Educational self-selection, tasks assignment and rising wage inequality*

A. Dupuy†

September 22, 2005

Abstract

In this paper, I develop a tasks assignment model with endogenous human capital formation. Workers are initially endowed with abilities of two types and select their education according to their abilities. The educational system transforms workers' abilities into skills of two types. To produce output firms assign tasks −which differ in both the type and the level of skills they require− to workers following Ricardo's principle of comparative advantage. I provide a family of closed form solutions of the model when tasks are distributed according to the Beta distribution and the output production function is of the well-known and broadly used CES form. Some calibrations of the model enable to link changes in the US wage distribution observed in the last 4 decades to changes in the degree of heterogeneity of workers' skills, the distribution of tasks and skill-biased technical change.

JEL Classification: D3, J21, J23 and J31.
Keywords: Endogenous human capital formation, tasks assignment, substitution, technical change and wage distribution.

*Preliminary draft. The author thanks participants at the 9th World Congress of the Econometric Society, 2005, London, Daron Acemoglu, Lex Borghans, Bart Golsteyn, Andries de Grip, Hans Heijke and Coen Teulings for helpful comments.

†Corresponding address: ROA, Maastricht University PO Box 616, NL-6200 MD, The Netherlands. Email: a.dupuy@roa.unimaas.nl.

1 Introduction

Assignment theory[1] focuses on the relationship between the distribution of workers' skills, the distribution of the skills required to perform the various jobs and the resulting income distribution. This relationship makes assignment models particularly attractive to analyze sources of wage inequality between and within educational categories through time and the substitutability between worker types. However, as acknowledged by a large body of empirical work,[2] the wage structure is very sensitive to technological and organizational changes and changes in the wage structure are complex and occur both between and within skill groups of workers. A challenge faced by assignment models is therefore to offer a convenient framework to analyze changes in the wage structure within and between skills groups of workers resulting from technological and organizational changes. As already acknowledged by Acemoglu (2002) to win this challenge, an assignment model requires a multi-dimensional skills framework.

> A major issue that most models discussed so far [see footnote 2 for an exhaustive list of these models] failed to address is the differential behavior of returns to schooling and residual inequality during the 1970s. I argue [...] that an explanation for this pattern requires models with multi-dimensional skills. Acemoglu (2002), p. 56.

As indicated in Table 1, the challenge assignment literature faces is that of deriving i) a model with two types of assignment i.e. assignment of workers to education and assignment of workers to jobs, ii) that has a *general* equilibrium −in order to account for the heterogeneity of both workers and jobs− in which iii) human capital formation is *endogenous*[3] that yields iv) *closed form* solutions and v) a *(well-) known* structure of substitution. This challenge is the main motivation of this paper. The model developed is a tasks assignment model with endogenous human capital formation. The supply side of the model is very similar to Roy's (1950 and 1951) self-selection model. Workers are initially endowed

[1] The assignment literature diverges from the matching theory, see Mortensen (1986) for instance, and search theory, (see Jovanovic (1979), Diamond (1981) and Pissarides (1984)) by assuming that workers have full knowledge of all employers' wage offers and that employers have full knowledge of all workers' abilities. In assignment models, it is assumed that free choice and competitive markets assign tasks to workers efficiently. Perfect competition ensures that workers in each group are rewarded according to the productivity of the marginal worker in the associated group. Free choice ensures that workers select their education according to their comparative advantage and tasks are assigned according to the structure of comparative advantage.

[2] See among others, Acemoglu (2002), Autor et al. (2003), Beaudry and Green (2003), Berman et al. (1994), Berman et al. (1998), Bound and Johnson (1992), Breshnahan (1999), Bresnahan et al. (2002), Brynjolfsson (1995), Card and Lemieux (2001), Katz and Murphy (1992), Krueger (1993) and Krusell et al. (2000).

[3] This can only be achieved if one considers a multi-skill structure.

with abilities of two types and select their education according to their abilities. Educational self-selection endogenizes human capital formation. The educational system transforms workers' abilities into skills of two types. The demand side of the model is very similar to Rosen's (1978) tasks assignment model. To produce output firms assign tasks, which differ in both the type and the level of skills they require, to workers following Ricardo's principle of comparative advantage. I provide a family of closed form solutions of the model when tasks are distributed according to the Beta distribution and the output production function is of the well-known and broadly used CES form. I Calibrate the model so as to fit changes in the US wage distribution in the last 4 decades. These calibrations enable to link changes in the between and within wage inequality to changes in the degree of heterogeneity of workers' skills, the distribution of tasks and skill-biased technical change.

In the literature, assignment models differ from each other with respect to i) the origin of heterogeneity in the labor market, workers and/or jobs, ii) whether a single or multiple-skill structure is used to differentiate workers and/or jobs, iii) whether human capital formation is endogenous or exogenous, iv) whether a partial or general equilibrium is considered, v) whether general equilibrium models yield closed form solutions and vi) whether the structure of substitution that results in equilibrium is (well-) known. Table 1 provides a review of the most influential assignment models −plus the model developed in this paper− and summarizes their respective characteristics. The self-selection model proposed by Roy (1950 and 1951) puts the emphasis on the *supply side* of the labor market by focusing on the heterogeneity of individuals.[4] Individuals are endowed with different abilities and choose a sector among a small number of sectors. The demand for workers and wage rates by sectors are exogenous to the model. An interesting feature of Roy's model is that it can be used to model endogenous human capital formation by accounting for individuals' educational choice.[5] Individuals choose their educational profile based on their initial abilities and the exogenous market wages associated with each educational background.[6]

In contrast to Roy's model, Rosen's (1978) tasks assignment model puts the emphasis on the *demand side* of the labor market by focusing on the heterogeneity of tasks. Tasks differ from one another by the levels of the various types of skills they require. Workers are grouped in a small number of homogeneous skill groups (educational categories) and the supply of workers and wage rates by skill groups are assumed to be exogenous. The main advantage of this model is that it offers a very convenient framework to analyze substitution between skill

[4]Roy's original model has two sectors only, namely fishing and hunting. Extensions of Roy's model to more than two sectors are provided in Heckman and Sedlacek (1985 and 1990) and Gould (2002).

[5]See Willis and Rosen (1979) .

[6]Note that Heckman et al. (1998) have recently extended Roy's model to allow for investment and endogenous skills prices determination. However, jobs remain absent in their model.

groups of workers. Rosen considers the demand for labor by modeling firms' *indirect* production function resulting from the assignment of tasks to workers that maximizes output, given exogenous wages. However, human capital formation, either exogenous like in Teulings (2002) or endogenous like in Roy's self-selection model, and workers heterogeneity within educational groups are not accounted for in Rosen's (1978) model.

General hedonic models, developed by Rosen (1974) (see also Lucas (1977)) are very appealing to model assignment in the labor market as they incorporate both heterogeneity in workers' skills and heterogeneity in the skills required in the various jobs. A well-known special case of hedonic models is the *allocation model* proposed by Tinbergen (1956). In the allocation model, workers as well as tasks are defined along a finite number of types of skills and the quantity of each type of skill workers have and tasks require are continuously distributed. Since in practice the supply and demand distributions do not coincide, the supply distribution has to be deformed so as to coincide with the demand distribution otherwise there will not be an equilibrium. The endogenous income scale serves to equilibrate the demand and supply distributions. Tinbergen's model is very general as it incorporates both sources of heterogeneity, but human capital formation is absent in the model. Moreover, a general drawback of hedonic models is that they do not give rise to closed-form solutions and even when they do, under very special properties (too restrictive in most applications, as shown by Ekeland et al. (2004)) like in Tinbergen (1956), the derivation of labor demand equations is very complicated as the type of the implicit aggregate function of production is undefined.

Teulings's (1995a and 1995b) assignment model offers an appealing solution to this problem by modelling the assignment of workers whose skills are ranked on a one dimensional scale to jobs differing by their complexity (one job scale) such that the output production function has the CES form. In a recent paper, Teulings (2005) extended his assignment model to account for exogenous human capital accumulation and DIstance Dependent Elasticity of Substitution (DIDES). Although Teulings's model offers a very convenient framework to study substitution between workers, it assumes that workers can be differentiated by one continuous variable[7] and does not allow to infer on workers' educational self-selection and therefore to endogenize human capital formation. Distinguishing between different types of skills has a second advantage compared to the assignment model proposed by Teulings (1995a, 1995b and 2005). In contrast to the one-to-one relationship between wages and skills in Teulings' model, an important feature of the model presented in this paper is that the distribution of wages of workers with education j will generally overlap to some extent with the distribu-

[7]The amount of one type of skills differentiates workers. Teulings's setting does not account for quality distinction between for instance manual and cognitive skills. This means that the skills of Albert Einstein, Sigmund Freud and Michael Jordan are ranked on one single skill scale which along with the single job scale determines their respective earnings.

tion of wages of workers with education k. This is a very interesting feature since in real data the distribution between educational groups of workers do overlap, i.e. workers in the upper tail of the wage distribution of high-school graduates have higher wages than workers in the lower tail of the wage distribution of college graduates.

The remaining structure of the paper is as follows. In the next section I present the model of tasks assignment with endogenous human capital formation. In section 3, the particular case where the output production function is of the well-known and broadly used CES type and tasks are distributed according to the general Beta type of distribution is investigated in details. Finally in section 4, I illustrate the implications of the model by calibrating the parameters so as to fit the main patterns of the wage distribution in the US in the last decades. The patterns of the wage distribution used here are those exposed in Acemoglu (2002). In particular I show that the model can be used to explain both the rise in the college skill-premium and the rise in overall wage inequality. The skill premium pattern is explained by technical and organizational changes that affect i) the distribution of tasks and therefore the demand for skills and ii) the efficiency of college and high-school graduates as well as the comparative advantage of college graduates via changes in the elasticity of substitution between both labor groups. The overall rise in wage inequality can be explained by i) the job polarization due to changes in the distribution of tasks and ii) increasing heterogeneity within educational groups.

2 A tasks assignment model with endogenous human capital formation

2.1 Setting of the model

Workers' abilities and skills

Suppose workers are endowed with initial abilities vector $a = \langle a_1, a_2 \rangle$ where a_1 and a_2 represent an individual's ability of type 1 and 2, for instance manual and cognitive abilities. In the first period, workers make an educational choice: for instance, whether to go to college or quit after high-school or whether to study economics or mathematics? The educational system transforms the vector of initial abilities into a vector of skills through an educational production matrix[8] E_k. The educational production matrix depends on the educational choice. A worker with initial abilities $\langle a_1, a_2 \rangle$ will have skills $\langle t_1, t_2 \rangle = aE_1 = \langle e_{11}a_1; e_{12}a_2 \rangle$ if she selects education 1 and skills $\langle t'_1, t'_2 \rangle = aE_2 = \langle e_{21}a_1; e_{22}a_2 \rangle$ if she selects education 2. For the sake of simplicity I assume that the educational system has a

[8]See Hartog (2001) for a similar approach of the role of the educational system.

common treatment effect, that is, it enhances abilities by a factor common to each individual, i.e. $\partial E_k/\partial a_j = 0$ for $\langle k, j\rangle = 1, 2$.[9] However, the educational system transforms initial abilities into skills in a non proportional way. A worker with initial abilities $\langle a_1, a_2\rangle$ will have skills $\langle t_1, t_2\rangle$ with $t_1/t_2 > a_1/a_2$ if she selects education 1 and skills with $t_1/t_2 < a_1/a_2$ if she selects education 2. Hence, $E_1 = \begin{pmatrix} e_{11} & 0 \\ 0 & e_{12} \end{pmatrix}$ with $e_{11} > e_{12} \geq 1$ and $E_2 = \begin{pmatrix} e_{21} & 0 \\ 0 & e_{22} \end{pmatrix}$ with $e_{22} > e_{21} \geq 1$. In order for educational choice to be non trivial, I further assume that education k develops abilities of type k relatively more than education j, i.e. $e_{kk} > e_{jk}\ \forall j, k$. Without this assumption, for instance with $e_{jj} > e_{jk} > e_{kk} > e_{kj} \geq 1$, educational choice would still be endogenous but education j would be a strictly dominant educational strategy so that all workers would select education j. Workers with relatively low ability of type k still seek to enhance their ability of type j through education j only now workers with relatively high ability of type k choose education j to enhance their ability of type k. Clearly, although human capital is still endogenous this case is less interesting.

It is important to bear in mind that changes in the educational system that lead to relative improvements in the production of skills of one type will play a non neutral role in the educational choice of workers. Indeed, given the distribution of initial abilities and relative wages, i.e. $\frac{w_1}{w_2} \equiv \mu$ where w_k is the pay-rate per unit of skills of type k supplied, an increase in e_{kk} will increase the potential skills k compared to the potential skill j of every workers and lead more workers to select education k. The model therefore allows to investigate the effects of exogenous changes in the educational production of skills on educational choices and hence wage distribution. For instance, the model could be used to evaluate the general equilibrium impact of the recent introduction of the "Literacy Hour" in primary English schools (see Machin and McNally (2005)) or Singapore's mathematics teaching method in the USA $-$launched in order to offset the relative poor numeracy score of American pupils$-$ on wage inequality.[10]

Tasks and productivity: the structure of comparative advantage

Consider an economy producing a composite commodity by means of the input of an infinite number of different tasks.[11] To produce output level Y a continuum of tasks v, $v \in (0, 1)$, needs to be performed. The distribution of

[9]Herewith a common treatment effect model is considered as opposed to a model in which responses to the treatment (education) are different across workers, for instance $\partial E_k/\partial a_k \lesseqgtr 0$. The common effect specification could however be easily dropped at this stage. See Heckman (1999), Heckman and Li (2004) and Heckman and Vytlacil (2005) for an exhaustive presentation of the evaluation of social programs in general equilibrium.

[10]In the distribution of students' achievements in mathematics at the age of 13 years old across countries the USA rank 19^{th} out of 38 countries whereas Singapore ranks first. See TIMSS (2000).

[11]Teulings (1995a, 1995b and 2002) use a similar economy.

tasks is exogenous and given by the density distribution of tasks $d(v)$ and cumulative distribution $F(v^*) = \int_0^{v^*} d(v)dv$. Each firm produces a single task and the markets for tasks are perfectly competitive and firms can enter freely into the respective markets. Therefore, in equilibrium a zero profit condition holds for firms, i.e. $\rho(v) \cdot p_k(t_k, v) - w(t_k) = 0$ where $\rho(v)$ is the price of task v, $p_k(t_k, v)$ is the productivity of workers with education k and skills t_k assigned to tasks v measured in units of output per worker and $w(t_k)$ is the cost of employing a worker supplying t_k units of skills of type k. From the self-selection of workers into education, the cost of employing a worker supplying t_k is equal to the pay-rate associated with skills of type k times the amount of skills supplied, i.e. $w(t_k) = w_k t_k$. Hence, the zero profit condition yields:

$$\rho(v)\frac{p_k(v, t_k)}{t_k} = w_k \; \forall k = 1, 2 \tag{1}$$

There is full employment so that each worker is assigned to a single task and each single task is assigned to one and only one worker. Output at task v can be produced by workers with different types and levels of skills. However, workers of different types and levels of skills differ in their productivity p_k. Without loss of generality, I assume that workers with education 1 have a comparative advantage in tasks v close to 0 and workers with education 2 have a comparative advantage in tasks v close to 1, i.e. $\partial p_1(t_1, v)/\partial v < 0$ and $\partial p_2(t_2, v)/\partial v > 0$. I assume further that the productivity of workers of each type of education increases with their skills, i.e. $\partial p_k(t_k, v)/\partial t_k > 0 \; \forall v$.[12] Moreover, among workers with education 1, those with high t_1 skills have a comparative advantage in tasks v close to 0 and among workers with education 2, those workers with high t_2 skills have a comparative advantage in tasks v close to 1, i.e. $\partial^2 p_1(t_1, v)/\partial t_1 \partial v \leq 0$ and $\partial^2 p_2(t_2, v)/\partial t_2 \partial v \geq 0$.

