The Minimum Wage and Inequality
– The Effects of Education and Technology

Zsófia L. Bárány*

This version: May 20, 2014

Abstract: In the last 30 years wage inequality increased steeply while real minimum wages fell. In this paper I demonstrate that a general equilibrium model with endogenous skill choice is required to correctly evaluate the implications of minimum wage changes. The minimum wage not only truncates the wage distribution, but also affects skill prices and therefore changes the incentives that people face when making educational decisions. The calibrated model suggests – in line with recent empirical literature – that even though minimum wages affect the bottom end of the wage distribution more, their impact on the top end is significant as well.

Keywords: Minimum wage, education, technology, wage inequality

JEL-Classification: E24, E65, J24, J31.

* I am grateful to Francesco Caselli and to Moshe Buchinsky, Mirko Draca, Alexander C. Lembcke, Aytek Malkhozov, Alan Manning, Albert Marcet, Guy Michaels, Barbara Petrongolo, Steve Pischke, Christian Siegel, Silvana Tenreyro and seminar participants at the LSE, Université de Montréal, Sciences Po, Bank of Italy, University of Essex, University of Mannheim, Cologne Macro Workshop, Winter Meeting Hungarian Academy of Sciences, EUI, Bristol University, UCLA for many discussions, comments and suggestions that have improved this paper.

Economics Department, Sciences Po, 28 rue des Saints-Pères, 75015, Paris, France.
Email: zsofia.barany@sciencespo.fr

1 Introduction

It is well documented that income inequality has drastically increased in the United States over the past 30 years along several dimensions: the skill premium, which is the relative wage of high- to low-skilled workers, as well as the gaps between different percentiles of the wage distribution increased significantly.[1] The steepest increase in wage inequality occurred during the 1980s, and as shown in Figure 1, at the same time the real value of the minimum wage fell substantially, and has remained below its original level since then.

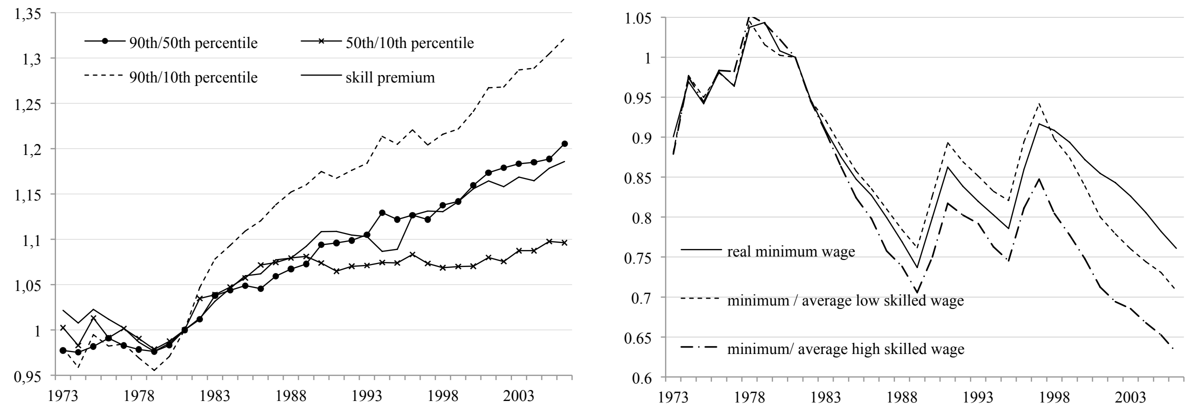

Figure 1: Changes in wage inequality and minimum wages

The value of all measures is normalized to be 1 in 1981, i.e. the graphs represent the percentage changes since then. Wages are calculated from Current Population Survey May Extracts and March Outgoing Rotation Groups supplements. The skill premium is the average wage of high- relative to low-skilled workers. The real hourly minimum wage is the federal minimum wage in 2000$, calculated using the consumer price index from the U.S. Bureau of Economic Analysis.

The changes in the structure of wages fueled an extensive debate on the forces driving them. One explanation focuses on changes in labor market institutions, and particularly, on a 30 percent decline in the real minimum wage that took place in the 1980s (DiNardo, Fortin, and Lemieux (1996), Lee (1999), Card and DiNardo (2002)). The general idea in these studies is that minimum wages truncate the distribution of wages, and only affect those, for who this wage is (close to) binding. However, the truncation of the wage distribution, by imposing a minimum wage, affects skill prices and therefore changes the incentives that people face when making educational decisions. Hence the consideration of general equilibrium effects is crucial in correctly understanding the implications of minimum wage changes.

This is the first paper, to my knowledge, that quantitatively assesses the contribution of falling minimum wages to increasing wage inequality in a general equilibrium model, where both the supply of high-skilled workers and the direction of technical change is endogenous.

[1]See for example Eckstein and Nagypál (2004), Goldin and Katz (2007) and 2008.

This paper also contributes to the debate about the driving forces behind the changes in the structure of wages. In general equilibrium there is a correlation between minimum wages and upper tail inequality, which, contrary to the claim of many, is not spurious. Indeed, the model provides a theoretical channel through which changes in minimum wages can affect inequality along the entire wage distribution. I develop a model which suggests that even though minimum wages affect the bottom end of the wage distribution more, their impact on the top end is significant as well.

Typically in the literature, either the supply of skills or the direction of technology is treated as exogenous. However, as the supply of skills affects the evolution of technology, and in turn technology affects educational decisions, there is a feedback mechanism between them. Therefore the endogenous determination of these two forces can either reinforce, or mitigate, the initial impact of minimum wages on inequality. If minimum wages exclude low productivity workers from the workforce, and education increases the productivity of workers, then the minimum wage alters the optimal education decisions. Consequently this implies a change in the skill composition of the labor force. Furthermore, the change in the supply of high- and low-skilled labor affects the returns to developing machines complementary to them, thereby altering the direction of technological change. Due to the links between the minimum wage, education, and technological change, the overall general equilibrium effects of changes in the minimum wage on inequality could be quite different from what is implied by a simple partial equilibrium reasoning.

I build a two sector growth model, where the labor supply and the available technologies in the two sectors are both endogenous, and there is a binding minimum wage in place. As in Acemoglu (1998), the production side is a two sector Schumpeterian model of endogenous growth, with more R&D spending going towards technologies that are complementary with the more abundant factor. I explicitly model the labor supply side: workers, who are heterogeneous in their ability and time cost of education, make educational decisions optimally. The productivity of workers depends on their ability and on their education: workers with higher innate ability are more productive, and acquiring education allows access to the high-skilled sector, where the productivity per unit of ability is higher in equilibrium. The productivity difference between the high- and the low-skilled sector depends on the relative technologies available to them. I solve for the balanced growth path and calibrate the model to the US economy in 1981 in order to compare the transitional dynamics with the observed patterns of wages in the US over the subsequent thirty years.

According to the model, the decline in the minimum wage accounts for almost one fifth of the observed increase in the 90/10 wage differential, and accounts for about one half of the increase in the 50/10 wage gap. In my model, the minimum wage also has some spill-over effects to the top end of the wage distribution, explaining almost one tenth of the increase in the 90/50 wage gap.

First, as the minimum wage decreases, previously unemployed low ability workers flow into the low-skilled workforce.[2] Second, the skill composition of the employed changes gradually. The inflow

[2]The effects of minimum wages on employment are debated in the empirical literature, see exchange between Neumark and Wascher (2000) and Card and Krueger (2000). I will discuss this in detail in Section 4.3.

of low-skilled workers to the pool of employed increases the skill premium, thus increasing the incentives for acquiring education at all ability levels. However, a lower minimum wage also makes it easier to find employment, reducing the role of education in avoiding unemployment for workers with very low ability. As a result of these two forces the average educational attainment of the workforce decreases. While educational attainment decreases at the lower end of the ability distribution, it increases at the top end.[3] Third, the ability composition of the labor aggregates changes, due to both the inflow from unemployment and the changing decision structure of skill acquisition. As the minimum wage decreases, lower ability workers flow into employment, thereby widening the range of abilities present among the employed. As both labor aggregates expand, the average ability in both the high- and the low-skilled sector decreases. Since more low-ability individuals enter the low-skilled labor force, the average ability in the low-skilled sector decreases more. This composition effect reinforces the initial increase in the observed skill premium.[4]

Finally, the direction of technology reacts to changes in the size of the low- and high-skilled labor aggregates. The direct effect of the minimum wage – the expansion of the low-skilled labor force – dominates, decreasing the relative supply of high-skilled labor. This implies that technology becomes less skill biased in the long run.

My paper is related to the debate about the driving forces behind the changes in the structure of wages. One of the leading explanations, supported by several empirical studies is skill-biased technical change (SBTC), which asserts that the relative demand for high-skilled workers has been continuously increasing since the 1980s (Katz and Murphy (1992), Juhn, Murphy, and Pierce (1993), Krueger (1993), Berman, Bound, and Griliches (1994), Autor, Katz, and Krueger (1998)).[5] Another popular explanation attributes much of the increase in wage inequality during the 1980s to the decline in the value of the minimum wage (DiNardo, Fortin, and Lemieux (1996), Lee (1999), Card and DiNardo (2002)). However, Lee (1999) also finds that the reduction in the minimum wage is correlated with rising inequality at the top end of the wage distribution. This is seen by many as a sign that the correlation between declining minimum wages and increasing inequality is mostly coincidental (Autor, Katz, and Kearney (2008), Autor, Manning, and Smith (2009)).

The model presented here provides theoretical support for the empirical finding that minimum wages affect inequality in the upper tail of the wage distribution: the minimum wage does not only

[3]The empirical evidence on the effects of minimum wages on enrollment is sensitive to whether the enrollment option is considered jointly with the employment option. The papers that consider these decisions jointly find that the effect of minimum wages is heterogeneous depending on the individual characteristics, such as ethnicity, race and family characteristics (Cunningham (1996), Ehrenberg and Marcus (1982)). Several papers find that higher minimum wages have a displacement effect, leading to an increase in the number of not enrolled/not employed teenagers (Mattila (1996)). I will discuss the relation between the predictions of my paper with the empirical findings on educational attainment in Section 4.3.

