



TAMPERE ECONOMIC WORKING PAPERS

ON THE OPTIMAL LIFETIME REDISTRIBUTION AND EQUALITY OF OPPORTUNITIES

Terhi Ravaska
Sanna Tenhunen
Matti Tuomala

Working Paper 100
February 2016

SCHOOL OF MANAGEMENT
FI-33014 UNIVERSITY OF TAMPERE, FINLAND

ISSN 1458-1191
ISBN 978-952-03-0071-5

On the optimal lifetime redistribution and equality of opportunities

Terhi Ravaska¹, Sanna Tenhunen², Matti Tuomala³

¹School of Management, University of Tampere, Finland (terhi.ravaska@uta.fi)

²Finnish Centre for Pensions, Finland

³School of Management, University of Tampere, Finland

February 2016

Abstract

In this paper we examine optimal lifetime redistribution policy in equality of opportunity framework (EoO). The dynamic complications are avoided by studying the optimal redistribution within a cohort. We characterise optimal redistribution policy when there are differences not only in individuals' productivities but also in their tastes towards the timing of consumption, i.e. some are patient and others impatient in consumption over the life-cycle and this preference together with productivity is non-observable to government. In the spirit of Roemer (1998) and Van de Gaer (1993) our approach applies a compromise between the principle of compensation and the principle of responsibility. We derive analytical expressions which describe the optimal distortion (upward or downward) in saving.

As the multidimensional problems become very complicated, to gain a better understanding, we also examine numerically the properties of an optimal lifetime redistribution policy. The numerical results show the implications of different structure of economy to optimal taxation policy. We find support for non-linear tax/pension program in which some types of individuals are taxed while some are subsidized. Numerical simulations show quite big differences in terms of the levels of marginal tax rates between utilitarian and EoO-cases, indicating that the optimal income taxation results are sensitive for the choice of social planner's goals. There are larger differences between types in working careers under equality of opportunity framework than in standard utilitarian case.

Key words: Optimal taxation, lifetime redistribution, heterogeneous time preferences, equality of opportunity

JEL classification: H21, H55, D71

Acknowledgments: OP Group Research Foundation and University of Tampere are gratefully acknowledged for financial support to Ravaska. Also funding from the Strategic Research Council (SRC) at the Academy of Finland (project no. 293120, "Work, Inequality, and Public Policy") is gratefully acknowledged. The authors thank seminar participants and commentators in IIPF 2015 and Nordic Workshop in Public Economics 2015.

Introduction

The assumption that differences in lifetime earnings can be completely explained by time preferences, is obviously an unrealistic one, just as it would be a simplification of reality to explain them as a result of differences in ability. A more realistic model should take both into account. There are well-known technical difficulties related to incentive constraints to study multidimensional optimal tax problems including both of the elements. Another problem is how to incorporate heterogeneous preferences into social welfare function (SWF) in analysing optimal tax policy. Social welfare functions can be quite straightforwardly parameterized when individuals have identical preferences represented by a utility function. In the case of one-dimensional population, there are two possible ways to observe differences in economic outcomes. Namely, if people have identical preferences but they differ in abilities, we are back in the Mirrlees model. In the opposite case the diversity of preferences is the sole source of inequality.

In the case of diversity in preferences some people would, however, say that if individuals have the same opportunities while their choices may differ, there is no ethical basis for redistributive taxation. According to this view individuals should be compensated for circumstances which they have no control of, such as their family background or disability at birth. On the other hand, individuals should be held responsible for circumstances which they can control, such as how many hours or weeks they work. Hence, no redistribution should take place based on such choices. The former is referred as the *principle of compensation* and the latter the *principle of responsibility* (see Fleurbaey (1994), Roemer (1998)). By the principle of compensation, it is fair to redistribute from high ability to low ability individuals. By the principle of responsibility, it is unfair to redistribute from the consumption lovers toward the leisure lovers. In the one dimensional population, those principles are easy to apply. For example, if individuals differ only according to their earnings ability (wage rate) and not in their preferences, then the principle of compensation reduces to a maximin criterion whereby the tax and transfer system should provide as much compensation as possible to the worst off people. If individuals differ solely in preferences, the principle of responsibility calls for no redistribution at all because everybody has the same opportunities. It would be unfair to redistribute based on tastes. The standard welfarist approach can obtain this result only in the case where social marginal utilities of net income are the same across individuals (absent transfers).

In multidimensional world the problem of choosing different utility functions for representing non-identical preferences is more complex. If individuals have different preferences, it isn't clear how to weight their utilities in a social welfare function. It can be argued that the fundamental distinction is not so much between earning abilities and preferences but between those factors which are beyond individual's control, and those which are purely a matter of individual choice. That means redistribution policies should aim to eliminate disparities in those matters which are beyond individual control, but should be neutral about those matters which are within their control. How to apply these two principles? There is a fundamental conflict between these principles. Namely, even in the world of perfect information with lump sum redistribution tools the government cannot generally satisfy these two principles at the same time.

This paper studies the optimal lifetime redistribution policy within a cohort with heterogeneity in earnings ability and preferences. The heterogeneity in preferences arise because tastes towards the timing of life-cycle consumption differs between individuals; some individuals are more present-oriented in consumption and so save less for the retirement period. These preferences are parametrized with different discount rates meaning that the more present-oriented consumer have a higher discount rate and thus a lower discount factor in his utility function. The differences in preferences differ from the myopic consumers because myopic consumers *ex-post* prefer saving more in the earlier periods and this justifies government interventions. However, in our paper there are true variation in tastes towards timing of consumption.

The differences in tastes raises the difficulty of how to choose the appropriate social welfare function. One way is to assume a paternalistic government as with myopic consumers (see Cremer et al. (2009), Tenhunen & Tuomala (2013)). With genuine differences in preferences there are some recent contributions that incorporate heterogeneous time preferences into optimal tax analysis while remaining agnostic about the appropriate cardinalization (see Cremer et al. (2009) Tenhunen & Tuomala (2010,2013)). In the spirit of Roemer (1998) and Van de Gaer (1993)¹our approach applies a compromise between the principle of compensation and the principle of responsibility. For individuals with same discount rates but different wage rates, the maximin criterion is applied. Thus we have a social ordering over each discount rate group. Then we aggregate over discount groups so that the minimum utility levels for different discount groups are averaged. The least well off of each preference groups are added together. In other words a zero aversion of inequality can be applied along the dimension of responsibility (in our case time preference) whereas a high aversion to inequality is acceptable along the dimension of circumstances (in our case skill).

Our model consists two periods, where individuals work only during the first period and decide how much to save for the second period. Our focus is on the distortions in savings decisions. The paper continues the research done by Tenhunen & Tuomala (2010,2013)² by introducing the framework of equality of opportunity. The equality of opportunity framework takes into account the principles of compensation and responsibility as noted above. Since the aim is to model an economy which multidimensional heterogeneity, the analytical results do no more reveal the signs of the distortions. For this reason numerical methods are used. Our results can also be interpreted in absence of private savings. In this case the second period consumption is publicly provided pension and thus we can extend our analysis into studying the optimal retirement plans in our modelling economy.

The structure of the paper is the following. In the first section we present the benchmark model where the time preference and ability are perfectly correlated. Then in section 3 we

¹Bossert (1995) and Fleurbaey (1994) have studied the idea of compensating inequalities due to circumstances only, while remaining other inequalities untouched.

²Tenhunen & Tuomala (2010) studied optimal life-time redistribution in 4-types setting where government objective is either utilitarian or paternalistic and consumer preferences are approximated with Cobb-Douglas utility function. Tenhunen & Tuomala (2013) studied how habit formation affects the optimal tax and pension scheme under heterogenous preferences.

extend the model to include three types, first by pooling the low ability types into one time preference group and in another case by pooling the high ability types into one time preference group. In section 4 we include all the four types in the model. Analytically the direction of distortions cannot be determined so in the end of each section we show with numerical simulations what kind of distortions occur for each type. Section 5 concludes.