The null profit condition alone is not sufficient to guarantee a stable equilibrium. To see this, consider two firms offering task v and v' respectively with $v < v'$. Firm v wants to recruit a worker with skills t_1 and following the null profit condition the firm should pay $w_1 t_1 = \rho(v)p_1(v, t_1)$. Similarly firm v' wants to recruit a worker with skills t_1' and should pay $w_1 t_1' = \rho(v')p_1(v', t_1')$. However, firm v could also try to recruit worker t_1' by offering a wage rate w_1' so that worker t_1' is willing to work at task v and firm v makes positive profits, i.e. $w_1 t_1 = \rho(v)p_k(v, t_k) > \rho(v)p_k(v, t_k') > w_1' t_1' > w_1 t_1' = \rho(v')p_k(v', t_k')$. To overcome this problem, the following two sets of first order conditions must be satisfied:

$$\rho(v)\frac{\partial p_k(v, t_k)}{\partial t_k} = w_k \; \forall k = 1, 2 \tag{2}$$

Plugging equation 1 into equation 2 yields:

[12] Within each educational group, high skilled workers have an asbolute advantage.

$$\frac{\partial p_k(v,t_k)}{\partial t_k} = \frac{p_k(v,t_k)}{t_k} \; \forall k = 1,2 \tag{3}$$

Note that it is straightforward from equation 3 that $p(v, t_k)$ could be any polynomial function of t_k.

By a similar argument, to ensure that firms do not switch tasks the following first order condition must be satisfied:

$$\begin{aligned} 0 &= \rho'(v)p_k(v,t_k) + \rho(v)\frac{\partial p_k(v,t_k)}{\partial v} \\ &\Leftrightarrow \\ \rho(v) &= -\rho'(v)\frac{p_k(v,t_k)}{\frac{\partial p_k(v,t_k)}{\partial v}} \end{aligned} \tag{4}$$

2.2 Endogenous human capital formation: educational self-selection

Individuals can choose between various education (no education being one of them) in order to transform their endowed abilities into productive skills so as to maximize their expected future income. Prior to this educational choice, individuals evaluate the streams of expected earnings associated to each educational possibility. Since future wages are unobserved this complicates the assignment problem considerably. Most authors, see Friedman and Kutznets (1945), Mincer (1974 and 1993), Willis and Rosen (1979) and Heckman et al. (1998) for instance, have assumed that workers have perfect foresight. With perfect foresight, workers anticipate every shocks taking place in the future and therefore never make forecasting errors. This assumption simplifies considerably the model as it leads to strict specialization of workers. Workers who selected education k supply skills of type k as I will show below.

In contrast, imperfect information in the form of limited foresight[13] opens new assignment possibilities as labor market shocks might be unexpected. For instance, think of a worker whose manual ability is as large as her cognitive ability. Suppose this worker can select a manual degree and end up with 10% more manual skills than cognitive skills or select a cognitive degree and end up with 10% more cognitive skills than manual skills. Furthermore, this worker *expects* to earn \$1 per unit of manual skills supplied and *expects* to earn \$1.2 per unit of cognitive skills supplied. Following her comparative advantage, this worker will select a cognitive degree to maximize her *expected* income. Suppose that at the time this worker graduates and enters the labor market, an unexpected

[13] See Siow (1984) for a comprhensive treatment of occupational choice under uncertainty.

shock occurs so that the pay-rate is \$1.2 per units of manual skills supplied and \$1 per unit of cognitive skills supplied. This worker's income will be maximized by supplying her manual skills full-time *rather* than her cognitive skills for which she studied (relative manual wage is $1.2 > 1.1$ relative manual skills). Hence, with limited foresight, some workers will not supply the skills they developed most during education, but rather supply their alternative skills. Obviously, this result holds as long as the difference between expected wages and actual wages is large enough to compensate relative skills differential *ex post* t_k/t_j, i.e. *ex ante* ability differential a_k/a_j times the contribution of education chosen to increase relative skills, e_{kk}/e_{kj}. Hence, if education increases one type of skills much more than the other type, few workers will be in a situation where supplying their alternative skills generates more income than supplying the skills they developed most during education. This means that the more unequally education enhances skills of the various types, the less restrictive the assumption of perfect foresight.

The limited foresight world deserves particular attention. This is on my personal agenda for future extensions of the tasks assignment model with endogenous human capital formation.[14] In this paper however, we follow the mainstream of the literature, e.g. Willis and Rosen (1979) and assume that individuals have perfect foresight so that expected earnings are by definition equal to actual earnings.

Specialization

Workers with skills $\langle t_1, t_2 \rangle$ can supply t_1 units of skills of type 1 for a share τ of their working time and supply t_2 units of skills of type 2 for a share $1 - \tau$ of their working time. Workers receive w_1 dollars per units of skills of type 1 they supply and w_2 dollars per unit of skills of type 2. Workers' income is therefore equal to $\tau w_1 t_1 + (1 - \tau) w_2 t_2$. Hence, workers will maximize income by specializing and supplying full time their skills of type 1 or their skills of type 2 according to $\frac{t_2}{t_1} \gtrless \mu$. This means that a worker with abilities $a = \langle a_1, a_2 \rangle$ so that $a_k e_{kk} w_k > a_j e_{jj} w_j$ selects education k and will never supply her skills of type j since $e_{kj} < e_{jj}$ and hence, $a_j e_{jj} w_j > a_j e_{kj} w_j$.[15] We have:

$$\begin{cases} \tau = 0 \text{ if } \frac{t_2}{t_1} > \mu \\ \tau = 1 \text{ if } \frac{t_2}{t_1} \leq \mu \end{cases}$$

Educational choice

[14] Note that some form of imperfect information in assignment models is treated in MacDonald (1982). MacDonald extends the comparative advantage model to labor markets with incomplete information about the types of workers.

[15] Note that w_k is the return to each unit of skill of type k supplied in the labor market. It also is the return to education k when poeple have perfect foresight. With perfect foresight, workers specialize and supply only one type of skills. Then workers with education k supply skills of type k so that w_k is equivelent to the return to educatuion k.

Workers with abilities $\langle a_1, a_2 \rangle$ will either select education 1 and supply $t_1 = e_{11}a_1$ or select education 2 and supply $t_2 = e_{22}a_2$. Workers select their education such as to maximize income. Educational choice is therefore governed by the principle of comparative advantage.[16] Workers with abilities $\langle a_1, a_2 \rangle$ therefore select education 1 if and only if $\frac{t_2}{t_1} = \frac{e_{22}a_2}{e_{11}a_1} < \mu$. Given an arbitrary relative price μ, all workers whose relative skills $\frac{t_2}{t_1}$ exceeds μ optimally select education 2 and all workers with comparative advantage less than μ select education 1. Workers with skills ratio equal to μ are indifferent between education 1 and 2 and are arbitrarily distributed to one or the other.

Defining $\xi(t_1, t_2)$ the density of workers whose skills are $\langle t_1, t_2^- \rangle$ if they select education 1 and $\langle t_1^-, t_2 \rangle$ if they select education 2,[17] the density of workers supplying t_1 is obtained by summing up all workers selecting education 1, that is all workers with $\frac{t_2}{t_1} < \mu$. Let $t_2^* = \mu t_1$. All workers with potential skills level $t_2 < t_2^*$ select education 1 and the others education 2. The density of workers supplying level t_1 of type 1 skills, i.e. $s_1(\mu, t_1)$, and the density of workers supplying level t_2 of type 2 skills, i.e. $s_2(\mu, t_2)$, is therefore defined parametrically by:

$$
\begin{aligned}
s_1(\mu, t_1) &= \int_{\underline{t_2}}^{t_2^*} \xi(t_1, t_2) dt_2 &(5)\\
s_2(\mu, t_2) &= 1 - s_1(\mu, t_1) &(6)
\end{aligned}
$$

where $\underline{t_k}$ is the minimum potential skills of type k in the workers population and $t_2^* = \mu t_1$.

In the economy, the supply of workers by education is given by:

$$
\begin{aligned}
S_1(\mu) &= \int_{\underline{t_1}}^{\overline{t_1}} s_1(\mu, t_1) dt_1 &(7)\\
S_2(\mu) &= \int_{\underline{t_2}}^{\overline{t_2}} s_2(\mu, t_2) dt_2 &(8)
\end{aligned}
$$

where $\overline{t_k}$ is the maximum potential skills of type k in the workers population.

[16] The theory of comparative advantage in labor markets was formally developed by Sattinger (1975) (see also Sattinger (1993) for a survey of assignment models and comparative advantage) and the presence of comparative advantage was later demonstrated empirically in Sattinger (1978 and 1980).

[17] With $t_1 > t_1^-$ and $t_2 > t_2^-$ by definition of E_1 and E_2.

As in Roy's self-selection model, the second moments of the density distribution $\xi(t_1, t_2)$ are determinant for the educational selection. Defining the variance of each type of potential skills j χ_j, the covariance χ_{12} and the correlation $\vartheta_{12} = \frac{\chi_{12}}{\chi_1 \chi_2}$, two cases can be distinguished.

1. $\frac{\chi_j}{\chi_k} > \vartheta_{12} > \frac{\chi_k}{\chi_j}$. There is hierarchical sorting.[18] Workers that select education j are drawn from the upper tail of the distribution of t_j (type of skills the less concentrated, i.e. $\chi_j > \chi_k$) whereas workers that select education k are drawn from the lower tail of the distribution of t_k.

2. $\frac{\chi_k}{\chi_j} > \frac{\chi_j}{\chi_k} > \vartheta_{12}$. There is no hierarchical sorting and workers that select education j tend to have high t_j skills, $\forall j = 1, 2$.

To illustrate graphically educational self-selection under hierarchical and non hierarchical sorting, it is helpful to introduce the notion of "worker's genre". A worker's genre regroups all workers of a particular level of skills of type 1 but different levels of skills of type 2.

Definition 1 *All N_α workers with potential skills $\left\langle \left\langle t_1^{\alpha_1}, t_2^{\alpha_1} \right\rangle, ..., \left\langle t_1^{\alpha_{N_\alpha}}, t_2^{\alpha_{N_\alpha}} \right\rangle \right\rangle$ so that $t_1^{\alpha_i} = t_1^{\alpha}$ for all $i = 1, ..., N_\alpha$ belong to genre α. Moreover, without loss of generality, it is assumed that $t_2^{\alpha_1} > t_2^{\alpha_2} > ... > t_2^{\alpha_{N_\alpha}}$*

In figure 1, I draw the joint density distribution $\xi(t_1, t_2)$ for $\chi_2 > \chi_1$ and $\frac{\chi_2}{\chi_1} > \vartheta_{12} > \frac{\chi_1}{\chi_2}$ so that there is hierarchical sorting and potential skills of type 2 are less concentrated than potential skills of type 1. The potential skills of type 2 are reported on the horizontal axis and the potential skills of type 1 on the vertical axis. Each horizontal line represents a different genre of workers. As drawn, workers of genre α possess the highest skills of type 1 ($t_1^{\alpha} = \overline{t_1}$) and workers of genre γ the lowest ($t_1^{\gamma} = \underline{t_1}$).[19] Note, however, that since there is a significant positive correlation between the potential level of skills of each type, workers of genre α will generally have more skills of type 2 than workers of another genre and workers of genre γ will generally have less potential skills of type 2 than workers of another genre. Although there might be some overlap in the skills of type 2 between workers of different genres. For instance, as shown in Figure 1, workers β_{N_β} have potentially more skills of type 1 than workers γ_1 but potentially less skills of type 2 whereas β_1 workers have more of both types of skills. The line of slope w_2/w_1 passing through the origin is the selection index. It corresponds to the ratio of skills of type 1 to skills of type 2 so that a worker

[18] In the extreme case with $\vartheta_{12} = 1$ and $\chi_j > \chi_k$, although we distinguish between skills of type k and j, there is only one type of skills. Moreover, the highest skilled workers of type k are clones of the lowest skilled workers of type j. This yields a one-to-one relationship between the level of skills and wages and the model looks like those of Teulings (1995a, 1995b and 2005).

[19] For the sake of completness, $t_2^{\alpha_1} = \overline{t_2}$ and $t_2^{\gamma_{N_\gamma}} = \underline{t_2}$.

is indifferent between selecting education 1 or 2. A worker whose potential skills point $\langle t_1, t_2 \rangle$ lies above the index line selects education 1 and supplies her skills of type 1 and vice versa. As drawn in Figure 1, at the equilibrium wage rate w_2/w_1, all workers of genre α, i.e. $\langle \alpha_1, ..., \alpha_{N_\alpha} \rangle$ will select education 2 and supply their skills of type 2, $t_2^{\alpha_1} > ... > t_2^{\alpha_{N_\alpha}}$ and all workers of genre γ, i.e. $\left\langle \gamma_1, ..., \gamma_{N_\gamma} \right\rangle$, will select education 1 and supply their skills of type 1, $t_1^{\gamma_1} = ... = t_1^{\gamma_{N_\gamma}}$. In contrast, for some other genres, workers will select education 1 while others will select education 2. For instance, while workers $\beta_2, ..., \beta_{N_\beta}$ will select education 1, workers β_1 are indifferent and might select education 2.

Note that workers that select education 2 are drawn from the right tail of the distribution of skills of type 2 (the less concentrated skill in the population) while workers that select education 1 are drawn from the lower tail of the distribution of skills of type 1 (the most concentrated in the population), there is hierarchical sorting. In contrast, in Figure 2 the joint distribution shows a negative correlation between types of skills. Educational selection results into workers with relatively high skills of type 1 selecting education 1 and workers of relatively high skills of type 2 selecting education 2.

Finally, similar to Roy's model, in the presence of comparative advantage the wage distribution will be skewed relative to the distribution of potential skills. In the case of equality of comparative advantage, when $\chi_1 = \chi_2$ and $\vartheta_{12} = 1$, workers will select education at random and the distribution of wages will map on the distribution of potential skills.

2.3 Tasks assignment and labor market equilibrium

Given the structure of comparative advantage, output will be maximized by assigning workers with education 1 to tasks $(0, \varepsilon)$ and workers with education 2 to tasks $(\varepsilon, 1)$ where ε is the marginal task in equilibrium.[20] This means that in Figure 1 and 2, a worker above the index line is assigned a task $v \in (0, \varepsilon)$ and a worker below the index line a task $v \in (\varepsilon, 1)$. Among workers with education 1, those with the highest level of skill 1 will be assigned to task 0 and so on until the marginal task ε is assigned to those workers with education 1 supplying the lowest level of skill 1. By symmetry, the tasks $(\varepsilon, 1)$ are assigned to workers with education 2. Workers with education 2 supplying the lowest level of skills of type 2 are assigned to task ε and so on until those workers with the highest level of skills 2 are assigned to task 1.

Note that, under hierarchical sorting, workers with education 2 supplying the lowest level of type 2 skills and workers with education 1 supplying the lowest level of type 1 skills will have different genres, i.e. all workers of the γ genre

[20]Firms are indifferent between assigning a worker with education 1 or education 2 at task ε.

and β_1 workers in the hierarchical sorting case in Figure 1, whereas in the non hierarchical sorting case, they have the same genre, i.e. β_{N_β} workers in Figure 2.