[4]Lemieux (2006) and Autor, Katz, and Kearney (2005) demonstrate that shifts in the education and experience composition play a significant role in the growth of residual inequality and in lower tail inequality. Carneiro and Lee (2011) shows that not correcting for the decline in the average quality of college graduates underestimates the increase in the skill premium. Several papers document that changes in employment or participation rates affect the evolution of inequality due to composition effects (Chandra (2003), Neal (2004), Mulligan and Rubinstein (2008), Olivetti and Petrongolo (2008)).

[5]Beaudry and Green (2005) find little support for ongoing skill-biased technological progress; in contrast, they show that changes in the ratio of human capital to physical capital conform to a model of technological adoption following a major change in technological opportunities.

affect those who earn wages close to it, but it affects the entire wage distribution. A lower minimum wage shifts the truncation point, and also alters the shape of the wage distribution.

Theoretical explanations for the the unprecedented increase in wage inequality during the 1980s either rely on exogenous skill-biased technical change or on an exogenously increasing relative supply of high-skilled workers. Heckman, Lochner, and Taber (1998), Caselli (1999), Galor and Moav (2000) and Ábrahám (2008) allow for endogenous skill formation and explore the effects of exogenous skill-biased technical change. Explanations for the skill-bias of technology rely on exogenous shifts in the relative labor supplies. Acemoglu (1998) and Kiley (1999) use the market size effect in research and development, while Krusell, Ohanian, Ríos-Rull, and Violante (2000) rely on capital-skill complementarity and an increasing supply of high-skilled labor to account for the path of the skill premium. To my knowledge this is the first paper where both the bias of technology and skill formation are endogenous.[6]

Another strand of literature that my paper is related to analyzes the effects of minimum wages on educational attainment. Cahuc and Michel (1996) show that if minimum wages increase unemployment risk for the unskilled, then more people will acquire education. Ravn and Sorensen (1999) assume that minimum wages reduce on-the-job training provided by employers, which gives rise to more schooling. The model of Agell and Lommerud (1997) is closer in spirit to the one presented here: education grants access to a higher-wage market, and the effect of minimum wages is heterogeneous across the ability of agents.

The paper is organized as follows: section 2 describes the model, section 3 the equilibrium and the steady state, section 4 provides the quantitative analysis, and finally section 5 concludes.

2 The Model

Time is infinite and discrete, indexed by $t = 0, 1, 2...$ The demographic structure is a perpetual youth overlapping generations model, as in Blanchard (1985). Individuals are heterogeneous in two aspects: in their time cost of acquiring education and in their innate ability.

Every individual has to decide whether to acquire education or not. Those who acquire education become high-skilled. In my calibration I identify the high-skilled as having attended college. Those who opt out from education remain low-skilled. Workers with high and low skills perform different tasks, are employed in different occupations, and produce different goods. The high-skilled sector includes skill-intensive occupations and production using high-skilled labor, while the low-skilled sector includes labor-intensive occupations and production using low-skilled labor. In equilibrium working in the high-skilled sector provides higher wages and greater protection from unemployment.

The government imposes a minimum wage in every period, and those who would receive a lower wage – depending on their skill and innate ability – cannot work and become unemployed. As soon as

[6]My paper more generally connects to the literature on the effects of labor market institutions on investments, which mainly focus on the differences in the European and American patterns (Beaudry and Green (2003), Alesina and Zeira (2006), Koeniger and Leonardi (2007)).

the minimum wage falls below their marginal productivity, they immediately become employed in the sector relevant to their skill.

There is a unique final good in this economy, which is used for consumption, the production of machines, and as an investment in R&D. It is produced by combining the two types of intermediate goods: one produced by the low- and the other by the high-skilled workers. Intermediate goods are produced in a perfectly competitive environment by the relevant labor and the machines developed for them.

Technological progress takes the form of quality improvements of machines that complement a specific type of labor, either high- or low-skilled. R&D firms can invest in developing new, higher quality machines. Innovators own a patent for machines and enjoy monopoly profits until it is replaced by a higher quality machine. There is free entry into the R&D sector, and more investment will be allocated to developing machines that are complementary with the more abundant labor type.

The economy is in a decentralized equilibrium at all times: all firms maximize their profits – either in perfect competition or as a monopoly – and individuals make educational decisions to maximize their lifetime income. I analyze how a permanent unexpected drop in the minimum wage affects the steady state and the transitional dynamics within this equilibrium framework.

2.1 Production and technological progress

The unique final good is produced in perfect competition by combining the two intermediate goods, which are not perfect substitutes:

$$Y = (Y_l^\rho + \gamma Y_h^\rho)^{\frac{1}{\rho}} \quad \text{for} \quad \rho < 1,$$

where Y_l and Y_h are the intermediate goods produced by the low- and high-skilled workers. Intermediate good production is also perfectly competitive in both sectors $s \in \{l, h\}$:

$$Y_s = \underbrace{\frac{1}{1-\beta}\int_0^1 q_s^j (X_s^j)^{1-\beta} dj}_{\equiv A_s} N_s^\beta = A_s N_s^\beta \quad \text{for} \quad \beta \in (0,1), \tag{1}$$

where N_s is the amount of effective labor employed[7], and A_s denotes the technology level, or productivity, in sector s, which is itself the result of firms' optimizing behavior. Firms decide the quantity, X_s^j, of a machine with quality q_s^j to use, for each machine line $j \in [0,1]$ available in sector s. I assume that machines fully depreciate during use. The optimal quantity of machine line j in sector s with quality q_s^j is

$$X_s^j = \left(\frac{p_s q_s^j}{\chi_s^j}\right)^{\frac{1}{\beta}} N_s \quad \text{for} \quad s = \{l, h\} \quad \text{and} \quad j \in [0,1], \tag{2}$$

where p_s is the price of intermediate good s and $\chi_s(j)$ is the price of this machine.

[7]The exact definition of N_s is specified in the labor supply section.

The price of each machine is determined in every period by the R&D firm which developed it, and which has perpetual monopoly rights over it. The firms, given the quality of their machine, q, set the price of their vintage, $\chi(q)$, to maximize current period profits. If quality differences between subsequent vintages of a given machine line are sufficiently large, then it can be shown that only the best quality machine, the leading vintage, is sold in equilibrium at its monopoly price.[8] Assuming that quality differences are large enough, and that the marginal cost of producing one unit of machine of quality q is q, the profit-maximizing price of the leading vintage is $\chi(q) = \frac{q}{1-\beta}$.

Using the monopoly price of machines, and the demand it implies for them from (2), we can express the equilibrium production of intermediate goods in sector s from (1) as

$$Y_s = (1-\beta)^{\frac{1-2\beta}{\beta}} p_s^{\frac{1-\beta}{\beta}} Q_s N_s, \tag{3}$$

where $Q_s = \int_0^1 q_s^j dj$ denotes the average quality of the leading vintages in sector s.

Technological advances are a discrete time version of Aghion and Howitt (1992). Investment in R&D are targeted at a specific line of machines in one of the two sectors, and produces a random sequence of innovations. Each innovation improves the quality of an existing line of machine by a fixed factor, $\overline{q} > 1$. The Poisson arrival rate of innovations for a firm k that invests z_k on improving a machine line is ηz_k. Denoting the total investments on a line by $\overline{z} \equiv \sum_k z_k$, the economy wide arrival rate of innovations in this line is $\eta \overline{z}$. Hence the probability that the quality of this machine line improves in one period is $(1 - e^{-\eta \overline{z}})$. I assume that whichever firm has the first successful innovation in a given period gets the patent for the leading vintage in that line of machines. Under these conditions the probability that firm k receives the patent is $(1 - e^{-\eta \overline{z}}) z_k / \overline{z}$. I assume that the marginal cost of investing in R&D to improve a line is proportional, with a factor B, to its quality.

The value of owning the leading vintage of a machine line of quality q in sector s at time t is the sum of the profit in period t and the expected value of owning the vintage of quality q in year $t+1$:

$$V_s(t,q) = q \underbrace{\beta(1-\beta)^{\frac{1-\beta}{\beta}} (p_s)^{\frac{1}{\beta}} N_s}_{\pi_s(t)} + \frac{e^{-\eta \overline{z}_s(t,q)} V_s(t+1,q)}{1+r} \quad \text{for} \quad s = \{l, h\}. \tag{4}$$

Where $\pi_s(t)$ is the period profit in sector s per unit of quality, given monopoly pricing and industry demand (2). If total R&D spending on this line is $\overline{z}_s(t,q)$, the probability that the quality q machine remains the leading vintage in period $t+1$ is $e^{-\eta \overline{z}_s(t,q)}$.[9] The value of owning the leading vintage in sector s depends on $N_s(t)$ and $p_s(t)$, as these affect the current period profits, for simplicity of notation the state variable t captures these in $V_s(t,q)$.

Free entry into the R&D sector implies that profit opportunities have to be exhausted: the expected return from R&D investment has to equal its cost. The fact that both the probability for a firm of having a

[8]The quality differences have to exceed $(1-\beta)^{-\frac{1-\beta}{\beta}}$. See Appendix for details.

