Wage Inequality and Education Policy with Skill-biased Technological Change in OG Setting

Árpád Ábrahám*†

November, 2001

Abstract

This paper analyzes the impact of different education policy regimes on efficiency and inequality assuming stochastic skilled-biased technological change. Wage dispersion is determined by the heterogeneity of skills by allowing for productivity differences due to education, ability and age. These features together with the usage of many overlapping generations enable us to follow several statistics of the wage distribution along stochastic technological change. The implications of the model are compared with some stylized facts on the time pattern of the U.S. wage distribution and human capital accumulation. The model perform reasonably well in reproducing a set of stylised facts. Our results highlight the fact that the introduction of many overlapping generations can be crucial for understanding long-run labour market dynamics. Finally, we evaluate the efficiency and distribution implications of different types of education policies. Two types of tuition subsidies are introduced together with linear and progressive taxation. Tuition subsidies generally improve individual and aggregate welfare and decrease inequality, however agents with different levels of ability and age would disagree on which policy regime should be chosen.

*Universitat Pompeu Fabra and London Business School; *Correspondence:* London Business School, Regent's Park, London NW1 4SA; *e-mail*: aabraham@london.edu

†I gratefully acknowledge the support and guidance of Albert Marcet and fruitful discussions with Jan Eeckhaut and Morten Ravn. The usual disclaimer holds.

1 Introduction

This paper studies the evolution of wage differentials and human capital accumulation in a model of many overlapping generations where all the generations contain a continuum of heterogenous agents facing stochastic skill-biased technological progress.

The issue of wage differentials during the last 40 years is particularly important because they are one of the most important determinants of income inequality. The consequences of the current technological development on inequality are still not clear. At the same time governments all over the world intervene in the education decision of individuals by subsidizing education. The instruments and the level of these education subsidies vary in a great extent among countries. Education subsidies are typically introduced because of efficiency and distribution reasons, in particular tuition subsidies are often thought to be the best instrument of reducing the recent increase in inequality and promoting economic growth at the same time. However, these transfers are typically financed by taxation which modifies both the incentives for human capital accumulation and the wealth distribution itself. Therefore the final impact of tuition subsidies is not clear neither on inequality or on the distribution of wealth and one has to study them in a general equilibrium framework in order to take into account all these forces.

If one wants to study the consequences of a given policy on some economic phenomena he has to do that in a framework which is able to reproduce the empirical regularities related to those phenomena. The issue of wage differentials due to skill differences has been studied in the empirical labour literature in a great extent both for the U.S. and for Europe. Some of the main results can be summarized by the following widely accepted stylised facts: directly and indirectly

- the wage premium of college graduates followed a cyclical pattern in the last 40 years
- labor supply of the better educated increased continuously even during the period when the wage premium decreased (the 1970´s)
- residual wage differentials (the part of wage dispersion not explained by age and education) has been increased
- the real wages of low-skilled employees decreased

Our objective in this paper is to build up a model which has the potential of explaining these stylised facts and can be applied to analyze the efficiency and distribution implications of education policy.

"Skill-biased" technological change is considered to be one of the most important forces behind the observed time pattern of wage inequality. It is widely accepted that developed economies experienced a technological change which favored individuals with better skills during the last decades. Empirical research supports the idea that more and more widespread use of computers and information technology in general, increased the demand for more skilled labor force (see for example Autor at al [1998]).

The above mentioned stylised facts seem to make it necessary to introduce a compound notion of skills into our model. We differentiate between three components of skills in this paper:

1. "innate" abilities, which are either determined by genetics and/or by social background

2. skills acquired through formal education

3. skills acquired informally: on-the-job training, learning by doing, experience

Educational attainment is chosen endogenously while ability and the life cycle pattern of productivity is given exogenously[1]. This specification of skills enables us to study the impact of educational policies on welfare of agents differing in their abilities, educational level and age at different stages of the technological progress.

Both the research objective and the basic structure of the model is closely related to the work of Heckman et al. [1998a, 1998b and 1999]. They build an "empirical dynamic general equilibrium model" of heterogeneous agents in OG setting, estimate its parameters on U.S. data and use it to evaluate the impact of tax and tuition policies on human capital accumulation. Our model is simpler in some dimensions (most importantly they make the accumulation of informal skills endogenous and allow for savings). However this simplification enables us to introduce uncertainty and solve for the stationary distribution and for the transition dynamics more rigorously. In

[1] The accumulation of informal skills could be considered endogenous as well (like in Heckman et al [1998a]), however since we want to concentrate on education policies we considered it exogenous for simplifying the analysis.

addition to this, we use a different specification for skill-biased technological change and have a different focus on the policy analysis. Apart from these differences we obtain remarkably similar results on both the predictive power of this type of models and on the efficiency and distribution implications of tuition policies (obviously within the limits of the comparability of the projects). On the one hand these results validate our simplifying assumptions to a some extent, on the other hand they help to identify some potentially fruitful extensions of our model.

Another potential advantage of our somewhat simpler model is that it enables us to focus on some specific mechanisms which we consider important for understanding the evolution of wage differentials and human capital accumulation. Under the assumption of perfectly flexible labor supply increasing demand for skilled labor could not lead to cyclical and long-lasting changes in the wage premia because relative supplies and relative wages would adjust to the demand shift instantaneously. However we see significant movements in the wage premia. That implies that labor supply of skills cannot adjust, at least not perfectly and/or not immediately. There is empirical evidence that human capital accumulation (college enrollment) reacts to the improvement of wage differentials, however the impact on the stock of human capital is lagged and last long (see Mincer [1994]). This implies that there should be some limitations to the adjustment of skilled labour supply what may cause the persistent behavior of wage differentials. We identify and use two straightforward ones of such limitations in this model:

(1) Only young people can choose their educational attainment, however they constitute only a small fraction of the total labor force. Therefore the change in the stock of human capital is mostly determined by the difference between the educational attainments of the entering and retiring birth cohort.

(2) The cost of acquiring higher education is positively correlated with ability, that is for less able people it can be very costly to acquire higher education.

Most of the labor force participants acquire their first degrees before reaching the age of 30 (actually before reaching 25 in the majority of the cases). Moreover, they only enter the labour force after completing their studies. This implies that it can take several cohorts for the labor supply to adjust to a relatively big shock. Therefore the temporary low level of human capital investment can cause an increase in wage differentials, however in the longer run labour supply will adjust which is going to the decrease wage differentials. Moreover, if we consider the ability distribution of new cohorts constant then increasing the proportion of people entering post-secondary education means that in average less able people are attending universities. It is more costly to obtain a degree for them, therefore higher wage premium is needed to give them incentive to enroll in universities. This enhances the increase in wage differentials due to the demand shift. However it is going to increase the heterogeneity among the college graduates and decrease among the rest of the population. Since college graduated earn more, this may result in increase in residual inequality. Finally, the average ability of unskilled labour decreases also, which could lead to the decrease of the average wages of this skill group.

The first limitation is present in the work of Heckman at al [1998a], however it is somewhat weakened because although agents cannot modify their educational decision they can still modify their human capital stock by undertaking informal training in later periods of their life. The second mechanism is not present in their model, because the direct cost of post-secondary education is independent of ability in their model.

The structure of the model, in particular the fact that many heterogeneous generations live together, allows us to follow the wage distribution along the technological progress. We can study not only the moments of the wage distribution but we can follow the wage movement of any particular agent in the distribution. This feature is particularly useful because it enables us to study the distributional consequences of different policies in great detail. We implement a similar policy analysis to Heckman at al [1999], but we compare the impact of *combinations of* different type of tax policies and tuition subsidies instead of studying them separately. We explore that tuition subsidies in general increase overall welfare and decrease inequality, however there is disagreement between agents with different ability and different age on which combinations should be chosen. More remarkably the policy regime which implies the highest aggregate welfare (lump sum subsidies combined with linear taxation) is not the most preferred policy of any of the age or ability groups. On the other hand, if subsidies are introduced on the onset of technological change they may increase inequality within the first generations.

We continue the presentation in Section 2 with a brief summary of the related empirical literature on U.S. and OECD wage differentials and review the related models of wage differentials and human capital accumulation. We describe our overlapping generation model in Section 3. The competitive equilibrium of the model is characterized in Section 4. Section 5 presents the calibration process, the way the parametrized expectation approach was applied to this problem and the simulation results are discussed and compared to the previously presented stylised facts. Education policies are introduced to the model and their efficiency and distributional implications are analyzed in Section 6. Finally, we conclude and indicate directions for future research in Section 7.

2 Related literature on wage differentials

2.1 Stylized facts

A vast amount of empirical research has taken place to explore the time pattern of wage differentials and the structure of labour supply since the early 90's. Most of them concentrated on the evolution of the wage distribution in the United States over the latest decades (60's to the 90's). There are many papers on these issues, however the one of Katz and Murphy [1992] and Juhn, Murphy and Pierce [1993] were probably the most influential ones. Most of the facts explored for the U.S. were established as well for a set of OECD countries by Berman, Bound and Machin [1998].

The most important findings of these papers (relevant to this study) can be summarized by the following list of stylized facts:

1. Return to the education (college/high school wage ratio) moved cyclically: it increased in 60's, decreased in the 70's and increased again in the 80's and 90's. However, during the overall period the return to college has increased significantly.

2. At the same time residual wage inequality (wage variation not explained by gender, experience and education) also increased continuously during this period. As a consequence its relative weight in the total variation of wages has increased. These two facts together led to increase in overall inequality in the long run.

3. The relative supply of more educated people (the share of them in the working force) and college enrollment increased throughout the whole time period. However, college enrollment had a stagnation period in the mid 70's.

4. The real wages of high school graduates decreased in spite of their decreasing relative supply.

These stylized facts give some intuition what components we should include in a model of wage differentials determined by skill differentials. Both educational and residual wage differentials are important, therefore we have to allow for both. Consequently, we need to have at least two dimensions of skills as production factors. We should allow for shock(s) which alter the relative return on the skilled and unskilled labor force. "Skill-biased" technological change will provide that feature. We need a model where educational decision is endogenous, but reacting "slowly " to shocks, to induce persistent changes in skill premia. Many overlapping generations provide this feature. In order to have a realistic model of human capital accumulation we have to allow for a life-cycle pattern of productivity for which we also need to introduce many generations living together.

2.2 Related theoretical literature

There is a considerable recent literature on the effect of technological change on earnings inequality and human capital formation. However, there is still no consensus about what is the common cause behind the four stylised facts enlisted previously. Krussel et al. [1997] show in a calibrated stochastic growth model with two types of labor and capital as factors of production that the observed pattern of educational wage differentials can be reproduced by skill-biased technological change. They require specific complementarity properties between capital and the two types of labor to obtain those results. Moreover, their infinitely lived representative agent model would not be appropriate to study the distributional consequences of technological change on education subsidies. Moreover, they consider the educational choice exogenous. As Heckman et al.[1999] show, this assumption (which is also typical in the empirically oriented demand-supply analysis of Katz and Murphy [1993] and Juhn et al.[1993]) can be very misleading especially when one wants to draw policy conclusions.

