-) If $\beta_{i-1} < F_{i-1}$ and $\beta_i > F_i$ then $IC_{i-1,i}$ and $IC_{i+1,i}$ binds. Optimality condition: $v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} = 1 - \frac{\mu_{i-1,i}}{\pi_i\delta^i}\left[1 - v'\left(\frac{y_i}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right] - \frac{\mu_{i+1,i}}{\pi_i\delta^i}\left[1 - v'\left(\frac{y_i}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$ where the distortion depends on agents $i-1$ and $i+1$. This is the only case that we do not know the sign of the distortion. If $IC_{i-1,i}$ is binding then agent $i-1$ is either FB or distorted upward. Then second term is a positive number. We could show that the result would hold even we do not have this positive term. We have $\frac{\mu_{i+1,i}}{\pi_i\delta^i} = \frac{\beta_i-F_i}{\pi_i\delta^i}$.

Result holds under $\frac{1-F_i}{1-F_{i-1}} \geq \frac{\beta_i-F_i}{\pi_i\delta^i}$

Suppose the opposite is true so; $\frac{1-F_i}{1-F_{i-1}} < \frac{\beta_i-F_i}{\pi_i\delta^i}$

rewrite the condition as:

$$[\pi_{i+1} + ... + \pi_N]\pi_i\delta^i < \pi_{i+1}[\pi_i + ... + \pi_N][1-\delta^i] + ... + \pi_N[\pi_i + ... + \pi_N][1-\delta^N]$$

by rearranging we have;

$$\pi_i\delta^i + ... + \pi_N\delta^N < \pi_i + ... + \pi_N - \frac{\pi_i}{[\pi_{i+1}+...+\pi_N]}[\pi_{i+1}\delta^{i+1} + ... + \pi_N\delta^N]$$

Which is not possible since $\beta_{i-1} < F_{i-1}$ implies $\pi_i\delta^i + ... + \pi_N\delta^N > \pi_i + ... + \pi_N$

5-) If $\beta_{i-1} < F_{i-1}$ and $\beta_i < F_i$ then $IC_{i-1,i}$ and $IC_{i,i+1}$ bind. Optimality condition:

$v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} = 1 - \frac{\mu_{i-1,i}}{\pi_i\delta^i + \mu_{i,i+1}}\left[1 - v'\left(\frac{y_i}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]$ where distortion depends on agent $i-1$, and $\frac{\mu_{i-1,i}}{\pi_i\delta^i + \mu_{i,i+1}} = \frac{F_{i-1}-\beta_{i-1}}{F_i-\beta_{i-1}}$

Since there is an upward distortion for the agent i, labour supply is higher than Rawlsian case. For the comparison between maximax and this possibility:

In maximax we have: $\left[1 - v'\left(\frac{y_i^M}{w_i}\right)\frac{1}{w_i}\right] = \frac{F_{i-1}}{F_i}\left[1 - v'\left(\frac{y_i^M}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]$

Since $\frac{F_{i-1}}{F_i} \geq \frac{F_{i-1}-\beta_{i-1}}{F_i-\beta_{i-1}}$, we have

$\frac{\left[1-v'\left(\frac{y_i^M}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_i^M}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]} \geq \frac{\left[1-v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_i}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}$ then by convexity of $v(.)$ we have $y_i^M \geq y_i$.

6-) If $\beta_{i-1} < F_{i-1}$ and $\beta_i = F_i$ then only $IC_{i-1,i}$ binds. Optimality condition:

$v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} = 1 - \frac{\mu_{i-1,i}}{\pi_i\delta^i}\left[1 - v'\left(\frac{y_i}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]$ where the distortion depends on agent $i-1$, and $\frac{\mu_{i-1,i}}{\pi_i\delta^i} = \frac{F_{i-1}-\beta_{i-1}}{\pi_i\delta^i}$

Since there is an upward distortion for the agent i, labour supply is higher than Rawlsian case.

For the maximax we have $\left[1 - v'\left(\frac{y_i^M}{w_i}\right)\frac{1}{w_i}\right] = \frac{F_{i-1}}{F_i}\left[1 - v'\left(\frac{y_i^M}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]$

So, if $\frac{F_{i-1}}{F_i} \geq \frac{F_{i-1}-\beta_{i-1}}{\pi_i\delta^i}$ result holds.

Suppose the opposite; $\frac{F_{i-1}}{F_i} < \frac{F_{i-1}-\beta_{i-1}}{\pi_i\delta^i}$

by using $\beta_i = F_i$, we can write the term as $[\pi_1 + ... + \pi_i] < \delta^i\pi_i$ which is not possible.

7-) If $\beta_{i-1} = F_{i-1}$ and $\beta_i > F_i$ then only $IC_{i+1,i}$ binds. Optimality condition:

$v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} = 1 - \frac{\mu_{i+1,i}}{\pi_i\delta^i}\left[1 - v'\left(\frac{y_i}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$ where distortion depends on agent $i+1$, and $\frac{\mu_{i+1,i}}{\pi_i\delta^i} = \frac{\beta_i-F_i}{\pi_i\delta^i}$

Result holds if $\frac{1-F_i}{1-F_{i-1}} \geq \frac{\beta_i-F_i}{\pi_i\delta^i}$

Suppose the opposite is true: $\frac{1-F_i}{1-F_{i-1}} < \frac{\beta_i-F_i}{\pi_i\delta^i}$

by using $\beta_{i-1} = F_{i-1}$, we can write the condition as: $\pi_i + ... + \pi_N < \delta^i\pi_i$. However this is not possible since $\beta_{i-1} = F_{i-1}$ implies $\pi_i\delta^i + ... + \pi_N\delta^N = \pi_i + ... + \pi_N$.