[9]With probability $1 - e^{-\eta \overline{z}_s(t,q)}$ a higher quality machine is invented. Given the assumption that $\overline{q} > (1-\beta)^{-\frac{1-\beta}{\beta}}$, only the leading vintage is demanded by the intermediate good producing firms, and this quality level becomes obsolete.

successful innovation and the cost of investing into R&D are proportional to the firm's total investment imply that free entry only pins down the total amount of R&D spending targeted at improving a given line.[10] The free entry condition for improving a line in sector s to quality q can be expressed as

$$\frac{E_t(V_s(t+1,q))}{q(1+r)} = \frac{B\overline{z}_s(t,q)}{1-e^{-\eta\overline{z}_s(t,q)}} \quad \text{for} \quad s=\{l,h\}. \tag{5}$$

Note that in equilibrium the total amount of R&D expenditure, $\overline{z}_s(t,q)$, depends on the expected value of owning a leading vintage in sector s with quality q, which in turn depends on the future demand for this machine line by firms which produce intermediate good s. These future demands depend on the labor force which is able to use these machines and the price of the intermediate good in the future. Thus R&D firms have to form expectations about future labor supplies and intermediate good prices.

Since there are a continuum of machine lines in a sector, the average quality of leading vintages in sector s evolves according to

$$Q_s(t+1) = \int_0^1 E_t(q_s^j(t+1))dj = \int_0^1 q_s^j(t)\left(1+[1-e^{-\eta\overline{z}_s^j(t,q_s^j(t))}](\overline{q}-1)\right)dj \quad \text{for} \quad s=\{l,h\}. \tag{6}$$

2.2 Labor supply

Every period a new generation of mass $1-\lambda$ is born, while the probability of surviving from period t to $t+1$ is λ. Let $f(a)$ and $g(c)$ respectively be the time invariant distribution of abilities, $a \in (0,\infty)$, and of the time cost of education, $c \in [0,\overline{c}]$.[11] Each worker i draws an ability a_i from $f(a)$, and a cost of education c_i from $g(c)$ at birth. These assumptions imply that both the size of the population, and the distribution of costs and abilities are constant over time.

Each individual can acquire education, but only in the first period of his life. Completing education takes a fraction c_i of the first period of individual i's life, during which time he cannot participate in the labor market.[12] An individual who completes education becomes high-skilled and has the option of working in the high-skilled sector for life. Those who choose not to acquire education, remain low-skilled and can start working in the period they are born as low-skilled. Due to perfect competition in intermediate good production, each worker is paid his marginal product.

The government imposes a minimum wage $\underline{w}(t)$ in every period. Workers cannot be paid less than the minimum wage, hence workers are unemployed in each period when their marginal productivity is below the current minimum wage.

I model innate ability as a factor that increases individual productivity. Each sufficiently productive worker supplies one unit of raw labor inelastically, which translates to a units of efficiency labor for someone with ability a. Using (3) the potential wage of a sector s worker with ability a can be expressed

[10] In the free entry condition it is assumed that each R&D firm is sufficiently small so that $\overline{z}_s^j(t)$ does not depend on $z_{s,k}^j(t)$. Assuming larger R&D firms would yield a similar result, with a more complicated expression.

[11] The choice of heterogeneous time cost and independent cost and ability distributions is explained in the Section 4.1.

[12] In the calibration exercise I set the length of a period to be five years.

as:

$$w_{s,a}(t) = a \underbrace{\beta(1-\beta)^{\frac{1-2\beta}{\beta}} p_s(t)^{\frac{1}{\beta}} Q_s(t)}_{\equiv w_s(t)} = a w_s(t) \quad \text{for} \;\; s = \{l, h\}. \tag{7}$$

A worker's efficiency units of labor can be separated from other factors determining his wage. This allows for a sectoral wage per efficiency unit of labor, $w_s(t)$, which depends on the price of the intermediate the sector produces and the average quality of machines in that sector.

Since wages are increasing in ability, there is a cutoff ability for both skill levels in every period below which people become unemployed. This threshold is:

$$\underline{a}_s(t) \equiv \frac{\underline{w}(t)}{\beta(1-\beta)^{\frac{1-2\beta}{\beta}} p_s(t)^{\frac{1}{\beta}} Q_s(t)} \quad \text{for} \;\; s = \{l, h\}. \tag{8}$$

Workers with skill level s and innate ability $a \geq \underline{a}_s(t)$ work in sector s in period t.[13] This implies that whether an individual is unemployed in a given period only depends on his ability, his education level, and the minimum wage. Let $\mathbf{1}(a \geq \underline{a}_s(t))$ be an indicator that takes the value one if a worker with skill s and ability a is able to work in period t.

I assume that individuals are risk-neutral and that they choose their education level to maximize:

$$\sum_{j=0}^{\infty} \left(\frac{\lambda}{1+r}\right)^j u_{t+j},$$

where u_{t+j} is consumption of the final good in period $t+j$, λ is the probability of staying alive until the next period, r is the discount rate, which has to equal to the interest rate due to linear utility. Linear utility also implies that maximizing the expected present value of lifetime consumption and income are equivalent. Hence individuals acquire education in the first period of their life if and only if the expected present value of their lifetime earnings is greater as high-skilled than as low-skilled.

The expected discounted value of lifetime earnings of a worker born in period t with ability a and cost c if high-skilled is:

$$\mathbf{W}_{h,a,c}(t) = a(1-c)w_h(t)\mathbf{1}(a \geq \underline{a}_h(t)) + a \sum_{s=t+1}^{\infty} \left(\frac{\lambda}{1+r}\right)^{s-t} w_h(s)\mathbf{1}(a \geq \underline{a}_h(s)) \tag{9}$$

Since acquiring education takes c fraction of the first period of an individual's life, he can only work in the remaining $1-c$ fraction of the first period. Therefore the lifetime earnings of a high-skilled individual are non-increasing in c: the more time he spends in school, the less time he has to earn money.

[13] If $w_h(t) < w_l(t)$, then high-skilled individuals with ability $a \in [\underline{a}_l(t), \underline{a}_h(t)]$ could work in the low-skilled sector rather than be unemployed. However, in equilibrium $w_h(t) > w_l(t)$ holds in all periods.

Similarly the expected discounted value of lifetime earnings of a low-skilled individual are:

$$\mathbf{W}_{l,a,c}(t) = a \sum_{s=t}^{\infty} \left(\frac{\lambda}{1+r} \right)^{s-t} w_l(s) \mathbf{1}(a \geq \underline{a}_l(s)). \tag{10}$$

The lifetime earnings of a low-skilled worker do not depend on c, while the earnings of a high-skilled worker is decreasing in it. This gives rise to a cutoff rule in c for acquiring education.

Lemma 1. *For every ability level a there exists a cutoff time cost for acquiring education, $c_a(t)$, such that for individuals born in period t with ability a acquiring education is optimal if and only if $c < c_a(t)$. If $a \geq \underline{a}_h(t)$, then this cutoff time cost is:*

$$c_a(t) = \frac{\sum_{s=t}^{\infty} \left(\frac{\lambda}{1+r} \right)^{s-t} [Q_h(s) p_h(s)^{\frac{1}{\beta}} \mathbf{1}(a \geq \underline{a}_h(s)) - Q_l(s) p_l(s)^{\frac{1}{\beta}} \mathbf{1}(a \geq \underline{a}_l(s))]}{Q_h(t) p_h(t)^{\frac{1}{\beta}}}$$

If $a < \underline{a}_h(t)$, then

$$c_a(t) = \begin{cases} \bar{c} & \textit{if} \quad \sum_{s=t}^{\infty} \left(\frac{\lambda}{1+r} \right)^{s-t} [Q_h(s) p_h(s)^{\frac{1}{\beta}} \mathbf{1}(a \geq \underline{a}_h(s)) - Q_l(s) p_l(s)^{\frac{1}{\beta}} \mathbf{1}(a \geq \underline{a}_l(s))] \geq 0 \\ 0 & \textit{if} \quad \sum_{s=t}^{\infty} \left(\frac{\lambda}{1+r} \right)^{s-t} [Q_h(s) p_h(s)^{\frac{1}{\beta}} \mathbf{1}(a \geq \underline{a}_h(s)) - Q_l(s) p_l(s)^{\frac{1}{\beta}} \mathbf{1}(a \geq \underline{a}_l(s))] < 0 \end{cases}$$

Proof. Use (7) in (9) and (10) □

Education is worth the investment for an individual with ability a and cost c if $\mathbf{W}_{h,a,c}(t) > \mathbf{W}_{l,a,c}(t)$. As described earlier, there are two channels through which education can increase lifetime earnings: either the wage per efficiency unit is higher for high-skilled than for low-skilled workers, or being high-skilled allows the individual to work while as low-skilled he would be unemployed. The second case arises when a is such that $aw_l(t) < \underline{w}(t) < aw_h(t)$, which also requires that $w_l(t) < w_h(t)$. Hence the following remark:

Remark 1. *To have high-skilled individuals in a generation born in period t, there has to be at least one period $s \geq t$, such that the wage per efficiency unit of labor is higher for the high-skilled than for the low-skilled: $w_l(s) < w_h(s)$.*

Assuming that the economy started at time zero with a cohort of measure one, and that their education cutoff cost is described by $c_a(0)$ for all a, the effective labor supplies in period t are:

$$\begin{aligned} N_l(t) = (1-\lambda) \sum_{j=0}^{t-1} \lambda^j \int_{\underline{a}_l(t)}^{\infty} \int_{c_a(t-j)}^{\bar{c}} a f(a) g(c) \mathbf{1}(a \geq \underline{a}_l(t)) dc da \\ + \lambda^t \int_{\underline{a}_l(t)}^{\infty} \int_{c_a(0)}^{\bar{c}} a f(a) g(c) \mathbf{1}(a \geq \underline{a}_l(t)) dc da, \end{aligned} \tag{11}$$

$$
\begin{aligned}
N_h(t) = (1-\lambda) & \int_{\underline{a}_h(t)}^{\infty} \int_0^{c_a(t)} a f(a) g(c)(1-c) \mathbf{1}(a \geq \underline{a}_h(t)) dcda \\
& + (1-\lambda) \sum_{j=1}^{t-1} \lambda^j \int_{\underline{a}_h(t)}^{\infty} \int_0^{c_a(t-j)} a f(a) g(c) \mathbf{1}(a \geq \underline{a}_h(t)) dcda \\
& + \lambda^t \int_{\underline{a}_h(t)}^{\infty} \int_0^{c_a(0)} a f(a) g(c) \mathbf{1}(a \geq \underline{a}_h(t)) dcda.
\end{aligned} \tag{12}
$$

3 Equilibrium

The economy is in a decentralized equilibrium at all times; that is, all firms maximize profits and all individuals maximize their lifetime utility given a sequence of minimum wages, or given an expectation of the sequence of minimum wages.