On the other hand, there are some attempts to analyze OG models with endogenous skill acquisition to get analytical results for inequality. An overlapping generation model was designed by Rios-Rull [1993] to analyze skill acquisition in general equilibrium business cycle framework with home-production. The shortcoming of that model is it does not allow for within group inequality and has just two generations. However, calibrating this model to the US economy gives still a good fit to some stylized facts on (between-group) inequalities and labor force participation in the stationary distribution. However, that model cannot really give predictions for a transition period of 40 years because of the assumption of two generations. On the other hand, in other paper Rios-Rull [1996] designed and calibrated a general equilibrium model with agents living for variable periods and was able to recover most of the well-known business cycle facts, however that model would become very difficult to handle if one introduces heterogeneity of agents and/or endogenized educational choice.

Another strand of literature allows for heterogeneity both in ability and education. Our paper is closely related to this literature. They use the fact that innate ability is an important factor for both educational choice and productivity. These ideas are present in Galor and Tsiddon [1997], Galor and Moav [2000], Rubinstein and Tsiddon [1998], Acemoglu [1998] and Caselli [1999]. These models are based on the fact that when technological change occurs innate abilities became more important, because agents have to cope with new (unexperienced and/or not "studied") situations, and more able people can do this more efficiently. With this structure they are generating both between and within-group inequalities. However they use only 2 overlapping generation typically, which is a problematic assumption if we actually want to explain facts occurring in 40 years, since this is the life-time of only one generation approximately. Moreover, these models miss the feature of "slow reaction" in our multigenerational model. Moreover, if technological change enhances the importance of ability and deteriorate the one of experience than the rate of return to experience should decrease when we see increase in residual inequality, however this is not confirmed by the empirical evidence (see for example Katz and Murphy [1992]).

Our model is similar to these models in its basic structure, however it is more macro-oriented in the sense that we can study quantitative implications of the model with respect to the dynamic path of wage distributions, what is not possible with those models. In addition to this we show that residual inequality may increase without imposing (exogenously) that ability is a more "important" production factor during technological transition.

However, our model is less quantitative than the one of Heckman et al.[1998a] because we do not have endogenous on the job training and because we do not allow for savings[2]. On the other hand, we have introduced uncertainty and our specification allows us to study the full rational expectations equilibrium along the transition towards the stationary distribu-

[2]They have also capital in their model, however their estimation results show, that the exclusion of capital (or keeping it constant) does not change the implications of their model regarding human capital formation and the path of wage differentials.

tion. They use "almost" perfect foresight equilibrium to study the transition between two steady states.[3] We use crucially different specification of skill-biased technological progress and we will show that both specifications have their advantages and shortcomings. Educational choice is modelled somewhat differently as well. Their model is empirical general equilibrium because they chose features which make direct estimation of the parameters possible. Our approach is more theoretical, we would like to build a minimal model which has the mechanisms we described above, has the potential to resolve the puzzle given by the four stylised facts and still provide a framework to compare different educational policy regimes in a meaningful way. Our policy analysis confirms the results they obtained, however we gain some more insights about distributional effects of different ways of subsidizing higher education.

3 The model

Agents We have a continuum of agents. Each agent lives for n periods. Cohort size is constant and normalised to 1 each period. Therefore the measure of the total population is always n.

Each agent decides between accumulating human capital ("going to university") or starting to work as an unskilled worker at the first period of her life cycle . If she went to the university from the next period on she is a skilled worker, otherwise she remains unskilled for her whole life. There are no "evening universities" and for simplicity there is no return from the working force to studying. This assumption is justified by the fact that most of college graduates obtain their first degree before they reach the age of 25.

Agents are heterogenous in their abilities, which we denote with γ. Ability has a time invariant distribution with density function $f(\gamma)$. Ability increases productivity and decreases the direct cost of education. Accumulating human capital has two types of costs: foregone labor income and "direct" educational costs, tuition. Educational costs $a(\gamma)$ are differing across agents, these differences are usually explained by the fact that it is easier to obtain a university degree for a more clever person than for a less intelligent person (that implies $a'(\gamma) < 0$). Another interpretation would be financial market imperfections, namely for the less able it is much harder to get fi-

[3]They assume that the agents have perfect foresight of the endogenoues variables, however they do not expect technology to change, although they model technological change as a deterministic time trend.

nance for their education. With this interpretation, $a(\gamma)$ could be the net present value of interest payments on student loans.[4]

Preferences We assume linear preferences with homogeneous time discounting: $R(c_1, ..., c_n) = \sum_{i=1}^{n} \delta^i c_i$, where $R(\cdot)$ is the life-time utility of the agent, δ is the is the discount factor and c_i measures the consumption of the agent in the ith period of her life. This formulation allows us to plug in labor income directly to the utility function, because risk neutrality implies that the only feasible equilibrium interest rate would be $1/\delta$, therefore the agents are indifferent between borrowing and consuming currently or postponing consumption. We should follow the accumulation of savings of each agent without this assumption. That would make the model much more complicated both analytically and numerically because it would at least double the number of state variables. This assumption is standard in the related OG literature (see for example Chari and Hopenhayn [1991] or Rubinstein and Tsiddon [1998]). We can interpret this specification as agents maximize their life-time permanent income and this decision is separate from their period by period savings-consumption decision.

The agents' problem We define $w_t^0(\gamma, i)$ as the unskilled and $w_t^1(\gamma, i)$ as the skilled wage rate of an agent with ability γ and age i at time t. Similarly, we denote the time t consumption of an agent with ability γ, born in period $t - i$ with $c_t(\gamma, i)$. $s_t(\gamma)$ refers to a choice variable which takes 1 if the agent with ability γ, born at period t decides to attend university and 0 otherwise. Having introduced this notation we can write down formally the decision problem of a newly born agent with ability γ at time t:

$$\max_{s_t(\gamma), \{c_{t+i}(\gamma, i)\}_{i=0}^{n-1}} E_t \left(\sum_{i=0}^{n-1} \delta^i c_{t+i}(\gamma, i) \right)$$

$$\texttt{if } s_t(\gamma) = 1 \quad \begin{cases} c_t(\gamma, 0) = -a(\gamma) & \\ c_t(\gamma, i) = w_t^1(\gamma, i) & i = 1, ..., n-1 \end{cases}$$

$$\texttt{if } s_t(\gamma) = 0 \quad c_t(\gamma, i) = w_t^0(\gamma, i) \qquad i = 0, ..., n-1$$

[4] We shall consider γ as the combination of intellectual ability and social background for this interpretation. We know that both increases productivity. They both make university degrees more accessible as well because ability decreases the time required for studying and social background provides better financial possibilities.

Technology We assume that there is one profit-maximizing firm each period using skilled and unskilled labour as production factors[5]. The production function is CES with skill-enhancing technological progress (A_t).

$$Y_t = f(L_t, A_t, H_t) = (\alpha L_t^\rho + (1-\alpha)(A_t H_t)^\rho)^{1/\rho} \tag{1}$$

L_t and H_t are the units of unskilled and skilled workers employed, respectively. The CES specification is useful because $\frac{1}{1-\rho}$ measures the demand elasticity of substitution between skilled and unskilled labor. Many empirical studies focused on the estimation of this parameter which is going to help us in calibrating the model. Also, since this is a constant return to scale technology in equilibrium there will be zero profit which makes the general equilibrium structure of the problem simpler (no dividends). A_t represents the skill-biased technological change. It measures the productivity difference between skilled and unskilled labor. We can interpret $A_t H_t$ as the measure of skilled workers in efficiency units. A_t is assumed to be lognormal with (non-stochastic) mean equal to $\overline{A}$ and it is assumed to follow a stationary autoregressive process:

$$\log(A_{t+1}) = (1-\psi)\log(\overline{A})\psi + \log(A_t) + \varepsilon_{t+1}, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \tag{2}$$

We will model technological progress as a stochastic transition from $A_0 \ll \overline{A}$ to the stationary distribution. A_0 is chosen in a way that the transition in average lasts for a sufficiently long period of time. Our way of modelling skill-biased technological change is not the only possible one. Note, that along the transition path this specification implies growth in output. Heckman et al.[1998a] assumes that technological change is a decreasing trend in α ($A_t = 1$), that is in their interpretation technological change is not progress but rather reallocation of resources. Their specification has the disadvantage that in the long run there is a very little impact on output and on wage differentials, that is if we take their results seriously we should expect a huge drop in wage differentials in the following decades[6]. On the other hand, our specification has the disadvantage that as we will see skill-biased

[5] Galor and Moav [2000] shows analytically that the introduction of capital as production factor together with the assumption of small open economy has no significant effect on the evolution wage differentials and on human capital accumulation. The estimations of a nested CES specification in Heckman et al. [1998a] supports this claim even if we assume closed economy.

[6] They admittedly only plot the first 30 periods of their simulation results which is very far from the long run steady state, partly because they assume that technological transition lasts only for 30 years then it stops.

technological change favors directly and indirectly the unskilled labour force as well.

Aggregation We have included only skilled and unskilled labour without specific correspondence to education, ability and age in the production function. We did this because we assume that the individual productivity endowments of agents are depending on ability, age and education, but within a given education group productivity units of different (with respect to ability and age) agents are prefect substitutes. We chose this particular form of skill aggregation because those skills (managerial, IT, etc.) gained most weight which require a college degree in the structure of labour demand (see Autor et al.[1998]). In addition to this, the empirical evidence by Hamermesh [1993] shows that individuals with the same schooling level but with different age are highly substitutable with each other. Finally, since we want to focus on education policies later this particular form seemed to be the most natural one to choose.

The assumption of perfect substitutability allows us using the following "linear aggregation schemes", where $p^0(\gamma,i)$ and $p^1(\gamma,i)$, denotes the efficiency unit endowments of an unskilled or skilled agent with ability γ and age i respectively:

$$L_t = \sum_{i=0}^{n-1} \int_{\gamma|s_{t-i}(\gamma)=0} p^0(\gamma,i) f(\gamma) d\gamma \tag{3}$$

$$H_t = \sum_{i=1}^{n} \int_{\gamma|s_{t-i}(\gamma)=1} p^1(\gamma,i) f(\gamma) d\gamma \tag{4}$$

The productivity endowments capture changing productivity along the life-cycle, that is $\{p^0(\gamma,i)\}_{i=0}^{n-1}$ represents the age-productivity profile of the unskilled agent with ability γ. The limits of the integrations are determined by individual optimization in equilibrium.