2. A benchmark model

2.1 Two types with a positive correlation between skill and discount factor

Unlike in the original Mirrlees model, we assume that individuals differ not only in productivity but also in time preference³. As a benchmark we use a simple two-type model, similar to the much used two-type model first introduced by Stern (1982) and Stiglitz (1982). Each individual has a skill level reflecting his wage rate, denoted by n , and differences in time preferences are represented by a discount factor, denoted δ . We denote low skill and low discount factor by the superscript L and high skill and high discount factor by the superscript H. The assumption of positive correlation implies that $\delta^L < \delta^H$. The proportion of individuals of type i in the population is N^i , with $\sum N^i = 1$.

As is well known, due to Atkinson & Stiglitz (1976) results, under a mild separability assumption income taxation does not need be supplemented by other taxes. Saez (2002) argues that the Atkinson-Stiglitz result of commodity taxes holds only when each individual has identical discount rates. He argues that individuals with higher earnings save relatively more which suggests that high-skilled individuals are more likely to have higher discount factors and thus there is role for taxing savings beside labour taxation. In this case, discount factor is positively correlated with productivity level. Diamond & Spinnewijn (2011) study heterogenous discount factors and capital taxation and show that perfect correlation between ability and discount factor isn't necessary required in order to derive the result of positive capital taxation for high-skilled type. However, this result is more robust if there is positive correlation between skill and discount preference.

In our benchmark model, the assumptions over preferences imply that high-skilled people discount the future less heavily and thus have higher savings rates. Given this, the statement that high-productivity people are more patient follows from the empirical correlation between savings rates and earnings. In Figure 1, based on Consumer Surveys in Finland over the period of 1976-2012, we see that there is a strong correlation between saving rates and income.

³Sandmo (1993) considers a case where people differ in preferences, but are endowed with the same resources. Tarkiainen & Tuomala (1999, 2007) also consider a continuum of taxpayers simultaneously distributed by skill and preferences for leisure and income.

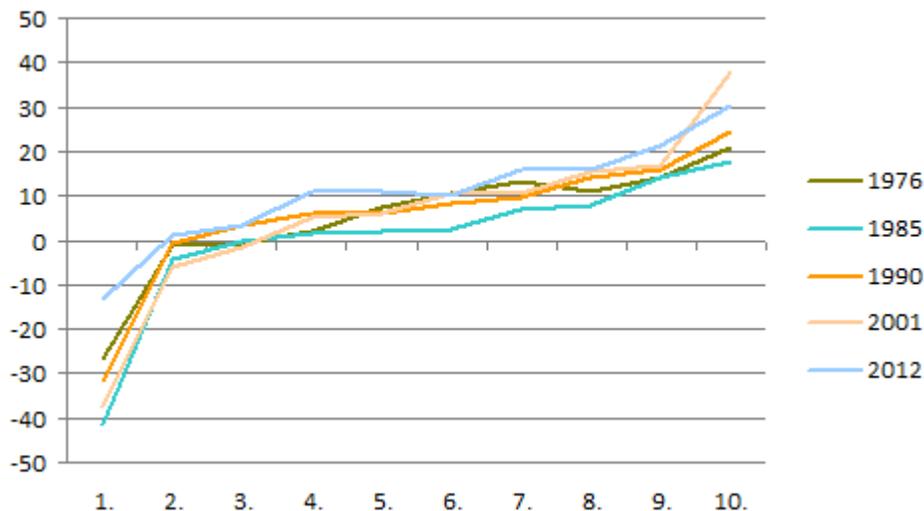


Figure 1: Average savings rates in different income deciles. Source: Riihela et al. (2015)

In light of this empirical evidence we take as a starting point a separable utility with positive correlation between discount factor and productivity⁴.

The life-time utility of an individual of type i is additive in the following way:

$$U^i = u(c^i) + \delta^i v(x^i) + \psi(1 - y^i), \quad (1)$$

where c and x denote consumption when young and old respectively and y is labour supply when young. Utility function is increasing in c and x and decreasing in y and it is strictly concave, i.e. $u', v', \psi' > 0$, $u'', v'', \psi'' < 0$. We also assume that all goods are normal.

To introduce returns to capital and the possible taxation thereof, it is necessary to consider a two-period model. Individuals are free to divide their first period income between consumption, c , and savings, s . Each unit of savings yields an additional $1 + \theta$ units of consumption in the second period after-tax income, x . As a further simplification we assume that the return on savings, θ , is fixed, which may be justified by assuming that we consider a small open economy facing a world capital market. Consumption in each period is given by $c^i = n^i y^i - T(n^i y^i) - s^i$ and $x^i = (1 + \theta)s^i$, $i = L, H$.

The government wishes to design a lifetime tax system that may redistribute income between individuals in the same cohort. There is asymmetric information in the sense that the tax authority is informed neither about individual skill levels, labour supply nor discount rates. It can only observe before-tax income, ny . In this setting, where tax on both earnings and savings income are available, we examine whether or not savings ought to be taxed. The separability assumption makes it possible to isolate the significance of variations in time preferences.

⁴Alternatively the same outcome could be reached by assuming homothetic preferences and linear Engel curves.

Assume that the government controls c^i , x^i and y^i directly. Alternatively, if we assume that there are no private savings, we have a model of labour income taxation in the first period and public provision of pension in the second period. In the case with perfect correlation between skills and discount factors Romer and Van de Gear approaches are equivalent with the maximin social welfare function. Here the government maximises the welfare of the worst-off group:

$$U^L = [u(c^L) + \delta^L v(x^L) + \psi(1 - y^L)], \quad (2)$$

subject to the revenue constraint

$$\sum N^i (n^i y^i - c^i - r x^i) = R, \quad (3)$$

where $r = \frac{1}{1+\theta}$ and the self-selection constraint⁵

$$u(c^H) + \delta^H v(x^H) + \psi(1 - y^H) \geq u(c^L) + \delta^H v(x^L) + \psi(1 - \frac{n^L}{n^H} y^L). \quad (4)$$

Multipliers λ and μ are attached respectively to the budget constraint and the self-selection constraint. The Lagrange function of the optimization problem is

$$\begin{aligned} \mathcal{L} = & N^L [u(c^L) + \delta^L v(x^L) + \psi(1 - y^L)] + \lambda [\sum N^i (n^i y^i - c^i - r x^i) - R] \\ & + \mu^{HL} [u(c^H) + \delta^H v(x^H) + \psi(1 - y^H) - u(c^L) - \delta^H v(x^L) - \psi(1 - \frac{n^L}{n^H} y^L)]. \end{aligned} \quad (5)$$

Our main interest is in the marginal taxation of savings⁶. For this purpose the first order conditions are written in the form $(\frac{u_c}{v_x})^i = \frac{\delta^i}{r}(1 - d^i)$, where the left hand side is individual i 's marginal rate of substitution between consumption in the first and in the second period and d^i is the distortion. A positive (negative) d^i implies that type i should have an implicit tax (subsidy) on savings. It is useful to define a relative difference in discount factors as $\Delta^{ij} \equiv \frac{\delta^i - \delta^j}{\delta^j}$ for any pair of discount factors. The first order conditions (presented in Appendix A equations A.1-A.6) imply that

$$\begin{aligned} d^L &= (\varphi^1 - 1)\Delta^{HL} \\ d^H &= 0, \end{aligned} \quad (6)$$

⁵The direction of the binding self-selection constraint is assumed to be, following the tradition in the one-dimensional two-type model, from high-skilled individual towards low-skilled individual. This pattern is also confirmed in the numerical simulations.

⁶In the numerical solution we also consider the marginal labour income tax rates. As has become conventional in the literature we may interpret the marginal rate of substitution between gross and net income as one minus the marginal income tax, $\frac{\psi'(\frac{ny}{n})}{nu_c} = 1 - T'(ny)$, which would be equivalent to the characterisation of the labour supply of an agent facing an income tax function $T'(ny)$. The marginal labour income tax rates satisfy the usual properties; $T(n^L y^L) > 0$ and $T(n^H y^H) = 0$.

where $\varphi^1 = \frac{N^L}{N^L - \mu^{HL}}$. The returns to savings of type i should not be taxed when d^i is zero. As $d^H = 0$, the optimal implicit marginal tax rate for the high-skill type is zero. When we assume, empirically plausibly, that $\delta^H > \delta^L$, we have $d^L > 0$ implying implicit taxation of savings for the low-skilled type. This is the same result as in utilitarian case by Diamond (2003).