Definition 1. *A **decentralized equilibrium** is given by a sequence of cutoff costs for education $\{c_a(t)\}_{t=1}^{\infty}$ for all $a \in [0, \infty]$, cutoff ability levels $\{\underline{a}_h(t), \underline{a}_l(t)\}_{t=1}^{\infty}$, effective labor supplies $\{N_h(t), N_l(t)\}_{t=1}^{\infty}$, intermediate good prices $\{p_h(t), p_l(t)\}_{t=1}^{\infty}$, leading vintage qualities $\{q_h^j(t), q_l^j(t)\}_{t=1}^{\infty}$ for $j \in [0,1]$, average qualities $\{Q_h(t), Q_l(t)\}_{t=1}^{\infty}$, investments into R&D $\{\bar{z}_h^j(t), \bar{z}_l^j(t)\}_{t=1}^{\infty}$ and values of owning the leading vintage $\{V_h^j(t,q), V_l^j(t,q)\}_{t=1}^{\infty}$ for $j \in [0,1]$ and $q > 0$, given $Q_h(0), Q_l(0)$, $c_a(0)$ for all $a \in [0, \infty]$, and $\{\underline{w}(t)\}_{t=1}^{\infty}$, such that the following conditions are satisfied:*

1. *the cutoff time cost for education is compatible with individuals maximizing the expected present value of lifetime earnings: $\{c_a(t)\}_{t=1}^{\infty}$ for all $a \in [0, \infty)$ are as in Lemma 1,*
2. *only those are employed whose marginal productivity is above the minimum wage: $\{\underline{a}_h(t), \underline{a}_l(t)\}_{t=0}^{\infty}$ satisfy* (8),
3. *effective labor supplies are compatible with individual education decisions, cutoff abilities for unemployment, and with $c_a(0)$, and are described by* (11) *and* (12),
4. *intermediate good prices are compatible with the profit-maximizing behavior of competitive firms, and the final good price normalized to one:*

$$
p_l(t) = \left(1 + \gamma^{1 + \frac{\beta}{(1-(1-\beta)\rho)}} \left(\frac{Q_h(t)}{Q_l(t)} \frac{N_h(t)}{N_l(t)} \right)^{\frac{\rho\beta}{(1-(1-\beta)\rho)}} \right)^{\frac{1-\rho}{\rho}}, \tag{13}
$$

$$
p_h(t) = \left(\gamma^{\frac{\beta}{(1-(1-\beta)\rho)}} \left(\frac{Q_h(t)}{Q_l(t)} \frac{N_h(t)}{N_l(t)} \right)^{-\frac{\rho\beta}{(1-(1-\beta)\rho)}} + \gamma \right)^{\frac{1-\rho}{\rho}}, \tag{14}
$$

5. *$\{V_s^j(t,q)\}_{t=0}^{\infty}$ is the expected discounted present value of profits from a machine of quality q in line j in sector s, as in* (4),
6. *for all lines j and both sectors s total R&D investment is compatible with free entry into the R&D sector:* (5) *is satisfied,*

7. *average quality in sector s is compatible with optimal R&D spending on all lines j in s and evolves according to* (6) *for given $Q_s(0)$.*

3.1 The balanced growth path

I analyze balanced growth paths (BGP), which are decentralized equilibria, where all variables are constant or grow at a constant rate.

In the Appendix I show that in the BGP total R&D spending on all lines within a sector are equal, $\overline{z}_s^{j*} = \overline{z}_s^*$ for $j \in [0,1]$ and $\overline{z}_s^*$ is given by:

$$\underbrace{\beta(1-\beta)^{\frac{1-\beta}{\beta}}(p_s^*)^{\frac{1}{\beta}}N_s^*}_{\pi_s^*} = B\overline{z}_s^* \frac{(1+r-e^{-\eta\overline{z}_s^*})}{1-e^{-\eta\overline{z}_s^*}} \quad \text{for} \quad s=\{l,h\}. \tag{15}$$

The above equation shows that R&D effort in a sector is increasing in the period profit from machine sales.[14] These profits are higher if the price of the intermediate produced by it, p_s^*, is higher, or if more effective labor, N_s^*, uses this technology.

Along the BGP relative quality in the two sectors, Q^*, has to be constant, which requires equal R&D spending in the two sectors: $\overline{z}_h^* = \overline{z}_l^* = \overline{z}^*$, and the growth rate of the economy is: $g^* = 1 + (\overline{q}-1)(1-e^{-\eta\overline{z}^*})$. From (15) R&D spending in the two sectors is equal if:

$$p^* = \frac{p_h^*}{p_l^*} = \left(\frac{N_h^*}{N_l^*}\right)^{-\beta}. \tag{16}$$

For profits per unit of quality to be equal from owning a leading vintage in both sectors to be equal, the relative price of the two intermediates has to depend negatively on the relative supply of high-skilled workers.

Combining the relative price of the two intermediates – implied by perfect competition in the final good production – with (16) and with the production of intermediate goods gives:

$$Q^* = \frac{Q_h^*}{Q_l^*} = \gamma^{\frac{1}{1-\rho}}\left(\frac{N_h^*}{N_l^*}\right)^{\frac{\beta\rho}{1-\rho}}. \tag{17}$$

The relative quality level depends positively on the relative abundance of high-skilled labor along the balanced growth path. With more high-skilled workers, an innovation in the high-skilled sector is more profitable. Hence technology is more skill-biased, Q^* is greater, if the relative supply of skills is higher.

[14] To see this, take the derivative:

$$\frac{\partial z^*\left(1+\frac{r}{1-e^{-\eta z^*}}\right)}{\partial z^*} = 1 + \frac{r}{1-e^{-\eta z^*}}\left(1 - \frac{\eta z^* e^{-\eta z^*}}{1-e^{-\eta z^*}}\right).$$

A sufficient condition for this derivative to be positive is $1 - \frac{\eta z^* e^{-\eta z^*}}{1-e^{-\eta z^*}} \geq 0$. This can be rearranged to the following inequality:

$$1 \geq e^{-\eta z^*}(1+\eta z^*).$$

For $z^* = 0$ this holds with equality, while the right hand side is decreasing in z^*. QED

The skill premium per efficiency unit of labor depends on the relative price of the intermediates and the relative quality in the two sectors. This can be seen from combining (7) with (16) and (17):

$$\frac{w_h^*(t)}{w_l^*(t)} = \left(\frac{p_h^*}{p_l^*}\right)^{\frac{1}{\beta}} \frac{Q_h^*(t)}{Q_l^*(t)} = \gamma^{\frac{1}{1-\rho}} \left(\frac{N_h^*}{N_l^*}\right)^{\frac{\beta\rho}{1-\rho}-1}. \tag{18}$$

The skill premium per efficiency unit of labor increases in both the relative price and the relative quality in the two sectors. Both of these ratios depend on the relative supply of high-skilled workers. On the one hand, if there are more high-skilled workers, high-skilled intermediate production is greater, other things being equal, depressing the relative price. On the other, more high-skilled workers implies a greater market for high-skilled complementary machines, thus leading to a higher relative quality, Q^*. Whether the price or the technology effect dominates depends on the elasticity of substitution between the two intermediates. The more substitutable the two goods are, the more relative technology responds, and the less does the relative price. If $\rho > 1/(1+\beta)$, then the steady state skill premium is increasing, while if $\rho < 1/(1+\beta)$, then it is decreasing in the relative supply of skills.[15]

Note that the skill premium per efficiency unit of labor is not the same as the empirically observed skill premium. The observed skill premium is the ratio of the average wages:

$$\frac{\overline{w}_h^*(t)}{\overline{w}_l^*(t)} = \frac{w_h^*(t)}{w_l^*(t)} \frac{\overline{a}_h^*}{\overline{a}_l^*},$$

where $\overline{a}_h^*$ is the average ability among the high-skilled and $\overline{a}_l^*$ is the average ability among the low-skilled. Therefore the observed skill premium depends on both the skill premium per efficiency unit of labor and the relative average quality of the two skill groups.

The threshold ability of unemployment for the low-skilled is defined in (8), combining this with steady state wages yields:

$$\underline{a}_l^* = \frac{\underline{w}(t)}{\beta(1-\beta)^{\frac{1-2\beta}{\beta}} (p_l^*)^{\frac{1}{\beta}} Q_l^*(t)}. \tag{19}$$

Note that for the existence of a BGP, it is required that the minimum wage grows at the same rate as the low-skilled wage per efficiency unit, g^*.[16] Since the growth in average quality is driving wage growth, let $\widetilde{w} \equiv \frac{\underline{w}(t)}{Q_l(t)}$ denote the *normalized minimum wage*, which has to be constant in the steady state. Using (13) and (16), the low-skilled cutoff ability for employment is given by:

$$\underline{a}_l^* = \frac{\widetilde{w}}{\beta(1-\beta)^{\frac{1-2\beta}{\beta}}} \left(1 + \gamma \left(\frac{N_h^*}{N_l^*}\right)^{\frac{\beta\rho}{1-\rho}}\right)^{-\frac{1-\rho}{\beta\rho}}. \tag{20}$$

[15] In the former case technology is said to be strongly biased, while in the latter it is said to be weakly biased. For an extensive discussion see Acemoglu (2007). Also see Bárány (2011) for a discussion of skill premium and skill supply outside the steady state.