Laws of motions Finally, the evolution of supply of skilled and unskilled labour is determined by the human capital accumulation decision of the agents. We define U_t as the proportion of agents enrolled in universities at period t (that is $U_t = \int_{\gamma|s_t(\gamma)=1} f(\gamma)d\gamma$). Then by our previous assumptions the laws of motions of the supply of skilled (h_t) and unskilled workers (l_t) are the following ones respectively:

$$h_{t+1} = h_t - U_{t-n+1} + U_t \tag{5}$$

$$l_t = l_{t-1} - (1 - U_{t-n}) + (1 - U_t) = l_t + U_{t-n} - U_t = n - (h_t + U_t) \tag{6}$$

The latter is true by the fact that every new-born generation has unit size and altogether n generations are living together.[7]

4 The competitive equilibrium

4.1 Definition of the competitive equilibrium

In this section we define the competitive equilibrium (CE) of our model. The CE involves threshold levels for enrolling university $\{\widehat{\gamma}_t\}_{t=0}^{\infty}$, where $\widehat{\gamma}_t$ refers to the ability level above which all agents decide to become skilled.[8] CE also includes sequences of measures of skilled and unskilled workers and university enrollment rates: $\{h_t\}_{t=0}^{\infty}$, $\{l_t\}_{t=0}^{\infty}$, $\{U_t\}_{t=0}^{\infty}$, where U_{-i} $(i = 1, ..., n-1)$ hence $h_0 = \sum_{i=1}^{n-1} U_{-i}$ are given. CE also defines sequences of wages for skilled and unskilled workers: $\left\{w_t^1(\gamma, i)\right\}_{t=0}^{\infty}$ and $\left\{w_t^0(\gamma, i)\right\}_{t=0}^{\infty}$ respectively.

The competitive equilibrium has to satisfy the following conditions:

(i) The wage sequences have to be compatible with profit maximizing behavior of firms. Since we have constant returns to scale, the unit payments of the different production factors should be equal to their marginal productivities. We have the following expressions for "skill prices" (by efficiency units) for the two different types of labor:

$$\overline{w}_t^1 = (\alpha L_t^\rho + (1-\alpha)(A_t H_t)^\rho)^{1/\rho - 1} (1-\alpha) A_t^\rho H_t^{\rho-1} \qquad \forall t \geq 0 \tag{7}$$

$$\overline{w}_t^0 = (\alpha L_t^\rho + (1-\alpha)(A_t H_t)^\rho)^{1/\rho - 1} \alpha L_t^{\rho-1} \qquad \forall t \geq 0 \tag{8}$$

[7] Note, that although (5) and (6) have a nice recursive structure it is not true for the efficiency units of skilled and unskilled labour (L_t and H_t) because we need to keep track of the educational attainment of all currently living generations ($\{U_{t-i}\}_{i=1}^{n-1}$) in order to update H_t.

[8] We cannot assure unique threshold levels which satisfy the requirements of competitive equilibrium in general. We will show that our specification will imply the existence and the uniqueness of the threshold "around" the non-stochastic steady state of the model.

We can solve for the individual wages as well, the linear aggregation scheme (perfect substitutability) implies:

$$w_t^s(\gamma, i) = \overline{w}_t^s p^s(\gamma, i) \tag{9}$$

(ii) The CE has to be consistent with the agents' optimal decisions concerning human capital accumulation. An agent with ability γ born at period t will attend university ($s_t(\gamma) = 1$) if the following condition is satisfied:

$$a(\gamma) + w_t^0(\gamma, 0) \leq E_t \left[\sum_{i=1}^{n-1} \delta^i \left(w_{t+i}^1(\gamma, i) - w_{t+i}^0(\gamma, i) \right) \right] \tag{10}$$

In other words an individual will go to university if the direct and opportunity cost of doing that is smaller than the expected gain of acquiring higher education. Assume for now, that (10) is strictly monotonous in γ and there exists a level γ' for which (10) is not satisfied. (Later we will discuss the general validity of this assumption.) This would imply the existence of threshold level of ability and $\widehat{\gamma}_t$ is given by (11).

$$a(\widehat{\gamma}_t) + w_t^0(\widehat{\gamma}_t, 0) - E_t \left[\sum_{i=1}^{n-1} \delta^i \left(w_{t+i}^1(\widehat{\gamma}_t, i) - w_{t+i}^0(\widehat{\gamma}_t, i) \right) \right] = 0 \quad \forall t \geq 0 \tag{11}$$

If we define $F(\gamma)$ as the cumulative distribution function of ability then $U_t = 1 - \Phi(\widehat{\gamma}_t)$.

(iii) The CE has to satisfy the laws of motions for the stocks of skilled and unskilled labor (5) and (6) with $\{U_{-i}\}_{i=1}^{n-1}$ given.

(iv) All markets clear. This is implicit in our model since wages are such that demand of both types of labour equals to its supply.

4.2 Characterization of the competitive equilibrium

We introduce some simplifying assumption and specific functional forms now. These assumptions are useful to get some insight analytically and they are also necessary for the later numerical solution of the model. We chose the distribution of ability being normal: $\gamma \sim N(\overline{\gamma}, \sigma_\gamma^2)$. We have chose age-productivity profiles separable in ability and age:

$$p^s(\gamma, i) = \gamma + P^s(i) \qquad i = 0, ..., n-1 \qquad s = 0, 1 \tag{12}$$

This assumption implies that ability and age are perfect substitutes. Separability makes the exposition and the calibration easier. However even if we introduced some complementarity between ability and experience, the main results would not change significantly.

(3) and (4) becomes more transparent with these two further assumptions:

$$L_t = \sum_{i=0}^{n-1} \left[F(\widehat{\gamma}_{t-i}) \left(E(\gamma \mid \gamma < \widehat{\gamma}_{t-i}) + P^0(i) \right) \right] = \tag{13}$$
$$= \sum_{i=0}^{n-1} \left[(1 - U_{t-i}) \left(\overline{\gamma}^0_{t-i} + P^0(i) \right) \right]$$

$$H_t = \sum_{i=1}^{n-1} \left[(1 - F(\widehat{\gamma}_{t-i})) \left(E(\gamma \mid \gamma > \widehat{\gamma}_{t-i}) + P^1(i) \right) \right] \tag{14}$$
$$= \sum_{i=1}^{n-1} \left[U_{t-i} \left(\overline{\gamma}^1_{t-i} + P^1(i) \right) \right]$$

Here $\overline{\gamma}^1_t$ and $\overline{\gamma}^0_t$ denotes the average ability of the skilled and unskilled labour force of the generation born at time t respectively.[9] (14) allows us to give a "law of motion" for H_t:

$$H_{t+1} = H_t + U_t \left(\overline{\gamma}^1_t + P^1(0) \right) - U_{t-n+1} \left(\overline{\gamma}^1_{t-n+1} + P^1(n-1) \right) + $$
$$+ \sum_{i=1}^{n-2} \left[U_{t-i} \left(P^1(i+1) - P^1(i) \right) \right] \tag{15}$$

It is clear from (15) that the change in the supply in educated labour is determined by three factors:

1. A new generation enters the skilled labour force with a given size (and hence average productivity) without experience.

[9] By the moments of truncated normal distribution we know that these averages have the following functional form ($\widetilde{\gamma}_t \equiv \frac{\widehat{\gamma}_t - \overline{\gamma}}{\sigma_\gamma}$)

$$\overline{\gamma}^1_t = \overline{\gamma} + \sigma \frac{\phi(\widetilde{\gamma}_t)}{1 - \Phi(\widetilde{\gamma}_t)} \qquad \overline{\gamma}^0_t = \overline{\gamma} - \sigma \frac{\phi(\widetilde{\gamma}_t)}{\Phi(\widetilde{\gamma}_t)}$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the density and distribution functions respectively of a standard normal random variable (see Johnson and Kotz [1970]).

2. The oldest generation lived previous period retires and "takes away" its educational attainment weighted by its high experience.

3. The productivity of all the other generations lived in the previous period is updated according to their age-productivity profile.

It is obvious from (15) that if we want to predict the evolution of skilled labour supply after time t we need to keep track of the educational attainments of all of the generations living at time t. (It is true for the supply of unskilled labour (see (6)) and consequently for the evolution of wages as well.)

Another functional assumption we made was: $a(\gamma) = C - \kappa\gamma$. This specification together with (12) implies that we can rewrite (11) in the following ways:

$$\begin{aligned} & C - \kappa\widehat{\gamma}_t + \overline{w}_t^0 \left(\widehat{\gamma}_t + P^0(0)\right) \qquad (16) \\ = \quad & E_t \left[\sum_{i=1}^{n-1} \delta^i \left(\overline{w}_{t+i}^1 \left(\widehat{\gamma}_t + P^1(i)\right) - \overline{w}_{t+i}^0 \left(\widehat{\gamma}_t + P^0(i)\right)\right)\right] \\ = \quad & E_t \left[\sum_{i=1}^{n-1} \delta^i \left(\overline{w}_{t+i}^1 P^1(i) - \overline{w}_{t+i}^0 P^0(i)\right)\right] + \widehat{\gamma}_t E_t \left[\sum_{i=1}^{n-1} \delta^i \left(\overline{w}_{t+i}^1 - \overline{w}_{t+i}^0 \right)\right] \end{aligned}$$

The expected gain from human capital accumulation can be decomposed to a fixed part which does not depend on any particular agent's ability level and to variable part which depends on the ability of the given agent (See 16). However, the supply of unskilled labour and consequently also current and future skill prices ($\overline{w}_{t+i}^s$) do depend on $\widehat{\gamma}_t$ (see (14) and (13)). Notice, that if this was not the case then (16) would be clearly monotonous in $\widehat{\gamma}_t$ (it would be linear) and because the support of $\widehat{\gamma}_t$ is the real line we would have existence and uniqueness trivially. This dependence makes it difficult to prove existence and uniqueness for all possible predetermined realizations of university threshold levels ($\{\widehat{\gamma}_{t-i}\}_{i=1}^{n-1}$). So far we had proven existence and uniqueness of the threshold of ability only for the non-stochastic steady state. We claim that if the variability of $\widehat{\gamma}_t$ is not very big in equilibrium the existence and uniqueness is guaranteed in the stochastic model as well.

Proposition 1 *Consider the non-stochastic steady state of our model. If* $0 < \rho < 1$ *then under fairly general conditions there exists a unique threshold level of ability* γ^* *which solves (16).*

Proof See Appendix A.

Conjecture 1 *For a given level of technology* $\widetilde{A}$, *there exist a* λ *for which if* $\left|\widehat{\gamma}_{t-i} - \gamma^*(\widetilde{A})\right| < \lambda$, $i = 1, .., n-1$ *then (16) has a unique solution, where we define* $\gamma^*(\widetilde{A})$ *as the threshold in the non-stochastic steady state with technology* $\widetilde{A}$.

Note that (34) in Appendix A is continuous in γ^*. Our conjecture has to be true because for sufficiently small deviations in the level of predetermined educational levels the single-crossing property which we explored in the non-stochastic steady state should be kept.We calibrated the model that the assumptions stated in Appendix A are satisfied, moreover we did not have problems of uniqueness and existence along the equilibrium path, this qualifies our conjecture. It seems that we can allow for reasonably big deviations.

5 The solution of the model

5.1 Calibration

We calibrated the model to the U.S. economy because most of the stylized facts about wage differentials are explored in detail for the U.S. Our general strategy was to use estimates for those parameters which can be observed or at least estimated from actual data.[10] We calibrated the rest of the parameters in a way that we have similar level of university enrollment and wage differentials to the observed ones in the non-stochastic steady state[11].