As a result of the two-dimensional heterogeneity, a tax on capital income is an effective way to relax an otherwise binding self-selection constraint. This is because even under separability the mimicker and the individual mimicked do not save the same amount. A high-skilled individual choosing to mimic low-skilled type values savings more than a low-skilled individual, since discounting of the future is less for the potential mimicker. Thus, taxing savings relaxes the self-selection constraint. Or put in another way: distortions generate second-order efficiency costs but first-order redistributive benefits.

2.2 Numerical simulations

Adding multidimensionality to a constrained optimization problem causes that the directions of the distortions cannot be determined from the analytical results. To gain information on the properties of the optimal redistribution policy we rely on numerical simulations⁷. These simulations reveal the binding incentive-compatibility constraints and so we can determine those allocations, that make the agents to reveal their true characteristics. However, solving multi-dimensional optimal tax problems numerically faces problems, especially when the number of the incentive-compatibility constraints increase.

In the numerical examples we assume the following separable form of utility function: $U^i = -\frac{1}{c^i} - \delta^i \frac{1}{x^i} - \frac{1}{1-y^i}$ (CES) and choose the following parameterization:

Fraction of individuals in each group	$N^i = 0.5$ for $i = L, H$
Discount factor	$\delta^L = 0.6, \delta^H = 0.8, r = 0.95$
Productivities (wages)	$n^L = 2, n^H = 3$

Table 1: Parameterization

No a priori assumptions of the binding self-selection constraints are made in the numerical simulations. In the benchmark model, with perfect positive correlation between the skill and time preferences, numerical simulation verify the assumption that the only binding self-selection constraint is type H considering mimicking type L⁸. Table 2 presents the optimal

⁷The numerical procedure is described in Tenhunen & Tuomala (2010) Appendix B.

⁸The slackness of the other self-selection constraints is also checked by calculating the difference in utilities when mimicking and when not.

consumption, labour and utility levels, the marginal tax rates for savings and labour and the replacement rate (x/ny). The additional results are presented in appendix B table B.1 and B.2.

maximin	c	x	y	U	T'	d	x/ny
Type L	0.67	0.41	0.38	-4.58	29.43	41.71	53.07
Type H	0.75	0.69	0.57	-4.79	0	0	40.62
utilitarian							
Type L	0.67	0.51	0.50	-4.67	6.30	6.72	51.20
Type H	0.79	0.72	0.54	-4.56	0	0	44.40

Table 2: Consumption in period 1 and 2 (c and x), labour supply (y), utility level (U), marginal tax rates on labour income and savings (T' and d), replacement rate (x/ny)

The results presented in table 2 shows that in an optimal solution the replacement rate decreases in earnings and there is a positive distortion i.e. marginal saving tax for the low-productivity type⁹. Compared to the case where there is no differences in preferences (and all have a high discount factor) the replacement rate in the current case is much lower for the low-productivity worker, because the level of the second period consumption in the case of lower discount factor is much smaller. However, as the labour supply y could also be interpreted as the length of career, it means that the different discount factors lead to much shorter career for the low-productivity worker compared to the case with the same discount factor. Also without heterogenous preferences, in the equality of opportunity case there would be no savings distortions when here the saving distortions are significant. The high-productivity worker is relatively less affected from introducing the differences in preferences to the economy.

Compared to the utilitarian social welfare function (weighted sum of low and high types' utility functions) the numerical simulations show that the levels of marginal tax rates differ significantly between the two types of objective functions. The average net taxes (shown in Appendix B) also show that there is a great difference between the equality of opportunity and utilitarian cases. In the maximin case, the government aims to improve the wellbeing of the low-productivity type by significant subsidies and by distorting the labour supply and savings decision relatively more than in the utilitarian case. Also if we interpret y as a length of career or retirement age, we notice that in maximin case the length of career is much shorter for the low-productivity worker but the labour supply decision for high-productivity worker is not much affected by changing the social preferences to maximin.

⁹In the case of perfect negative correlation, numerical simulations shows that L's distortion is a marginal subsidy. See figure 2 in section 4 for results with varying correlation.

3. Equality of opportunity and 3 types

In this and the next sections we generalise the previous model by giving up the assumption that productivity and time preferences are perfectly correlated. In general, there are now four types of individuals who differ both in productivity and time preferences numbered as in Table 3 ¹⁰.

	low-skilled, n^L	high-skilled, n^H
low delta, δ^L	Type 1	Type 3
high delta, δ^H	Type 2	Type 4

Table 3: Types of individuals

3.1 Low-ability types have same time preference

To maintain the tractability, we first simplify the model further by assuming that there are actually only three types. First we explore the case where the low-productivity types all have a low discount factor, δ^L , and are indexed as type 1. It can be justifiable to think that the low-productivity types are pooled together either because they have homogenous preferences towards savings or because their saving ability is constrained i.e. heavier time discounting is forced. High-productivity types with a low discount factor are denoted as type 3 and high-productivity worker with a high discount factor as type 4. We also assume that utility is given by Eq. (1), so we have the the same additively separable form as in the two-types case.

With more than two types there are several possibilities for mimicking. Which of the self-selection constraints bind depends on the interaction between individual preferences and the distributional preferences of the government, which hinge on the time preferences and skill level. A priori no binding constraints are forced but they are determined in the numerical simulation. To shorten the notation, the analytical results are shown only with the binding constraints.

When the government aims to the equality of opportunity, the government's problem reduces to maximising the welfare of the low-ability type as

$$N^1[u(c^1) + \delta^L v(x^1) + \psi(1 - y^1)] \quad (7)$$

subject to budget constraint and, without any assumptions of the mimicking behaviour, there are six possible self-selection constraints

$$u(c^i) + \delta^i v(x^i) + \psi(1 - y^i) \geq u^{ij}(c^j) + \delta^i v^{ij}(x^j) + \psi^{ij}(1 - \frac{n^j}{n^i} y^j) \text{ for } i, j = 1, 3, 4 \text{ and } i \neq j.$$

¹⁰In the case of one-dimensional heterogeneity types are ordered usually with respect to their income, consumption or utilities but in a two-dimensional world the ordering is not self-evidently clear.

As before to solve the distortions for savings we rewrite the first order conditions (presented in appendix A equations A.16-24) in form $\frac{u_x^i}{v_x^i} = \frac{\delta^i}{r}(1 - d^i)$, $i = 1, 3, 4$, where the distortions are given by (exploiting the binding incentive constraints from numerical solution)

$$\begin{aligned} d^1 &= 0 \\ d^3 &= \frac{\mu^{43}}{\mu^{31} - \mu^{43}} \Delta^{HL} \\ d^4 &= 0. \end{aligned} \tag{8}$$

The saving decision of type 3, high-productivity impatient individual, is now distorted and thus the "no distortion at the top" (with respect to skill) result holds no more. The sign of the distortion cannot be determined from the analytical results but numerical solution provides the binding incentive-compatibility constraints and their levels.

The solution for numerical simulation¹¹ is given in table 4 and additional results are presented in appendix B table B.3 and B.4. The binding constraints are (3,1) and (4,3). Also the effects of changing the parameter values are shown in subsection 3.3. Numerical solutions show that analytically ambiguous sign of the distortion d^3 is positive, implying a tax at the margin. The distortion for impatient high-productivity worker is helping to relax the incentive for patient high-productivity worker to mimick type 3.

equality of opportunity	c	x	y	U	T'	d	x/ny
Type 1	0.59	0.47	0.41	-4.67	37.07	0	57.44
Type 3	0.83	0.56	0.52	-4.35	0	27.38	36.05
Type 4	0.77	0.70	0.56	-4.71	0	0	41.92
utilitarian							
Type 1	0.65	0.51	0.50	-4.73	8.88	0	51.00
Type 3	0.83	0.64	0.52	-4.22	0	7.73	41.26
Type 4	0.80	0.73	0.54	-4.53	0	0	44.97

Table 4: Consumption in period 1 and 2, labour supply, utility level, marginal tax rates on labour income and savings, replacement rates; low-skilled grouped together

In this 3-type model the impatient high-productivity worker is relatively better off even though his the replacement rate is much lower than for the other types. This occurs as he consumes more during the first period as he prefers. Compared to the model where all agents have identical preferences about the timing of the consumption, including impatient workers in the model increases the labour supply of low-productivity workers.