[16] If the minimum wage was growing at a slower(faster) rate than the low skilled wage per efficiency unit, then $\underline{a}_l^*$ would be falling(rising) over time, which would lead to a changing supply of high- and low-skilled labor.

Given $\underline{a}_l^*$, and using (8) and (18) the cutoff ability for the high-skilled is given by:

$$\underline{a}_h^* = \underline{a}_l^* \gamma^{\frac{-1}{1-\rho}} \left(\frac{N_h^*}{N_l^*}\right)^{1-\frac{\beta\rho}{1-\rho}}. \tag{21}$$

As pointed out earlier, the skill premium is greater than one, implying that the threshold ability for unemployment for the low-skilled is higher than the threshold ability for the high-skilled: $\underline{a}_h^* < \underline{a}_l^*$. Acquiring skills through education, for instance learning how to use different machines, increases workers' productivity and protects them from unemployment. Therefore, education allows people with low ability to increase their marginal productivity above the minimum wage, and to find employment.

In the steady state everyone has a constant employment status: they are either unemployed or employed in the low- or high-skilled sector. Moreover, depending on their innate ability, a, everyone falls into one of the following categories: $a < \underline{a}_h^*$, $a \in [\underline{a}_h^*, \underline{a}_l^*)$ or $a \geq \underline{a}_l^*$.

Consider an individual with ability $a < \underline{a}_h^*$. Since such an individual is unemployed regardless of his skills, the optimal decision is not to acquire education.

Now consider an individual with ability $a \in [\underline{a}_h^*, \underline{a}_l^*)$. If he does not acquire education, he becomes unemployed and earns zero income in every period. On the other hand, by completing his studies he earns the high-skilled wage. Since the opportunity cost of education is zero in this case, acquiring education to become high-skilled is the optimal decision, regardless of the individual's time-cost of education. This implies that at very low ability levels education is a substitute for ability; people can make up for their low ability by acquiring education, this way pushing their marginal productivity above the threshold.

Finally, consider an individual with ability $a \geq \underline{a}_l^*$, who is always employed regardless of his skill level. Such an individual acquires education if the present value of his earnings as high-skilled (9) exceed his present value earnings as low-skilled (10).

Result 1. *Every individual with ability $a \geq \underline{a}_l^*$ born in period t acquires education if his cost $c < c^*$, where c^* is the* cutoff time cost *implicitly defined by:*

$$c^* = \frac{1 - \frac{w_l^*(t)}{w_h^*(t)}}{1 - \frac{g^*\lambda}{1+r}} \tag{22}$$

Proof. Building on Lemma 1 and using that in equilibrium $\mathbf{1}(a \geq \underline{a}_l^*) = 1$ for all $k \geq 0$, for $s = l, h$, and $a \geq \underline{a}_l^*$, implies that $c_a(t)$ is independent of a in this range. Using the fact that wages in both sectors grow at a constant rate g^*, and that the skill premium, $w_h^*(t)/w_l^*(t)$ is constant, $c^*(t) = c^*$ is constant and given by (22). □

The threshold time cost for acquiring education and consequently the fraction of high-skilled workers depends positively on the skill premium and on the growth rate of the average qualities.[17] A higher

[17]The growth rate depends on the equilibrium amount of R&D investment $\overline{z}^*$, which is an increasing function of $(p_s^*)^{\frac{1}{\beta}} N_s^* = \left((N_l^*)^{\frac{\beta\rho}{1-\rho}} + \gamma (N_h^*)^{\frac{\beta\rho}{1-\rho}}\right)^{\frac{1-\rho}{\beta\rho}}$.

skill premium implies a greater per period gain from working as high-skilled, hence c^* is increasing in $w_h^*(t)/w_l^*(t)$. The growth rate of wages also increases the threshold time cost; if wages grow at a higher rate, then for a given skill premium, future gains are greater.

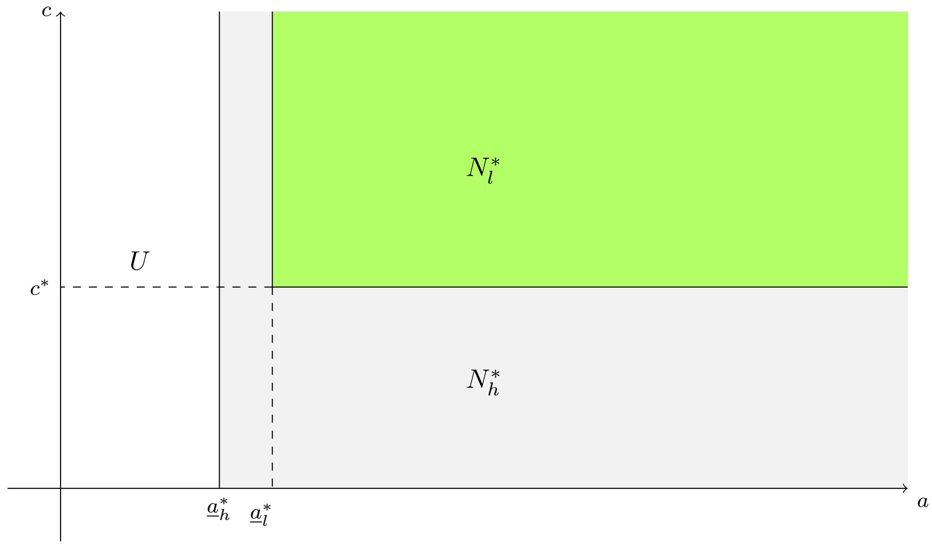

Figure 2: Optimal education and employment status

The horizontal axis represents the support of the ability distribution, and the vertical axis represents the support of the cost distribution.

Figure 2 depicts educational choices in the steady state. Individuals with ability lower than $\underline{a}_h^*$ are unemployed and do not acquire education (U). Between the two thresholds, $\underline{a}_h^* \leq a < \underline{a}_l^*$, everyone acquires education and becomes high-skilled to avoid unemployment. Finally individuals with ability above $\underline{a}_l^*$ acquire skills if their time cost is below c^*. The steady state supply of high- and low-skilled labor is:

$$
\begin{aligned}
N_h^* = & \left((1-\lambda)\int_0^{\overline{c}}(1-c)g(c)dc+\lambda\right)\int_{\underline{a}_h^*}^{\underline{a}_l^*} af(a)da \\
& + \left((1-\lambda)\int_0^{c^*}(1-c)g(c)dc+\lambda G(c^*)\right)\int_{\underline{a}_l^*}^{\infty} af(a)da, \\
N_l^* = & (1-G(c^*))\int_{\underline{a}_l^*}^{\infty} af(a)da.
\end{aligned}
$$

This completes the description of the steady state: the three cutoff values, $\underline{a}_h^*, \underline{a}_l^*$ and c^*, determine the effective labor supplies, N_h^* and N_l^*. In turn, the effective labor supplies determine every other variable in the economy in steady state.

4 Quantitative Analysis

This paper emphasizes that permanent changes in the minimum wage affect educational choices and technological change, and through this impact the shape of the wage distribution, and inequality. To assess the importance of these channels, I study the contribution of the decline in the real US federal minimum wages since the 1980s to the evolution of wage inequality.

4.1 Calibration

In order to separate the effects of the minimum wage from other forces that shaped wage inequality since the 1970s, I assume that the only exogenous change was in the level of the minimum wage. The minimum wage in real terms was pretty stable in the decade prior to 1981, while from 1982 onwards its value was not adjusted by inflation, and hence its real value started declining. For this reason, I assume in the calibration and in the quantitative exercise that the economy was initially in a steady state, and I calibrate the initial steady state to match the data from 1981.

Table 1 summarizes the calibrated parameter values. The first three parameters are taken from the literature; β, which is the share of labor in production is set to $2/3$, while the annual interest rate is set to 5 per cent. Since in the model, people can spend only a fraction of their first period studying, I set one period to correspond to five years. The per period interest rate is then $1.05^5 - 1$. The probability of survival, λ, is chosen such that individuals in expectation spend 45 years working and studying.

Table 1: Calibrated parameters

β	r	λ	$\bar{c}$	σ	ρ	γ	η	$\bar{q}$	B	$\widetilde{w}$
$2/3$	$1.05^5 - 1$	$8/9$	0.82	0.73	0.9	1.15	0.25	2.08	0.15	0.4

The second set of parameters describe the distribution of costs of education and abilities. Since abilities and education costs are not directly observable, I combine equilibrium conditions of the model with aggregate employment status and individual-level observable characteristics such as wages, education levels and age to estimate these distributions.

For the estimation of the cost and ability distribution I use the May and Outgoing Rotation Group supplements of the Current Population Survey for 1981 (MORG CPS). For the construction of the fraction of unemployed, studying, low-skilled and high-skilled workers, those 18-65 year olds who are either studying, unemployed, or employed are kept in the sample. The employed are divided into high- and low-skilled based on college education: those who attended college are high-skilled, those who did not are low-skilled. Only full time, full year, 18-65 year old workers are kept in the sample for the wage data. In order to capture only the effects of education and underlying ability, I use a cleaned measure of wage. This measure is the exponent of the residuals generated from regressing log hourly wages on age, age square, sex and race.