The tasks assignment results in a mapping function g_1 that associates to each task $v \in (0, \varepsilon)$ a single value of skills t_1 depending on the relative wage rate μ, i.e. $t_1 = g_1(\mu, v)$, and a mapping function g_2 that associates to each task $v \in (\varepsilon, 1)$ a single value of skills t_2, i.e. $t_2 = g_2(\mu, v)$.[21] Moreover, under non hierarchical sorting we have $g_1(\mu, \varepsilon) = g_2(\mu, \varepsilon)$. In Figure 3, I illustrate graphically the task assignment under hierarchical sorting as presented in Figure 1. The horizontal axis represents the tasks v and the vertical axis the productivity of the various workers, $\alpha_1, ..., \alpha_{N_\alpha}, \beta_1, ..., \gamma_{N_\gamma}$. The dashed curves represent the productivity of the various workers when supplying their type 1 skills (downward sloping) or their type 2 skills (upward sloping). The assignment of tasks, represented by bullets in Figure 3, takes place as follows. Among those workers that select education 2 and supply their skills of type 2, workers α_1 have the highest level of skills. These workers are therefore assigned to task 1. Moving from 1 to the left, the following tasks are then assigned to α_2, α_3 etc. As drawn, workers α_{N_α} are indifferent between education 1 and 2 and are therefore distributed at random between education 1 and 2. Those that select education 2 are assigned to the next task close to $1-$ the others, as drawn in Figure 3, have the highest type 1 skills and are assigned to task $0-$. The marginal task is assigned to those workers with education 2 supplying the lowest level of skills of type 2, i.e. β_1 workers and those workers with education 1 supplying the lowest level of skills of type 1, i.e. workers of the γ genre. By symmetry, tasks to the left of the marginal task are assigned to workers with education 1. Moving from the left to the right and connecting the various bullets yields the mapping functions g_1 and g_2.

To show how the assignment equilibrates the supply of and demand for skills, I split the interval of tasks assigned to workers with education 1, $v \in (0, \varepsilon)$ into N intervals of equal length $\Delta_i = \Delta = \frac{\varepsilon}{N}$ for $i = 1, ..., N$. Labor demand in interval i, that is the number of jobs to be filled in the interval i, is $\int_{(i-1)\Delta}^{i\Delta} d(v)dv$ where $d(v)$ is the density distribution of tasks. Therefore, to equilibrate each tasks interval firms will assign the $\int_0^{\Delta} d(v)dv$ most skilled workers to the first interval, the following $\int_{\Delta}^{2\Delta} d(v)dv$ most skilled workers to the second interval and so on and so forth until the last interval is filled with the $\int_{(N-1)\Delta}^{N\Delta} d(v)dv$ least skilled workers. Note however that, since the density distribution of workers by level of skills needs not to correspond to the density distribution of tasks, the skill differences between the most and least skilled workers in each interval needs not be the same. For instance, suppose that skills are normally distributed among workers with education 1 and tasks are uniformly distributed on $(0, 1)$. The skill differential will first decrease (upper tail of a normal distribution is

[21] See Sattinger (1975). Functions g_i, $i = 1, 2$ play the same role as the function $h(g)$ in Sattinger (1975) p. 356 where g is workers' ability (single scale) and $h(g)$ the difficulty (single scale) of the task performed by workers with ability g in equilibrium.

thinner than the upper tail of a uniform distribution) and increase once the median skilled worker has been assigned. To put it in Tinbergen's terms "the supply distribution of skills has to be deformed so as to coincide with the demand distribution otherwise there will not be an equilibrium." The assignment of tasks to workers deforms (stretches) the density distribution of skills of each type such as to make it fit the distribution of tasks in equilibrium.

Therefore, to match the density of workers' skills with the density of tasks for both education 1 and 2, firms adjust the length of each of the N skills interval, $\Delta t_{1,i}$ and $\Delta t_{2,i}$ for $i = 1, ..., N$, so that:

$$\int_{(i-1)\Delta}^{i\Delta} d(v)dv = \int_{\overline{t_1}-i\Delta t_{1,i}}^{\overline{t_1}-(i-1)\Delta t_{1,i}} s_1(\mu, t_1)dt_1 \text{ for } i = 1, ..., N \tag{9}$$

$$\int_{1-i\Delta}^{1-(i-1)\Delta} d(v)dv = \int_{\overline{t_2}-i\Delta t_{2,i}}^{\overline{t_2}-(i-1)\Delta t_{2,i}} s_2(\mu, t_2)dt_2 \text{ for } i = 1, ..., N \tag{10}$$

As the number of intervals N increases to infinity, the length of the interval Δ tends to infinitesimal values, i.e. tends to dv, so that the density distribution of workers' skills is deformed to fit the density distribution of tasks. A small interval of skills Δt_k of mass $d(v)$ is associated to each infinitesimal interval of task dv. The size of each of these skills intervals is actually derived from the functions g_1 and g_2 by simply deriving with respect to v, $\frac{dt_k}{dv} \equiv \frac{\partial g_k}{\partial v}$. The density of workers in the interval of tasks dv is equal to $d(v)$ and hence, the density of workers in the interval dt_k is equal to $d(v)$, which by definition is equal to $s_k(\mu, t_k)$. Therefore, the assignment of workers to tasks equilibrates the supply of workers with education k to the demand for workers with education k:

$$S_1(\mu) = \int_{g_1(\mu,\varepsilon)}^{g_1(\mu,0)} s_1(\mu, t_1)dt_1 \equiv \int_0^{\varepsilon} d(v)dv = F(\varepsilon) = D_1(\varepsilon) \tag{11}$$

$$S_2(\mu) = \int_{g_2(\mu,\varepsilon)}^{g_2(\mu,1)} s_2(\mu, t_2)dt_2 \equiv \int_{\varepsilon}^{1} d(v)dv = 1 - F(\varepsilon) = D_2(\varepsilon) \tag{12}$$

where $F(x) = \Pr(v < x)$ is the cumulative distribution function of tasks.

An important implication of this result is that given the distribution of tasks, one could either make assumptions on the distribution of skills and derive the assignment functions g_k, $k = 1, 2$, or make assumptions on the shape of g_k and derive the shape of the resulting skills distributions s_k, $k = 1, 2$.

Using the function g_k, the demand for workers with education k per unit of output is directly derived by integrating the number of workers per unit of output for each skills type (the inverse of the productivity of workers) over the spectrum of tasks. The demand for workers with education 1 and 2 per unit of output is then given by:

$$\frac{D_1(\mu,\varepsilon)}{Y} = \int_0^{\varepsilon} \frac{d(v)}{p_1(g_1(\mu,v),v)} dv \tag{13}$$

$$\frac{D_2(\mu,\varepsilon)}{Y} = \int_{\varepsilon}^{1} \frac{d(v)}{p_2(g_2(\mu,v),v)} dv \tag{14}$$

Solving for the marginal task ε one can derive, under invertable conditions, the shape of the production function linking the output level Y to the employment level of workers with education 1 and 2, $Y = h(L_1, L_2)$.

Equilibrium is achieved by equalizing the relative supply of labor by education with the relative demand for labor by education and noting that the ratio of equation 11 by equation 12 should be equal to the ratio of equation 13 by equation 14 in equilibrium.

$$\frac{S_1(\mu)}{S_2(\mu)} \equiv \frac{D_1(\varepsilon)}{D_2(\varepsilon)} = \frac{D_1(\mu,\varepsilon)}{D_2(\mu,\varepsilon)} \tag{15}$$

Moreover, firms are indifferent between assigning the worker with education 1 supplying the lowest skills of type 1, $t_{1,\varepsilon}$, or the worker with education 2 supplying the lowest skills of type 2, $t_{2,\varepsilon}$ to the marginal task ε. Therefore, the relative wage rate is equal to the ratio of productivity of the worker supplying $t_{1,\varepsilon}$ and assigned at task ε to the productivity of the worker supplying $t_{2,\varepsilon}$ assigned at task ε divided by the ratio of relative skills:[22]

$$\begin{aligned} 0 &= \ln \frac{\rho(\varepsilon)}{\rho(\varepsilon)} + \ln \frac{p_1(t_{1,\varepsilon},\varepsilon)}{p_2(t_{2,\varepsilon},\varepsilon)} - \left(\ln \mu + \ln \frac{t_{1,\varepsilon}}{t_{2,\varepsilon}}\right) \\ &\Leftrightarrow \\ \ln \mu &= \ln \frac{p_1(t_{1,\varepsilon},\varepsilon)}{p_2(t_{2,\varepsilon},\varepsilon)} - \ln \frac{t_{1,\varepsilon}}{t_{2,\varepsilon}} \end{aligned} \tag{16}$$

2.4 Wage distribution

Overall wage distribution

In contrast to the assignment model proposed by Teulings (1995a, 1995b and 2005) where there is a one-to-one relationship between wages and skills, an

[22]Note that the relative wage rate can be indifferently derived from the null profit condition or the first order condition defined by equation 2 are also satisfied.

important feature of the model presented in this paper is that the distribution of wages of workers with education j will (generally)[23] overlap to some extent with the distribution of wages of workers with education k. This is a very interesting feature since in real data the distribution between educational groups of workers do overlap, i.e. workers in the upper tail of the wage distribution of high-school graduates have higher wages than workers in the lower tail of the wage distribution of college graduates.

Since the distribution of wages of both types of workers overlap, to derive the overall wage distribution, with cumulative distribution $F_w(w^*)$, one needs to derive the skills of type 2, say t_2^*, a worker with education 2 needs to supply in order to get the same wage as a worker with education 1 and skills t_1 and the other way around. One needs to separate two cases: the case where $w_1 t_{1,\varepsilon} < w_2 t_{2,\varepsilon}$ and the case where $w_1 t_{1,\varepsilon} \geq w_2 t_{2,\varepsilon}$. In the case where $w_1 t_{1,\varepsilon} \geq w_2 t_{2,\varepsilon}$, since the least skilled worker with education 1 gets $w_1 t_{1,\varepsilon}$, only workers with education 2 earn less than $w_1 t_{1,\varepsilon}$. To earn this wage, a worker with education 2 needs to supply skills[24] $t_{2,\varepsilon}^* = \mu t_{1,\varepsilon}$ or $v_{2,\varepsilon}^* = 1 - g_2^{-1}(\mu t_{1,\varepsilon})$. Workers with both types of education earns wages above the threshold $w_1 t_{1,\varepsilon}$ though in different proportion. From this result, I derive the distribution of wages, for $w^* = w_1 t_1 \equiv w_2 t_2^*$, as follows:

$$F_w(w^*) = \begin{cases} F(v_2^*) - F(\varepsilon) \text{ if } v_2^* < v_{2,\varepsilon}^* \\ \underbrace{F(\varepsilon) - F(v_1)}_{\text{Education 1}} + \underbrace{F(v_2^*) - F(\varepsilon)}_{+\text{ Education 2}} = \underbrace{F(v_2^*) - F(v_1)}_{=\text{ All workers}} \text{ if } v_2^* \geq v_{2,\varepsilon}^* \end{cases} \tag{17}$$

The same type of result can be derived in the case where $w_1 t_{1,\varepsilon} < w_2 t_{2,\varepsilon}$. In this case, $t_{1,\varepsilon}^* = \frac{1}{\mu} t_{2,\varepsilon}$ or equivalently,[25] $v_{1,\varepsilon}^* = 1 - g_1^{-1}(\frac{1}{\mu} t_{2,\varepsilon})$. Workers with both types of education earns wages above the threshold $w_2 t_{2,\varepsilon}$ though in different proportion. From this result, I derive the distribution of wages, for $w^* = w_1 t_1^* \equiv w_2 t_2$:

$$F_w(w^*) = \begin{cases} F(\varepsilon) - F(v_1^*) \text{ if } v_1^* > v_{1,\varepsilon}^* \\ \underbrace{F(\varepsilon) - F(v_1^*)}_{\text{Education 1}} + \underbrace{F(v_2) - F(\varepsilon)}_{+\text{ Education 2}} = \underbrace{F(v_2) - F(v_1^*)}_{=\text{ All workers}} \text{ if } v_1^* \leq v_{1,\varepsilon}^* \end{cases} \tag{18}$$

[23] See footnote 18.

[24] In the CES case presented later in this paper we have $v_{2,\varepsilon}^* = 1 - g_2^{-1}(\mu t_{1,\varepsilon}) = 1 - \varepsilon \left(\mu \frac{\beta_2}{\beta_1}\right)^{-1/\alpha_2}$.

[25] $v_{1,\varepsilon}^* = 1 - g_1^{-1}(\frac{1}{\mu} t_{2,\varepsilon}) = (1-\varepsilon)\left(\frac{1}{\mu}\frac{\beta_1}{\beta_2}\right)^{-1/\alpha_1}$ for the CES case.

where $v_1^* = 1 - g_1^{-1}(\frac{1}{\mu}t_2)$.

Residual wage distribution

In the literature, an other measure of wage inequality often used is the so-called within-group wage inequality. Since this measure is based on the residuals of earnings regressions, it is also called residuals inequality. This measure refers to the wage inequality once adjusted for observable differences like education. In the model, residual wage inequality can be inferred directly from the distribution of tasks once we control for wage differences due to educational choice.

To derive the residual wage distribution, with cumulative distribution $F_{rw}(w^r)$, one needs to derive the skills of type 2, say t_2^r, a worker with education 2 needs to supply in order to get the same wage as a worker with education 1 and skills t_1 and the other way around once correcting for the skill premium μ, that is multiplying the wage of workers with education 2 by μ. One needs to separate two cases: the case where $t_{1,\varepsilon} < t_{2,\varepsilon}$ and the case where $t_{1,\varepsilon} \geq t_{2,\varepsilon}$. In the case where $t_{1,\varepsilon} \geq t_{2,\varepsilon}$, since the least skilled worker with education 1 gets $t_{1,\varepsilon}$, only workers with education 2 earn less than $t_{1,\varepsilon}$. To earn this wage, a worker with education 2 needs to supply skills[26] $t_{2,\varepsilon}^r = t_{1,\varepsilon}$ or $v_{2,\varepsilon}^r = 1 - g_2^{-1}(t_{1,\varepsilon})$. Workers with both types of education earns wages above the threshold $t_{1,\varepsilon}$ though in different proportion. From this result, I derive the residual distribution of wages, for $w^r = t_1 \equiv t_2^r$, as follows:

$$F_{rw}(w^r) = \begin{cases} F(v_2^r) - F(\varepsilon) \text{ if } v_2^r < v_{2,\varepsilon}^r \\ \underbrace{F(\varepsilon) - F(v_1)}_{\text{Education 1}} + \underbrace{F(v_2^r) - F(\varepsilon)}_{+\text{ Education 2}} = \underbrace{F(v_2^r) - F(v_1)}_{=\text{ All workers}} \text{ if } v_2^r \geq v_{2,\varepsilon}^r \end{cases} \tag{19}$$

The same type of result can be derived in the case where $t_{1,\varepsilon} < t_{2,\varepsilon}$. In this case, $t_{1,\varepsilon}^r = t_{2,\varepsilon}$ or equivalently,[27] $v_{1,\varepsilon}^r = 1 - g_1^{-1}(t_{2,\varepsilon})$. Workers with both types of education earns wages above the threshold $t_{2,\varepsilon}$ though in different proportion. From this result, I derive the distribution of wages, for $w^r = t_1^r \equiv t_2$:

[26] $v_{2,\varepsilon}^r = 1 - g_2^{-1}(t_{1,\varepsilon}) = 1 - \varepsilon \left(\frac{\beta_2}{\beta_1}\right)^{-1/\alpha_2}$ for the CES case.