¹¹Due to solvability problem the results with CES utility functions are given with parametric values $N^1 = 0.5$, $N^3 = 0.254$ and $N^4 = 0.246$.

In the utilitarian case (analytical results presented in appendix A equations A.26-34 and additional numerical results in table B.5) the main difference from the equality of opportunity model is that the labour supply or the length of career are closer between types. Intuitively in the equality of opportunity model the low-skilled type is better off with the cost of high-skilled type. More importantly the numerical simulations show that the differences between the two cases in the levels of marginal tax rates are significant. The average net taxes (shown in Appendix B) also show that there is a significant difference between the two social objectives. Interestingly adding one more type reduces the subsidy relative to income for low type and tax for high type in the equality of opportunity case but increases the relative subsidy in the utilitarian case.

3.2 High-ability types have same time preference

Alternatively suppose the kind of 3-type model where the high-productivity individuals have the same time preference. Now type 1 has individuals with low discount factor and low productivity, type 2 include individuals with high discount factor and low productivity and type 4 includes all high-productivity individuals with high discount factor. Now the government in turn maximises the equality of opportunity objective function

$$N^1[u(c^1) + \delta^L v(x^1) + \psi(1 - y^1)] + N^2[u(c^2) + \delta^H v(x^2) + \psi(1 - y^2)] \quad (9)$$

subject to same budget constraint and self-selection constraints as earlier. The first order conditions are presented in appendix A (equations A.37-45). In this case the distortions are (exploiting the binding incentive constraints from numerical solution)

$$\begin{aligned} d^1 &= \frac{\mu^{21} + \mu^{41}}{N^1 - \mu^{21} - \mu^{41}} \Delta_{HL} \\ d^2 &= d^4 = 0 \end{aligned} \quad (10)$$

Now only the saving decision of the impatient low-productivity type is distorted in order to prevent the other types from mimicking impatient type.

The binding incentive constraints when high-productivity types are pooled together are (2,1), (4,1) and (4,2). Table 5 presents the results with equality of opportunity and utilitarian social welfare function (additional results are found in appendix B table B.4 and B.6 and sensitivity analysis with different parameter values are shown in next subsection). The numerical solutions reveal that type 1's distortion is positive, so there is a tax at margin and this distortion is quite significant. However, comparing this distortion to the case of perfect correlation between preferences and productivities it shows up not as large. It seems that introducing some patient low-productivity workers to the model facilitate to soften the distortion for the impatient ones in a significant way.

equality of opportunity	c	x	y	U	T'	d	x/ny
Type 1	0.68	0.43	0.41	-4.57	24.73	37.08	52.6
Type 2	0.56	0.51	0.41	-5.04	43.3	0	63.2
Type 4	0.76	0.7	0.56	-4.73	0	0	41.7
utilitarian							
Type 1	0.67	0.52	0.51	-4.68	3.7	5.32	51.0
Type 2	0.62	0.57	0.51	-5.07	12.4	0	55.7
Type 4	0.79	0.73	0.54	-4.55	0	0	44.7

Table 5: Consumption in period 1 and 2, labour supply, utility level, marginal tax rates on labour income and savings, replacement rates; high-skilled grouped together

In the equality of opportunity case it is also noticeable that in the optimal solution type 1 and 2 work the same amount, thus the different time preference only leads to different division of consumption between periods. Same pattern applies also to the utilitarian case. Here the pattern of replacement rates implies again progressive pension system, and also indicates that in optimum the government distorts the patient low-productivity type's work incentives significantly. As before the most striking difference between the different social goals are in terms of the levels of distortions and thereof the length of the career or retirement age. The average tax rates (shown in Appendix B) also support this difference. In utilitarian case the labour supply is closer to similar between types.

We can conclude from both of the 3-type cases that the saving decision of the less patient individual are distorted in margin. In the case where high-productivity types have the same time preference, the low-productivity types' labour decisions are the same but due to the difference in their time preference for consumption, the overall utility differs.

3.3 Comparative statistics

In order to see what kind of effects the parameters have for the numerical results, in this section we let them vary. Table 6 presents the case for varying discount factor in the case of pooling the low-productivity workers. In benchmark those are set to 0.6 and 0.8 respectively for low- and high-productivity type. The distortion for saving is getting smaller as the discount factor is getting closer to the higher discount factor. In absence of private savings the replacement rates are increasing for types 1 and 3 the closer we get to the higher discount factor. When the discount factors are the same for low- productivity and high-productivity workers, there is no distortion for the saving's decision. The big distortions in the equality of opportunity case can be thus accounted for differences in time preferences. Also the simulation suggest that 0.6 is a lower bound with respect to the time preference for our model, as below that the solver either does not find a solution or there are no binding constraints.

δ^L	0.6			0.65			0.7			0.75		
Type	1	3	4	1	3	4	1	3	4	1	3	4
U	-4.7	-4.4	-4.7	-4.8	-4.4	-4.7	-4.9	-4.5	-4.7	-4.9	-4.6	-4.7
T'	37	0	0	38	0	0	38	0	0	38	0	0
d	0	27	0	0	19	0	0	13	0	0	6	0
x/ny	57	36	42	59	38	42	60	39	42	61	41	43

Table 6: Comparison of the solutions with different discount rates, low-ability types pooled.

Another case is to vary the sizes of the different types. Table 7 present these comparisons. First thing to notice is that in the equality of opportunity optimization a small size of the low-productivity workers means relatively more redistribution to their end. The replacement rates for high-productivity types are nearly the same in every case. The large marginal labour tax rates make sure that the high-productivity types do not mimick the low-productivity type. The distortion for savings is not much affected by the varying the sizes of the groups. Another thing to notice is that as the amount of low-productivity workers increase, with CES function, it is impossible to find a solution with equal group sizes of the high-productivity type.

n^i	0.2	0.4	0.4	0.25	0.375	0.375	0.3	0.35	0.35
Type	1	3	4	1	3	4	1	3	4
U	-4.3	-4.3	-4.6	-4.4	-4.3	-4.6	-4.5	-4.3	-4.7
T'	48	0	0	40	0	0	38	0	0
d	0	28	0	0	28	0	0	28	0
x/ny	137	37	43	100	37	43	81	37	43
n^i	0.35	0.325	0.325	0.4	0.29	0.31	0.45	0.295	0.255
Type	1	3	4	1	3	4	1	3	4
U	-4.5	-4.3	-4.7	-4.6	-4.3	-4.7	-4.6	-4.3	-4.7
T'	35	0	0	32	0	0	30	0	0
d	0	28	0	0	26	0	0	24	0
x/ny	72	36	42	65	36	42	60	36	43

Table 7: Comparison of the solutions with varying group sizes, low-ability types pooled.

In the second case, where the high-productivity types are pooled together, solving the maximization problem with different discount factors is much more complex. In fact, with CES function, in order to see what kind of effects the discount rates have for the results. The size of the groups needs to be modified, the sizes of the groups in this exercise are set to $N^1 = 0.2$, $N^2 = 0.3$ and $N^3 = 0.5$. Table 8 presents the results which show that the replacement rates

are increasing for type 1 when his patience increases but this does not significantly affect other types' replacement rates. Also the size of the discount factor seems to have quite significant effect on the level of the marginal savings tax.

δ^L	0.6			0.65			0.7		
Type	1	2	4	1	2	4	1	2	4
U	-4.6	-5	-4.7	-4.7	-5	4.7	-4.8	-5	-4.7
T'	24	42	0	28	41	0	31	40	0
d	36	0	0	29	0	0	20	0	0
x/ny	53	63	42	55	63	43	57	62	42

Table 8: Comparison of the solutions with varying discount factors, high-ability types pooled.