Proof of Proposition 10:

Table A1 presents the counterpart of Table 2 for bunching case supposing agent 3 and 4 bunched and we have 6 agents.

Table A1: Binding IC Constraints and Sign of Distortions for Bunching Case

Cases	***Distortion***	***Agent 1***	***Agent 2***	***Agent 3 and 4***	***Agent 5***	***Agent 6***
1	Downward	$IC_{2,1}$	$IC_{2,1}, IC_{3,2}$	$(IC_{4,3}, IC_{3,4}) - IC_{3,2}, IC_{5,4}$	$IC_{5,4}, IC_{6,5}$	–
7	Downward	–	$IC_{3,2}$	$(IC_{4,3}, IC_{3,4}) - IC_{5,4}$	$IC_{6,5}$	–
5	Upward	–	$IC_{1,2}, IC_{2,3}$	$(IC_{4,3}, IC_{3,4}) - IC_{2,3}, IC_{4,5}$	$IC_{4,5}, IC_{5,6}$	$IC_{5,6}$
6	Upward	–	$IC_{1,2}$	$(IC_{4,3}, IC_{3,4}) - IC_{2,3}$	$IC_{4,5}$	–
4	Ambiguous	–	$IC_{1,2}, IC_{3,2}$	$(IC_{4,3}, IC_{3,4}) - IC_{2,3}, IC_{5,4}$	$IC_{4,5}, IC_{6,5}$	–
2	Undistorted	$IC_{1,2}$	$IC_{2,1}, IC_{2,3}$	$(IC_{4,3}, IC_{3,4}) - IC_{3,2}, IC_{4,5}$	$IC_{5,4}, IC_{5,6}$	$IC_{6,5}$
3	Undistorted	–	$IC_{2,1}$	$(IC_{4,3}, IC_{3,4}) - IC_{3,2}$	$IC_{5,4}$	–
8	Undistorted	–	$IC_{2,3}$	$(IC_{4,3}, IC_{3,4}) - IC_{4,5}$	$IC_{5,6}$	–
9	Undistorted	*None*	*None*	$(IC_{4,3}, IC_{3,4})$	*None*	*None*

For notational convenience suppose that agents i and $i+1$ are bunched at the optimum. Then we will have 9 possible cases for binding IC constraints. Each case leads to less downward distortion compared to Rawlsian, and less upward distortion compared to Maximax. This leads to the result that marginal tax rates are highest in Rawlsian SWF and lowest in maximax case. General optimality condition is as follows:

$$\left[1 - v'\left(\frac{y_b}{w_i}\right)\frac{1}{w_i}\right] = \frac{\mu_{i-1,i}}{\pi_i\delta^i+\mu_{i,i-1}}\left[1 - v'\left(\frac{y_b}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right] - \frac{\pi_{i+1}\delta^{i+1}+\mu_{i+1,i+2}}{\pi_i\delta^i+\mu_{i,i-1}}\left[1 - v'\left(\frac{y_b}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right] + \frac{\mu_{i+2,i+1}}{\pi_i\delta^i+\mu_{i,i-1}}\left[1 - v'\left(\frac{y_b}{w_{i+2}}\right)\frac{1}{w_{i+2}}\right]$$

All possible case are;

1-) $IC_{i+1,i} - IC_{i,i+1} - IC_{i+2,i+1}$

$$\left[1 - v'\left(\frac{y_b}{w_i}\right)\frac{1}{w_i}\right] = \frac{\mu_{i+2,i+1}}{\pi_i\delta^i}\left[1 - v'\left(\frac{y_b}{w_{i+2}}\right)\frac{1}{w_{i+2}}\right] - \frac{\pi_{i+1}\delta^{i+1}}{\pi_i\delta^i}\left[1 - v'\left(\frac{y_b}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$$

2-) $IC_{i+1,i} - IC_{i,i+1} - IC_{i+1,i+2}$

$$\left[1 - v'\left(\frac{y_b}{w_i}\right)\frac{1}{w_i}\right] = -\frac{\pi_{i+1}\delta^{i+1}-\mu_{i+1,i+2}}{\pi_i\delta^i}\left[1 - v'\left(\frac{y_b}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$$

3-) $IC_{i+1,i} - IC_{i,i+1} - IC_{i-1,i}$

$$\left[1 - v'\left(\frac{y_b}{w_i}\right)\frac{1}{w_i}\right] = \frac{\mu_{i-1,i}}{\pi_i\delta^i}\left[1 - v'\left(\frac{y_b}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right] - \frac{\pi_{i+1}\delta^{i+1}}{\pi_i\delta^i}\left[1 - v'\left(\frac{y_b}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$$

4-) $IC_{i+1,i} - IC_{i,i+1} - IC_{i,i-1}$

$$\left[1 - v'\left(\frac{y_b}{w_i}\right)\frac{1}{w_i}\right] = -\frac{\pi_{i+1}\delta^{i+1}}{\pi_i\delta^i+\mu_{i,i-1}}\left[1 - v'\left(\frac{y_b}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$$

5-) $IC_{i+1,i} - IC_{i,i+1} - IC_{i+1,i+2} - IC_{i-1,i}$ Maximax is a special case of this one.

$$\left[1-v'\left(\frac{y_b}{w_i}\right)\frac{1}{w_i}\right]=\frac{\mu_{i-1,i}}{\pi_i\delta^i}\left[1-v'\left(\frac{y_b}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]-\frac{\pi_{i+1}\delta^{i+1}+\mu_{i+1,i+2}}{\pi_i\delta^i}\left[1-v'\left(\frac{y_b}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$$