[27] $v_{1,\varepsilon}^r = 1 - g_1^{-1}(t_{2,\varepsilon}) = (1 - \varepsilon) \left(\frac{\beta_1}{\beta_2}\right)^{-1/\alpha_1}$ for the CES case.

$$F_{rw}(w^r) = \begin{cases} F(\varepsilon) - F(v_1^r) \text{ if } v_1^r > v_{1,\varepsilon}^r \\ \underbrace{F(\varepsilon) - F(v_1^r)}_{\text{Education } 1} + \underbrace{F(v_2) - F(\varepsilon)}_{+\text{ Education } 2} = \underbrace{F(v_2) - F(v_1^r)}_{=\text{ All workers}} \text{ if } v_1^r \leq v_{1,\varepsilon}^r \end{cases} \tag{20}$$

where $v_1^r = 1 - g_1^{-1}(t_2)$.

3 The case where h has the CES form

A general drawback of hedonic models is that the function h does not exist. As indicated in Table 1, among general equilibrium models, only Teulings (1995 and 2005) propose (well-) known functional forms, i.e. Constant Elasticity of Substitution CES and DIstance Dependent Elasticity of Substitution. In models with exogenous human capital formation like Teulings (2005), the functional form of the skills distribution is imposed for it is convenient and not too restrictive. However, educational choice implies that the distribution of endowed abilities is endogenously distorted to form the distribution of skills. This distortion is very hard to model with well-know probability distributions without imposing very strong assumptions (distortion affects the mean and variance but not the skewness or Kurtosis for instance as in Tinbergen (1956) and Teulings (1995,2005)). To get around this issue, I impose the shape of the production function and the distribution of tasks and *derive* the resulting distributions of abilities and skills rather than imposing restrictions on the functional form of the distribution of both abilities and tasks. Unlike hedonic models with exogenous human capital formation that infer the function h given the shape of the distribution of abilities, I infer the shape of the distribution abilities given the function h.

3.1 Density of skills and supply of labor by education

Suppose that the function assigning tasks v to skills of type 1 is such that $t_1 = g_1(\mu, v) = \frac{v^{-\alpha_1}}{\beta_1(\mu)}$ with $\alpha_1 \geq 0$ and the function assigning tasks v to skills of type 2 is such that $t_2 = g_2(\mu, v) = \frac{(1-v)^{-\alpha_2}}{\beta_2(\mu)}$ where $\alpha_2 \geq 0$. Note that the parameters α_k indicate the percentage increase in the type 1 (respectively type 2) skills of workers assigned to a task situated 1 percent more to the left (respectively to the right), i.e. $\frac{\partial \ln t_1}{\partial \ln v} = -\alpha_1$ and $\frac{\partial \ln t_2}{\partial \ln(1-v)} = -\alpha_2$.[28] The parameter α_k therefore measures the degree of heterogeneity between workers with same education. The educational selection rule, i.e. $\frac{t_2}{t_1} > \mu$ select education 2, indicates that the minimum level

[28]Note that for $\alpha_1 = \alpha_2 = 0$, workers with education k have homogenous skills $t_k = \frac{1}{\beta_k(\mu)}$. Therefore, for $\alpha_1 = \alpha_2 = 0$, and $\beta_k(\mu) = c_k$ independent of μ, the model reduces to Rosen's (1978) task assignment model.

of skills of type 2 of workers who select education 2, $t_2^* = \mu t_1$ increases with the relative wage rate. An increase in the relative wage rate will lead workers with relatively lower level of type 1 skills to select education 1, $\frac{\partial t_1}{\partial \mu} > 0$, and workers with relatively higher level of type 2 skills to select education 2, $\frac{\partial t_2}{\partial \mu} < 0$. This implies that $\beta_1^{'} < 0$ and $\beta_2' > 0$.

I assume further that the distribution of tasks follows a Beta distribution on $(0, 1)$ and has density function $d(v) = Av^{d_1}(1-v)^{d_2}$ with $d_j > 0$ and $A = 1/B(d_1+1, d_2+1)$ and $B(.)$ is the Beta function and cumulative distribution $F(v^*) = \Pr(v < v^*) = \int_0^{v^*} d(v)dv$. The mean task is given by $E[v] = \frac{d_1+1}{2+d_1+d_2}$ ($E[v] = \frac{1}{2}$ when $d_1 = d_2$) and the variance equals $V[v] = \frac{(d_1+1)(d_2+1)}{(2+d_1+d_2)(3+d_1+d_2)}$, with $\frac{\partial V[v]}{\partial d_k} < 0$. Moreover, the distribution is skewed toward 0 when $d_1 > d_2$ and vice versa. The Beta distribution is appealing because it ranges between 0 and 1, it has only two parameters, and its shape is flexible. If $d_1 > 0$ and $d_2 > 0$ the distribution is unimodal. If $d_1 = d_2 = d$ and $d = 0$ tasks are uniformly distributed. Moreover, for $d > 1$ the Beta distribution and the normal distribution with average $\frac{1}{2}$ and variance equal to $\frac{(d_1+1)(d_2+1)}{(2+d_1+d_2)(3+d_1+d_2)}$ look alike.

The total supply of workers with education 1 and 2 is equal to the density of tasks between 0 and ε, i.e. ε given the distribution of tasks, and the density of tasks between ε and 1 respectively.

$$
\begin{aligned}
S_1(\mu) &= \int_{g_1(\mu,\varepsilon)}^{g_1(\mu,0)} s_1(\mu, t_1)dt_1 \equiv \int_0^{\varepsilon} d(v)dv = F(\varepsilon) = D_1(\varepsilon) && (21) \\
S_2(\mu) &= \int_{g_2(\mu,\varepsilon)}^{g_2(\mu,1)} s_2(\mu, t_2)dt_2 \equiv \int_{\varepsilon}^{1} d(v)dv = 1 - F(\varepsilon) = D_2(\varepsilon) && (22)
\end{aligned}
$$

3.2 Productivity and demand for labor by education

Suppose further that the productivity of workers with skills k at task v is given by:

$$
\begin{aligned}
p_1(t_1, v) &= b_1 t_1 (1-v)^{d_2} v^{1-a+\alpha_1+d_1} && (23) \\
p_2(t_2, v) &= b_2 t_2 (1-v)^{1-a+\alpha_2+d_2} v^{d_1} && (24)
\end{aligned}
$$

where b_k are parameters indicating the efficiency units of workers with skills k and a is the parameter that governs the comparative advantage structure in the economy as will be shown below.

Note that the productivity functions are linear in skills so that they satisfy the first order conditions in equation 2.

In equilibrium using the assignment functions $t_k = g_k(\mu, v)$ we have:

$$p_1(g_1(\mu, v), v) = \frac{b_1}{\beta_1(\mu)}(1-v)^{d_2} v^{1-a+d_1} \quad (25)$$

$$p_2(g_2(\mu, v), v) = \frac{b_2}{\beta_2(\mu)}(1-v)^{1-a+d_2} v^{d_1} \quad (26)$$

The productivity of workers with education 1 and skills t_1 relative to workers with education 2 and skills t_2 at task v expressed in logarithm is equal to:

$$\ln \frac{p_1(t_1, v)}{p_2(t_2, v)} = \ln \frac{b_1}{b_2} + \ln \frac{\beta_2(\mu)}{\beta_1(\mu)} + (1-a) \ln \frac{v}{1-v} \quad (27)$$

Note that although the productivity of workers with education 1 and 2 depend on the distribution parameters d_1 and d_2 the relative productivity does not. The parameter a indicates the curvature of the distribution of relative productivity. The smaller a, the smaller the comparative advantage of workers with education 2 in tasks close to 1 and the larger the comparative advantage of workers with education 1 in tasks close to 0. When a tends to 1, the relative productivity of workers with education 1 is constant in every tasks, there is equity of comparative advantage. Note that equations 23 and 24 satisfy $\partial p_k(t_k, v)/\partial t_k > 0$. Equations 23 and 24 satisfy $\partial p_1(t_1, v)/\partial v < 0$ and $\partial p_2(t_2, v)/\partial v > 0$ if and only if $a > 1 + \alpha_1 + d_1$ and $a > 1 + \alpha_2 + d_2$. To satisfy the comparative advantage structure I therefore impose $a > 1 + \max_k(\alpha_k + d_k)$

Given the structural form in equation 23 and 24, the employment of workers with education k per unit of output reads as:

$$\frac{D_1(\mu, \varepsilon)}{Y} = \int_0^{\varepsilon} \frac{d(v)}{p_1(v)} dv = A \int_0^{\varepsilon} \frac{\beta_1(\mu)}{b_1} v^{a-1} dv = A \frac{\beta_1(\mu)}{b_1} \varepsilon^a \quad (28)$$

$$\frac{D_2(\mu, \varepsilon)}{Y} = \int_{\varepsilon}^{1} \frac{d(v)}{p_2(v)} dv = A \int_{\varepsilon}^{1} \frac{\beta_2(\mu)}{b_2} (1-v)^{a-1} dv = A \frac{\beta_2(\mu)}{b_2} (1-\varepsilon)^a \quad (29)$$

Solving the system for the marginal task ε yields:

$$\varepsilon = \left(\frac{D_1(\mu, \varepsilon)}{Y} \frac{1}{A} \frac{b_1}{\beta_1(\mu)} \right)^{\frac{1}{a}} = 1 - \left(\frac{D_2(\mu, \varepsilon)}{Y} \frac{1}{A} \frac{b_2}{\beta_2(\mu)} \right)^{\frac{1}{a}}$$

$$\Leftrightarrow$$

$$Y = \frac{1}{A} \left[\left(\frac{b_1}{\beta_1(\mu)} D_1(\mu, \varepsilon) \right)^{\frac{1}{a}} + \left(\frac{b_2}{\beta_2(\mu)} D_2(\mu, \varepsilon) \right)^{\frac{1}{a}} \right]^a \quad (30)$$

Equation 30 reads as a CES production function. The parameter a indicating the curvature of the relative productivity of workers in the various tasks is related to the elasticity of substitution between labor of type 1 and 2, $\sigma = \frac{a}{a-1}$. Given the functional form in equation 23 and 24 and $t_1 = g_1(\mu, v) = \frac{v^{-\alpha_1}}{\beta_1(\mu)}$ and $t_2 = g_2(\mu, v) = \frac{(1-v)^{-\alpha_2}}{\beta_2(\mu)}$ the production function is a CES function.

The terms $b_k/\beta_k(\mu)$ in equation 30 indicate the efficiency units of labor with education k. The efficiency units depend therefore on the relative wage rate of workers with education 1. As the equilibrium relative wage μ increases, workers with relatively low potential skills of type 1 find it profitable to select education 1. Hence, we should expect a decrease in the efficiency units of labor with education 1. However, equation 30 indicates that an increase in the relative wage rate of workers with education 1 increases the efficiency of these workers. The reason for this counter intuitive result is that, in equilibrium, there is a negative relationship between the marginal task ε and the relative wage rate μ.[29] As the relative wage rate increases, the marginal task shifts to the left, toward 0, where workers with education 1 have their comparative advantage. Although workers with education 1 have lower level of type 1 skills in average after the wage shift, workers with education 1 are now assigned to tasks at which they are relatively more productive. Hence, the increase in the relative wage rate of workers with education 1 affects the efficiency units of workers with education 1 via two channels: the supply channel and the assignment channel. The supply channel has a negative impact on efficiency as workers with relatively lower levels of type 1 skills select education 1. The assignment channel has a positive impact as workers with education 1 are now assigned to tasks in which they are relatively more productive. The magnitude of the assignment effect is larger than the magnitude of the supply effect.

3.3 Equilibrium

Equilibrium is achieved by equalizing the relative supply of labor by education with the relative demand for labor by education. The assignment of workers to tasks ensures equilibrium by definition of the assignment functions, g_1 and g_2. Since firms are indifferent between assigning the least skilled worker with education 1 or the least skilled worker with education 2 to the marginal task, the relative wage rate μ is so that:

$$\ln \frac{p_1(t_{1,\varepsilon}, \varepsilon)}{p_2(t_{2,\varepsilon}, \varepsilon)} = \ln \mu + \ln \frac{t_{1,\varepsilon}}{t_{2,\varepsilon}} \tag{31}$$

[29] I show this result below in equation 32.

After some simplifications and rearranging yields:

$$\begin{aligned}\ln\mu &= \ln\frac{b_1}{b_2} + (\alpha_1 + 1 - a)\ln\frac{\varepsilon}{1-\varepsilon} + (\alpha_1 - \alpha_2)\ln(1-\varepsilon) &(32)\\ &= \ln\frac{b_1}{b_2} + (\alpha_2 + 1 - a)\ln\frac{\varepsilon}{1-\varepsilon} - (\alpha_2 - \alpha_1)\ln\varepsilon &(33)\end{aligned}$$

Employment equilibrium is given by:

$$\begin{aligned}D_1(\mu,\varepsilon) &= F(\varepsilon) \text{ and } D_2(\mu,\varepsilon) = 1 - F(\varepsilon) &(34)\\ &\Leftrightarrow \\ A\frac{\beta_1(\mu)}{b_1}\varepsilon^a Y &= F(\varepsilon) \text{ and } A\frac{\beta_2(\mu)}{b_2}(1-\varepsilon)^a Y = 1 - F(\varepsilon)\\ &\Leftrightarrow \\ \frac{F(\varepsilon)}{1-F(\varepsilon)} &= \frac{b_2}{b_1}\frac{\beta_1(\mu)}{\beta_2(\mu)}\left(\frac{\varepsilon}{1-\varepsilon}\right)^a\end{aligned}$$

Substituting equation 32 into equation 34 and solving for ε yields the marginal task ε^* that equilibrates demand and supply. In equilibrium, relative wage rate, μ^*, output per worker, Y^*, and employment of workers with education k, L_k^*, are given by:

$$\begin{aligned}\mu^* &= \frac{b_1}{b_2}\left(\frac{\varepsilon^*}{1-\varepsilon^*}\right)^{\alpha_1+1-a}(1-\varepsilon^*)^{\alpha_1-\alpha_2} &(35)\\ Y^* &= \frac{b_1(\varepsilon^*)^{-a}}{A\beta_1(\mu^*)}F(\varepsilon^*) &(36)\\ L_1^* &= F(\varepsilon^*) &(37)\\ L_2^* &= 1 - F(\varepsilon^*) &(38)\\ \rho(v) &= \begin{matrix}\frac{w_1}{b_1}(1-v)^{-d_2}v^{a-1-\alpha_1-d_1} & \text{if } v < \varepsilon\\ \frac{w_2}{b_2}v^{-d_1}(1-v)^{a-1-\alpha_2-d_2} & \text{if } v \geq \varepsilon\end{matrix} &(39)\end{aligned}$$

Since w_1 and w_2 are not identified separately, i.e. only their ratio μ is identified, the price function is not identified. Identification requires to solve the first order differential equation defined by 4 which given the mapping functions $g_k(v)$ simplifies to:

$$\rho(v) = -\rho'(v)\frac{p_k(v,g_k(v))}{\frac{\partial p_k(v,g_k(v))}{\partial v}}$$

The solution of this first order differential equation is then straightforward, $\rho(v) = \frac{w_1 g_1(v)}{p_1(v,g_1(v))}$ on $v \in (0,\varepsilon^*)$ and $\rho(v) = \frac{w_2 g_2(v)}{p_2(v,g_2(v))}$ elsewhere.