4. Equality of opportunity and 4 types

Finally we include all four types in our model. Now in the economy there are both impatient and patient high-productivity workers as well as impatient and patient low-productivity workers, labelled as in table 3. In order to study optimal taxation in the equality of opportunity framework we make a compromise between the principles of compensation and responsibility by firstly computing the minimum within each responsibility group (discount rates here) and then applying the utilitarian criterion. This means that for individuals with the same discount rates but different wage rates the maximin criterion is applied and thus we have a social ordering over each discount group. Then these minimum numbers are added together¹². Now we have

$$N^1[u(c^1) + \delta^L v(x^1) + \psi(1 - y^1)] + N^2[u(c^2) + \delta^H v(x^2) + \psi(1 - y^2)]. \quad (11)$$

The government maximizes now (11) subject to self-selection constraints (without any assumptions of the mimicking behaviour there are twelve possible self-selection constraints) given by $u(c^i) + \delta^i v(x^i) + \psi(1 - y^i) \geq u^{ij}(c^j) + \delta^i v^{ij}(x^j) + \psi^{ij}(1 - \frac{n^j y^j}{n^i})$ for $i, j = 1, \dots, 4$ and $i \neq j$, and the production constraint $\sum N^i(n^i y^i - c^i - r x^i) - R \geq 0$.

Specifications with CES function in this 4 types case turn out to be harder to solve numerically than the earlier cases. For example the case of zero correlation between ability and time preference ends up with no binding constraints. When the correlation gets closer to -1 or 1, the problem is solvable and these binding constraints ((3,1), (3,2), (4,2) and (4,3)) are exploited in deriving the first order conditions (Appendix A equations A.47-58) and distortions. The

¹²Or we can first calculate the average utility in each skill group and then apply the maximin criterion to such average figures.

distortions are

$$\begin{aligned}
 d^1 &= 0 \\
 d^2 &= \frac{\mu^{32}}{N^2 - \mu^{32} - \mu^{42}} \Delta^{LH} \\
 d^3 &= \frac{\mu^{43}}{\mu^{32} + \mu^{31} - \mu^{43}} \Delta^{HL} \\
 d^4 &= 0.
 \end{aligned} \tag{12}$$

Figure 2 shows that the problematic points in numerical simulations arise when we come closer to the zero correlation (i.e. group sizes are equal for all 4 types). This is intuitive because moving to either end of the correlation line we come closer to the two-type kind of economy which is generally easier to solve. Studying the optimal bundles in the case that is as close as possible to zero correlation we can conclude that interestingly the labour supply of impatient and patient low-productivity worker is nearly the same. In the equality of opportunity case this occurs because the saving decision of the patient individual is heavily subsidized and this leads to the situation that type 2's replacement rate is also high. Introducing the impatient high-productivity worker to the model leads to significantly higher marginal tax rates on labour compared to the three-type model. Interpreting the results without private savings, it can be noticed that the pension system is progressive, i.e. the replacement rates are smaller for high-productivity workers.

In the case of utilitarian social welfare function the binding self-selection constraints are (3,1), (3,2) and (4,3). Compared to the equality of opportunity case, the distortions are significantly smaller and the labour supply of the low-productivity type is greater which is in line with the earlier results.

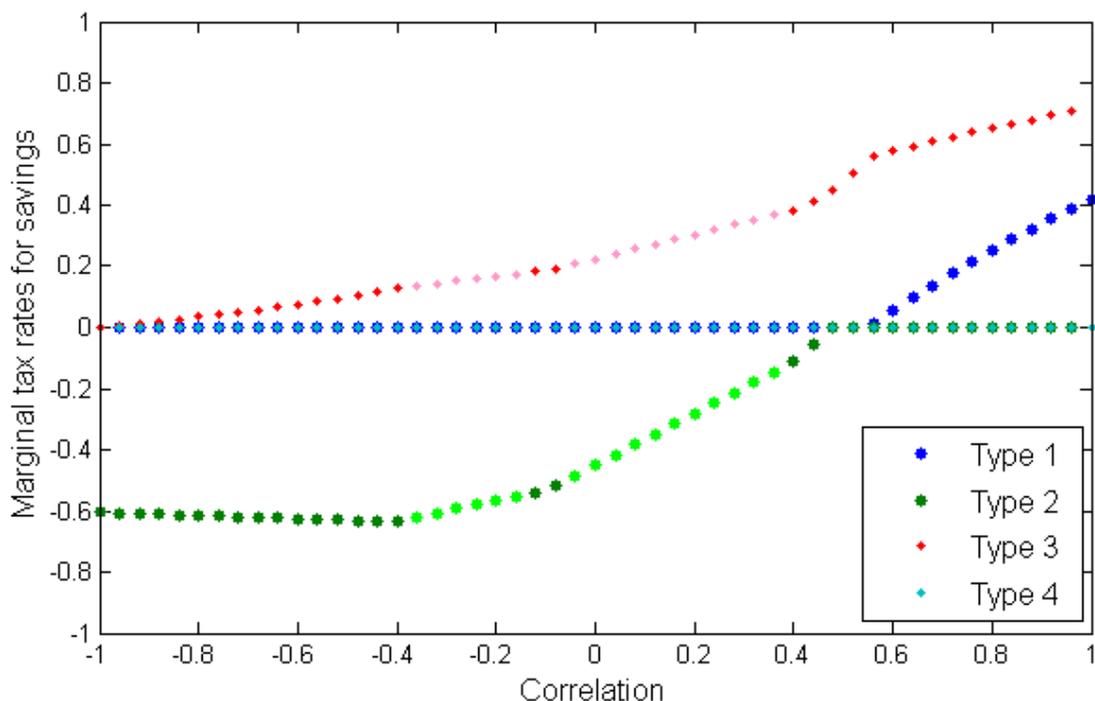


Figure 2: Marginal tax on saving and correlation between time preference and ability. Linearly interpolated values are drawn with lighter colour.

In figure 2 we show the marginal distortion of savings for each type with different correlation between preferences and skills. These results also provide information of the robustness of the earlier results with different distribution of types. The lighter values are linearly interpolated values as the program cannot solve the optimization problem in certain cases. In the figure we can see that the savings decision of type 2 is heavily subsidised when there is stronger negative correlation between time preference and ability. On the other end of the correlation line, type 3's saving decision is heavily distorted. Type 4 is undistorted in every correlation. Compared to Tenhunen&Tuomala (2010) which studies the similar kind of economy but with utilitarian government, the shape of the curve is similar but here the absolute values of marginal taxes and subsidies are greater. Also here the tax for type 1 become non-zero with slightly smaller correlations. To summarize briefly the results about the marginal labour income tax rates we can say that for different correlations both low-skilled types are taxed on margin in the optimal result. The levels of these taxes are relatively high and stable in different correlations points.

5. Conclusion

In this paper we have continued and extended the work of Tenhunen & Tuomala (2013, 2010) by introducing the equality of opportunity framework. Instead of solely examining an economy where the government is utilitarian, we have considered an economy where the social preferences aim to maximize the welfare of those who have lower productivity while not redistributing on the basis of preferences, which are individual's own responsibility. The multidimensionality stemming from the differences in preferences and productivities makes it unfeasible to fully satisfy these principle of responsibility and compensation at the same time but we have offered a one way to derive a compromise of the principles into the social welfare maximization problem.

In the context of two-period model we studied mainly the savings distortions and the optimal redistribution policy within a cohort. The results are derived analytically and numerically in several three and four productivity -types settings. The three types settings are somewhat easier to solve compared to the four-type models and they can also been used as an insight for the robustness of the four-types results. The numerical solutions help to reveal the sign of the distortion (upwards or downwards) in saving and labour supply behavior compared to the first-best case. The numerical results are compared to the utilitarian case to determine how the objective function affects the results. We have implicitly assumed that the government can commit to a lifetime tax in order to carry out the optimal redistribution policy.

In the lifetime context of labour supply we find that retirement age (length of career) is much lower (shorter) in the equality of opportunity case. The different goals of the government in welfare maximization causes a large differences in the implicit distortions and labour market outcomes. This demonstrates that goventment's goals do not only have secondary effects on the optimal taxation results but have an important role in consideration of the levels of distortions. The results indicate that irrespective of goals of the government the pension system is progressive, i.e. the replacement rates decrease with income.