6-) $IC_{i+1,i}-IC_{i,i+1}-IC_{i+1,i+2}-IC_{i,i-1}$

$$\left[1-v'\left(\frac{y_b}{w_i}\right)\frac{1}{w_i}\right]=-\frac{\pi_{i+1}\delta^{i+1}+\mu_{i+1,i+2}}{\pi_i\delta^i+\mu_{i,i-1}}\left[1-v'\left(\frac{y_b}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$$

7-) $IC_{i+1,i}-IC_{i,i+1}-IC_{i+2,i+1}-IC_{i-1,i}$

$$\left[1-v'\left(\frac{y_b}{w_i}\right)\frac{1}{w_i}\right]=\frac{\mu_{i-1,i}}{\pi_i\delta^i}\left[1-v'\left(\frac{y_b}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]-\frac{\pi_{i+1}\delta^{i+1}}{\pi_i\delta^i}\left[1-v'\left(\frac{y_b}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]+\frac{\mu_{i+2,i+1}}{\pi_i\delta^i}\left[1-v'\left(\frac{y_b}{w_{i+2}}\right)\frac{1}{w_{i+2}}\right]$$

8-) $IC_{i+1,i}-IC_{i,i+1}-IC_{i+2,i+1}-IC_{i,i-1}$

Rawlsian is a special case of this one.

$$\left[1-v'\left(\frac{y_b}{w_i}\right)\frac{1}{w_i}\right]=\frac{\mu_{i+2,i+1}}{\pi_i\delta^i+\mu_{i,i-1}}\left[1-v'\left(\frac{y_b}{w_{i+2}}\right)\frac{1}{w_{i+2}}\right]-\frac{\pi_{i+1}\delta^{i+1}}{\pi_i\delta^i+\mu_{i,i-1}}\left[1-v'\left(\frac{y_b}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$$

9-) $IC_{i+1,i}-IC_{i,i+1}$

$$\left[1-v'\left(\frac{y_b}{w_i}\right)\frac{1}{w_i}\right]=-\frac{\pi_{i+1}\delta^{i+1}}{\pi_i\delta^i}\left[1-v'\left(\frac{y_b}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]$$

We showed cases 5 and 8 above in the text. So need the check other cases one by one.

1-) $IC_{i+1,i}-IC_{i,i+1}-IC_{i+2,i+1}$

Rewrite Rawlsian case as

$$\frac{\left[1-v'\left(\frac{y_b^R}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_b^R}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}=\frac{\mu_{i+2,i+1}}{\pi_i\delta^i+\mu_{i,i-1}}\frac{\left[1-v'\left(\frac{y_b^R}{w_{i+2}}\right)\frac{1}{w_{i+2}}\right]}{\left[1-v'\left(\frac{y_b^R}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}-\frac{\pi_{i+1}\delta^{i+1}}{\pi_i\delta^i+\mu_{i,i-1}}$$

$$=\frac{1-F_{i+1}}{1-F_{i-1}}\frac{\left[1-v'\frac{y_b^R}{w_{i+2}}\frac{1}{w_{i+2}}\right]}{\left[1-v'\left(\frac{y_b^R}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}$$

$IC_{i+1i}-IC_{ii+1}-IC_{i+2i+1}$ are binding, we have:

$$\frac{\left[1-v'\left(\frac{y_b^1}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_b^1}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}=\frac{\mu_{i+2,i+1}}{\pi_i\delta^i}\frac{\left[1-v'\left(\frac{y_b^1}{w_{i+2}}\right)\frac{1}{w_{i+2}}\right]}{\left[1-v'\left(\frac{y_b^1}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}-\frac{\pi_{i+1}\delta^{i+1}}{\pi_i\delta^i}$$

$$=\frac{\beta_{i+1}-F_{i+1}}{\pi_i\delta^i}\frac{\left[1-v'\left(\frac{y_b^1}{w_{i+2}}\right)\frac{1}{w_{i+2}}\right]}{\left[1-v'\left(\frac{y_b^1}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}-\frac{\pi_{i+1}\delta^{i+1}}{\pi_i\delta^i}$$

There are two possible cases for agents i and $i+1$. Both of them could be downward distorted or while agent i is upward distorted agent $i+1$ could be downward distorted. If agent i is upward distorted and $i+1$ is downward distorted we have;

$$\frac{\left[1-v'\left(\frac{y_b^1}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_b^1}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}<\left[1-v'\left(\frac{y_b^R}{w_i}\right)\frac{1}{w_i}\right]\left[1-v'\left(\frac{y_b^R}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]\text{ which implies } y_b^1>y_b^R.$$