Figure 3 offers a good starting point for identifying the distribution of abilities and costs of education. The graph plots the density function of cleaned hourly wages for the high- and low-skilled workers. A striking feature in the figure is the significant overlap between the wages of the two educational groups. An appropriate distribution, therefore, must reproduce this pattern.[18]

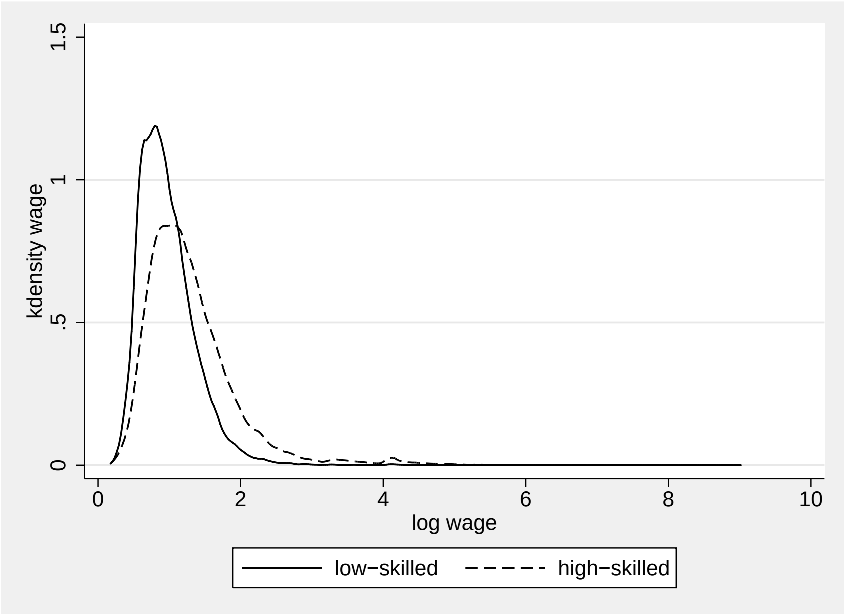

Figure 3: Hourly wages of the high- and low-skilled in 1981

Wages are calculated from the CPS MORG supplements from 1981. Only full time full year workers between the age of 18 and 65 are kept in the sample. Wages are the exponent of the residuals from regressing log hourly wage on age, age square, sex and race, person weights are used. Those who attended college are high-skilled, everyone else is low-skilled. The lines represent the Epanechnikov kernel function estimate produced by Stata, using the optimal bandwidth.

In general there are two components to the cost of education: a time cost and a consumption cost. Both these costs could be thought of as homogeneous or heterogeneous across individuals.[19] However, if the costs were purely consumption costs, then individuals with higher ability would acquire more education, leading to non-overlapping wage distributions of high- and low-skilled individuals. Therefore in the calibration and in the numerical results I assume that the cost of education is purely an idiosyncratic time cost.[20]

I assume a uniform time cost distribution on $[0, \overline{c}]$, with $\overline{c} \leq 1$, allowing a maximum of five years for studies if $\overline{c} = 1$. I assume that ability is log-normally distributed.[21] Since all variables of interest in the steady state calibration and in the quantitative assessment of the transition are invariant to the mean of

[18]This pattern is present in all years considered.

[19]The time cost arises because a person can work part-time at most while studying. The consumption cost is due to tuition fees and other expenses. Heterogeneity in consumption costs can be a reduced form of for example, a model with credit constraints and differential endowments.

[20]In the Appendix I discuss the case of a homogeneous cost, and the case of a distribution of consumption and time costs in detail.

[21]The results are robust to alternative distributions.

Table 2: Moments

Moment	Data	Model
L_u	0.0705	0.1023
L_l	0.5339	0.4923
L_h	0.3547	0.3964
$\overline{w}_h/\overline{w}_l$	1.3344	1.0518
$\overline{w}/w_{50}$	1.1072	1.2942
w_{90}/w_{50}	1.7060	2.4252
w_{50}/w_{10}	1.7006	2.0778
$\overline{w}_h/\overline{w}$	1.1796	1.0280
g	0.062	0.0798

the ability distribution, I normalize this mean to be one, i.e. $\mu = -\frac{1}{2}\sigma^2$.[22]

In the model, the wage of an individual with ability a_i and education s is $w_s(a_i) = a_i w_s$, implying a sectorial average wage of $\overline{w}_s = \overline{a}_s w_s$, where $\overline{a}_s$ is the average ability among those with education s. Therefore, an individual's ability relative to the average ability in his education group is equal to his wage relative to the average wage in that sector:

$$\frac{a_i}{\overline{a}_s} = \frac{w_s(a_i)}{\overline{w}_s} \equiv \widetilde{a}_{s,i}.$$

Since the education and wages of every respondent in the sample are recorded, I can infer *relative ability*, $\widetilde{a}_{s,i}$, from the data.

If the distribution of time costs and abilities is known, cutoff values for unemployment, $\underline{a}_h^*, \underline{a}_l^*$ and time cost c^* can be found by matching the fractions of unemployed, low- and high-skilled workers (constructed from the 1981 MORG CPS).[23] The thresholds $\underline{a}_h^*, \underline{a}_l^*$ and c^*, and the parameters of the ability and cost distributions are sufficient to calculate the average ability in both education groups, $\overline{a}_h, \overline{a}_l$ (see Figure 2 and the Appendix).

Multiplying the relative ability of a person by the average ability in his education group gives his ability level:

$$a_i = \frac{a_i}{\overline{a}_s}\overline{a}_s = \frac{w_s(a_i)}{\overline{w}_s}\overline{a}_s.$$

Once the ability of all individuals is backed out from the steady state conditions of the model, I calculate the likelihood of observing the sample of wage and education pairs. I maximize the likelihood by choosing parameters σ and $\overline{c}$. [24]

I calibrate the remaining parameters to minimize the weighted distance between moments of the

[22] Furthermore, in any model, where agents are heterogeneous in ability, the mean of the ability distribution and the technology level are not separable along any observable measure. Since this setup does not require the absolute level of technology, or the mean of the ability distribution for any quantity of interest, this normalization is without loss of generality.

[23] In the calibration I do not make a distinction in the educational attainment of the unemployed. In the steady state, only those who will be employed in the future should acquire education. In the data, half of the unemployed have some college education.

[24] The details of the maximization can be found in the Appendix.

initial steady state and the same moments from the data.[25] I use three types of moments: moments that describe the skill-composition and fraction of unemployed in the economy, those that describe the wage distribution, and those that reflect the R&D process. Moments of the first type are important to match, as most of the movement in the model comes from changes in these aggregates. These data are taken from the CPS MORG supplements between 1973-1981, and the selection criteria are the same as for calculating these shares for the maximum likelihood. The second type is also crucial, since I analyze the effects of minimum wages on inequality. The wage data used to construct these moments are also calculated from the 1973-1981 MORG CPS, and the selection of the sample and the treatment of the wages is the same as for Figure 3. Finally, matching the growth rate of GDP per worker, which is governed by the R&D process, determines the responsiveness of technology.[26] The moments and the fit of the model are summarized in Table 2.[27] The results presented in the following sections are all qualitatively unchanged for different parameter values, while the quantitative results are also fairly robust.

4.2 Transition

To analyze the role of the drop in minimum wages during the 1980s in the increasing inequality since then, I feed in the actual changes in the real minimum wage in 5-year steps, and assume that the final steady state features a 25 per cent lower normalized minimum wage than the initial pre-1981 steady state, as shown in Figure 4.[28]

I consider a sequence of unanticipated changes in the level of the normalized minimum wage. I assume that the economy is initially in a steady state, and the level of the normalized minimum wage unexpectedly changes several times. Those born in the period of the first shock have to form expectations about the future path of the minimum wage, and also about the future wages. I assume that they are myopic about the minimum wage, they think its level will stay the same from now on, but they are sophisticated in that they form correct expectations about how future generations of workers and

[25] The weight of the ith moment is the estimated standard deviation of the ith moment in the data for the period 1973-1981.

[26] The yearly growth rate of real GDP per worker is calculated by using yearly real GDP data from the Bureau of Economic Analysis, from yearly population data from the Census Bureau and from yearly employment to population ratios from the Bureau of Labor Statistics. All data was downloaded from the St Louis Fed. This is transformed into 5-yearly growth rates.

[27] I run a grid search over the set of parameter values and find the set that globally minimizes the distance from the moments. The set of gridpoints is around 200000, with $B \in [0.15, 0.3], \eta \in [0.05, 0.3], \gamma \in [0.85, 1.3], \overline{q} \in [1.78, 2.08], \rho \in [0.64, 0.98], \widetilde{w} \in [0.3, 0.72]$.

[28] By 1986 the minimum wage fell to 90 per cent of its pre-1981 value, by 1991 to 82 per cent, then it increased to 86 percent and remained at that level until 2001, and by 2006 it fell further to 75 per cent of its pre-1981 value. The change in the normalized minimum wage is not necessarily the same as the change in the minimum wage compared to the average low- or high-skilled wage, but Figure 4 shows that it is sufficiently close. Using the normalized minimum wage implies:

$$\widetilde{w}_1 \equiv \frac{\underline{w}(t)}{Q_l(t)} = \underline{a}_l(t)\beta(1-\beta)^{\frac{1-2\beta}{\beta}}(p_l(t))^{\frac{1}{\beta}},$$

while using the minimum wage compared to the average low-skilled wage implies:

$$\widetilde{w}_2 \equiv \frac{\underline{w}(t)}{\overline{w}_l(t)} = \underline{a}_l(t)\overline{a}_l(t).$$

These clearly do not imply the same dynamics for $\underline{a}_l(t)$, but since the magnitude of the change in both $p_l(t)$ and $\overline{a}_l(t)$ is small, their effect will be dominated by the drop in $\widetilde{w}$ throughout the transition.