3.4 Closed form solutions

Further functional specifications are needed before being able to derive sets of closed form solutions. I assume $\beta_1(\mu) = 1/\beta_2(\mu) = \mu^{-1/2\delta}$ with $\delta > 0$ and $\alpha_1 = \alpha_2$. This yields the following equilibrium condition:

$$\begin{aligned}\frac{F(\varepsilon)}{1-F(\varepsilon)} &= \left(\frac{b_2}{b_1}\right)^{1+\delta}\left(\frac{\varepsilon}{1-\varepsilon}\right)^{a-\delta(\alpha_1+1-a)} \\ &= \Gamma\left(\frac{\varepsilon}{1-\varepsilon}\right)^{\Lambda}\end{aligned} \tag{40}$$

where $\Gamma = \left(\frac{b_2}{b_1}\right)^{1+\delta} \in (0;\infty)$ and $\Lambda = a - \delta(\alpha_1 + 1 - a) \in [1;\infty)$

I derive closed form solutions for different shapes of the tasks distribution, i.e. symmetric distributions (uniform, inverted-U shape and normal look alike distribution), distributions skewed to the right and distributions skewed to the left.

Symmetric distributions: $d_1 = d_2 = d$

The first example of symmetric distribution is when tasks are uniformly distributed on v, i.e. $d = 0$. This yields the cumulative density $F(\varepsilon) = \varepsilon$ which once substituted in (40) and rearranged gives:

$$\varepsilon^* = \frac{\Gamma^{\frac{1}{1-\Lambda}}}{1+\Gamma^{\frac{1}{1-\Lambda}}}$$

When $d = 1$ the distribution of tasks has an Inverted-U shape and cumulative density function $F(\varepsilon) = 3\varepsilon^2 - 2\varepsilon^3$. Substituting in (40) yields:

$$\frac{3\varepsilon^2 - 2\varepsilon^3}{1 - 3\varepsilon^2 + 2\varepsilon^3} = \Gamma\left(\frac{\varepsilon}{1-\varepsilon}\right)^{\Lambda}$$

Closed form solutions for this equality exist for small positive integer values of $\Lambda = a - \delta(\alpha_1 + 1 - a) > 0$. For $\Lambda = 1$ we have $\varepsilon^* = \frac{3-\Gamma+\left(9-14\Gamma+9\Gamma^2\right)^{1/2}}{4(1-\Gamma)}$ for all $\Gamma > 0$, for $\Lambda = 2$, we have: $\varepsilon^* = \begin{cases} 0 \text{ if } \Gamma \geq 3 \\ \frac{3-\Gamma}{2(1+\Gamma)} \text{ if } 1/3 < \Gamma < 3 \\ 1 \text{ if } \Gamma \leq 1/3 \end{cases}$ and for $\Lambda = 3$ we have $\varepsilon^* = \begin{cases} \frac{5+\Gamma-\left(1+34\Gamma+\Gamma^2\right)^{1/2}}{4(1-\Gamma)} \text{ if } \Gamma < 23.172 \\ 1 \text{ otherwise} \end{cases}$.

As noted earlier, when $d > 1$ the distribution of tasks looks like a normal distribution with mean $1/2$ and variance $\frac{(d+1)^2}{2(1+d)(3+2d)}$. For $d = 2$ cumulative density function is $F(\varepsilon) = \varepsilon^3(10 - 15\varepsilon + 6\varepsilon^2)$ and once substituted into (40) yields:

$$\frac{\varepsilon^3(10 - 15\varepsilon + 6\varepsilon^2)}{1 - \varepsilon^3(10 - 15\varepsilon + 6\varepsilon^2)} = \Gamma\left(\frac{\varepsilon}{1-\varepsilon}\right)^\Lambda$$

Closed form solutions for this equality exist for small positive integer values of $\Lambda = a - \delta(\alpha_1 + 1 - a) > 0$. For $\Lambda = 3$, for instance we have: $\varepsilon^* = \begin{cases} 0 \text{ if } \Gamma \geq 10 \\ \frac{15+3\Gamma-3^{1/2}\left(118\Gamma-5-5\Gamma^2\right)^{1/2}}{12(1-\Gamma)} \text{ if } 1/10 < \Gamma < 10 \\ 1 \text{ if } \Gamma \leq 1/10 \end{cases}$.

Distribution skewed to the left: $d_1 > d_2$

For instance, $d_1 = 1$ and $d_2 = 0$ yields $F(\varepsilon) = \varepsilon^2$ which once substituted in (40) gives:

$$\frac{\varepsilon^2}{1-\varepsilon^2} = \Gamma\left(\frac{\varepsilon}{1-\varepsilon}\right)^\Lambda$$

Closed form solutions for this equality exist for small positive integer values of $\Lambda = a - \delta(\alpha_1 + 1 - a) > 0$. For $\Lambda = 1$ we have $\varepsilon^* = \begin{cases} \frac{\Gamma}{1-\Gamma} \text{ if } \Gamma<1/2 \\ 1 \text{ otherwise} \end{cases}$, for $\Lambda = 2$, we have: $\varepsilon^* = \begin{cases} \frac{1-\Gamma}{1+\Gamma} \text{ if } 0<\Gamma<1 \\ 0 \text{ otherwise} \end{cases}$ and for $\Lambda = 3$ we have $\varepsilon^* = \frac{\Gamma+2-\Gamma^{1/2}(8+\Gamma)^{1/2}}{2(1-\Gamma)}$ for all $\Gamma > 0$.

Distribution skewed to the right: $d_2 > d_1$

For instance $d_1 = 0; d_2 = 1$ yields $F(\varepsilon) = 2\varepsilon - \varepsilon^2$ which once substituted in (40), gives:

$$\frac{2\varepsilon - \varepsilon^2}{1 - 2\varepsilon + \varepsilon^2} = \Gamma\left(\frac{\varepsilon}{1-\varepsilon}\right)^\Lambda$$

Closed form solutions for this equality also exist for small positive integer values of $\Lambda = a - \delta(\alpha_1 + 1 - a) > 0$. For $\Lambda = 1$ we have $\varepsilon^* = \begin{cases} \frac{2-\Gamma}{1-\Gamma} \text{ if } 2<\Gamma \\ 0 \text{ otherwise} \end{cases}$, for $\Lambda = 2$, we have $\varepsilon^* = \begin{cases} \frac{2}{1+\Gamma} \text{ if } 1<\Gamma \\ 1 \text{ otherwise} \end{cases}$ and for $\Lambda = 3$ we have $\varepsilon^* = \frac{3-(1+8\Gamma)^{1/2}}{2(1-\Gamma)}$ for all $\Gamma > 0$.

3.5 Comparative statics

I am now interested in the effects of i) a shift in the relative efficiency units of workers with education 2, i.e. $\frac{b_2}{b_1}$, ii) a shift in the elasticity of substitution between labor types, i.e. a and iii) a shift in the degree of heterogeneity between workers with same education, i.e. α on the equilibrium task ε, educational employment and relative wage μ. For the remaining of this part I assume that $\beta_1(\mu) = \mu^{-\frac{1}{2}\delta}$ and $\beta_2(\mu) = \mu^{\frac{1}{2}\delta}$ with $\delta \geq 0$.[30]

i) Upward shift in the relative efficiency units of workers with education 2: $\frac{b_2}{b_1}$

Substituting equation 35 in the equilibrium employment equation 34 and taking log forms yields:

$$\begin{aligned}
\ln \frac{F(\varepsilon)}{1-F(\varepsilon)} &= (1+\delta)\ln \frac{b_2}{b_1} + (a - \delta(\alpha_1 + 1 - a)) \ln \frac{\varepsilon}{1-\varepsilon} - \qquad (41)\\
&\quad \delta\,(\alpha_1 - \alpha_2)\ln(1-\varepsilon)\\
&= (1+\delta)\ln \frac{b_2}{b_1} + (a - \delta(\alpha_2 + 1 - a)) \ln \frac{\varepsilon}{1-\varepsilon} + \qquad (42)\\
&\quad \delta\,(\alpha_2 - \alpha_1)\ln \varepsilon
\end{aligned}$$

While the Left Hand Side is not affected by a change in the efficiency units at constant marginal task, an increase in the efficiency units of workers with education 2 compared to workers with education 1 increases the Right Hand Side of 41 at constant marginal task.

$$\frac{\partial RHS(\varepsilon^*)}{\partial \ln \frac{b_2}{b_1}} = (1+\delta) > 0$$

The RHS of equation 41 increases while the LHS remains constant at constant marginal task. To restore equilibrium, the marginal task has to shift. If for a given percentage change in the marginal task, the LHS changes slower than the RHS, then the equilibrium marginal task will shift downward and vice versa. Deriving both, the RHS and the LHS of equation 41 with respect to ε yields:

[30] Note that for $\alpha_k = 0$ for $k = 1, 2$ and $\delta = 0$ the model reduces to Rosen's task assignment model.

$$
\begin{aligned}
\frac{\partial LHS(\varepsilon^*)}{\partial \varepsilon} &= \frac{d(\varepsilon^*)}{F(\varepsilon^*)\,(1-F(\varepsilon^*))} > 0 \\
& \text{and} \\
\frac{\partial RHS(\varepsilon^*)}{\partial \varepsilon} &= (a-\delta(\alpha_1+1-a))\,\frac{1}{\varepsilon^*\,(1-\varepsilon^*)} + \delta\,(\alpha_1-\alpha_2)\,\frac{1}{1-\varepsilon^*} > 0 \\
&= (a-\delta(\alpha_2+1-a))\,\frac{1}{\varepsilon^*\,(1-\varepsilon^*)} + \delta\,(\alpha_2-\alpha_1)\,\frac{1}{\varepsilon^*} > 0
\end{aligned}
$$

In appendix A, I show that $\frac{\partial LHS(\varepsilon^*)}{\partial \varepsilon} < \frac{\partial RHS(\varepsilon^*)}{\partial \varepsilon}$ for all $\varepsilon^* \in (0,1)$ and all $a > 1+\max_k(\alpha_k+d_k)$ and therefore $\frac{\partial \varepsilon}{\partial \ln \frac{b_2}{b_1}} < 0$. Hence, to reach a new equilibrium the marginal task will have to decrease and shift toward 0.

The properties of the density function guarantee that employment of workers with education 1 will decrease and employment of workers with education 2 will increase, i.e. $\frac{\partial F(x)}{\partial x} = d(x) > 0$.

Using equation 32, I investigate the effect of a shift in the relative efficiency units of workers with education 2 on the relative wage rate.

$$
\begin{aligned}
\frac{\partial \ln \mu}{\partial \ln \frac{b_2}{b_1}} &= -1 + \frac{\partial \ln \mu}{\partial \varepsilon}\frac{\partial \varepsilon}{\partial \ln \frac{b_2}{b_1}} \\
&= -1 + \left((\alpha_1+1-a)\,\frac{1}{\varepsilon\,(1-\varepsilon)} + (\alpha_1-\alpha_2)\,\frac{1}{1-\varepsilon} \right) \frac{\partial \varepsilon}{\partial \ln \frac{b_2}{b_1}}
\end{aligned}
$$

There are two effects of opposite direction. A 1% increase in the efficiency units of labor with education 2 will directly reduce the relative wage rate of workers with education 1 by 1%. However, the shift in efficiency units will also decrease the marginal task in equilibrium, $\frac{\partial \varepsilon^*}{\partial \frac{b_2}{b_1}} < 0$. Note that there is a negative relationship between the marginal task and the relative wage rate. Indeed, we have:

$$
\begin{aligned}
\frac{\partial \ln \mu}{\partial \varepsilon} &= \left((\alpha_1+1-a)\,\frac{1}{\varepsilon\,(1-\varepsilon)} + (\alpha_1-\alpha_2)\,\frac{1}{1-\varepsilon} \right) \\
&= \left((\alpha_2+1-a)\,\frac{1}{\varepsilon\,(1-\varepsilon)} + (\alpha_2-\alpha_1)\,\frac{1}{\varepsilon} \right) < 0 \\
\text{since } a &> 1+\max(\alpha_k, d_k) > \alpha_k + 1 \text{ for all } k
\end{aligned}
$$

Therefore, the downward shift in the marginal task induced by the upward shift in the relative efficiency units of workers with education 2 will drive the wage rate up. Which effect dominates depends on the magnitude of the change in the marginal task with respect to a 1% change in the relative efficiency units, i.e. $\frac{\partial\varepsilon}{\partial \ln \frac{b_2}{b_1}}$.

ii) Upward shift in a: decrease in the elasticity of substitution between skilled and unskilled labor σ.

I turn to the effect of an increase in a on the equilibrium. For the sake of simplicity I use equation 41.

$$\ln \frac{F(\varepsilon)}{1-F(\varepsilon)} = (1+\delta)\ln \frac{b_2}{b_1} + (a-\delta(\alpha_1+1-a))\ln \frac{\varepsilon}{1-\varepsilon} - \delta\,(\alpha_1-\alpha_2)\ln(1-\varepsilon) \quad (43)$$

Deriving the RHS of equation 43 with respect to a gives:

$$\frac{\partial RHS}{\partial a} = (1+\delta)\ln \frac{\varepsilon}{1-\varepsilon} > 0 \quad (44)$$

An increase in a will shift the RHS up while the LHS remains constant ceteris paribus. To restore equilibrium, the marginal task will have to shift. Since the LHS of equation 43 decreases slower than the RHS as the marginal task decreases, as shown in Appendix A, the equilibrium marginal task decreases when a increases, i.e. $\frac{\partial\varepsilon}{\partial a} < 0$.

A decrease in the employment of workers with education 1 and an increase in the employment of workers with education 2 will follow from the downward shift in the marginal task.

Using equation 32, I investigate the effect of a shift in a on the relative wage rate.

$$\begin{aligned}
\frac{\partial \ln \mu}{\partial a} &= -\ln \frac{\varepsilon}{1-\varepsilon} + \frac{\partial \ln \mu}{\partial \varepsilon}\frac{\partial \varepsilon}{\partial a} \\
&= -\ln \frac{\varepsilon}{1-\varepsilon} + \left((\alpha_1+1-a)\frac{1}{\varepsilon\,(1-\varepsilon)} + (\alpha_1-\alpha_2)\frac{1}{1-\varepsilon}\right)\frac{\partial \varepsilon}{\partial a}
\end{aligned}$$

The first term indicates that an increase in a will increase or decrease the relative wage rate depending on whether the initial equilibrium task is smaller or greater than a half. However, the increase in a will also lead to a decrease in the marginal task in equilibrium, $\frac{\partial\varepsilon^*}{\partial a} < 0$. Since there is a negative relationship between the marginal task and the relative wage rate, the downward shift in the marginal task associated to the shift in a will drive the wage rate up. To summarize, if the initial marginal task is smaller than a half, an increase in a will increase the relative wage rate. However, if the initial marginal task is larger than a half, the relative wage rate will decrease via the direct effect of an increase in a but increase because the increase in a also shifts the marginal task downward which increases the relative wage rate. Which effect dominates depends on the magnitude of the change in the marginal task with respect to a given change in a, i.e. $\frac{\partial \varepsilon}{\partial a}$.

iii) Upward shift in α_k

I turn to the effect of an increase in the degree of skills heterogeneity within educational groups on the equilibrium. I derive the RHS of equation 43 with respect to α_k and obtain.

$$\frac{\partial RHS}{\partial \alpha_1} = -\delta \ln \varepsilon > 0 \tag{45}$$

$$\frac{\partial RHS}{\partial \alpha_2} = \delta \ln (1-\varepsilon) < 0 \tag{46}$$

An increase in α_1, respectively α_2, will shift the RHS up (down) while the LHS remains constant at constant marginal task. Once again, to restore equilibrium, the marginal task will have to shift. Since the RHS of equation 43 increases faster than the LHS as the marginal tasks increases, the new equilibrium marginal task will be situated to the left, i.e. $\frac{\partial \varepsilon^*}{\partial \alpha_1} < 0$, respectively to the right, i.e. $\frac{\partial \varepsilon^*}{\partial \alpha_2} > 0$, of the old marginal task when α_1, respectively α_2, increases.