References

- Atkinson, A. and Stiglitz, J. (1976). The design of tax structure: direct versus indirect taxation. *Journal of Public economics*, 6, 55-75.
- Bossert, W. (1995). Redistribution mechanisms based on individual characteristics. *Mathematical Social Sciences*, volume 29, issue 1, 1–17.
- Cremer, H., De Donder, P., Maldonado, D. and Pestieau, P. (2009). Forced saving, redistribution and non-linear social security scheme. *Southern Economic Journal*, 76(1), 86-98.
- Diamond, P. A. (2003). *Taxation, incomplete markets, and social security*. The 2000 Munich lectures. Cambridge, MA: MIT Press.
- Fleurbaey, M. (1994). On fair compensation. *Theory and Decision*, 36, 277–307.
- Riihelä, M., Sullström, R., and Tuomala, M. (2015) Veropolitiikka huipputulojen ja -varallisuuden taustalla -Onko Pikettyn kuvaama kehitys nähtävissä Suomessa? Book edited by Taimio, Heikki: *Hyvinvointivaltio 2010-luvulla; Mitä kello on lyönyt?*
- Roemer, J. E. (1998). *Equality of opportunity*. Harvard: University Press.
- Saez, E. (2002). The desirability of commodity taxation under non-linear income taxation and heterogenous tastes. *Journal of Public economics*, 83, 217-230.
- Sandmo, A. (1993). Optimal Redistribution when Tastes Differ. *Finanz Archiv*, 50(2), 149-163.
- Stern, N. (1982). Optimum taxation with errors in administration. *Journal of Public economics*, 17, 181-211.
- Stiglitz, J. (1982). Self-Selection and Pareto efficient taxation. *Journal of Public economics*, 17, 213-240.
- Tarkiainen, R. and Tuomala, M. (1999). Optimal Nonlinear Income Taxation with a Two-Dimensional Population; A Computational Approach. *Computational Economics*, 13, 1-16.
- Tarkiainen, R and Tuomala, M. (2007). On optimal income taxation with heterogeneous work preferences. *International Journal of Economic Theory*, 3, 35-46.

Tenhunen, S. and Tuomala, M. (2010). On Optimal Lifetime Redistribution Policy. *Journal of Public Economic Theory*, vol, 12(1), 171-198.

Tenhunen, S. and Tuomala, M. (2013). On the design of an optimal non-linear tax/pension system with habit formation. *International tax and public finance*, vol, 20, issue 3, 485-512.

Van de Gaer, D. (1993). Equality of opportunity and investments in human capital. *Ph.D.Thesis, Katholieke Universiteit Leuven*.

A First order conditions

To shorten the notation, we denote the partial derivatives as follows:

$$\frac{du(c^i)}{dc^i} = u_c^i, \frac{dv(x^i)}{dx^i} = v_x^i \text{ and } \frac{d\psi(1-y^i)}{dy^i} = \psi'^i$$

Two types

The first-order conditions with respect to c^i , x^i and y^i , $i = L, H$ from the Lagrange function given in Eq. (5) are

$$N^L u_c^L - \lambda N^L - \mu^{HL} u_c^L = 0 \quad (\text{A.1})$$

$$N^L \delta^L v_x^L - \lambda r N^L - \mu^{HL} \delta^H v_x^L = 0 \quad (\text{A.2})$$

$$-N^L \psi'^L + \lambda N^L n^L + \mu^{HL} \frac{n^L}{n^H} \psi'^L = 0 \quad (\text{A.3})$$

$$-\lambda N^H + \mu^{HL} u_c^H = 0 \quad (\text{A.4})$$

$$-\lambda r N^H + \mu^{HL} \delta^H v_x^H = 0 \quad (\text{A.5})$$

$$\lambda N^H n^H - \mu^{HL} \psi'^H = 0 \quad (\text{A.6})$$

In the paternalistic case the equation (A.2) is replaced by

$$N^L \delta^g v_x^L - \lambda r N^L - \mu^{HL} \delta^H v_x^L = 0 \quad (\text{A.7})$$

In the utilitarian case Lagrange is

$$\begin{aligned} \mathcal{L} = & \sum N^i [u(c^i) + \delta^i v(x^i) + \psi(1 - y^i)] + \lambda [\sum N^i (n^i y^i - c^i - r x^i) - R] \\ & + \mu^{HL} [u(c^H) + \delta^H v(x^H) + \psi(1 - y^H) - u(c^L) - \delta^H v(x^L) - \psi(1 - \frac{n^L}{n^H} y^L)] \end{aligned} \quad (\text{A.8})$$

and first order conditions:

$$N^L u_c^L - \lambda N^L - \mu^{HL} u_c^L = 0 \quad (\text{A.9})$$

$$N^L \delta^L v_x^L - \lambda r N^L - \mu^{HL} \delta^H v_x^L = 0 \quad (\text{A.10})$$

$$N^L \psi'^L - \lambda N^L n^L - \mu^{HL} \frac{n^L}{n^H} \psi'^L = 0 \quad (\text{A.11})$$

$$N^H u_c^H - \lambda N^H + \mu^{HL} u_c^H = 0 \quad (\text{A.12})$$

$$N^H \delta^H v_x^L - \lambda r N^H + \mu^{HL} \delta^H v_x^H = 0 \quad (\text{A.13})$$

$$N^H \psi'^H - \lambda N^H n^H - \mu^{HL} \psi'^H = 0 \quad (\text{A.14})$$

Three types: Low-productivity workers pooled

Using the information of the binding self-selection constraints provided by numerical solution.

the Lagrange function in the case of maximin objective function can be written as

$$\begin{aligned} \mathcal{L} = & N^1[u(c^L) + \delta^L v(x^L) + \psi(1 - y^L)] + \lambda[\sum N^i(n^i y^i - c^i - r x^i) - R] \\ & + \mu^{43}[u(c^4) + \delta^H v(x^4) + \psi(1 - y^4) - u(c^3) - \delta^H v(x^3) - \psi(1 - y^3)] \\ & + \mu^{31}[u(c^3) + \delta^L v(x^3) + \psi(1 - y^3) - u(c^1) - \delta^L v(x^1) - \psi(1 - \frac{n^L}{n^H} y^1)] \end{aligned} \quad (\text{A.15})$$

The first order condition with respect to c^i , x^i and y^i , $i = 1, 3, 4$ are given by

$$N^1 u_c^1 - \lambda N^1 - \mu^{31} u_c^1 = 0 \quad (\text{A.16})$$

$$N^1 \delta^L v_x^1 - \lambda r N^1 - \mu^{31} \delta^L v_x^{31} = 0 \quad (\text{A.17})$$

$$N^1 \psi' - \lambda N^1 n^L - \mu^{31} \frac{n^L}{n^H} \psi'^{31} = 0 \quad (\text{A.18})$$

$$-\lambda N^3 - \mu^{43} u_c^{43} + \mu^{31} u_c^3 = 0 \quad (\text{A.19})$$

$$-\lambda r N^3 - \mu^{43} \delta^H v_x^{43} + \mu^{31} \delta^L v_x^3 = 0 \quad (\text{A.20})$$

$$-\lambda N^3 n^H - \mu^{43} \psi'^{43} + \mu^{31} \psi'^{31} = 0 \quad (\text{A.21})$$

$$-\lambda N^4 + \mu^{43} u_c^4 = 0 \quad (\text{A.22})$$

$$-\lambda r N^4 + \mu^{43} \delta^H v_x^4 = 0 \quad (\text{A.23})$$

$$-\lambda N^4 n^H + \mu^{43} \psi' = 0 \quad (\text{A.24})$$

In the utilitarian case, the Lagrange function with binding incentive constraints can be written as

$$\begin{aligned} L = & \sum N^i[u(c^i) + \delta^i v(x^i) + \psi(1 - y^i)] + \lambda[\sum N^i(n^i y^i - c^i - r x^i) - R] \\ & + \mu^{43}[u(c^4) + \delta^H v(x^4) + \psi(1 - y^4) - u(c^3) - \delta^H v(x^3) - \psi(1 - y^3)] \\ & + \mu^{31}[u(c^3) + \delta^L v(x^3) + \psi(1 - y^3) - u(c^1) - \delta^L v(x^1) - \psi(1 - \frac{n^L}{n^H} y^1)]. \end{aligned} \quad (\text{A.25})$$