Table 1 contains the parameter values we used.

Table 1

Parameters for calibration

n	α	ρ	δ	$\overline{\gamma}$	σ_γ	κ	C	σ_ε	ψ	$\overline{A}$	A_0
9	0.55	0.29	0.96^5	0.2	0.1	1	0.6	0.05	0.95^5	1.5	0.5

We assumed that individuals live for 9 periods because the length of one university career is around five years and an average person spends 45

[10] We have not adopted as sophisticated calibration as Heckman et al [1998a] and [1999]. Their work shows some directions through which our calibration can be improved.

[11] The non-stochastic steady state is defined and some of its properties are explored in Appendix A.

years with working and studying after completing high school. Since we considered the length of one period being 5 years, the natural choice for the discount factor δ was 0.96^5, because 0.96 is often used as a discount factor in models calibrated for annual data.

As we have noticed $1/(1-\rho)$ measures the elasticity of substitution between skilled and unskilled labor. Katz and Murphy [1992] estimates this elasticity being 1.41 and there seems to be a general consensus that true elasticity has to be around this value (Heckman et al.[1998a] estimated a similar numerical value for their similar model).

The age-productivity profiles $\{P^0(i)\}_{i=0}^{n-1}$ and $\{P^1(i)\}_{i=1}^{n-1}$ were calibrated to the age-earning profiles estimated by Welch [1979].[12] More precisely, the average unskilled and skilled person in the non-stochastic steady state have the same age-earning profile as the estimated ones.

The moments of the ability distribution (γ and σ_γ) together with the parameters of the educational cost (C and κ) were chosen in a way that in the non-stochastic steady state of the model we would have college enrollment around 27 % which is close to the percentage of college graduates among the young generation in the 90's (enrollment was higher during these periods, however since we do not have college drop-outs in the model we need to consider only the ones who completed college). These parameters were chosen in a way that we could expect significant increase in enrollment along the transition path, that is the university attainment would be around 16 percent in the non-stochastic steady state with A_0, which was the university enrollment in the U.S. in the 60's.

There is no straightforward way of estimating the parameters of the skill-biased technological change we have in the model.[13] Therefore we chose a value (0.95^5) for the persistence parameter (ψ) which is the adjustment of the typical annual persistent parameter for skill-neutral technological change. Then we chose σ_ε in a way, that the transition between A_0 and $\overline{A}$ takes place during 15 periods in average. This would enable us to interpret our transition dynamics as a transition dynamics during an interval of 75 years.

[12] Both profiles are quadratic in experience and they have an early career spline. The advantage of these estimates is that they were estimated in a way that the effect of cohort size is filtered out.

[13] Katz and Murphy [1992] and Heckman et al [1998a] estimated the parameters of a different specification of skill-biased technological change. Their results cannot be directly used for calibrating our model.

5.2 The PEA algorithm

We used the parametrized expectation approach (PEA) (see for example in Lorenzoni and Marcet [1998]) to solve the model numerically. Notice, that the necessary and sufficient information at time t can be described by the following set of variables:[14] $S_t = \left[\{U_{t-n+i}\}_{t-n}^{t-2}, H_t, A_t\right]$. Therefore $\widehat{\gamma}_t$ is a function of S_t. For this reason the expectation in (16) can be written in the following way

$$\begin{aligned} E_t\left[\sum_{i=1}^{n-1} \delta^i \left(\overline{w}^1_{t+i}\left(\widehat{\gamma}_t + P^1(i)\right) - \overline{w}^0_{t+i}\left(\widehat{\gamma}_t + P^0(i)\right)\right) \mid S_t\right] &= \\ \Psi(\widehat{\gamma}_t(S_t), S_t) &= \Lambda(S_t) \end{aligned} \tag{17}$$

(17) highlights that S_t is a sufficient statistics for predicting the expected wage differential for the marginal agent in equilibrium. This implies that if we knew $\Lambda(S_t)$ (16) would become a nonlinear equation of $\widehat{\gamma}_t$ (non-linear because $\overline{w}^0_t$ depends on $\widehat{\gamma}_t$ through L_t). This observation motivated the numerical method we had chosen. If we approximate $\Lambda(S_t)$ we can recover $\widehat{\gamma}_t$ (16) relatively easily, and then all the other variables are determined analytically. Moreover, PEA has advantages compared to other methods in this case because we have many state variables. The implementation of the algorithm is explained in Appendix B.

Since one of our purposes is to discuss the impact of technological progress on the distribution of wages we want to find the solution for the (stochastic) transition dynamics from an initial state characterized by lower level of technology to the stationary distribution. We applied PEA to solve for the transition dynamics similarly to Marcet and Marimon [1992]. The implementation of the algorithm for the transition dynamics is explained in Appendix C.

5.3 Discussion of the results

On Figure 1 in Appendix D we summarize the results of our simulations. We study one particular transition path towards the stationary distribution of

[14] This set of variables is not unique, we could use a sequence of $\widehat{\gamma}_t$ or H_t instead of U_t, they would provide the same information. However, it is crucial that the information set describes the current state of technology and the complete educational composition of the currently living population.

A_t. This path represents the average behavior along the transition because we set $\varepsilon_t = 0$ throughout the transition.[15]

We have plotted the key variables of the analysis. Note, that in this model wage differentials are better measured as a difference than a ratio, because this is what determines the human capital accumulation decision (see (16)). Our measure of wage differentials is the following:[16]

$$\begin{aligned}\Delta_t &= E_{\gamma i}\left(w_t^1(\gamma,i)\right) - E_{\gamma i}\left(w_t^0(\gamma,i)\right) = \\ &= \frac{\overline{w}_t^1 \sum_{i=1}^{n-1} U_{t-i}\left(\overline{\gamma}_{t-i}^1 + P^1(i)\right)}{\sum_{i=1}^{n-1} U_{t-i}} - \frac{\overline{w}_t^0 \sum_{i=0}^{n-1} (1-U_{t-i})\left(\overline{\gamma}_{t-i}^0 + P^0(i)\right)}{\sum_{i=0}^{n-1} (1-U_{t-i})}\end{aligned}$$

Δ_t is simply the difference between average wages of educated and uneducated workers.[17]

Figure 1 here

The average unskilled wage is given by $E\left(w_t^0(\gamma,i)\right)$, while we defined the relative supply of skilled labour as $H_t/(H_t + L_t)$. Finally, we have defined residual inequality ($\widetilde{\sigma}_t$) as the standard deviation of wages after controlling for wage dispersion due to education and age.

$$\widetilde{\sigma}_t \equiv \sqrt{\frac{\sum_{i=0}^{n-1} (1-U_{t-i})\, var_\gamma\left(w_t^0(\gamma,i)\right) + \sum_{i=1}^{n-1} U_{t-i} var_\gamma\left(w_t^1(\gamma,i)\right)}{n - U_t}}$$

Note, that by (5) and (6) labour force participation is given by $n - U_t$. Note also, that by our assumptions the within-group variances ($var_\gamma\left(w_t^s(\gamma,i)\right) =$

[15] Another way of performing this analysis would be to take the average of many short run simulations in every point in time along the transition. Although $E(\varepsilon_t) = 0$, we might obtain significantly different results from this exercise, because uncertainty itself plays important role along the transition. However, the results we obtained using the two methods were very similar.

[16] For expositional convinience we define $E_\gamma\left[x_t^s(\gamma,i)\right]$ as the average of variable x_t^s across abilities for a given i. Similarly $E_{\gamma i}\left[x_t^s(\gamma,i)\right]$ is the average of x_t^s across age and ability. Finally, $E\left[x_t^s(\gamma,i)\right]$ indicates the average of x_t^s over time for given γ and i.

[17] A different definition of wage differentials (based upon regression analysis and/or controlling for age) would not change the results significantly.

$E_\gamma\left[\left(w_t^s(\gamma,i)-E\gamma\left(w_t^s(\gamma,i)\right)\right)^2\right]$) are given by the following expressions

$$var_\gamma\left(w_t^1(\gamma,i)\right)=\left(\overline{w}_t^1\right)^2E\left[\left(\gamma-\overline{\gamma}_{t-i}^1\right)^2\mid\gamma>\widehat{\gamma}_{t-i}\right]=\left(\overline{w}_t^1\right)^2\xi^2(\widehat{\gamma}_{t-i})$$

$$var_\gamma\left(w_t^0(\gamma,i)\right)=\left(\overline{w}_t^0\right)^2E\left[\left(\gamma-\overline{\gamma}_{t-i}^0\right)^2\mid\gamma<\widehat{\gamma}_{t-i}\right]=\left(\overline{w}_t^0\right)^2\zeta^2(\widehat{\gamma}_{t-i})$$

Here we have defined $\xi(\widehat{\gamma}_{t-i})$ and $\zeta(\widehat{\gamma}_{t-i})$ as the standard deviations of a truncated normal variable where the truncation is $\gamma>\widehat{\gamma}_{t-i}$ and $\gamma<\widehat{\gamma}_{t-i}$ respectively.[18] Therefore, we can rewrite $\widetilde{\sigma}_t$ in a simpler form.

$$\widetilde{\sigma}_t\equiv\sqrt{\frac{\left(\overline{w}_t^0\right)^2\sum_{i=0}^{n-1}\left(1-U_{t-i}\right)\zeta^2(\widehat{\gamma}_{t-i})+\left(\overline{w}_t^1\right)^2\sum_{i=1}^{n-1}U_{t-i}\xi^2(\widehat{\gamma}_{t-i})}{n-U_t}}\tag{18}$$

In general, the results are rather satisfactory they reproduce reasonably well the joint behavior of labour supply, college wage premium, residual inequality and university enrollment (to a lesser extent), assuming only smooth skilled-biased technological change. However, enrollment proved to be too volatile and unskilled real wages do not show decreasing pattern. Bellow, we try to give some intuition for these results and highlight the factors which may be missing from the model and could be a potential explanation for the failures.

We can see on the first panel of Figure 1, that we simulated "smooth" technological transition towards the stationary distribution of this economy. However, even with this smooth technological change we see that the behavior of enrollment and wage differentials has a non-monotonic time evolution in the last 10 periods of the transition.[19] This immediately highlights the importance of the "overshooting" property of human capital accumulation.

[18]. If we normalize $\widehat{\gamma}_{t-i}$ ($\widetilde{\gamma}_{t-i}=\frac{\widehat{\gamma}_{t-i}-\overline{\gamma}}{\sigma_\gamma}$), then we can write the standard deviations in the following form:

$$\xi^2(\widehat{\gamma}_{t-i})=\sigma_\gamma^2\left[1-\frac{\phi(\widetilde{\gamma}_{t-i})}{1-\Phi(\widetilde{\gamma}_{t-i})}\left(\frac{\phi(\widetilde{\gamma}_{t-i})}{1-\Phi(\widetilde{\gamma}_{t-i})}-\widetilde{\gamma}_{t-i}\right)\right]$$

$$\zeta^2(\widehat{\gamma}_{t-i})=\sigma_\gamma^2\left[1-\frac{\phi(\widetilde{\gamma}_{t-i})}{\Phi(\widetilde{\gamma}_{t-i})}\left(\frac{\phi(\widetilde{\gamma}_{t-i})}{\Phi(\widetilde{\gamma}_{t-i})}+\widetilde{\gamma}_{t-i}\right)\right]$$

Where $\phi(\cdot)$ and $\Phi(\cdot)$ are the density and cumulative distribution functions of a standard normal variable.