It follows that an increase in the degree of heterogeneity of workers' skills within education k decreases employment of workers with education k and increases employment of workers with other education.

Using equation 32, I investigate the effect of an upward shift in α_1 on the relative wage rate.

$$\begin{aligned}
\frac{\partial \ln \mu}{\partial \alpha_1} &= \ln \varepsilon + \frac{\partial \ln \mu}{\partial \varepsilon}\frac{\partial \varepsilon}{\partial \alpha_1} \\
&= \ln \varepsilon + \left((\alpha_1 + 1 - a) \frac{1}{\varepsilon (1-\varepsilon)} + (\alpha_1 - \alpha_2) \frac{1}{1-\varepsilon} \right) \frac{\partial \varepsilon}{\partial \alpha_1} > 0
\end{aligned}$$

The first term indicates that the direct effect of an increase in α_1 will be to decrease the relative wage rate of workers with education 1, $\ln \varepsilon < 0$. However, the increase in α_1 will also decrease the marginal task in equilibrium, $\frac{\partial \varepsilon^*}{\partial \alpha_1} < 0$. This increase in the marginal task will drive the wage rate up. Hence, an increase in the degree of heterogeneity of workers' skills within education 1 will increase or decrease the relative wage rate of workers with education 1 depending on the magnitude of $\frac{\partial \varepsilon}{\partial \alpha_1}$.

Similarly, using equation 32, I investigate the effect of an upward shift in α_2 on the relative wage rate.

$$\begin{aligned}\frac{\partial \ln \mu}{\partial \alpha_2} &= -\ln(1-\varepsilon) + \frac{\partial \ln \mu}{\partial \varepsilon}\frac{\partial \varepsilon}{\partial \alpha_2} \\ &= -\ln(1-\varepsilon) + \left((\alpha_1 + 1 - a)\frac{1}{\varepsilon(1-\varepsilon)} + (\alpha_1 - \alpha_2)\frac{1}{1-\varepsilon}\right)\frac{\partial \varepsilon}{\partial \alpha_2}\end{aligned}$$

The first term indicates that an increase in α_2 will increase the relative wage rate of workers with education 1, $-\ln(1-\varepsilon) > 0$. However, the increase in α_2 will also increase the marginal task in equilibrium, $\frac{\partial \varepsilon^*}{\partial \alpha_2} > 0$. This increase in the marginal task will drive the wage rate down. Hence, an increase in the degree of heterogeneity of workers' skills within education 2 will increase or decrease the relative wage rate of workers with education 1 depending on the relative magnitude of $\frac{\partial \varepsilon}{\partial \alpha_2}$.

4 Implications

To illustrate the implications of the model, I calibrate the various parameters of the model to fit as close as possible to the features of the wage distribution observed in the US in the last decades. The stylized facts of the wage distribution I use are directly taken from Acemoglu (2002) and can be summarized in the four following facts:

1. Small increase in the skill premium of 2 percent between 1963 and 1980 and sharp increase of 25 percent between 1980 and 1996
2. Relative supply of college graduates has increased (constantly) almost twofold between 1963 and 1996
3. Overall wage inequality increased sharply since 1970, and the workers at the 1^{st} decile saw their wage decline between 1980 and 1996 while the wage of workers at the 9^{th} decile increased
4. Residual wage inequality increased steadily at the bottom and top of the distribution even in the 70s when skill-premium stagnated.

In what follows I aim at replicating these stylized facts by calibrating the parameters of the model. As part of my goal is to explain patterns of the skill premium, that is the college to high-school graduates wage differential, the educational choice workers are facing in my model is: go to college, choose education 2 or leave school after high-school, education 1. Each worker is endowed with a particular level of strength and a particular level of intelligence and can choose to either develop her strength and do not go to college or develop her intelligence and go to college, as in Willis and Rosen (1979) and Willis (1986). The relative wage rate μ can therefore be seen as the skill-premium referred to in the related literature. Tasks close to 0 are tasks that require a relative high level of manual skills whereas tasks close to 1 require high cognitive skills. The task $\frac{1}{2}$ is the "anybody can do it" task and therefore does not require any particular type of skills.

Wage distribution in 1963: baseline parameters

Since several parameters have a direct link with parameters often estimated in the literature, these parameters should therefore be calibrated so as to fit in the range of estimates found in the literature. The parameter $\sigma = \frac{a}{a-1}$ is the well-known elasticity of substitution between college and high-school graduates and should be around 1.44 as found by Katz and Murphy (1992) or even 1.9 as found by Acemoglu (2002). The parameter a should therefore be ranging between 2 and 4.

The relative efficiency parameter of college graduates, given by $\frac{b_2}{b_1}$, has grown at an average rate of 3.3 percent per year as reported by Katz and Murphy (1992) or 3.6 percent as reported by Heckman et al. (1998) when controlling for heterogeneity and endogenous human capital formation. However, some authors, Acemoglu (2002) and Krusell et al. (2000) among others, have argued in favor of an acceleration in SBTC in the 80s. I choose the later specification and consider an average rate of 4.9 percent per year[31] between 1963 and 1996 but with first a yearly rate of 2 percent between 1963 and 1980 and an acceleration to a yearly rate of 8 percent between 1980 and 1996.

I calibrate the other parameters so as to fit the relative supply of college graduates, the skill-premium and the 1^{st}, 5^{th} and 9^{th} decile of the wage distribution reported by Acemoglu (2002) for the years 1963, 1980 and 1996 respectively. The results are reported in Table 2.

In my 1963 baseline scenario, the distribution of tasks is relatively symmetric $d_1 \approx d_2 \approx 1$. The shape of the tasks distribution is drawn in (full line in) Figure 4. The vertical lines indicate the equilibrium marginal task in the respective years (full line for 1963, small dash for 1980 and long dash for 1996. The relative efficiency units parameter of college graduates is set equal to 0.27 and the elasticity of substitution parameter is set to 1.55. The (log) skill premium is equal to 0.35 and the relative supply of college graduates equal to 0.24. As calibrated, the workers in the first decile earns 42 percent less than the median worker and 101 percent less than the top 1 decile.

[31] Note that lower growth rates, of the same magnitude as Katz and Murphy (1992) for instance, did replicate the skill-premium and relative skill supply schedule adequalty but failed to replicate the overall and residual wage inequality. For instance, for rates lower than 4 percent, the model failed to replicate the decrease in the wage earned by workers at the 1^{st} decile between 1980 and 1996.

Wage distribution in 1980

The period 1963-1980 saw a sharp increase in the relatively supply of college graduates from 0.24 to 0.41 whereas the skill premium remained fairly stable − slight increase from 0.36 to 0.38 −. Since the skill-premium remained stable at the same time that the relative supply increased this means that the relative demand and thereby the relative efficiency must have increased too. This increase is captured by an increase of b_2 compared to b_1, from 0.27 to 0.38 which is equivalent to a annual increase of 2 percent. Although changes in the efficiency parameter along with changes in the distribution of tasks provide a very good fit of the skill-premium and relative supply for 1980, in order to fit the overall and residual wage inequality more parameters need to be re-calibrated.

To fit changes in the overall wage distribution, I increased α_1 from 0.83 to 1.1 which corresponds to an increase in the degree of heterogeneity in skills among unskilled workers. Similarly α_2 increased from 0.25 to 0.3 which indicates an increase in the degree of heterogeneity in skills among skilled workers. The intuition for an increase in the degree of heterogeneity among workers is the following. Under the pressure of technological change, tasks at both ends of the tasks distribution demand more specific skill. For instance, an extra 0.62 log units[32] of manual skills are required at task 0.1 in 1980 compared to 1963 everything else constant. Similarly, 0.12 extra units of cognitive skills are required at task 0.9 in 1980. This argument corresponds to the argument used by Acemoglu (2002) in order to fit residual wage inequality within a two–skill index model: "Under the plausible assumption that more skilled workers within each education group also benefit from skill-biased technical progress, technical change spurred by the increase in the supply of educated workers will immediately benefit workers with more unobserved skills, raising within-group inequality", p. 59.

As shown by the comparative statics, an increase in α_1 increases L_2. To keep the relative supply of college graduates equal to 0.41, d_2 has to increase relatively to d_1 which skews the distribution relatively to the right and increases the number of tasks in the left part of the distribution, see the short dashed line in Figure 4. Note that the equilibrium marginal task shifts to the left from 0.72 to 0.59 indicating that the tasks in the middle of the distribution have become more demanding in terms of cognitive skills. However, the new distribution of tasks has more tasks in the middle of the tasks spectrum and less tasks in the right tail.

Moreover, the increase in the parameters d_k and α_k pushes the parameter a upward[33] and therefore decreases the elasticity of substitution σ to 1.36. This result fits with empirical evidence provided in Dupuy and Marey (2004). Dupuy and Marey (2004) reveal the presence of two structural breaks in the Katz and Murphy (1992) model, one in 1978, characterized by a decrease in the elasticity

[32] $\Delta_{63-80} \ln t_{1,v=0.1} = -(1.1 - 0.83) \times \log 0.9 = 0.62$

[33] Indeed, we have $a > \max_k(d_k + \alpha_k) + 1$.

of substitution and one in 1990 characterized by an increase in the elasticity of substitution. The intuition for the change in the elasticity of substitution is that technical and organizational changes affect the relative productivity of workers differently in the various tasks and therefore affect workers' comparative advantage.

As indicated in Table 2, the calibration presented here reproduced quite well the increase in the three deciles between 1963 and 1980 as well as the increase in wage inequality at both the lower and upper part of the wage distribution as indicated by the figures presented in Acemoglu (2002). Moreover, the model is able to predict the increased residual inequality observed at the top of the wage distribution as well as at the bottom of the wage distribution between 1963 and 1980 in the US, as indicated in the last three rows of Table 2. Note however, that in contrast to the observed residual inequality, the model predicts a larger increase in the top distribution than in the bottom distribution.

Wage distribution in 1996

To match the 1996 situation, I calibrated the model so that the growth rate in efficiency units of college graduates accelerates to a yearly rate of 8 percent. Moreover, the tasks distribution shifts so that it becomes more skewed to the left. However, there are also less tasks found in the middle of the spectrum and more tasks are found in both tails. This feature corresponds to recent evidence provided by Autor et al. (2003) that technology can replace labor in routine tasks (either cognitive or manual) but not in non-routine tasks (either cognitive or manual) leading to job polarization (see Goos and Manning (2004)). In turns, the decrease in the parameter d_k gives room to the elasticity of substitution parameter to rise above its 1963 level to 1.62 consistently with Dupuy and Marey's (2004) finding. Given this new set of parameters, the model predicts a (log) skill premium of 0.58 and relative supply of college graduates of 0.68 which is very close to the figures reported by Acemoglu (2002) for 1996. For the period 1980-1996, the model predicts a decrease in the wage of workers at the 1^{st} decile, i.e. -6% and workers at the 5^{th} decile, -7%, as well as an increase in the wage of workers at the 9^{th} decile of 9% conform to Acemoglu (2002). All together the results indicate a widening of the wage inequality in the lower part of the wage distribution as well as in the upper part. Note that the model captures very well the general evolution of the 1^{st}, 5^{th} and 9^{th} deciles in the period 1963-1996, and is especially relevant as it fits the observed drop in the wage of workers at the first decile between 1980 and 1996 within a skill-biased technical change framework. However, even though the wage of workers at the first decile drops, it remains larger in 1996 than in 1963 unlike actual data indicates.

To summarize, the tasks assignment model with endogenous human capital formation developed in this paper can be used to explain changes in the US wage distribution in the last decades. The skill premium pattern is explained by technical and organizational changes that affects i) the distribution of tasks and therefore the demand for skills and ii) the efficiency of college and high-school graduates as well as the comparative advantage of college graduates via changes in the elasticity of substitution between both labor groups. The overall and residual rise in wage inequality can be explained by i) the job polarization due to changes in the distribution of tasks and ii) increasing heterogeneity within educational groups due to increasing demand for specific skills in tasks at both ends of the tasks distribution.

References

Acemoglu, D. (2002): "Technical Change, Inequality and the Labor Market," *Journal of Economic Literature*, 40(1), 7–72.

Autor, D., F. Levy, and R. Murnane (2003): "The Skill Content of Recent Technological Change: An Empirical Exploration," *Quarterly Journal of Economics*, 118(4).

Beaudry, P., and D. Green (2003): "Wages and Employment in the United States and Germany: What Explains the Differences?," *American Economic Review*, 93(3), 573–602.

Berman, E., J. Bound, and Z. Griliches (1994): "Changes in the Demand for Skilled Labor Within US Manufacturing: Evidence from Annual Survey of Manufactures," *Quarterly Journal of Economics*, 109(1), 367–97.

Berman, E., J. Bound, and S. Machin (1998): "Implications of Skill-Biased Technological Change: International Evidence," *Quarterly Journal of Economics*, 113(4), 1245–79.

Bound, J., and G. Johnson (1992): "Changes in the Structure of Wages During the 1980's: An Evaluation of Alternative Explanations," *American Economic Review*, 82(3), 371–92.

Bresnahan, T. (1999): "Computerisation and Wage Dispersion: An Analytical Reinterpretation," *Economic Journal*, 109(456), 390–415.

Bresnahan, T., E. Brynjolfsson, and L. Hitt (2002): "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence," *Quarterly Journal of Economics*, 117(1), 339–76.

Brynjolfsson, E., and L. Hitt (1995): "Information Technology as a Factor of Production: The Role of Differences Among Firms," *Economics of Innovation and New Technology*, 3(3-4), 183–99.

Card, D., and T. Lemieux (2001): "Can Falling Supply Explain the Rising Return to College for Younger Men? A Cohort-Based Analysis," *Quarterly Journal of Economics*, 116(2), 705–46.

Diamond, P. (1981): "Mobility Costs, Frictional Unemployment and Efficiency," *Journal of Political Economy*, 89, 798–812.

Dupuy, A., and P. Marey (2004): "Shifts and Twists in the Relative Productivity of Skilled Labor: Reconciling Accelerated SBTC with the Productivity Slowdown," *mimeo*.

Ekeland, I., J. J. Heckman, and L. Nesheim (2004): "Identification and Estimation of Hedonic Models," *Journal of Political Economy*, 112(2), S60–S109.

Friedman, M., and S. Kuznets (1945): *Income from Independent Professional Practice*. New York: National Bureau of Economic Research.

Goos, M., and A. Manning (2004): "Lousy and Lovely Jobs: The Rising Polarization of Work in Britain," *mimeo*.

Gould, E. (2002): "Rising Wage Inequality, Comparative Advantage, and the Growing Importance of General Skills in the United States," *Journal of Labor Economics*, 20(1), 105–47.

Hartog, J. (2001): "On Human Capital and Individual Capabilities," *Scholar Working Paper Series*, WP 22.

Heckman, J., and G. Sedlacek (1985): "Heterogeneity, Aggregation and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market," *Journal of Political Economy*, 93(6), 1077–125.