The first order condition with respect to c^i , x^i and y^i , $i = 1, 3, 4$ are given by

$$N^1 u_c^1 - \lambda N^1 - \mu^{31} u_c^1 = 0 \quad (\text{A.26})$$

$$N^1 \delta^L v_x^1 - \lambda r N^1 - \mu^{31} \delta^L v_x^{31} = 0 \quad (\text{A.27})$$

$$N^1 \psi'^1 - \lambda N^1 n^L - \mu^{31} \frac{n^L}{n^H} \psi'^{31} = 0 \quad (\text{A.28})$$

$$N^3 u_c^3 - \lambda N^3 - \mu^{43} u_c^{43} + \mu^{31} u_c^3 = 0 \quad (\text{A.29})$$

$$N^3 \delta^L v_x^3 - \lambda r N^3 - \mu^{43} \delta^H v_x^{43} + \mu^{31} \delta^L v_x^3 = 0 \quad (\text{A.30})$$

$$N^3 \psi'^3 - \lambda N^3 n^H - \mu^{43} \psi'^{43} + \mu^{31} \psi'^{31} = 0 \quad (\text{A.31})$$

$$N^4 u_c^4 - \lambda N^4 + \mu^{43} u_c^4 = 0 \quad (\text{A.32})$$

$$N^4 \delta^H v_x^4 - \lambda r N^4 + \mu^{43} \delta^H v_x^4 = 0 \quad (\text{A.33})$$

$$N^4 \psi'^4 - \lambda N^4 n^H + \mu^{43} \psi'^{43} = 0 \quad (\text{A.34})$$

The distortions in this case are

$$\begin{aligned} d^1 &= 0 \\ d^3 &= \frac{\mu^{43}}{N^3 + \mu^{31} - \mu^{43}} \Delta^{HL} \\ d^4 &= 0. \end{aligned} \quad (\text{A.35})$$

Three types: high-productivity workers pooled

Using the information of the binding self-selection constraints provided by numerical solution. the Lagrange function can be written as

$$\begin{aligned} \mathcal{L} &= N^1 [u(c^1) + \delta^L v(x^1) + \psi(1 - y^1)] + N^2 [u(c^2) + \delta^H v(x^2) + \psi(1 - y^2)] \\ &+ \lambda \left[\sum_{i=1}^n N^i (n^i y^i - c^i - r x^i) - R \right] + \mu^{21} [u(c^2) + \delta^H v(x^2) + \psi(1 - y^2) - u(c^1) - \delta^H v(x^1) + \psi(1 - y^1)] \\ &\quad + \mu^{41} \left[u(c^4) + \delta^H v(x^4) + \psi(1 - y^4) - u(c^1) - \delta^H v(x^1) + \psi(1 - \frac{n^L}{n^H} y^1) \right] \\ &\quad + \mu^{42} \left[u(c^4) + \delta^H v(x^4) + \psi(1 - y^4) - u(c^2) - \delta^H v(x^2) + \psi(1 - \frac{n^L}{n^H} y^2) \right] \end{aligned} \quad (\text{A.36})$$

The first order conditions with respect c^i , x^i and y^i , $i = 1, 2, 4$:

$$N^1 u_c^1 - \lambda N^1 - \mu^{21} u_c^1 - \mu^{41} u_c^1 = 0 \quad (\text{A.37})$$

$$N^1 \delta^L v_x^1 - \lambda r N^1 - \mu^{21} \delta^H v_x^1 - \mu^{41} \delta^H v_x^1 = 0 \quad (\text{A.38})$$

$$-N^1 \psi_y^1 + \lambda N^1 n^L + \mu^{21} \psi_y^1 + \mu^{41} \frac{n^L}{n^H} \psi_y^1 = 0 \quad (\text{A.39})$$

$$N^2 u_c^2 - \lambda N^2 + \mu^{21} u_c^2 - \mu^{42} u_c^2 = 0 \quad (\text{A.40})$$

$$N^2 \delta^H v_x^2 - \lambda r N^2 + \mu^{21} \delta^H v_x^2 - \mu^{42} \delta^H v_x^2 = 0 \quad (\text{A.41})$$

$$-N^2 \psi_y^2 + \lambda N^2 n^L - \mu^{21} \psi_y^2 + \mu^{42} \frac{n^L}{n^H} \psi_y^2 = 0 \quad (\text{A.42})$$

$$-\lambda N^4 + \mu^{42} u_c^4 + \mu^{41} u_c^4 = 0 \quad (\text{A.43})$$

$$-\lambda r N^4 + \mu^{42} \delta^H v_x^4 + \mu^{41} \delta^H v_x^4 = 0 \quad (\text{A.44})$$

$$-\lambda N^4 n^H - \mu^{42} \psi_y^4 - \mu^{41} \psi_y^4 = 0 \quad (\text{A.45})$$

Four types

$$\begin{aligned} \mathcal{L} = & N^1 [u(c^1) + \delta^L v(x^1) + \psi(1 - y^1)] + N^2 [u(c^2) + \delta^H v(x^2) + \psi(1 - y^2)] \\ & + \lambda \left[\sum_{i=1}^n N^i (n^i y^i - c^i - r x^i) - R \right] \\ & + \mu^{31} \left[u(c^3) + \delta^L v(x^3) + \psi(1 - y^3) - u(c^1) - \delta^L v(x^1) + \psi \left(1 - \frac{n^L}{n^H} y^1 \right) \right] \\ & + \mu^{32} \left[u(c^3) + \delta^L v(x^3) + \psi(1 - y^3) - u(c^2) - \delta^L v(x^2) + \psi \left(1 - \frac{n^L}{n^H} y^2 \right) \right] \\ & + \mu^{42} \left[u(c^4) + \delta^H v(x^4) + \psi(1 - y^4) - u(c^2) - \delta^H v(x^2) + \psi \left(1 - \frac{n^L}{n^H} y^2 \right) \right] \\ & + \mu^{43} \left[u(c^4) + \delta^H v(x^4) + \psi(1 - y^4) - u(c^3) - \delta^L v(x^3) + \psi(1 - y^3) \right] \end{aligned} \quad (\text{A.46})$$

The first order conditions with respect to c^i , x^i , and y^i for $i = 1, 2, 3, 4$ are

$$N^1 u_c^1 - \lambda N^1 - \mu^{31} u_c^1 = 0 \quad (\text{A.47})$$

$$N^1 \delta^L v_x^1 - \lambda r N^1 - \mu^{31} \delta^L v_x^1 = 0 \quad (\text{A.48})$$

$$-N^1 \psi_y^1 + \lambda N^1 n^L + \mu^{31} \frac{n^L}{n^H} \psi_y^1 = 0 \quad (\text{A.49})$$

$$N^2 u_c^2 - \lambda N^2 - \mu^{32} u_c^2 - \mu^{42} u_c^2 = 0 \quad (\text{A.50})$$

$$N^2 \delta^H v_x^2 - \lambda r N^2 - \mu^{32} \delta^L v_x^2 - \mu^{42} \delta^H v_x^2 = 0 \quad (\text{A.51})$$

$$-N^2 \psi_y^2 + \lambda N^2 n^L + \mu^{32} \frac{n^L}{n^H} \psi_y^2 + \mu^{42} \frac{n^L}{n^H} \psi_y^2 = 0 \quad (\text{A.52})$$

$$-\lambda N^3 + \mu^{31} u_c^3 + \mu^{32} u_c^3 - \mu^{43} u_c^3 = 0 \quad (\text{A.53})$$

$$-\lambda r N^3 + \mu^{31} \delta^L v_x^3 + \mu^{32} \delta^L v_x^3 - \mu^{43} \delta^H v_x^3 = 0 \quad (\text{A.54})$$

$$-\lambda N^3 n^H - \mu^{31} \frac{n^L}{n^H} \psi_y^3 - \mu^{32} \frac{n^L}{n^H} \psi_y^3 + \mu^{43} \psi_y^3 = 0 \quad (\text{A.55})$$

$$-\lambda N^4 + \mu^{42} u_c^4 + \mu^{43} u_c^4 = 0 \quad (\text{A.56})$$

$$-\lambda r N^4 + \mu^{42} \delta^H v_x^4 + \mu^{43} \delta^H v_x^4 = 0 \quad (\text{A.57})$$

$$-\lambda N^4 n^H - \mu^{42} \psi_y^4 - \mu^{43} \psi_y^4 = 0 \quad (\text{A.58})$$