[19]We will concentrate on these periods because the behavior in the fisrt periods also depends on the initialization of human capital stock.

Agents have finite horizon, they do not take into account the impact of their decision on the welfare of other agents. This explains the cyclical behavior during the last 10 periods. The early generations invest in human capital in a great extent, which leads to a temporary overflow of high-skilled people and that pushes down their marginal productivity and consequently their wage premium. This will imply a decrease in enrollment, since in the close future wage premium is not increasing, however the new generations entering the market are still better educated than the ones retiring, therefore the relative supply of educated labour still increases.

At the same time residual wage differentials are increasing. We can obtain some intuition why did it happen from (18), residual dispersion depends on the levels of marginal productivities ($\overline{w}_t^s$). This is natural because there is positive relationship between ability and productivity therefore the dispersion of ability maps to the dispersion of wages through productivities. We know that the technological change alone would improve the productivity of both skilled and unskilled labour, therefore we expect an increase in residual inequality. It is obvious from (18) that residual inequality depends on other variables as well. Ceteris paribus it depends positively on U_t, because higher enrollment decreases both the weight of unskilled young people whose productivity ($\overline{w}_t^0$) is lower and the within group inequality ($\zeta(\widehat{\gamma}_t)$) in that group as well. The impact of the relevant other threshold levels ($\{\widehat{\gamma}_{t-i}\}_{i=1}^{n-1}$) is more complicated. It is easy to show that $(1-\Phi(\widehat{\gamma}_{t-i}))\,\xi(\widehat{\gamma}_{t-i})+\Phi(\widehat{\gamma}_{t-i})\zeta(\widehat{\gamma}_{t-i})$ is decreasing in the absolute value of $\widehat{\gamma}_{t-i}$. Therefore if enrollment is increasing (but it is still bellow 50 percent) then this force by itself is decreasing residual inequality.[20] It is clear from the last panel of Figure 1 that the first two impacts are dominating, therefore we could reproduce the increasing patterns in residual inequality.

However, we have not so appealing results for the average unskilled wages. Again, there are several forces working in opposite directions here

[20] We face the problem of self-selection here. The fact that more able people become educated with a higher probability biases down our estimate on the impact of the ability on wages. This impact is present in most of the empirical papers because they consider educational choice exogeneous.

(see (19)).

$$E_{\gamma i}\left(w_t^0(\gamma,i)\right) \tag{19}$$

$$= \left[\left(\alpha L_t^{\rho}+(1-\alpha)(A_t H_t)^{\rho}\right)^{1/\rho-1}\alpha L_t^{\rho-1}\right]\frac{\sum\limits_{i=0}^{n-1}(1-U_{t-i})\left(\overline{\gamma}_{t-i}^0+P^0(i)\right)}{\sum\limits_{i=0}^{n-1}(1-U_{t-i})}$$

Technological change implies increase in enrollment (at least in the long run) which in turn implies deterioration on the average ability level of the unskilled ($\overline{\gamma}_{t-i}^0$), this impact itself would imply decrease in average unskilled wages. However, technological change has a direct impact on the marginal productivity of unskilled labour as well. The direct effect itself has two components, first, since the two inputs are complementary then an increase in A_t will increase $\overline{w}_t^0$. Secondly, decreasing L_t implies that the marginal productivity of unskilled labour increases because it becomes more scarce. It is clear from our figure that the two direct forces dominate the decrease in average ability. If we chose another specification of skill-biased technological change where technological change has a direct negative impact on the productivity on unskilled labour we might overcome this shortcoming. Another potential explanation for not capturing the decreasing pattern of unskilled wages can be immigration, in most of Western countries there was a significant level of unskilled immigration, which provided always an abundance of unskilled labour. This could offset the increasing scarcity of unskilled labour, therefore it could make the decreasing average ability being the dominant force.[21]

Another less appealing feature of our model is that university enrollment reacts too strongly, we see cycles with a much higher amplitude than in real data. We can invoke too important reasons for this behavior. Obviously, college enrollment is not only motivated by earning prospective but also by several other factors, like the consumption value of education or signalling better productivity. We exclude these possibilities, however these effect may stabilize the "overshooting" behavior of enrollment in some extent, because their movement can be independent of wage differentials.

Finally, cohort size itself influences the evolution of wage differentials (see Welch [1979]) and there were significant differences between the sizes of different U.S. birth cohorts in the. A baby boom cohort can react "much

[21] Actually, increase in foreign trade has a similar impact because it devaluates domestic unskilled labour.

better" to technological improvements just because there are more people available for higher education. At the same time, they provide a big increase of unskilled labour as well, this may explain why we see smaller changes in enrollment rates in data than in our model and it may also help to keep the marginal productivity of unskilled labour down. That is the introduction of exogenous changes in labour supply may increase the "predictive power" of our model.[22]

Finally, educational choices are also influenced by economic policy on education and taxation and we will analyze the impact of those policies in the next section.

6 Policy analysis

There is empirical evidence that expected wage differentials are one of the most important determinants of decisions regarding educational choice (Mincer [1994]). At the same time, higher education is highly subsidized in most of the developed countries. There are two types of arguments behind educational subsidies. First, they increase the educational attainment of the population, this would increase the total skill level of the economy, which would imply higher production. That is education is increasing (output) efficiency. These arguments are typically stronger in the presence of skill-biased technological change because insufficient human capital accumulation might represent a barrier on economic growth.

Another argument for subsidizing higher education is that it makes income distribution more favorable because some individuals may obtain education who cannot obtain it before. Actually, educational subsidies are promoted often in order to stop the growing inequality in society.

However these subsidies must be financed by the government in some way. They imply higher taxation and therefore redistribution between individuals. The generations and/or individuals who finance the subsidies are generally different from the ones who benefit from them. A (more able) segement of the young generation receives the subsidy which is financed by the tax payments of older generations. There is another transfer if taxation is

[22]These exogenous factors can be considered endogenous as well. The size of cohort could be endogenised in this model in flavor of Becker at al [1990] by endogenizing fertility. Immigration is partly endogenous as well, people will immigrate only if their income (net of living costs) is higher in the host country than at home. Heckman et al. [1998a] introduced these two factors in a simple exogenous manner and shown that they indeed improve predictive power.

universal (everybody has to pay taxes), less able people are financing partly the education of more able people, although they do not benefit from the subsidies (at least not directly). Progressive taxation is (partly) promoted in many countries having high education subsidies in order to decrease this type of redistribution. Moreover, taxation has an impact not only on intergenerational and intragenerational distribution, but it has a direct impact on educational choice because it modifies the costs and benefits of education.

Therefore it is not obvious what is the final impact of these policies on efficiency, on the welfare of different agents and on overall inequality.

We have to take into account an additional factor in the presence of skill-biased technological change. Older generations had already decided their human capital accumulation, therefore their decision is not influenced by the current technological transition, however their labour market situation is influenced by the transition, their wage is going to depend on human capital accumulation decision of other agents and on the level of technology itself. Moreover they are also influenced directly by taxation.

We can observe different scenarios in the real world, in some countries higher education is subsidized less, however taxation is also smaller (see for example U.S.). In other (mostly European) countries education is subsidized more but taxation is higher and more progressive as well. Another difference is between different systems is the level of screening before entering education. In many countries admission exams are strict and better performance on the admission exams implies admission to better universities. We will study not only the impact of tuition subsidies on efficiency and inequality but we will compare different scenarios with respect to the type of subsidies and taxation.

6.1 Modelling policies

As we pointed out there are several possible implementation of higher educational policies. The tuition subsidy can be lump-sum (G), that is the cost of education becomes $C - \kappa\gamma - G$. We could interpret this as any individual is entitled for the same level of tuition subsidy given that he enrols in university. Another possibility is that the subsidy is proportional to ability. In this case the educational cost is going to be: $C - \kappa(1 + g)\gamma$, that is the subsidy is given by $g\gamma$ for an agent with ability γ. We model the selection process to universities by admission exams with this specification. Candidates may sorted between universities based upon their performances (more able people can go to universities which are cheaper and better) and a minimum threshold for entering education may be required. This system may be combined with a state scholarship system which depends on performance during the studies. This type of subsidies favors more able people in a greater extent.

On the taxation side we will also have two different options. In the first case everybody has to pay a uniform linear tax rate τ_t^l, in the second case the government levies a (linear) tax τ_t^p only on the educated labor force (this is an extreme approximation of progressive taxation).[23] We assume balanced budget for the government each period, that is total subsidies are equal to total tax revenues each period. We could interpret this assumption as the government offers some subsidy on education, then young generations decide university enrollment taking into account the offered subsidies and current and (expected) future taxation. At the same time the government sets a tax rate what is sufficient to finance the total amount of subsidies for the given period taking into account the individuals choice concerning university enrollment.

We will study four different scenarios, by combining t7he two possible types of tuition subsidies with the two possible types of taxation. The key equation what changes from our benchmark model is (11), the equation defining the threshold agent. We have an additional equation to the benchmark model given by the budget constraint of the government. These equations are going to be the following ones in the four possible cases.

Case A: Lump-sum subsidy and linear taxation:

$$C-\kappa\widehat{\gamma}_t-G+(1-\tau_t^l)w_t^0(\widehat{\gamma}_t,0) = E_t\left[\sum_{i=1}^{n-1}\delta^i(1-\tau_{t+i}^l)\left(w_{t+i}^1(\widehat{\gamma}_t,i)-w_{t+i}^0(\widehat{\gamma}_t,i)\right)\right] \tag{20}$$

$$GU_t = \tau_t^l\left[\sum_{i=0}^{n-1}\int_{-\infty}^{\widehat{\gamma}_{t-i}} w_t^0(\gamma,i)f(\gamma)d\gamma + \sum_{i=1}^{n-1}\int_{\widehat{\gamma}_{t-i}}^{\infty} w_t^1(\gamma,i)f(\gamma)d\gamma\right] \tag{21}$$

[23] Although this seems to be a very extreme way of intoducing progressive taxation, it has a nice feature in our model. First, note that in our model the only reason for taxation is financing the education subsidies. Therefore this specification would satisfy those who are worried about that educational subsides are unfair, because unskilled people do not benefit from them, however they have to contribute to its funding.