Heckman, J., and G. Sedlacek (1990): "Self-Selection and the Distribution of Hourly Wages," *Journal of Labor Economics*, 8(1), S329–63.

Heckman, J. J. (1999): "Accounting for Heterogeneity, Diversity, and General Equilibriumin Evaluating Social Programs," Discussion Paper 7230, National Bureau of Economic Research, Inc.

Heckman, J. J., and X. Li (2004): "Selection bias, comparative advantage and heterogeneous returns to education: evidence from China in 2000," *Pacific Economic Review*, 9(3), 155–171.

Heckman, J. J., L. Lochner, and C. Taber (1998): "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents," *Review of Economic Dynamics*, 1, 1–58.

HECKMAN, J. J., AND E. VYTLACIL (2005): "Structural Equations, Treatment Effects, and Econometric Policy Evaluation," *Econometrica*, 73(3), 669–738.

JOVANOVIC, B. (1979): "Job Matching and the Theory of Turnover," *Journal of Political Economy*, 87, 972–90.

KATZ, L., AND K. MURPHY (1992): "Changes in Relative Wages, 1963-1987: Supply and Demand Factors," *Quarterly Journal of Economics*, 107(1), 35–78.

KRUEGER, A. (1993): "How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984-1989.," *Quarterly Journal of Economics*, 110, 33–60.

KRUSELL, P., L. OHANIAN, J.-V. RÍOS-RULL, AND G. VIOLANTE (2000): "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis," *Econometrica*, 68(5), 1029–53.

LUCAS, R. (1977): "Hedonic Wage Equations and Psychic Wages in the Returns to Schooling," *American Economic Review*, 67(4), 549–58.

MACDONALD, G. (1982): "A Market Equilibrium Theory of Job Assignment and Sequential Accumulation of Information," *American Economic Review*, 72(5), 1038–55.

MACHIN, S., AND S. MCNALLY (2005): "The Literacy Hour," *mimeo.*

MINCER, J. (1974): *Schooling Experience and Earnings.* New York: Columbia University Press.

——— (1993): *Studies in Human Capital.* Aldershot: Edward Elgar.

MORTENSEN, D. (1986): "Job Search and Labor Market Analysis," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and R. Layard, vol. 2. Amsterdam: Northe Holland.

PISSARIDES, C. (1984): "Search Intensity, Job Advertising, and Efficiency," *Journal of Labor Economics*, 2, 128–43.

ROSEN, S. (1974): "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *Journal of Political Economy*, 82, 34–55.

ROSEN, S. (1978): "Substitution and the Division of Labor," *Economica*, 45, 235–50.

ROY, A. (1950): "The Distribution of Earnings and of Individual Output," *Economic Journal*, 60(3), 489–505.

——— (1951): “Some Thoughts on the Distribution of Earnings,” *Oxford Economic Papers*, 3(1), 135–46.

SATTINGER, M. (1975): “Comparative Advantage and the Distribution of Earnings and Abilities,” *Econometrica*, 43(3), 455–68.

——— (1978): “Comparative Advantage in Individuals,” *Review of Economics and Statistics*, 60(2), 259–67.

——— (1980): *Capital and the Distribution of Labor Earnings*. Amsterdam: North-Holland.

——— (1993): “Assignment Models of the Distribution of Earnings,” *Journal of Economic Literature*, 31(2), 831–80.

SIOW, A. (1984): “Occupational Choice under Uncertainty,” *Econometrica*, 52(3), 631–45.

TEULINGS, C. (1995a): “A Generalized Assignment Model of Workers to Jobs for the U.S. Economy,” *Tinbergen Institute, Discussion Paper*, TI 95-67,.

——— (1995b): “The Wage Distribution in a Model of the Assignment of Skills to Jobs,” *Journal of Political Economy*, 103(2), 180–215.

——— (2005): “Comparative Advantage, Relative Wages, and the Accumulation of Human Capital,” *Journal of Political Economy*, 113(2), 425–61.

TIMSS (2000): *International Mathematics Report*. The International Study Center, Lynch School of Education and Boston College.

TINBERGEN, J. (1956): “On the Theory of Income Distribution,” *Weltwirtschaftliches Archiv*, 77(1), 156–75.

WILLIS, R. (1986): “Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions,” in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and R. Layard, pp. 525–602. Amsterdam: North-Holland.

WILLIS, R., AND S. ROSEN (1979): “Education and Self-Selection,” *Journal of Political Economy*, 87(5), S7–S36.

Appendix A: proof of $\frac{\partial LHS(\varepsilon^*)}{\partial\varepsilon} < \frac{\partial RHS(\varepsilon^*)}{\partial\varepsilon}$ for all $\varepsilon^* \in (0,1)$ and all $a > 1 + \max_k(\alpha_k + d_k)$ (Not completed yet)

Proof. Note that:

$$\begin{aligned}
\frac{\partial RHS(\varepsilon^*)}{\partial\varepsilon} &= (a - \delta(\alpha_1 + 1 - a))\frac{1}{\varepsilon^*(1-\varepsilon^*)} + \delta(\alpha_1 - \alpha_2)\frac{1}{1-\varepsilon^*} \\
&\geq (a - \delta(\alpha_1 + 1 - a))\frac{1}{\varepsilon^*(1-\varepsilon^*)} > 0 \\
\text{for } \alpha_1 &\geq \alpha_2 \\
&\text{or} \\
\frac{\partial RHS(\varepsilon^*)}{\partial\varepsilon} &= (a - \delta(\alpha_2 + 1 - a))\frac{1}{\varepsilon^*(1-\varepsilon^*)} + \delta(\alpha_2 - \alpha_1)\frac{1}{\varepsilon^*} \\
&> (a - \delta(\alpha_2 + 1 - a))\frac{1}{\varepsilon^*(1-\varepsilon^*)} > 0 \\
\text{for } \alpha_2 &> \alpha_1
\end{aligned}$$

Therefore, to prove that $\frac{\partial LHS(\varepsilon^*)}{\partial\varepsilon} < \frac{\partial RHS(\varepsilon^*)}{\partial\varepsilon}$ for all $\varepsilon^* \in (0,1)$ and all $a > 1 + \max_k(\alpha_k + d_k)$, it is enough to prove, for $\alpha = \max\alpha_k \geq 0$, that:

$$\frac{\partial LHS(\varepsilon^*)}{\partial\varepsilon} < (a - \delta(\alpha + 1 - a))\frac{1}{\varepsilon^*(1-\varepsilon^*)}$$

I first rearrange the inequality using $d(\varepsilon) = A\varepsilon^{d_1}(1-\varepsilon)^{d_2}$ and obtain:

$$A\frac{\varepsilon^{d_1+1}(1-\varepsilon)^{d_2+1}}{F(\varepsilon)(1-F(\varepsilon))} < (a - \delta(\alpha + 1 - a)) \tag{47}$$

Note that $(a - \delta(\alpha + 1 - a)) > a$.

To check the validity of inequality 47, I define the function $G(\varepsilon) = AG_a(\varepsilon)G_b(\varepsilon)$ with $G_a(\varepsilon) = \frac{\varepsilon^{d_1+1}}{F(\varepsilon)}$ and $G_b(\varepsilon) = \frac{(1-\varepsilon)^{d_2+1}}{1-F(\varepsilon)}$.

Note that $G'_a(\varepsilon) = \frac{(d_1+1)\varepsilon^{d_1}F(\varepsilon) - \varepsilon^{d_1+1}d(\varepsilon)}{F^2(\varepsilon)} = \frac{G_a(\varepsilon)}{\varepsilon}\left(d_1 + 1 - \frac{\varepsilon d(\varepsilon)}{F(\varepsilon)}\right)$. G_a is strictly positive so that the sign of G'_a is equal to the sign of $d_1 + 1 - \frac{\varepsilon d(\varepsilon)}{F(\varepsilon)}$. We have $d_1 + 1 > 0$. The first derivative of $\frac{\varepsilon d(\varepsilon)}{F(\varepsilon)}$ is equal to $\frac{d(\varepsilon)}{F(\varepsilon)}\left[1 + d_1 - d_2\frac{\varepsilon}{1-\varepsilon} - \frac{\varepsilon d(\varepsilon)}{F(\varepsilon)}\right]$, the sign of which is given by the sign of $\left[1 + d_1 - d_2\frac{\varepsilon}{1-\varepsilon} - \frac{\varepsilon d(\varepsilon)}{F(\varepsilon)}\right]$. The limit at 0 is given by:

$$\begin{aligned}
\lim_{\varepsilon\to 0}\left[1+d_1-d_2\frac{\varepsilon}{1-\varepsilon}-\frac{\varepsilon d(\varepsilon)}{F(\varepsilon)}\right] &= 1+d_1-\lim_{\varepsilon\to 0}\frac{\varepsilon d(\varepsilon)}{F(\varepsilon)}\\
&\quad \text{using l'Hôpital's rule}\\
&= 1+d_1-\lim_{\varepsilon\to 0}\frac{d(\varepsilon)+\varepsilon d'(\varepsilon)}{d(\varepsilon)}\\
&= 1+d_1-\lim_{\varepsilon\to 0}1+\frac{d_1(1-\varepsilon)-d_2\varepsilon}{1-\varepsilon}\\
&= 1+d_1-(1+d_1)=0
\end{aligned}$$

Note that $d_2\frac{\varepsilon}{1-\varepsilon}$ is strictly increasing on $(0,1)$.

- Assume that $\frac{\varepsilon d(\varepsilon)}{F(\varepsilon)}$ strictly increases on $(0,1)$. The following inequality holds $\left[1+d_1-d_2\frac{\varepsilon}{1-\varepsilon}-\frac{\varepsilon d(\varepsilon)}{F(\varepsilon)}\right]<0$ which indicates that $\frac{\varepsilon d(\varepsilon)}{F(\varepsilon)}$ is strictly decreasing and contradicts our initial assumption. Therefore, $\frac{\varepsilon d(\varepsilon)}{F(\varepsilon)}$ cannot be strictly decreasing.
- Assume instead that $\frac{\varepsilon d(\varepsilon)}{F(\varepsilon)}$ decreases at least as fast as $d_2\frac{\varepsilon}{1-\varepsilon}$ increases. Then $\left[1+d_1-d_2\frac{\varepsilon}{1-\varepsilon}-\frac{\varepsilon d(\varepsilon)}{F(\varepsilon)}\right]\geq 0$ which violates again the initial assumption that $\frac{\varepsilon d(\varepsilon)}{F(\varepsilon)}$ decreases.
- Hence, the only possibility is for $\frac{\varepsilon d(\varepsilon)}{F(\varepsilon)}$ to be strictly decreasing on $(0,1)$ but at a pace lower than the pace at which $d_2\frac{\varepsilon}{1-\varepsilon}$ increases.

That $\frac{\varepsilon d(\varepsilon)}{F(\varepsilon)}$ decreases on $\varepsilon\in(0,1)$ and $\lim_{\varepsilon\to 0}\frac{\varepsilon d(\varepsilon)}{F(\varepsilon)}=d_1+1$ yields that $G'_a(\varepsilon)>0$ for $1\geq\varepsilon>0$ for all $d_1\geq 0$ and $d_2\geq 0$. Moreover, using l'Hôpital's rule, $\lim_{\varepsilon\to 0}G_a(\varepsilon)=\lim_{\varepsilon\to 0}\frac{\varepsilon^{d_1+1}}{F(\varepsilon)}=\lim_{\varepsilon\to 0}(d_1+1)\frac{\varepsilon^{d_1}}{d(\varepsilon)}=\lim_{\varepsilon\to 0}\frac{d_1+1}{A(1-\varepsilon)^{d_2}}=\frac{d_1+1}{A}$, with $0<\frac{d_1+1}{A}<1$ and $G_a(1)=1$. The function $G_a(\varepsilon)$ is therefore strictly increasing on $(0,1)$ from $\frac{d_1+1}{A}$ to 1.

Similarly, we can show that:

$G'_b(\varepsilon)=\frac{-(d_2+1)(1-\varepsilon)^{d_2}(1-F(\varepsilon))+(1-\varepsilon)^{d_2+1}d(\varepsilon)}{(1-F(\varepsilon))^2}=\frac{G_b(\varepsilon)}{1-\varepsilon}\left(\frac{1-\varepsilon}{1-F(\varepsilon)}d(\varepsilon)-d_2-1\right)<0$ for all $\varepsilon\in(0,1)$ and all $d_1\geq 0$ and $d_2\geq 0$. Using l 'Hôpital's rule, $\lim_{\varepsilon\to 1}G_b(\varepsilon)=\lim_{\varepsilon\to 1}\frac{(1-\varepsilon)^{d_1+1}}{1-F(\varepsilon)}=\lim_{\varepsilon\to 1}(d_2+1)\frac{(1-\varepsilon)^{d_1}}{d(\varepsilon)}=\lim_{\varepsilon\to 1}\frac{d_2+1}{A\varepsilon^{d_2}}=\frac{d_2+1}{A}$, with $0<\frac{d_2+1}{A}<1$ and $G_b(0)=1$. The function $G_b(\varepsilon)$ is therefore strictly decreasing on $(0,1)$ from 1 to $\frac{d_2+1}{A}$.

Moreover, it can be shown that the functions G_a and G_b are convex on $\varepsilon\in(0,1)$. Note that:

$$\begin{aligned} G_a''(\varepsilon) &= G_a'(\varepsilon)H(\varepsilon) + G_a H'(\varepsilon) \\ &= G_a(\varepsilon)\left(H^2(\varepsilon) + H'(\varepsilon)\right) \end{aligned}$$

where $H(\varepsilon) = \frac{1}{\varepsilon}(d_1 + 1 - \frac{\varepsilon d(\varepsilon)}{F(\varepsilon)})$.

Since $G_a > 0$ on $(0,1)$ for all $d_k \geq 0$, a necessary and sufficient condition for $G_a'' \geq 0$ on $(0,1)$ is $H^2 + H' \geq 0$. We already have shown that $(d_1 + 1 - \frac{\varepsilon d(\varepsilon)}{F(\varepsilon)})$.

Indeed, $G_a'' \geq 0$ and $G_b'' \leq 0$ on $\varepsilon \in (0,1)$ and for all $d_1 \geq 0$ and $d_2 \geq 0$.

still need to prove that G_a and G_b are convex.

Since the functions G_a and G_b are convex, strictly increasing and strictly decreasing respectively, we know that $G(\varepsilon) = AG_a(\varepsilon)G_b(\varepsilon)$ is $U{-}shaped$ on $(0,1)$. Therefore, $\max G(\varepsilon) = 1 + \max_k d_k \leq 1 + \alpha + \max_k d_k < a < (a - \delta(\alpha + 1 - a))$ so that inequality 47 is always satisfied.

Hence, $\frac{\partial LHS(\varepsilon^*)}{\partial \varepsilon} < \frac{\partial RHS(\varepsilon^*)}{\partial \varepsilon}$ for all $\varepsilon^* \in (0,1)$ and all $a > 1 + \max_k(\alpha_k, d_k)$. $\blacksquare$

Table 1: Review of Assignment models and their characteristics.