In the paternalistic case the problem is to solve

$$\begin{aligned} \mathcal{L} = & N^1 [u(c^1) + \delta^g v(x^1) + \psi(1 - y^1) + N^2 [u(c^2) + \delta^g v(x^2) + \psi(1 - y^2)]] \\ & + \lambda \left[\sum_{i=1}^n N^i (n^i y^i - c^i - r x^i) - R \right] + \mu^{12} \left[u(c^1) + \delta^L v(x^1) + \psi(1 - y^1) - u(c^2) - \delta^H v(x^2) + \psi(1 - y^2) \right] \\ & + \mu^{21} \left[u(c^2) + \delta^H v(x^2) + \psi(1 - y^2) - u(c^1) - \delta^H v(x^1) + \psi(1 - y^1) \right] \\ & + \mu^{31} \left[u(c^3) + \delta^L v(x^3) + \psi(1 - y^3) - u(c^1) - \delta^L v(x^1) + \psi(1 - \frac{n^L}{n^H} y^1) \right] \\ & + \mu^{32} \left[u(c^3) + \delta^L v(x^3) + \psi(1 - y^3) - u(c^2) - \delta^L v(x^2) + \psi(1 - \frac{n^L}{n^H} y^2) \right] \\ & + \mu^{42} \left[u(c^4) + \delta^H v(x^4) + \psi(1 - y^4) - u(c^2) - \delta^H v(x^2) + \psi(1 - \frac{n^L}{n^H} y^2) \right] \\ & + \mu^{43} \left[u(c^4) + \delta^H v(x^4) + \psi(1 - y^4) - u(c^3) - \delta^L v(x^3) + \psi(1 - y^3) \right]. \end{aligned} \quad (\text{A.59})$$

The first order conditions with respect to c^i , x^i , and y^i for $i = 1, 2, 3, 4$ are

$$N^1 u_c^1 - \lambda N^1 + \mu^{12} u_c^1 - \mu^{21} u_c^1 - \mu^{31} u_c^1 = 0 \quad (\text{A.60})$$

$$N^1 \delta^g v_x^1 - \lambda r N^1 + \mu^{12} \delta^L v_x^1 - \mu^{21} \delta^H v_x^1 - \mu^{31} \delta^L v_x^1 = 0 \quad (\text{A.61})$$

$$-N^1 \psi_y'^1 + \lambda N^1 n^L - \mu^{12} \psi_y'^1 + \mu^{21} \psi_y'^1 + \mu^{31} \frac{n^L}{n^H} \psi_y'^1 = 0 \quad (\text{A.62})$$

$$N^2 u_c^2 - \lambda N^2 - \mu^{12} u_c^2 + \mu^{21} u_c^2 - \mu^{32} u_c^2 - \mu^{42} u_c^2 = 0 \quad (\text{A.63})$$

$$N^2 \delta^g v_x^2 - \lambda r N^2 - \mu^{12} \delta^L v_x^2 + \mu^{21} \delta^H v_x^2 - \mu^{32} \delta^L v_x^2 - \mu^{42} \delta^H v_x^2 = 0 \quad (\text{A.64})$$

$$-N^2 \psi_y'^2 + \lambda N^2 n^L + \mu^{12} \psi_y'^2 - \mu^{21} \psi_y'^2 + \mu^{32} \frac{n^L}{n^H} \psi_y'^2 + \mu^{42} \frac{n^L}{n^H} \psi_y'^2 = 0 \quad (\text{A.65})$$

$$-\lambda N^3 + \mu^{31} u_c^3 + \mu^{32} u_c^3 - \mu^{43} u_c^3 = 0 \quad (\text{A.66})$$

$$-\lambda r N^3 + \mu^{31} \delta^L v_x^3 + \mu^{32} \delta^L v_x^3 - \mu^{43} \delta^H v_x^3 = 0 \quad (\text{A.67})$$

$$-\lambda N^3 n^H - \mu^{31} \frac{n^L}{n^H} \psi_y'^3 - \mu^{32} \frac{n^L}{n^H} \psi_y'^3 + \mu^{43} \psi_y'^3 = 0 \quad (\text{A.68})$$

$$-\lambda N^4 + \mu^{42} u_c^4 + \mu^{43} u_c^4 = 0 \quad (\text{A.69})$$

$$-\lambda r N^4 + \mu^{42} \delta^H v_x^4 + \mu^{43} \delta^H v_x^4 = 0 \quad (\text{A.70})$$

$$-\lambda N^4 n^H - \mu^{42} \psi_y'^4 - \mu^{43} \psi_y'^4 = 0 \quad (\text{A.71})$$

B Additional results from the numerical simulations

λ	μ^{LH}	μ^{HL}		Type L	Type H
0.984	0 (-4.53)	0.278	Average tax rate	-40.5	17.0

Table B.1: Lagrange multipliers and average tax rates for two-type model, maximin case. Binding constraints in optimum are bolded. For non-binding constraint the value of the constraint ($U^{ij} - U^i$) is given in paranthesis.

λ	μ^{LH}	μ^{HL}		Type L	Type H
1.87	0 (-2.69)	0.08	Average tax rate	-16.5	8.1

Table B.2: Lagrange multipliers and average tax rates for two-type model in utilitarian case. Binding constraints in optimum are bolded. For non-binding constraint the value of the constraint ($U^{ij} - U^i$) is given in paranthesis.

λ	μ^{13}	μ^{14}	μ^{31}	μ^{34}	μ^{41}	μ^{43}		Type 1	Type 3	Type 4
1.02	0 (-2.15)	0 (-3.6)	0.33	0 (-0.07)	0 (-0.07)	0.15	Average tax rate	-26.4	12.7	14.6

Table B.3: Lagrange multipliers and average tax rates for three-type model in maximin case. Pooling of low-ability types. Binding constraints in optimum are bolded. For non-binding constraint the value of the constraint ($U^{ij} - U^i$) is given in paranthesis.

λ	μ^{12}	μ^{14}	μ^{21}	μ^{24}	μ^{41}	μ^{42}		Type 1	Type 2	Type 4
1.03	0 (-0.07)	0 (-3.8)	0.02	0 (-3.3)	0.12	0.18	Average tax rate	-32.7	-27.4	15.2

Table B.4: Lagrange multipliers and average tax rates for three-type model in maximin case. Pooling of high-ability types. Binding constraints in optimum are bolded. For non-binding constraint the value of the constraint ($U^{ij} - U^i$) is given in paranthesis.

λ	μ^{13}	μ^{14}	μ^{31}	μ^{34}	μ^{41}	μ^{43}		Type 1	Type 3	Type 4
1.85	0 (-1.97)	0 (-2.65)	0.11	0 (-0.04)	0 (-0.08)	0.04	Average tax rate	-17.9	3.4	3.7

Table B.5: Lagrange multipliers and average tax rates for three-type model in utilitarian case. Pooling of low-ability types. Binding constraints in optimum are bolded. For non-binding constraint the value of the constraint ($U^{ij} - U^i$) is given in paranthesis.

λ	μ^{12}	μ^{14}	μ^{21}	μ^{24}	μ^{41}	μ^{42}		Type 1	Type 2	Type 4
1.92	0 (-0.03)	0 (-2.8)	0.02	0 (-2.65)	0.12	0.18	Average tax rate	-19.6	-19.3	4.7

Table B.6: Lagrange multipliers and average tax rates for three-type model in utilitarian case. Pooling of high-ability types. Binding constraints in optimum are bolded. For non-binding constraint the value of the constraint ($U^{ij} - U^i$) is given in paranthesis.