If both agents are downward distorted then its better to rewrite the conditions as;

$$\frac{\left[1-v'\left(\frac{y_b^1}{w_{i+2}}\right)\frac{1}{w_{i+2}}\right]}{\left[1-v'\left(\frac{y_b^1}{w_i}\right)\frac{1}{w_i}\right]}=\frac{\pi_i\delta^i}{\mu_{i+2i+1}}+\frac{\pi_{i+1}\delta^{i+1}}{\mu_{i+2,i+1}}\frac{\left[1-v'\left(\frac{y_b^1}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^1}{w_i}\right)\frac{1}{w_i}\right]}=\frac{\pi_i\delta^i+\pi_{i+1}\delta^{i+1}*A}{\beta_{i+1}-F_{i+1}}$$ where $A>1$ and Rawlsian;

$$\frac{\left[1-v'\left(\frac{y_b^R}{w_{i+2}}\right)\frac{1}{w_{i+2}}\right]}{\left[1-v'\left(\frac{y_b^R}{w_i}\right)\frac{1}{w_i}\right]} = \frac{\pi_i\delta^i+\mu_{i,i-1}}{\mu_{i+2,i+1}} + \frac{\pi_{i+1}\delta^{i+1}}{\mu_{i+2,i+1}}\frac{\left[1-v'\left(\frac{y_b^R}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^R}{w_i}\right)\frac{1}{w_i}\right]} = \frac{1-F_{i-1}}{1-F_{i+1}}$$

if $\frac{\pi_i\delta^i+\pi_{i+1}\delta^{i+1}}{\beta_{i+1}-F_{i+1}} \geq \frac{1-F_{i-1}}{1-F_{i+1}}$ then we have $y_b^{'} \geq y_b^R$.

Suppose not, and we have $\frac{\pi_i\delta^i+\pi_{i+1}\delta^{i+1}}{\beta_{i+1}-F_{i+1}} < \frac{1-F_{i-1}}{1-F_{i+1}}$

This implies $\frac{\pi_i\delta^i+\pi_{i+1}\delta^{i+1}}{\beta_{i+1}-F_{i+1}} < \frac{\pi_i+...+\pi_N}{\pi_{i+2}+...+\pi_N}$

$[\pi_i\delta^i + \pi_{i+1}\delta^{i+1}][\pi_{i+2} + ... + \pi_N] < [\beta_{i+1} - F_{i+1}][\pi_i + ... + \pi_N]$

$0 < [\beta_{i-1} - F_{i+1}][\pi_i + ... + \pi_N] + [\pi_i\delta^i + \pi_{i+1}\delta^{i+1}][\pi_i + \pi_{i+1}]$

since $\beta_{i-1} = F_{i-1}$ rewrite as

$0 < [\pi_i + \pi_{i+1}][\pi_i\delta^i + \pi_{i+1}\delta^{i+1}] - [\pi_i + \pi_{i+1}][\pi_i + ... + \pi_N]$

$0 < [\pi_i + \pi_{i+1}][\beta_{i+1} - 1]$ which is impossible.

To compare with maximax case; we have;

$$\frac{\left[1-v'\left(\frac{y_b^1}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^1}{w_i}\right)\frac{1}{w_i}\right]} = \frac{\mu_{i+2,i+1}}{\pi_i\delta^i}\frac{\left[1-v'\left(\frac{y_b^1}{w_{i+2}}\right)\frac{1}{w_{i+2}}\right]}{\left[1-v'\left(\frac{y_b^1}{w_i}\right)\frac{1}{w_i}\right]} - \frac{\pi_i\delta^i}{\pi_{i+1}\delta^{i+1}}$$

and for maximax;

$$\frac{\left[1-v'\left(\frac{y_b^M}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^M}{w_i}\right)\frac{1}{w_i}\right]} = -\frac{\pi_i\delta^i}{\pi_{i+1}\delta^{i+1}+\mu_{i+1,i+2}} + \frac{\mu_{i-1,i}}{\pi_{i+1}\delta^{i+1}+\mu_{i+1,i+2}}\frac{\left[1-v'\left(\frac{y_b^M}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[1-v'\left(\frac{y_b^M}{w_i}\right)\frac{1}{w_i}\right]}$$

$$= \frac{F_{i-1}}{F_{i+1}}\frac{\left[1-v'\left(\frac{y_b^M}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[1-v'\left(\frac{y_b^M}{w_i}\right)\frac{1}{w_i}\right]}$$

Again we need to check possible distortion cases for agents i and $i+1$. If both of them downward distorted then result trivially holds as in maximax there is an upward distortion for both agents. If agent i is upward distorted and agent $i+1$ is downward distorted then we have

$$\frac{\left[1-v'\left(\frac{y_b^1}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^1}{w_i}\right)\frac{1}{w_i}\right]} < \frac{\left[1-v'\left(\frac{y_b^M}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^M}{w_i}\right)\frac{1}{w_i}\right]} \text{ which implies } y_b^M > y_b^1.$$

2-) $IC_{i+1,i} - IC_{i,i+1} - IC_{i+1,i+2}$

$$\frac{\left[1-v'\left(\frac{y_b^2}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_b^2}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} = -\frac{\pi_{i+1}\delta^{i+1}+\mu_{i+1,i+2}}{\pi_i\delta^i}$$

Since this expression is negative we have;

$$\frac{\left[1-v'\left(\frac{y_b^R}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_b^R}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} > \frac{\left[1-v'\left(\frac{y_b^2}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_b^2}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} \text{ which implies } y_b^2 > y_b^R.$$

Similarly for the maximax case, rewrite the condition as;

$$\frac{\left[1-v'\left(\frac{y_b^2}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^2}{w_i}\right)\frac{1}{w_i}\right]} = -\frac{\pi_i\delta^i}{\pi_{i+1}\delta^{i+1}+\mu_{i+1,i+2}}$$

again since it is a negative value, we have

$$\frac{\left[1-v'\left(\frac{y_b^M}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^M}{w_i}\right)\frac{1}{w_i}\right]} > \frac{\left[1-v'\left(\frac{y_b^2}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^2}{w_i}\right)\frac{1}{w_i}\right]} \text{ which implies } y_b^M > y_b^2.$$