Case B: Lump-sum subsidy and progressive taxation:

$$C - \kappa\widehat{\gamma}_t - G + w_t^0(\widehat{\gamma}_t, 0) = E_t\left[\sum_{i=1}^{n-1}\delta^i\left((1-\tau_{t+i}^p)w_{t+i}^1(\widehat{\gamma}_t, i) - w_{t+i}^0(\widehat{\gamma}_t, i)\right)\right] \quad (22)$$

$$GU_t = \tau_t^p\sum_{i=1}^{n-1}\int_{\widehat{\gamma}_{t-i}}^{\infty} w_t^1(\gamma, i)f(\gamma)d\gamma \quad (23)$$

Case C: Proportional subsidy and linear taxation:

$$C-(1+g)\kappa\widehat{\gamma}_t+(1-\tau_t^l)w_t^0(\widehat{\gamma}_t, 0) = E_t\left[\sum_{i=1}^{n-1}\delta^i(1-\tau_{t+i}^l)\left(w_{t+i}^1(\widehat{\gamma}_t, i) - w_{t+i}^0(\widehat{\gamma}_t, i)\right)\right] \quad (24)$$

$$g\kappa\int_{\widehat{\gamma}_t}^{\infty}\gamma f(\gamma)d\gamma = \tau_t^l\left[\sum_{i=0}^{n-1}\int_{-\infty}^{\widehat{\gamma}_{t-i}} w_t^0(\gamma, i)f(\gamma)d\gamma + \sum_{i=1}^{n-1}\int_{\widehat{\gamma}_{t-i}}^{\infty} w_t^1(\gamma, i)f(\gamma)d\gamma\right] \quad (25)$$

Case D: Proportional subsidy and progressive taxation:

$$C - (1+g)\kappa\widehat{\gamma}_t + w_t^0(\widehat{\gamma}_t, 0) = E_t\left[\sum_{i=1}^{n-1}\delta^i\left((1-\tau_{t+i}^p)w_{t+i}^1(\widehat{\gamma}_t, i) - w_{t+i}^0(\widehat{\gamma}_t, i)\right)\right] \quad (26)$$

$$g\kappa\int_{\widehat{\gamma}_t}^{\infty}\gamma f(\gamma)d\gamma = \tau_t^p\sum_{i=1}^{n-1}\int_{\widehat{\gamma}_{t-i}}^{\infty} w_t^1(\gamma, i)f(\gamma)d\gamma \quad (27)$$

We can see that the competitive solution discussed before is a special case of all the four cases with the subsidies (G or g) equal to 0. For this reason it is easy to compare the results to the ones of the benchmark model.[24]

Notice, that we can follow the same solution procedure in all four cases. The government's budget gives us τ_t^l or τ_t^p as a function of $\widehat{\gamma}_t$ (and S_t), then we can use the modified version of (17) to determine $\widehat{\gamma}_t$ as before. Notice, that there is no need to add new state variables. This would not be true in the case of unbalanced government budget or with time-varying tuition subsidies.

[24]This property also helps us to calculate the numerical solution of the models with policies by using homotopy from the solution of the benchmark model.

Calibration We had to calibrate G and g in this model. We chose $G = 0.1$, which implied that the governmental contribution to the educational sector is in average 1.42 percent of total output in the case of lump-sum subsidies and linear taxes (Case A). We chose g to be comparable with G in the following sense. If the average skilled person at the benchmark model ($E\left(\overline{\gamma}_t^1\right)$) goes to university she receives the same amount of subsidy with both lump-sum or proportional subsidies:

$$g: \qquad G = E\left[\overline{\gamma}_t^1\right] g$$

6.2 Efficiency and distribution

We solved the model under all the four scenarios and calculated some measures of efficiency and inequality both in the stationary distribution and along the transition. The stationary distribution is important for the policy analysis because in this case, every agents' enrollment decision is taken under the tax regime therefore we filter out both the problem of the initial generation and technological transition which is obviously present in the case of the transition. Table 2 and 3 in Appendix D presents our results.

Table 2 here

Apart from presenting the time averages of some aggregate statistics we have presented the average life-time utility of three different agents. We have chosen three particular agents from the benchmark model. The average unskilled agent ($\gamma^U = E\left[\overline{\gamma}_t^0\right]$), the average threshold agent ($\gamma^T = E\left[\widehat{\gamma}_t\right]$) and the average skilled agent ($\gamma^S = E\left[\overline{\gamma}_t^1\right]$). The advantage of this choice is that these agents have similar behavior in all the five cases, that is agent γ^U and γ^S choose to be unskilled and skilled respectively in the vast majority of time periods under all of the five scenarios. On the other hand, the decision of agent γ^T varies a lot in all five cases depending on the state of the economy. Therefore welfare changes in the case of agent γ^U and γ^S are mostly due to redistribution by taxation and by the subsidy itself. The welfare changes of agent γ^T are also due to the fact that the introduction of policies modifies significantly his educational choice.

The results highlight the fact that education subsidies are actually beneficial both for efficiency considerations and for equality considerations in the long run. Notice, that output is higher in the stationary economy only because we "convert" some unskilled people to skilled. We are still at the range of the production function where this change increases production, therefore higher average enrollment obviously implies higher average output. However if technology does not change that would imply increase in the marginal productivity (and skill price) of unskilled people and decrease in the marginal productivity of the skilled people. Therefore we can expect decrease in wage differentials and this decrease has to be correlated positively with the increase in enrollment. These results are all confirmed in the upper panel of Table 2. However, different policies imply different magnitude of these changes. It is easier to understand the differences if we see how the welfare of different agents vary between the benchmark case and the different policies.

The welfare of the agents depend not on gross wages but on net wages therefore obviously taxation has an important impact. In addition to this, their welfare changes if they receive a subsidy, because they receive it when they are young and they have to pay taxes during subsequent periods. The overall effect on welfare is the sum of all of these forces.

Probably the most surprising observation is that the unskilled agents are better off in all cases after introducing the policies. The reason is simple, in all cases we see increase in the size of skilled labour which implies less unskilled labour, therefore they become more scarce hence their marginal productivity ($\overline{w}_t^0$) increases. Obviously, the magnitude of this impact depends positively on the magnitude of the increase in skilled labour and on progressiveness of taxation. Linear taxation implies higher increase in skilled labour but higher taxation for unskilled people as well. The first impact seems to be stronger, because even the unskilled agent would prefer linear taxation independently from the form of subsidies because that has higher impact on labour supply.

Skilled labour is influenced in three ways by tuition subsidies, they receive a benefit when they are young which makes them better off, on the other hand they have to pay taxes from their future income and they become more abundant, which makes them worse off. If the subsidy is proportional they become less abundant and have to pay lower taxes that's why they would prefer this type of subsidies. However agent γ^T is going to receive smaller subsidy with the proportional subsidy, therefore he prefers always the lump-sum type. Proportional subsidies have smaller effect, simply because they modify less the decision problem of the marginal agent. They are less costly, that is they imply smaller taxation, what is better for both skilled and unskilled labour. Therefore we could see a coalition between the two ends of the distribution in this economy because both the average unskilled employees and average skilled employees prefer proportional subsidies and linear taxation what is the least preferred option for the marginal agent. However his most preferred policy (lump-sum subsidies and progressive taxation) is the least preferred for the other agents. Finally, note that all the policies decrease inequality as well. We obtain this result if we measure inequality as the wage premium, the standard deviation of wages or the relative position of unskilled people.

Table 3 here

The figures for the transition dynamics in Table 3 show that the efficiency and distributional implications of the different policy regimes do not change too much compared to the stationary distribution. The effects of the policies on the aggregate variables are very similar to the stationary distribution both in their magnitude and order in the four cases. However now we have a new dimension of distributional impacts given by the date of the birth of the agent. We see welfare gains quite similar to the stationary distribution for agents who were born at the stage of the transition relatively close to the stationary distribution. However, welfare gains are smaller and very similar across agents with different ability for the generation who was born at the beginning of the technological change. Actually, in all of the cases educational policies will increase wage inequality within this generation. We have this result, because although unskilled labour becomes more scarce as before but enrollment is too high relative to the level of technology (see Figure 1), which implies higher taxation and less production. This is the manifestation of the "overshooting" property of human capital accumulation.

The picture is very different for a generation who has already decided the level of human capital accumulation before he policy reform. Unskilled people still benefit some (especially if they do not have to pay taxes) because they become more scarce, but this benefit is smaller than the one of the other generations. However, skilled agents will suffer from both of the negative impacts of the policy: they are more abundant which pushes their marginal productivity down and they have to pay taxes. On the other hand, they do not benefit from the subsidy, since they do not receive it.

We can summarize our results as educational subsidies are increasing efficiency because they provide intergenerational transfers which attracts more people to universities. However, the impact measured in size and distribution is quite across different policies. Somewhat surprisingly, agents whose educational decision is not influenced by the policy ex ante will prefer proportional subsidies and linear taxation independently of their actual decision. However, the ones who may change skill level due to the policy would prefer a very different policy, lump-sum subsidies and progressive taxation.

If we consider the effect of these policies along the transition we could see that welfare gains are much smaller for less able people in the generation who were born at the onset of technological change. These gains are bigger for more able members of this generation, therefore subsidies may increase inequality initially. Finally, in the case of generations born before the start of technological change and consequently before the introduction of the subsidies, less able people still benefit some, however educated people will lose because the subsidies attract more "competition" for them and because they have to pay taxes.

7 Conclusion and directions for further research

We presented a stochastic model of wage differentials and human capital accumulation in the presence of skill-biased technological progress in this paper. We have introduced many heterogenous generations living together and explored the consequences of limited adjustment of skilled labour supply on wage differentials and on human capital accumulation. We show that this model has the potential to reproduce a set of well-established stylised fact for the dynamics of wage differentials and labor supply. We solve the model numerically for the stationary distribution and along skill-biased technological transition. The model performed reasonably well in reproducing stylised facts regarding the joint dynamic behavior of wage premium, relative supply of skilled labour and residual inequality.

These results indicate that the fact the human capital accumulation may have an "overshooting" property (due to the finite horizon of individual decisions and because of the limited possibility of supply adjustment) is very important to understand the persistent and non-monotonic pattern of wage inequality. The current version of the model could not reproduce the decrease in real wages of unskilled labour. The introduction of immigration and/or varying cohort size (enabling us to introduce the "baby boom" effect) may help us to overcome this failure. We can not only improve the predictive power of our model with these changes but check whether our results are robust with respect to these modifications.

We used the model to compare different education policy regimes in the stationary distribution and along the transition. We have learned that educational subsidies do increase welfare and decrease overall inequality in our model in the long run, because they provide intergenerational transfers to individuals and by this they make education cheaper. However different agents prefer different policies depending on their relative position in the ability distribution, for this reason although there is a clear order of policies according production efficiency there is no agreement in the society about which one to choose. Moreover, the introduction of education policies may increase inequality within generations which are born on the onset of technological change. Finally, high-skilled individuals born before the start of technological change and before the introduction of educational policies will lose welfare after introducing these policies.

In the light of these results we would like to analyze what is the "optimal" way of subsidizing higher education. Unfortunately, it is not obvious how one should define optimality in a heterogeneous OG setting. We could define it as the particular combination of tuition subsidies and taxation which leads to the highest possible production without making any particular agent worse off. Obviously another definitions of "optimality" could take into account the distribution of wealth more directly.

Finally, skill-biased technological change, in addition to the the previous facts, caused decreasing labour force participation as well in many countries. We have not studied this impact in this model but with the introduction of home production or minimum wage requirement we could introduce it to our model as well.