Models: *Author* *Year*	*Roy* * 1951	*Tinbergen* 1956	*Rosen* 1974	*Sattinger* 1975	*Rosen* 1978	*Teulings* 1995	*Teulings* 2005	*This paper*
Characteristics								
Heterogeneity and skills								
Workers	Yes	Yes	Yes	Yes	Discrete	Yes	Yes	Yes
Jobs	Discrete	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Multi-skill	Yes	Yes	No	No	Yes	No	No	Yes
Human capital formation								
Endogenous	Yes	No	No	No	No	No	No	Yes
Equilibrium and solution								
General/Partial	Partial	General	General	General	Partial	General	General	General
Closed form solutions	_	Yes	No	No	_	Yes	Yes	Yes
Structure of substitution	Unknown	Unknown	Unknown	Unknown	CES	CES	DIDES	CES

$*$ Reference to Willis and Rosen's (1979) application of Roy's model

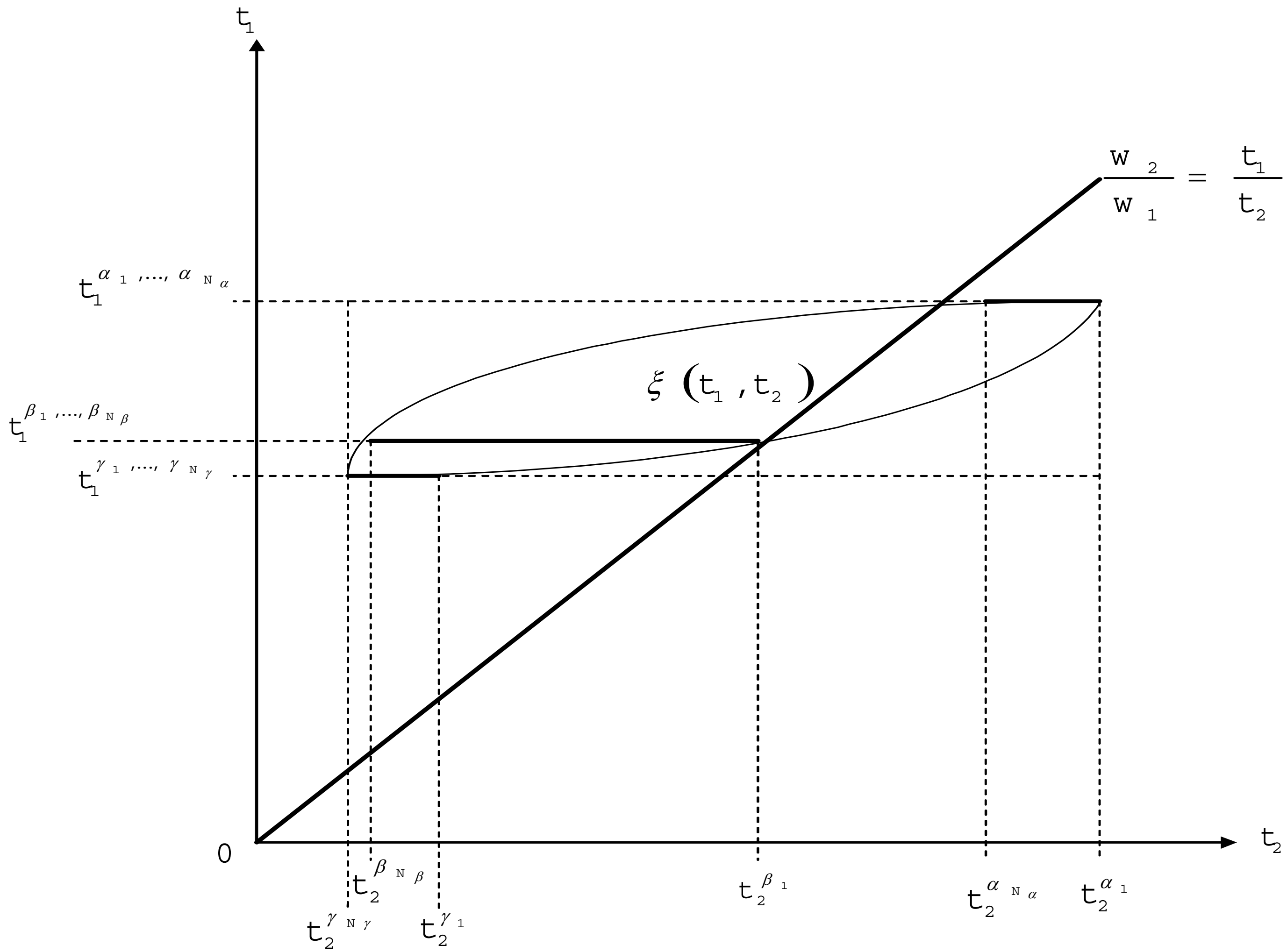

Figure 1: Educational self-selection with positive hierarchical sorting.

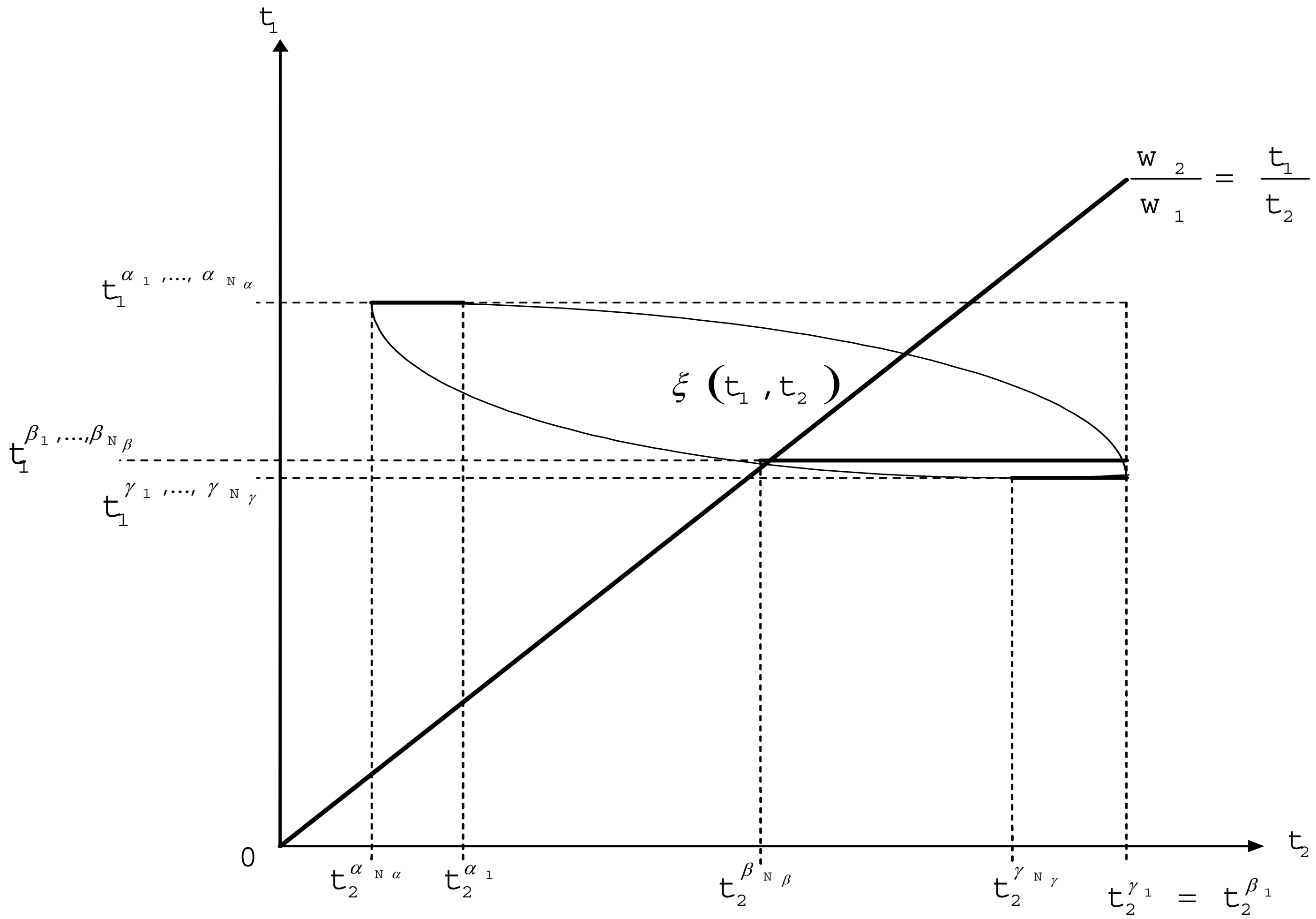

Figure 2: Educational self-selection without hierarchical sorting.

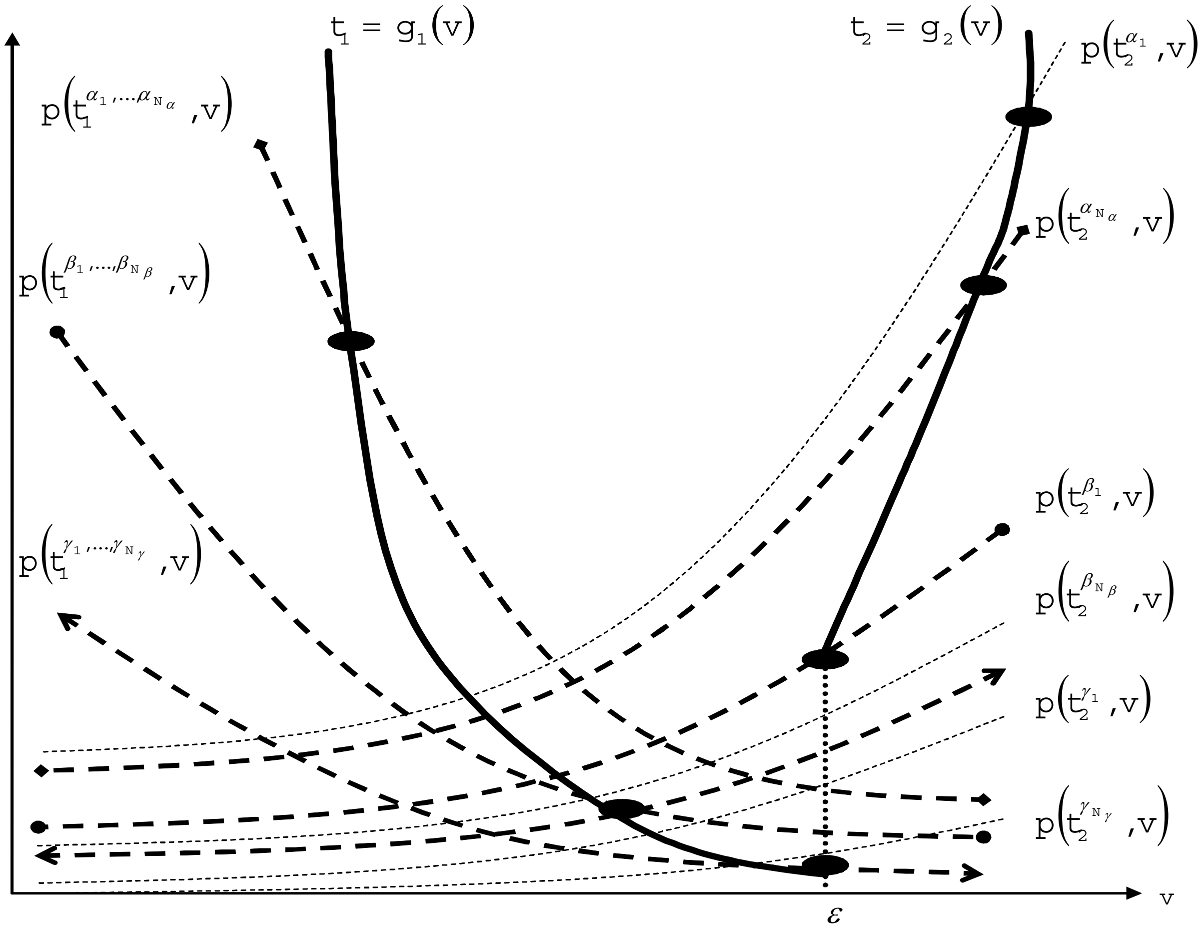

Figure 3: Tasks assignment with hierarchical sorting.

Table 2: Wage distribution in the US 1963-1996.

Parameters	Model 63	Data* 63	Model 80	Model 80	Data* 80	Model 96	Model 96	Data* 96
b_1	1.00		1.00			1.00		
b_2 $(= \frac{b_2}{b_1})$	0.27	100	0.38	141	174	1.30	481	365
α_1	0.83		1.10			0.90		
α_2	0.25		0.30			0.25		
δ	-0.15		-0.65			-0.40		
d_1	0.95		1.25			0.20		
d_2	0.97		1.60			0.70		
a	2.81		3.80			2.61		
$\sigma = \frac{a}{a-1}$	1.55		1.36			1.62		
Equilibrium								
ε^*	0.72		0.59			0.47		
$\ln \mu$	0.35	(0.36)	0.37		0.37	0.58		0.60
$\frac{L_S}{L_U}$	0.24	(0.22)	0.41		0.42	0.68		0.70
Wage distribution Deciles in logs								
1^{st}	-0.03	100	0.10	113	110	0.04	107	90
5^{th}	0.39	100	0.57	118	125	0.50	111	115
9^{th}	0.98	100	1.22	124	130	1.31	133	140
Residual inequality Deciles in logs								
$9^{th} - 5^{th}$	0.77	100	0.97	120	106	1.31	154	115
$5^{th} - 1^{st}$	0.27	100	0.42	115	107	0.48	121	115
$9^{th} - 1^{st}$	1.04	100	1.39	135	113	1.79	175	130

$*$Source: Acemoglu (2002)

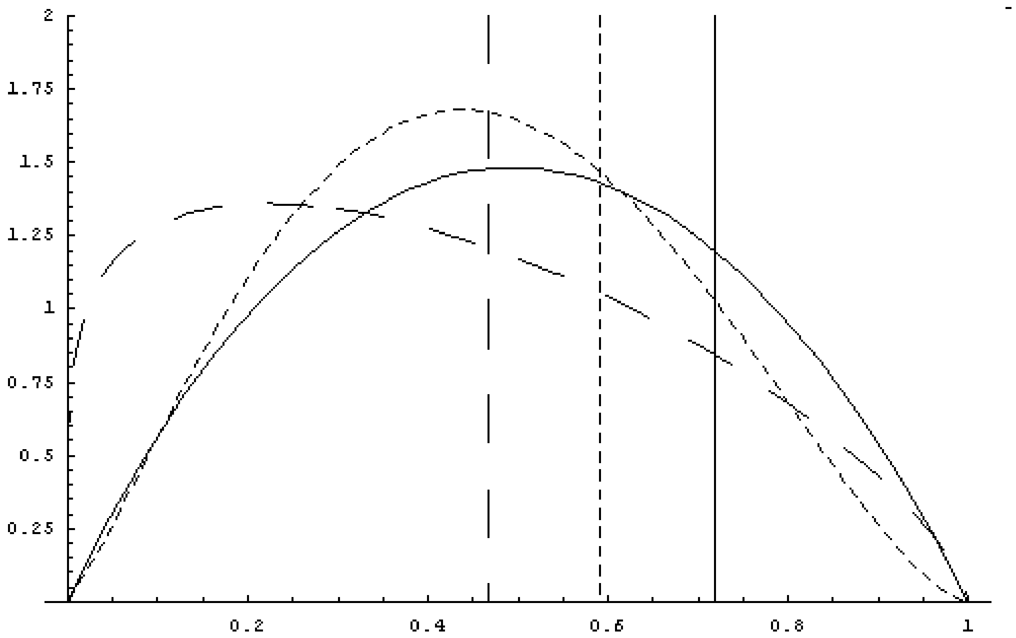

Figure 4: Evolution of the distribution of tasks in the US: 1963 full line, 1980 short dash and 1996 long dash.