3-) $IC_{i+1,i} - IC_{i,i+1} - IC_{i-1,i}$

$$\frac{\left[1-v'\left(\frac{y_b^3}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_b^3}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} = \frac{\mu_{i-1,i}}{\pi_i\delta^i}\frac{\left[1-v'\left(\frac{y_b^3}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[1-v'\left(\frac{y_b^3}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} - \frac{\pi_{i+1}\delta^{i+1}}{\pi_i\delta^i}$$

There are two possibilities. Both i and $i+1$ are upward distorted, or agent i is upward distorted and agent $i+1$ is downward distorted. If both of them are upward distorted then there is nothing to discuss as in Rawlsian case both of them are distorted downwards. If agent i is upward distorted and agent $i+1$ is downward distorted then we have;

$$\frac{\left[1-v'\left(\frac{y_b^R}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_b^R}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} > \frac{\left[1-v'\left(\frac{y_b^3}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_b^3}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} \text{ which implies } y_b^3 > y_b^R.$$

For Maximax case we have the following conditions;

$$\frac{\left[1-v'\left(\frac{y_b^M}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^M}{w_i}\right)\frac{1}{w_i}\right]} = -\frac{\pi_i\delta^i}{\pi_{i+1}\delta^{i+1}+\mu_{i+1,i+2}} + \frac{\mu_{i-1,i}}{\pi_{i+1}\delta^{i+1}+\mu_{i+1,i+2}}\frac{\left[1-v'\left(\frac{y_b^M}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[1-v'\left(\frac{y_b^M}{w_i}\right)\frac{1}{w_i}\right]}$$

$$= \frac{\mu_{i-1,i}}{\pi_{i+1}\delta^{i+1}+\mu_{i+1,i+2}}\frac{\left[1-v'\left(\frac{y_b^M}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[1-v'\left(\frac{y_b^M}{w_i}\right)\frac{1}{w_i}\right]}$$

and

$$\frac{\left[1-v'\left(\frac{y_b^3}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^3}{w_i}\right)\frac{1}{w_i}\right]} = \frac{\mu_{i-1,i}}{\pi_{i+1}\delta^{i+1}}\frac{\left[1-v'\left(\frac{y_b^3}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[1-v'\left(\frac{y_b^3}{w_i}\right)\frac{1}{w_i}\right]} - \frac{\pi_i\delta^i}{\pi_{i+1}\delta^{i+1}}$$

we need to check two possible cases. If agent i is upward and agent $i+1$ is downward distorted than we have;

$$\frac{\left[1-v'\left(\frac{y_b^3}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^3}{w_i}\right)\frac{1}{w_i}\right]} < \frac{\left[1-v'\left(\frac{y_b^M}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^M}{w_i}\right)\frac{1}{w_i}\right]} \text{ which implies } y_b^M > y_b^3.$$

If both of them are upward distorted then rewrite the conditions as;

$$\frac{\left[1-v'\left(\frac{y_b^M}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[1-v'\left(\frac{y_b^M}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} = \frac{F_{i+1}}{F_{i-1}}$$

and

$$\frac{\left[1-v'\left(\frac{y_b^3}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[1-v'\left(\frac{y_b^3}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} = \frac{\pi_{i+1}\delta^{i+1}}{\mu_{i-1,i}} + \frac{\pi_i\delta^i}{\mu_{i-1,i}}\frac{\left[1-v'\left(\frac{y_b^3}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_b^3}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} = \frac{\pi_{i+1}\delta^{i+1}+\pi_i\delta^i*A}{\mu_{i-1,i}} \text{ where } A > 1.$$

Suppose$A = 1$, then we should have $\frac{\pi_{i+1}\delta^{i+1}+\pi_i\delta^i}{F_{i-1}-\beta_{i-1}} > \frac{F_{i+1}}{F_{i-1}}$ for the result.

Rewriting and manipulations give that;

$F_{i-1}(\pi_{i+1}\delta^{i+1} + \pi_i\delta^i) > F_{i+1}F_{i-1} - F_{i+1}\beta_{i-1}$

Since we have $F_{i+1} = \beta_{i+1}$, it is equal to

$F_{i-1}(\pi_{i+1}\delta^{i+1} + \pi_i\delta^i) > \beta_{i+1}F_{i-1} - F_{i+1}\beta_{i-1}$ which implies

$0 > \beta_{i-1}[F_{i-1} - F_{i+1}]$, and this condition always holds as $F_{i+1} > F_{i-1}$.

4-) $IC_{i+1,i} - IC_{i,i+1} - IC_{i,i-1}$

From FOC we have;

$$\frac{\left[1-v'\left(\frac{y_b^4}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_b^4}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} = -\frac{\pi_{i+1}\delta^{i+1}}{\pi_i\delta^i+\mu_{i,i-1}}$$

so agent i is upward distorted and agent $i+1$ is downward distorted.

For Rawlsian we have

$$\frac{\left[1-v'\left(\frac{y_b^R}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_b^R}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} = \frac{\mu_{i+2,i+1}}{\pi_i\delta^i+\mu_{i,i-1}}\frac{\left[1-v'\left(\frac{y_b^R}{w_{i+2}}\right)\frac{1}{w_{i+2}}\right]}{\left[1-v'\left(\frac{y_b^R}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} - \frac{\pi_{i+1}\delta^{i+1}}{\pi_i\delta^i+\mu_{i,i-1}} = \frac{1-F_{i+1}}{1-F_{i-1}}\frac{\left[1-v'\left(\frac{y_b^R}{w_{i+2}}\right)\frac{1}{w_{i+2}}\right]}{\left[1-v'\left(\frac{y_b^R}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}$$

$$\frac{\left[1-v'\left(\frac{y_b^R}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_b^R}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} > \frac{\left[1-v'\left(\frac{y_b^4}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_b^4}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}$$ which implies $y_b^4 > y_b^R$.