References

[1] Acemoglu, D.: "Why do new technologies complement skills? Directed technological change and wage inequality", *Quarterly Journal of Economics,* November 1998

[2] Autor, D.H., L.F, Katz and A.B. Krueger: "Computing inequality: Have computers changed the labor market?", *Quarterly Journal of Economics,* November 1998

[3] Becker, G.S., K.M. Murphy and R. Tamura: "Human capital, fertility, and economic growth", *Journal of Political Economy,* 1990, October

[4] Berman, E., J. Bound and S. Machin: "Implications of skill-biased technological change: international evidence", *Quarterly Journal of Economics,* November 1998

[5] Caselli F.: "Technological revolutions", *American Economic Review,* 1999, 78-102

[6] Chari, V.V. and H. Hopenhayn: "Vintage capital, growth, and the diffusion of new technology", *Journal of Political Economy,* 1991, no. 6

[7] Galor O. and O. Moav: "Ability-biased technological transition, wage inequality, and economic growth", *Quarterly Journal of Economics,* May 2000

[8] Galor O. and D. Tsiddon: "Technological progress, mobility and growth, *American Economic Review,* 1997, 363-382.

[9] Hamermesh, D., *Labor Demand,* Princeton: Princeton University Press, 1993.

[10] Heckman J.J., L. Lochner and C. Taber: "Explaining rising wage inequality: explorations with a dynamic general equlibrium model of labour earnings with heterogeneous agents", *NBER Working Paper 6384*, 1998a

[11] Heckman J.J., L. Lochner and C. Taber: "Tax policy and human capital formation", *NBER Working Paper 6462*, 1998b

[12] Heckman J.J., L. Lochner and C. Taber: "General equilibrium cost benefit analysis of education and tax policies", *NBER Working Paper 6881*, 1999

[13] Juhn, C., K.M. Murphy and B. Pierce: "Wage inequality and the rise in returns to skill", *Journal of Political economy,* 1993, no. 3

[14] Katz, L.F. and K.M. Murphy: "Changes in relative wages, 1963-1987: supply and demand factors", *Quarterly Journal of Economics,* 1992, no. 1.

[15] Johnson, N.L. and S. Kotz: Continuous Univariate Distributions", *John Wiley & Sons,* New York, 1970

[16] Krussel P., Ohanian L.E., Rios-Rull J.V. and Violante G.L.: "Capital-skill complementarity and inequality: A macroeconomic analysis", *Federal Reserve Bank of Minneapolis Research department Staff report #236,* 1997

[17] Lorenzoni, G. and A. Marcet: "Parametrized Expectation Approach; Some practical issues", in: *Marimon, R. and A. Scott (eds.): Computational Methods for the Study of Dynamic Economies,* Oxford University Press, 1998

[18] Marcet, A. and R. Marimon: "Communication, Commitment, and Growth", *Journal of Economic Theory,* **58,** 1992, pp. 219-249.

[19] Mincer J.: "Investment in U.S. education and training", *NBER Working Paper 4844*, 1994

[20] Rios-Rull J.V.: "Working in the market, working at home, and the acquisition of skills: A general equilibrium approach, *American Economic Review,* 1993, 893-907.

[21] Rios-Rull J.V.: "Life-cycle economies and aggregate fluctuations", *Review of Economic Studies,* 1996, 465-489.

[22] Rubinstein Y. and D. Tsiddon: "Coping with technological progress: The role of ability in making inequality so persistent", *mimeo,* 1998

[23] Welch, F.:"Effects of cohort size on earnings: The baby boom babies' financial bust", Journal of Political Economy, 1979, S65-S97

A Appendix: Proof of Proposition 1

We fix the level of technology to $\overline{A}$ in the non-stochastic steady state (SS) of the model. Then the SS is given by the property that the educational decision is determined by the same threshold level of ability γ^* each period. Note that, then the supply of the two types of the labour is given by the following expressions:

$$L^* = nF(\gamma^*)\left(E\left[\gamma \mid \gamma < \gamma^*\right]\right) + F(\gamma^*)\sum_{i=0}^{n-1} P^0(i) \tag{28}$$

$$\begin{aligned} H^* &= (n-1)\left(1 - F(\gamma^*)\right)\left(E\left[\gamma \mid \gamma > \gamma^*\right]\right) + \left(1 - F(\gamma^*)\right)\sum_{i=1}^{n-1} P^1(i) = \\ &= (n-1)\int_{\gamma^*}^{\infty} \gamma f(\gamma)d\gamma + \left(1 - F(\gamma^*)\right)\sum_{i=1}^{n-1} P^1(i) \end{aligned} \tag{29}$$

Note that, $\frac{\partial L^*}{\partial \gamma^*} > 0$ and if $\gamma^* > 0$ then $\frac{\partial H^*}{\partial \gamma^*} < 0$. Skill prices are constant as well:

$$\overline{w}^{1*} = \left(\alpha L^{*\rho} + (1-\alpha)\left(\overline{A}H^*\right)^{\rho}\right)^{1/\rho - 1}(1-\alpha)\overline{A}^{\rho}H^{*\rho-1} \tag{30}$$

$$\overline{w}^{0*} = \left(\alpha L^{*\rho} + (1-\alpha)\left(\overline{A}H^*\right)^{\rho}\right)^{1/\rho - 1}\alpha L^{*\rho-1} \qquad \forall t \geq 0 \tag{31}$$

Since skill prices are equal to marginal productivities we can use the concavity of the CES production function ($F_{HH} < 0$, $F_{LL} < 0$) together with our assumption of complementarity between two production factor ($0 < \rho < 1 \rightarrow F_{LH} > 0$) in order to obtain the impact of γ^* on skill prices (under the assumption that $\gamma^* > 0$)

$$\frac{\partial \overline{w}^{1*}}{\partial \gamma^*} = F_{HH}\frac{\partial H^*}{\partial \gamma^*} + F_{LH}\frac{\partial L^*}{\partial \gamma^*} > 0 \tag{32}$$

$$\frac{\partial \overline{w}^{0*}}{\partial \gamma^*} = F_{LL}\frac{\partial L^*}{\partial \gamma^*} + F_{LH}\frac{\partial H^*}{\partial \gamma^*} < 0 \tag{33}$$

Let us rewrite (16) in the non-stochastic steady state and define $V(\gamma^*)$ as the net benefit function of attending university

$$V(\gamma^*) \equiv \gamma^*\left(\overline{\delta}\overline{w}^{1*} - (\overline{\delta}+1)\overline{w}^{0*}\right) + \kappa\gamma^* + \overline{w}^{1*}\overline{P}^1 - \overline{w}^{0*}\overline{P}^0 - C = 0 \tag{34}$$

Where we define $\overline{\delta} = \sum_{i=1}^{n-1}\delta^i$, $\overline{P}^1 = \sum_{i=1}^{n-1}\delta^i P^1(i)$ and $\overline{P}^0 = \sum_{i=0}^{n-1}\delta^i P^0(i)$. Note that equations (28)-(31) imply that $V(\cdot)$ is continuous on the interval $(\underline{\gamma}, +\infty)$. Here we defined $\underline{\gamma}$ as the ability level which would imply $L^* = 0$.

$$\underline{\gamma}: \quad F(\underline{\gamma})\left(nE\left[\gamma \mid \gamma < \underline{\gamma}\right] + \sum_{i=0}^{n-1}P^0(i)\right) = 0 \tag{35}$$

The existence of $\underline{\gamma}$ is guaranteed because the support of the normal distribution is the real line. We have to prove that $V(\cdot)$ has a unique solution on $(\underline{\gamma}, +\infty)$.

Existence Note since $0 < \rho < 1$ $\lim_{\gamma^* \to \underline{\gamma}} \overline{w}^{0*} = \infty$, on the other hand $\overline{w}^{1*}$ and γ^* remains bounded at the limit therefore $\lim_{\gamma^* \to \underline{\gamma}} V(\gamma^*) = -\infty$.

Notice also that $\lim_{\gamma^* \to \infty} H^* = 0$ and therefore $\lim_{\gamma^* \to \infty} \overline{w}^{1*} = \infty$ and similarly to the previous case $\overline{w}^{1*}$ remains bounded at the limit and (34) implies $\lim_{\gamma^* \to \infty} V(\gamma^*) = \infty$. Since $V(\gamma^*)$ is continuous these two limit results imply the existence of a solution to $V(\cdot)$ by the intermediate value theorem.

Uniqueness Let us study $V'(\gamma^*)$ in order to prove uniqueness.

$$V'(\gamma^*) = \left(\overline{\delta}\overline{w}^{1*} - (\overline{\delta}+1)\overline{w}^{0*}\right) + \kappa + \frac{\partial \overline{w}^{1*}}{\partial \gamma^*}\left(\overline{\delta}\gamma^* + \overline{P}^1\right) - \frac{\partial \overline{w}^{0*}}{\partial \gamma^*}\left(\left(\overline{\delta}+1\right)\gamma^* + \overline{P}^0\right) \tag{36}$$

Note, that if $\gamma^* > 0$ and $\left(\overline{\delta}\overline{w}^{1*} - (\overline{\delta}+1)\overline{w}^{0*}\right) > 0$ then $V'(\gamma^*) > 0$ as well. We can define γ' as the level of ability which imply that the discounted sum of skill prices are equal to each other[25]:

$$\gamma' : \quad \overline{\delta}\overline{w}^{1*}(\gamma') - (\overline{\delta}+1)\overline{w}^{0*}(\gamma') = 0 \tag{37}$$

Now, we will give two sufficient conditions which will guarantee uniqueness of γ^*.

Assumption 1 $\gamma' > 0$ and $\kappa\gamma' + \overline{w}^{1*}(\gamma')\overline{P}^1 - \overline{w}^{0*}(\gamma')\overline{P}^0 - C < 0$.

Assumption 2 If $\underline{\gamma} < 0$ then it is sufficiently close to 0.

The first assumption says that age-productivity profiles and educational costs are such that agent γ' would get negative benefits from going to university. Note first, that by (32) and (33) for any $\gamma > \gamma'$ we have $\overline{\delta}\overline{w}^{1*} - (\overline{\delta}+1)\overline{w}^{0*} > 0$ and therefore $V'(\gamma) > 0$. Note second, that by Assumption 1 we have $V(\gamma') < 0$. Finally recall $\lim\limits_{\gamma\to\infty} V(\gamma) = \infty$. These three observations together with the continuity of $V(\cdot)$ imply that there is a unique solution to (34) on the interval (γ', ∞).