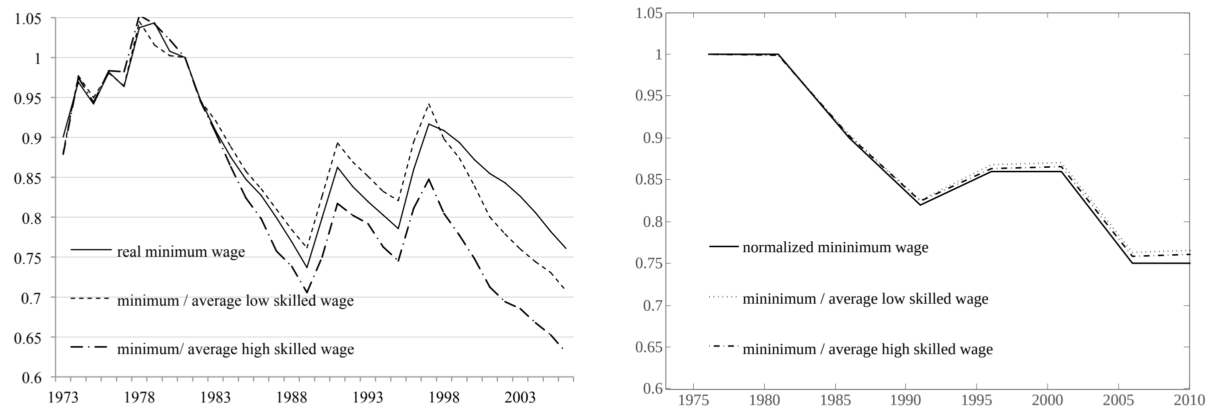

Figure 4: Actual and simulated minimum wage paths

The left panel shows the path of the actual minimum wage (calculated the same way as in Figure 1). The right panel shows the path of the normalized minimum wage, which is the exogenous variation driving the changes in the model, the minimum wage relative to the average low- and high-skilled wages are the outcome of agents' optimal reaction to path of the minimum wage.

R&D firms would respond to the minimum wage that they incorrectly believe to be stable from now on. Those born in the period of the second unexpected change in the level of the minimum wage are also myopic, but only about the level of the minimum wage and so on.

Figure 5 shows the transitional path of the main variables from the initial steady state to the new one, during which the economy is in a decentralized equilibrium.[29] As the top left panel shows, the lower minimum wage lowers the employment threshold for both the high- and the low-skilled, closely following the path of the minimum wage, thus lowering the average ability in both sectors. The path of the cutoff time cost for acquiring education is shown in the middle panel of the first row. This threshold c^* increases and overshoots in steps and then settles at its new steady state value, which is higher than the original. Through endogenous R&D, the increase in the supply of effective labor raises the growth rate of the economy, thus increasing the incentives to acquire education, resulting in a higher steady state cutoff cost of education. The two left panels in the bottom row show the path of the low- and high-skilled efficiency units of employed labor. Both immediately start to increase as the employment cutoffs fall. More employment in both sectors leads to more innovation and thus higher TFP growth rates, as shown in the top right panel. TFP in the low-skilled sector grows slightly more initially than in the high-skilled one, as employment in the low-skilled sector increases more, as shown in the bottom right panel.

The shift in the cutoffs lead to two types of changes: in the employment status of individuals and in the education decisions. The latter only affects the *new generations*: those born in the period when the minimum wage falls, and in subsequent generations. This is because the option of acquiring education is only available at birth.[30]

[29] The transition path is simulated using a forward shooting algorithm: given the path of the normalized minimum wage, the path of the three cutoffs: $\underline{a}_l$, $\underline{a}_h$ and c^* is guessed, and these are iterated on until these cutoffs are indeed optimal and the markets clear in each period. See Appendix for details.

[30] Allowing individuals to retrain themselves in later periods would not change the steady state, it would only speed up the

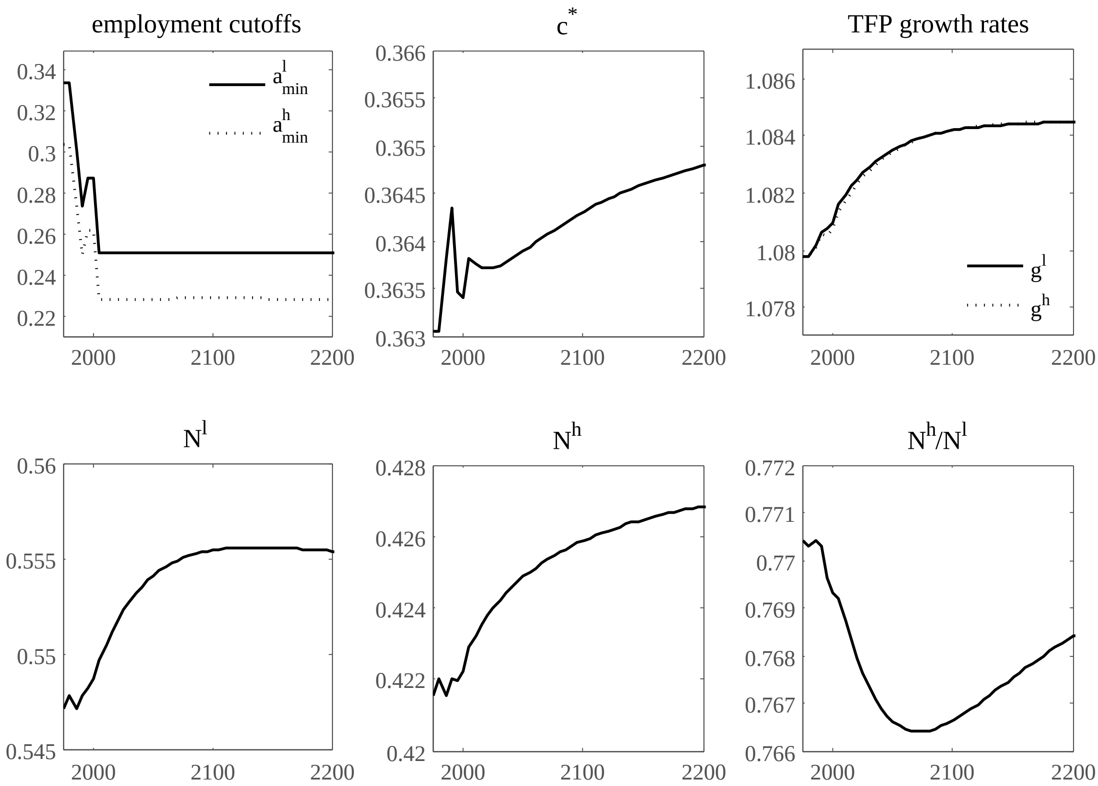

Figure 5: Transition of the main variables

The horizontal axis denotes the year, with the changes in the normalized minimum wage occurring after 1981.

Three of the variables with empirically observable counterparts are the unemployment (or non-employment), the relative supply of high- to low-skilled raw labor, L_h/L_l, and the average skill premium, $\overline{w}_h/\overline{w}_l$. The top panels of Figure 6 show the actual path of these variables, while the bottom panels show their simulated paths in the data. The model does relatively well in matching the path of unemployment (and non-employment) except for the increase in unemployment in the early 1980s. This Figure also shows that the model cannot explain the dramatic increase in the relative supply of high-skilled labor that is observed in the data. Though the supply of high-skilled workers increases according to the model, the supply of low-skilled increases even more, leading to an overall modest fall in the relative supply. The paths of the relative raw labor supplies and of the relative efficient labor supplies (in Figure 5) are quite different. This difference is due to the difference in ability between those who join the low-skilled and the high-skilled labor market. The measure of people joining the low-skilled workforce is larger than the measure of those joining the high-skilled workforce, reflected in the clear, though modest, overall decline in the relative supply of raw high-skilled labor. On the other hand, the average ability of those joining the high-skilled workforce is higher than the average ability of those joining the low-skilled. This is demonstrated by the even smaller long run decline in the relative supply of high-skilled effective labor. This implies that there are significant compositional changes in both the high-skilled and the low-skilled workforce. The average ability in both sectors decreases, but

transition process.

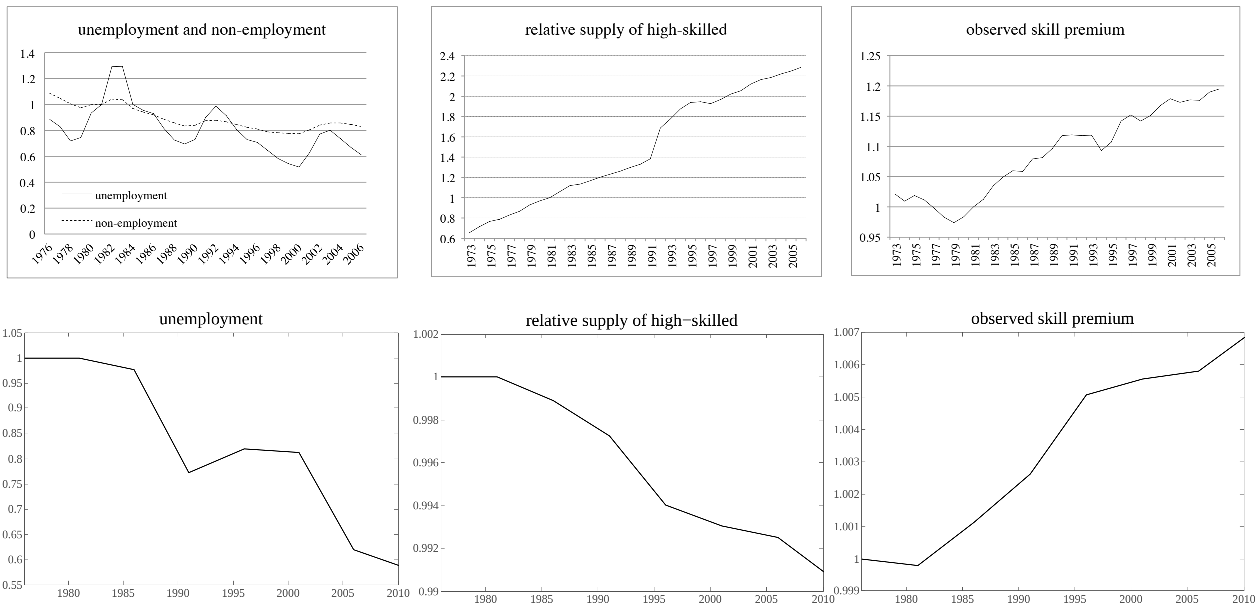

Figure 6: The path of unemployment, the relative high-skilled supply, and the skill premium
The top panels show the actual paths of un and non-employment, the relative high-skilled labor employed and the observed skill premium (all calculated the same way as for the calibration). The bottom panels show the same paths from the simulation. All paths are shown as indices, with their value normalized to 1 in 1981.