For Maximax we have;

$$\frac{\left[1-v'\left(\frac{y_b^4}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^4}{w_i}\right)\frac{1}{w_i}\right]} = -\frac{\pi_i\delta^i+\mu_{i,i-1}}{\pi_{i+1}\delta^{i+1}}$$

and Maximax condition;

$$\frac{\left[1-v'\left(\frac{y_b^M}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^M}{w_i}\right)\frac{1}{w_i}\right]} = -\frac{\pi_i\delta^i}{\pi_{i+1}\delta^{i+1}+\mu_{i+1,i+2}} + \frac{\mu_{i-1,i}}{\pi_{i+1}\delta^{i+1}+\mu_{i+1,i+2}}\frac{\left[1-v'\left(\frac{y_b^M}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[1-v'\left(\frac{y_b^M}{w_i}\right)\frac{1}{w_i}\right]}$$

$$= \frac{F_{i-1}}{F_{i+1}}\frac{\left[1-v'\left(\frac{y_b^M}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[1-v'\left(\frac{y_b^M}{w_i}\right)\frac{1}{w_i}\right]}$$

$$\frac{\left[1-v'\left(\frac{y_b^M}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\frac{y_b^M}{w_i}\frac{1}{w_i}\right]} > \frac{\left[1-v'\left(\frac{y_b^4}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^4}{w_i}\right)\frac{1}{w_i}\right]}$$ which implies that $y_b^M > y_b^4$.

6-) $IC_{i+1,i} - IC_{i,i+1} - IC_{i+1,i+2} - IC_{i,i-1}$

$$\frac{\left[1-v'\left(\frac{y_b^6}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_b^6}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} = -\frac{\pi_{i+1}\delta^{i+1}+\mu_{i+1,i+2}}{\pi_i\delta^i+\mu_{i,i-1}}$$

since it is a negative term we have;

$$\frac{\left[1-v'\left(\frac{y_b^R}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_b^R}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} > \frac{\left[1-v'\left(\frac{y_b^6}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_b^6}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}$$ which implies that $y_b^6 > y_b^R$.

Similarly for the Maximax case we have;

$$\frac{\left[1-v'\left(\frac{y_b^M}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^M}{w_i}\right)\frac{1}{w_i}\right]} > \frac{\left[1-v'\left(\frac{y_b^6}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^6}{w_i}\right)\frac{1}{w_i}\right]} \text{ which implies that } y_b^M > y_b^6.$$

7-) $IC_{i+1,i} - IC_{i,i+1} - IC_{i+2,i+1} - IC_{i-1,i}$

$$\frac{\left[1-v'\left(\frac{y_b^7}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_b^7}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} = \frac{\mu_{i+2,i+1}}{\pi_i\delta^i}\frac{\left[1-v'\left(\frac{y_b^7}{w_{i+2}}\right)\frac{1}{w_{i+2}}\right]}{\left[1-v'\left(\frac{y_b^7}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} - \frac{\pi_{i+1}\delta^{i+1}}{\pi_i\delta^i} + \frac{\mu_{i-1,i}}{\pi_i\delta^i}\frac{\left[1-v'\left(\frac{y_b^7}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[1-v'\left(\frac{y_b^7}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}$$

In this case we have three possibilities for agents i and $i+1$. Both of them could be downward distorted, or both of them could be upward distorted. Third possibility is that agent i is upward distorted and agent $i+1$ is downward distorted. If both of the are upward distorted there is nothing to discuss. If agent i is upward distorted and agent $i+1$ is downward distorted we have;

$$\frac{\left[1-v'\left(\frac{y_b^R}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_b^R}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} > \frac{\left[1-v'\left(\frac{y_b^7}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_b^7}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} \text{ which implies } y_b^R < y_b^7.$$

If both of the agents are downward distorted then rewrite the conditions as follows;

$$\frac{\left[1-v'\left(\frac{y_b^R}{w_{i+2}}\right)\frac{1}{w_{i+2}}\right]}{\left[1-v'\left(\frac{y_b^R}{w_i}\right)\frac{1}{w_i}\right]} = \frac{1-F_{i-1}}{1-F_{i+1}} \text{ which is positive for sure.}$$

$$\frac{\left[1-v'\left(\frac{y_b^7}{w_{i+2}}\right)\frac{1}{w_{i+2}}\right]}{\left[1-v'\left(\frac{y_b^7}{w_i}\right)\frac{1}{w_i}\right]} = \frac{\pi_i\delta^i}{\mu_{i+2,i+1}} + \frac{\pi_{i+1}\delta^{i+1}}{\mu_{i+2,i+1}}\frac{\left[1-v'\left(\frac{y_b^7}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^7}{w_i}\right)\frac{1}{w_i}\right]} - \frac{\mu_{i-1,i}}{\mu_{i+2,i+1}}\frac{\left[1-v'\left(\frac{y_b^7}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[1-v'\left(\frac{y_b^7}{w_i}\right)\frac{1}{w_i}\right]}$$

we should have the following condition for the result;

$$\frac{\pi_i\delta^i}{\mu_{i+2,i+1}} + \frac{\pi_{i+1}\delta^{i+1}}{\mu_{i+2,i+1}}\frac{\left[1-v'\left(\frac{y_b^7}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^7}{w_i}\right)\frac{1}{w_i}\right]} - \frac{\mu_{i-1,i}}{\mu_{i+2,i+1}}\frac{\left[1-v'\left(\frac{y_b^7}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[1-v'\left(\frac{y_b^7}{w_i}\right)\frac{1}{w_i}\right]} > \frac{1-F_{i-1}}{1-F_{i+1}}$$

$$\frac{\pi_i\delta^i}{\mu_{i+2,i+1}} + \frac{\pi_{i+1}\delta^{i+1}}{\mu_{i+2,i+1}}A - \frac{\mu_{i-1,i}}{\mu_{i+2,i+1}}\frac{\left[1-v'\left(\frac{y_b^7}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[1-v'\left(\frac{y_b^7}{w_i}\right)\frac{1}{w_i}\right]} > \frac{1-F_{i-1}}{1-F_{i+1}}$$