Now we will prove that for $\forall\gamma \in (\underline{\gamma}, \gamma')$ we have $V(\gamma) < 0$. Note first, that by (32) and (33) for $\forall\gamma \in (0, \gamma')$ we have $\overline{\delta}\overline{w}^{1*} - (\overline{\delta}+1)\overline{w}^{0*} < 0$. Note second, by (34), (32), (33) and by Assumption 1 this implies that $V(\gamma) < 0$. If $\underline{\gamma} > 0$, we are done. If $\underline{\gamma} < 0$, note that because of continuity of $V(\cdot)$ and because $\lim\limits_{\gamma^*\to\underline{\gamma}} V(\gamma^*) = -\infty$ $\exists\lambda$ such that $\forall\gamma \in (\underline{\gamma}, \underline{\gamma}+\lambda)$ we have $V(\gamma) < 0$. Assumption 2 says that $\underline{\gamma} + \lambda \geq 0$. Note that we proved that under Assumptions 1 and 2 we have that $V(\gamma) < 0$ on $(\underline{\gamma}, \gamma']$ and it has a unique solution on (γ', ∞). Therefore it has a unique solution on $(\underline{\gamma}, \infty)$.[26]

[25]Note, that γ' is given by the following expression

$$\frac{(1-\alpha)\overline{A}^{\rho}}{\alpha}\left(\frac{H^*(\gamma')}{L^*(\gamma')}\right)^{\rho-1} = \frac{\overline{\delta}+1}{\overline{\delta}}$$

[26]Note, that Assumption 1 is sufficient but not neccesary but Assumption 2 is necceseary and sufficient as well. It is not very hard to satify either of the assumptions. Assumption

B Appendix: The PEA algorithm in the stationary distribution

We chose exponential-linear polynomial approximation[27] of $\Lambda(S_t)$ in the following way:

$$\Lambda(S_t) \approx PEA(\beta_0, \beta; S_t) \equiv \beta_0 e^{\beta' S_t}$$

We have implemented the algorithm in the following steps.

Step 1 Simulate a series of the stochastic shock of length T: $\{A_t\}_{t=1}^T$, $A_0 = \overline{A}$. Select an initial parametrization (β_0, β).

Step 2 Simulate the model with the usage of $\{A_t\}_{t=0}^T$ and (β_0, β). This step involves the solution of the non-linear equation (38) for $\widehat{\gamma}_t(\beta_0, \beta)$ for $t = 0, .., T$.

$$C - \kappa\widehat{\gamma}_t(\beta_0, \beta) + \overline{w}_t^0 \left(\widehat{\gamma}_t(\beta_0, \beta), S_t\right) \left(\widehat{\gamma}_t(\beta_0, \beta) + P^0(i)\right) - PEA(\beta_0, \beta; S_t) = 0 \tag{38}$$

We calculate $U_t(\beta_0, \beta)$, $L_t(\beta_0, \beta)$, $H_{t+1}(\beta_0, \beta)$ and consequently $\overline{w}_t^1(\beta_0, \beta)$ and $\overline{w}_t^0(\beta_0, \beta)$ with $\widehat{\gamma}_t(\beta_0, \beta)$. Then we calculate the ex post realizations of the expectations in (16). Call this $\{X_t(\beta_0, \beta)\}_{t=0}^T$.

Step 3 Estimate the parameter vector (β_0', β') such that

$$(\beta_0', \beta') = \arg \min_{(\widetilde{\beta}_0, \widetilde{\beta})} \sum_{t=0}^{T} \left[X_t(\beta_0, \beta) - PEA(\widetilde{\beta}_0, \widetilde{\beta}; S_t'(\beta_0, \beta))\right]^2 \tag{39}$$

In this way (β_0', β') is the best predictor of the ex post realization given the chosen functional form and the simulated series.

---

1 implies intuitively that the main incentive for going to university is coming from higher skill prices and not because of "higher" slope of life-cycle pattern of productivity. Finally Assumption 2 guarantees only that for agents with negative abilities it is not worth going to university.

[27]Liear approximation turned out to be sufficient because there is "multicollinearity" between our state variables, therefore a low degree polynomial can fit reasonably well a non-linear function.

Step 4 If $(\beta_0', \beta') = (\beta_0, \beta)$ the solution procedure finished, we found the best parametrization of our expectations and consequently we found the solution to our model. If $(\beta_0', \beta') \neq (\beta_0, \beta)$ go to step 2 and use initial parameters which are some linear combinations of (β_0', β') and (β_0, β) (usually we have to put much higher weight to the initial parameter vector to avoid diverging paths).

C Appendix: The PEA algorithm along the transition

Assume that we already solved for the parameters of the stationary distribution (β_0, β) as explained in Appendix B. We have to modify some of the steps of the procedure described above to find the appropriate parameters for the transition path.

Step 1b Simulate M independent realizations of the stochastic shock with length T': $\left\{\{A_{m,t}\}_{t=1}^{T'}\right\}_{m=1}^{M}$, with $A_0 \ll \overline{A}$ fixed. Select an initial parametrization (β_0^o, β^o). Here T' should be relatively small, but long enough to assure that by the end of this period we are in the stationary distribution.

Step 2b Simulate model with usage of $\left\{\{A_{m,t}\}_{t=1}^{T'}\right\}_{m=1}^{M}$ and (β_0^o, β^o). Solve for (38) like in the previous algorithm, however in the last periods of the simulation we use (β_0, β) instead of (β_0^o, β^o) to insure that our out-steady state dynamics is consistent with the steady state distribution when we are actually there.

Step 3b and 4b We have to proceed with Step 3 and Step 4 as in the stationary case.

D Appendix: Figures and Tables

Figure 1

Simulated times series for selected variables along the transition

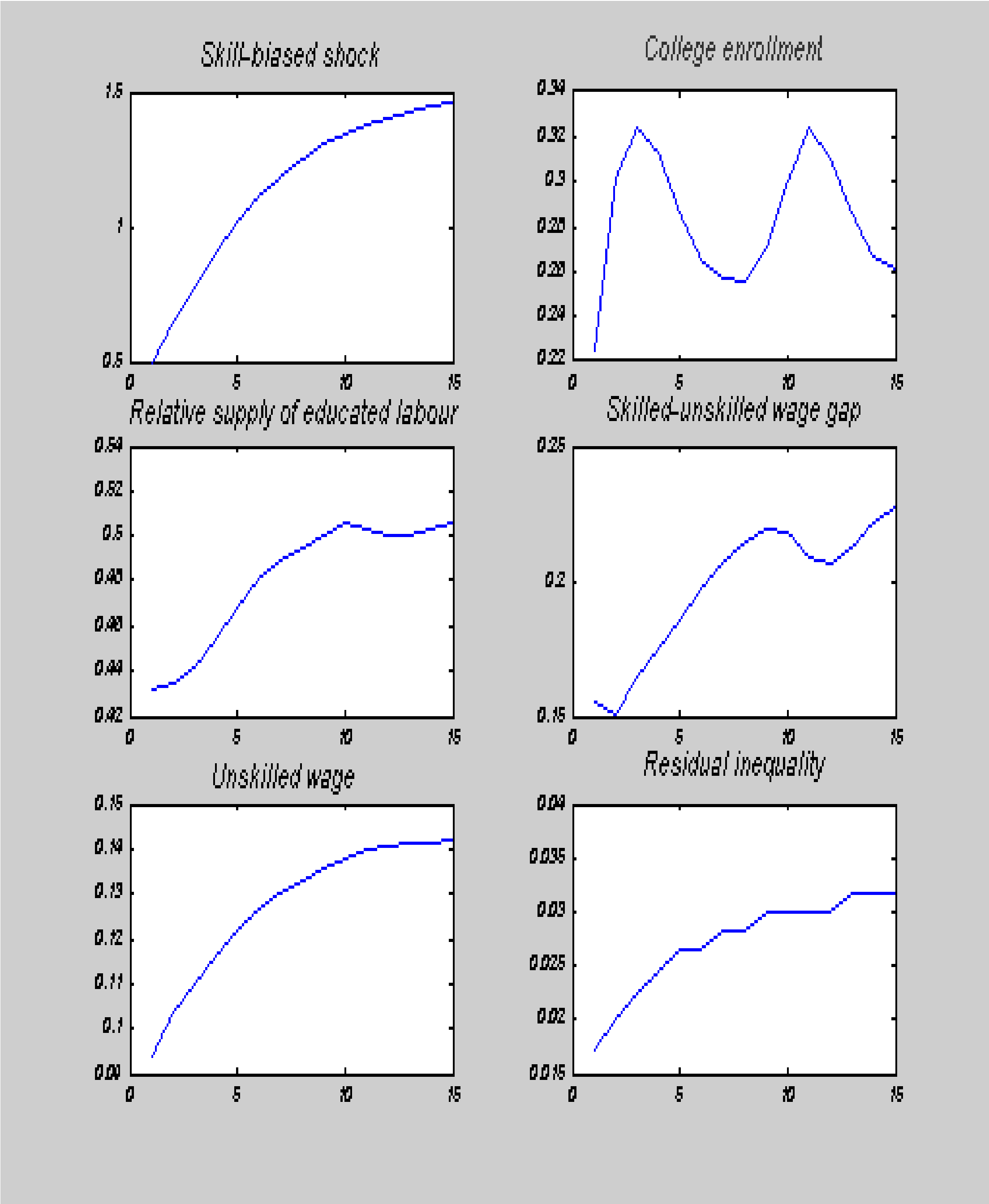

Table 2: Efficiency and distribution implications of different education policy regimes in the stationary distribution

	Benchmark	Policies (% deviations from the benchmark case)			
		Case A	Case B	Case C	Case D
Output	1.956	1.022	0.716	0.818	0.460
Enrolment	0.264	1.800	1.200	1.400	0.800
Wage premium	0.249	-7.631	-5.221	-5.622	-3.213
Std. Dev. Of wages	0.135	-3.704	-2.963	-2.963	-1.481
Tax rate		1.420	2.350	1.120	2.310
		Life-time utilities			
Unskilled average person	0.728	6.044	3.434	7.143	4.808
Threshold person	0.967	2.068	2.172	1.448	1.448
Educated average person	1.171	1.281	1.196	2.220	2.135

Case A: Lump-sum subsidies and linear taxation
Case B: Lump-sum subsidies and progressive taxation
Case C: Proportional subsidies and linear taxation
Case D: Proportional subsidies and progressive taxation

Table 3: Efficiency and distribution implications of different education policy regimes along the transition

	Benchmark	Policies (% deviations from the benchmark case)			
		A	B	C	D
Output	1.693	1.063	0.709	0.827	0.473
Enrolment	0.257	1.8	1.1	1.5	0.8
Wage premium	0.219	-7.306	-4.566	-5.936	-3.196
Std. Dev. Of wages	0.115	-3.478	-2.609	-2.609	-1.739
Tax rate		1.310	2.910	1.290	2.890
Born at –3		Life-time utilities			
Unskilled average person	0.256	0.781	1.563	0.391	1.172
Threshold person	0.356	0.562	1.124	0.000	0.843
Educated average person	0.732	-4.508	-5.328	-3.825	-4.781
Born at 0					
Unskilled average person	0.527	1.328	1.518	0.569	1.139
Threshold person	0.753	1.195	1.594	0.531	1.062
Educated average person	0.915	1.749	1.530	2.951	2.514
Born at 7					
Unskilled average person	0.650	7.385	4.308	6.769	4.000
Threshold person	0.911	2.195	2.195	1.427	1.427
Educated average person	1.106	1.266	1.266	2.351	2.260