$A > 1$ and the last term on the left hand side is positive. So suppose $A = 1$ and we do not have the last term, then;

$\frac{\pi_i\delta^i+\pi_{i+1}\delta^{i+1}}{\beta_{i+1}-F_{i+1}} > \frac{1-F_{i-1}}{1-F_{i+1}}$ suppose not and we have $\frac{\pi_i\delta^i+\pi_{i+1}\delta^{i+1}}{\beta_{i+1}-F_{i+1}} < \frac{\pi_i+...+\pi_N}{\pi_{i+2}+...+\pi_N}$

$$[\pi_i\delta^i + \pi_{i+1}\delta^{i+1}][\pi_{i+2} + ... + \pi_N] < [\beta_{i+1} - F_{i+1}][\pi_i + ... + \pi_N]$$

$$0 < \pi_1\delta[\pi_i + ... + \pi_N] + \pi_{i-1}\delta^{i-1}[\pi_i + ... + \pi_N] + \pi_i\delta^i[\pi_i + \pi_{i+1}]$$
$$+\pi_{i+1}\delta^{i+1}[\pi_i + \pi_{i+1}] - F_{i+1}[\pi_i + ... + \pi_N]$$

$$0 < [\pi_1\delta^1 + \pi_{i-1}\delta^{i-1}][\pi_i + ... + \pi_N] + [\pi_i\delta^i + \pi_{i+1}\delta^{i+1}][\pi_i + \pi_{i+1}] - F_{i+1}[\pi_i + ... + \pi_N]$$

$$0 < [\pi_i + \pi_{i+1}][\beta_{i+1} - 1] + [\pi_{i+2} + ... + \pi_N][\beta_{i-1} - F_{i-1}]$$

which is impossible since $\beta_{i+1} < 1$ and $\beta_{i-1} < F_{i-1}$.

For Maximax case we have;

$$\frac{\left[1-v'\left(\frac{y_b^M}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^M}{w_i}\right)\frac{1}{w_i}\right]} = -\frac{\pi_i\delta^i}{\pi_{i+1}\delta^{i+1}+\mu_{i+1,i+2}} + \frac{\mu_{i-1,i}}{\pi_{i+1}\delta^{i+1}+\mu_{i+1,i+2}}\frac{\left[1-v'\left(\frac{y_b^M}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[1-v'\left(\frac{y_b^M}{w_i}\right)\frac{1}{w_i}\right]}$$

$$= \frac{F_{i-1}}{F_{i+1}}\frac{\left[1-v'\left(\frac{y_b^M}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[1-v'\left(\frac{y_b^M}{w_i}\right)\frac{1}{w_i}\right]}$$

$$\frac{\left[1-v'\left(\frac{y_b^7}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^7}{w_i}\right)\frac{1}{w_i}\right]} = \frac{\mu_{i-1,i}}{\pi_{i+1}\delta^{i+1}}\frac{\left[1-v'\left(\frac{y_b^7}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[1-v'\left(\frac{y_b^7}{w_i}\right)\frac{1}{w_i}\right]} + \frac{\mu_{i+2,i+1}}{\pi_{i+1}\delta^{i+1}}\frac{\left[1-v'\left(\frac{y_b^7}{w_{i+2}}\right)\frac{1}{w_{i+2}}\right]}{\left[1-v'\left(\frac{y_b^7}{w_i}\right)\frac{1}{w_i}\right]} - \frac{\pi_i\delta^i}{\pi_{i+1}\delta^{i+1}}$$

again we have 3 possibilities. If both of them are downward distorted result holds trivially as in Maximax case both of them are distorted upwards. If agent i is upward and agent $i+1$ is downward distorted then we have;

$$\frac{\left[1-v'\left(\frac{y_b^M}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^M}{w_i}\right)\frac{1}{w_i}\right]} > \frac{\left[1-v'\left(\frac{y_b^7}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^7}{w_i}\right)\frac{1}{w_i}\right]}$$ which implies $y_b^M > y_b^7$.

If both agents i and $i+1$ is upward distorted, then rewrite the equations as;

$$\frac{\left[1-v'\left(\frac{y_b^M}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[1-v'\left(\frac{y_b^M}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} = \frac{F_{i+1}}{F_{i-1}}$$

$$\frac{\left[1-v'\left(\frac{y_b^7}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[1-v'\left(\frac{y_b^7}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} = \frac{\pi_i\delta^i}{\mu_{i-1,i}}\frac{\left[1-v'\left(\frac{y_b^7}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_b^7}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} + \frac{\pi_{i+1}\delta^{i+1}}{\mu_{i-1,i}} - \frac{\mu_{i+2,i+1}}{\mu_{i-1,i}}\frac{\left[1-v'\left(\frac{y_b^7}{w_{i+2}}\right)\frac{1}{w_{i+2}}\right]}{\left[1-v'\left(\frac{y_b^7}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}$$

we should have; $\frac{\left[1-v'\left(\frac{y_b^7}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[1-v'\left(\frac{y_b^7}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} > \frac{\left[1-v'\left(\frac{y_b^M}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[1-v'\left(\frac{y_b^M}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}$

$$\frac{\pi_i\delta^i}{\mu_{i-1,i}}\frac{\left[1-v'\left(\frac{y_b}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_b}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} + \frac{\pi_{i+1}\delta^{i+1}}{\mu_{i-1,i}} - \frac{\mu_{i+2,i+1}}{\mu_{i-1,i}}\frac{\left[1-v'\left(\frac{y_b}{w_{i+2}}\right)\frac{1}{w_{i+2}}\right]}{\left[1-v'\left(\frac{y_b}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} > \frac{F_{i+1}}{F_{i-1}}$$

$$\frac{\pi_i\delta^i}{\mu_{i-1,i}}A + \frac{\pi_{i+1}\delta^{i+1}}{\mu_{i-1,i}} - \frac{\mu_{i+2,i+1}}{\mu_{i-1,i}}\frac{\left[1-v'\left(\frac{y_b}{w_{i+2}}\right)\frac{1}{w_{i+2}}\right]}{\left[1-v'\left(\frac{y_b}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} > \frac{F_{i+1}}{F_{i-1}}$$

$A > 1$and the last term on the left hand side is positive. So suppose $A = 1$ and we do not have the last term, then we have;

$$\frac{\pi_i\delta^i+\pi_{i+1}\delta^{i+1}}{F_{i-1}-\beta_{i-1}} > \frac{F_{i+1}}{F_{i-1}}$$

$F_{i-1}(\pi_i\delta^i + \pi_{i+1}\delta^{i+1}) > F_{i+1}(F_{i-1} - \beta_{i-1})$ Since $\beta_{i+1} > F_{i+1}$ rewrite as

$F_{i-1}(\pi_i\delta^i + \pi_{i+1}\delta^{i+1}) > \beta_{i+1}(F_{i-1} - \beta_{i-1})$

$0 > \beta_{i-1}(F_{i-1} - \beta_{i+1})$ which holds for sure.

9-) $IC_{i+1,i} - IC_{i,i+1}$

$$\frac{\left[1-v'\left(\frac{y_b^9}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_b^9}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} = -\frac{\pi_{i+1}\delta^{i+1}}{\pi_i\delta^i}$$

since it is negative we have;

$$\frac{\left[1-v'\left(\frac{y_b^R}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_b^R}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} > \frac{\left[1-v'\left(\frac{y_b^9}{w_i}\right)\frac{1}{w_i}\right]}{\left[1-v'\left(\frac{y_b^9}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} \text{ which implies that } y_b^9 > y_b^R.$$

Similarly for the Maximax case we have;

$$\frac{\left[1-v'\left(\frac{y_b^M}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^M}{w_i}\right)\frac{1}{w_i}\right]} > \frac{\left[1-v'\left(\frac{y_b^9}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[1-v'\left(\frac{y_b^9}{w_i}\right)\frac{1}{w_i}\right]} \text{ which implies that } y_b^M > y_b^9.$$

Proof of Proposition 12:

The condition comes from the relation of social weight δ^i with the marginal tax rates $T_i^{'}$ and $T_{i-1}^{'}$. From the proposition we have;

$$\delta^i = \frac{\left[1-v'\left(\frac{y_i}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]}{\left[v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i}-v'\left(\frac{y_i}{w_{i+1}}\right)\frac{1}{w_{i+1}}\right]} - \frac{\pi_{i-1}}{\pi_i}\left\{\frac{\left[1-v'\left(\frac{y_{i-1}}{w_{i-1}}\right)\frac{1}{w_{i-1}}\right]}{\left[v'\left(\frac{y_{i-1}}{w_{i-1}}\right)\frac{1}{w_{i-1}}-v'\left(\frac{y_{i-1}}{w_i}\right)\frac{1}{w_i}\right]}\right\}$$

we also have;

$$v'\left(\frac{y_i}{w_i}\right)\frac{1}{w_i} = 1 - T_i^{'}(y_i)$$
$$y_i = w_i\left[v^{'-1}\left[w_i(1-T_i^{'}(y_i))\right]\right]$$
$$y_{i-1} = w_{i-1}\left[v^{'-1}\left[w_{i-1}(1-T_{i-1}^{'}(y_{i-1}))\right]\right]$$

Plugging the terms yields;

$$\delta^i = \frac{\left[1-v'\left\{\left[v^{'-1}\left[w_i(1-T_i^{'}(y_i))\right]\right]\frac{w_i}{w_{i+1}}\right\}\frac{1}{w_{i+1}}\right]}{\left[1-T_i^{'}(y_i)-v'\left\{\left[v^{'-1}\left[w_i(1-T_i^{'}(y_i))\right]\right]\frac{w_i}{w_{i+1}}\right\}\frac{1}{w_{i+1}}\right]}$$
$$-\frac{\pi_{i-1}}{\pi_i}\left\{\frac{T_{i-1}^{'}(y_{i-1})}{\left[1-T_{i-1}^{'}(y_{i-1})-v'\left\{\left[v^{'-1}\left[w_{i-1}(1-T_{i-1}^{'}(y_{i-1}))\right]\right]\frac{w_{i-1}}{w_i}\right\}\frac{1}{w_i}\right]}\right\}$$