it decreases relatively more among the low-skilled. In terms of the observed skill premium, the model qualitatively matches the increase, though in terms of magnitude it explains very little. In the model the observed skill premium increases due to the increase in the relative average ability of the high-skilled workers. Note that the skill premium per efficiency unit of labor is not the same as the average skill premium. The average skill premium is the ratio of the average wages:

$$\frac{\overline{w}_h(t)}{\overline{w}_l(t)} = \frac{w_h}{w_l}\frac{\overline{a}_h(t)}{\overline{a}_l(t)},$$

where $\overline{a}_h(t)$ is the average ability among the high-skilled and $\overline{a}_l(t)$ is the average ability among the low-skilled. Therefore the average skill premium can change for two reasons: due to a change in the skill premium per efficiency unit of labor and in the relative average ability in the two skill groups. The skill premium per efficiency unit of labor actually falls slightly, as the relative supply of efficiency units of high-skilled workers falls, leading to a slightly unskill-biased technical change. Note also that the raw relative skill supply has been increasing continuously in the US even before the 1980s. These observations suggest – not surprisingly – that there are other forces driving the joint increase in the skill supply and the skill premium throughout this period.

The widening wage inequality is well captured by the increasing gap between the wages of workers at the 90th, 50th and 10th percentile of the wage distribution.[31] Figure 7 shows the actual path on the left

[31] Several authors have emphasized the importance to distinguish between lifetime and cross-sectional inequality (Flinn (2002), Bowlus and Robin (2004)). In the main text I only discuss the impact on cross-sectional inequality measures. In this model, since there are no lifetime income dynamics and the employment status of individuals does not fluctuate, the effect on lifetime

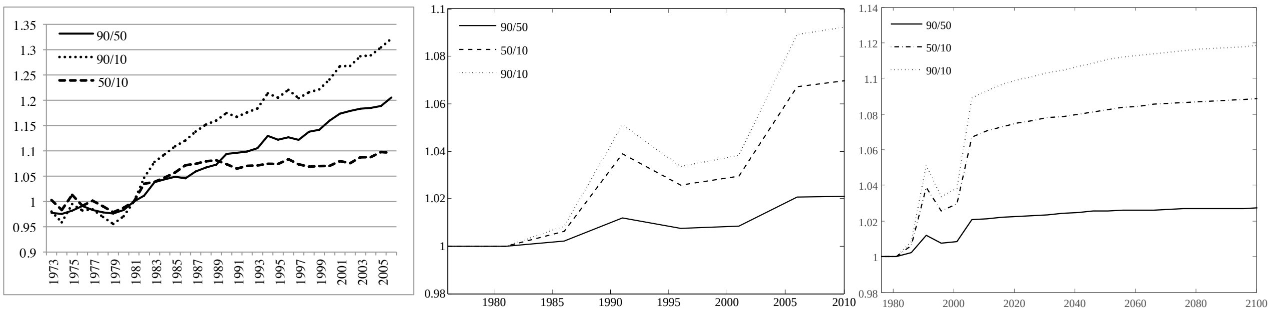

Figure 7: Actual and simulated path of the wage gaps

The left panel shows the actual path of the relative wage gaps (calculated the same way as for the calibration), while the middle and right panel show their simulated paths on different time horizons. All paths are shown as indices, with their value normalized to 1 in 1981.

and the simulated path on middle and on the right. The model does quite well both qualitatively and quantitatively in terms of these wage gaps. These wage gaps increase due to two factors: changes in the educational decisions, and changes in the ability composition of the high- and low-skilled workforce. Compositional forces always put an upward pressure on inequality. One component is the widening range of abilities present on the labor market. As the normalized minimum wage falls, the threshold abilities for employment decrease, increasing the range of abilities present on the labor market. As the range of abilities widens, the gap between the ability level at the 90th percentile gets further away from the ability level at the 50th percentile, which gets further from the 10th percentile. Another component is the changing ratio of high- to low-skilled workers at every percentile in the wage distribution. The fraction of high-skilled workers among the top 10 percent of earners increases, while their ratio at the bottom 10 percent decreases. The right panel shows that the transition takes a very long time, and inequality keeps increasing even after minimum wages have stabilized: newer cohorts with different optimal educational decisions gradually replace the older cohorts.

All three wage gaps increase during the transition, with a slight reduction when the minimum wage increases. The 90/10 wage differential increases the most, while the 90/50 increases the least. This is expected, since most of the compositional changes affect the lower end of the wage distribution. Note, however, that the change in the minimum wage causes the top end of the wage distribution to widen as well. This is mostly due to the compositional changes both in ability and in skill levels, which affect the position of the 90th percentile and the 50th percentile earner differentially.

Figure 8 contrasts the model's predictions to the observed pattern in the data. It shows the contribution of the changing real minimum wage to the change in these measures of inequality over time. According to the model the contribution of the fall in minimum wages to the widening of the wage gaps is: on average about one fifth of the observed increase in the 90/10 wage differential, about half of the increase in the 50/10 wage gap, and almost one tenth of the increase in the 90/50 wage gap.

inequality is similar to the effect on cross-sectional inequality, see graph in the Appendix.

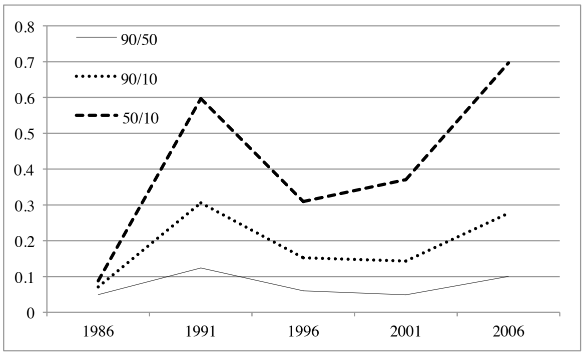

Figure 8: The dynamic contribution of minimum wages to the widening wage distribution
The lines depict the fraction of the measured change in (the given measure of) wage inequality (compared to 1981) that the model predicts following a fall in the minimum wage comparable to the data.

4.3 Relation to empirical studies

The main assumption of the model is that labor markets are perfectly competitive. This assumption implies that as minimum wages fall, employment and employment opportunities increase, which leads to adjustments in educational decisions. Direct empirical evidence on the main assumption of the model, on whether labor markets are perfectly competitive, is not available. However, we can confront the predictions of the model in terms of employment and education effects with the available empirical evidence.

The direct effect of minimum wage changes in the model is a change in employment and in employment opportunities. Most of the literature up to the late 1980s used time series variation in the federal minimum wage and aggregate data, and supported the conventional view that minimum wages reduce employment among teenagers (Brown, Gilroy, and Kohen (1982)). From the early 1990s a new line of research appeared, which uses natural experiments, state-level variation in minimum wages and in economic conditions, this became known as the 'new minimum wage research'. Several of these studies – challenging the conventional view – found that minimum wages have no, or small positive effects on employment, while others confirmed the conventional view that minimum wages reduce employment.[32] The new literature has proven quite controversial, with employment effects more contested, as shown by the exchange in Card, Katz, and Krueger (1994), Card and Krueger (2000) and Neumark and Wascher (1994), (2000). Brown (1999) and Neumark and Wascher (2006) provide an assessment of this literature. Both of these surveys highlight two reasons why the findings might be different. One reason might be the time frame of the minimum wage effects. The studies that find no, or positive, employment effects usually rely on case studies, and their estimates rely on short run differences in the data.

[32]For example some studies that found no or small positive employment effects are: Wellington (1991), Card (1992a), (1992b), Katz and Krueger (1992), Card and Krueger (1994), Machin and Manning (1994), Card and Krueger (1995); and some that found disemployment effects: Neumark and Wascher (1992), Deere, Murphy, and Welch (1995), Currie and Fallick (1996), Baker, Benjamin, and Stanger (1999).

The studies that find negative employment effects also incorporate lagged minimum wages, and point out that employment responds to long run, permanent changes in the minimum wage. The discrepancy in findings might suggest that low-skilled labor demand is not very elastic in the short run, perhaps due to sunk costs in production (Brown (1999)).[33] Regardless of the mechanism, if minimum wages affect employment opportunities at least in the long-run, then according to the model the education decisions are affected, which in turn alters the distribution of wages and the returns to education.[34]

The second reason for the discrepancy between the findings of the new minimum wage research highlighted by these surveys is the treatment of enrollment. The studies that control for school enrollment tend to confirm the disemployment effect of minimum wages. In light of the model presented here, the treatment of enrollment together with employment is crucial, as changes in minimum wages affect both decisions. The model predicts that higher minimum wages reduce overall educational attainment in the population: the lowest ability individuals become non-enrolled and non-employed, since even getting education will not guarantee them a job, while the higher ability individuals leave school for employment, as working, conditional on having a job is more worthwhile. The studies that look at the joint effects of the minimum wage on employment and enrollment tend to confirm the predictions of the model presented here. Neumark and Wascher (1995), (1996) and (2003) find that higher minimum wages lead to a higher probability that teenagers leave school, that non-enrolled teenagers become non-enrolled and non-employed, and to a lower probability that non-enrolled non-employed teenagers find a job. Turner and Demiralp (2001) confirm that at higher minimum wages more teenagers leave school, and that the fraction of non-employed non-enrolled students increase among non-whites. In general these studies suggest that there is a labor-labor substitution: as minimum wages increase employers substitute towards higher ability individuals, who might drop out of school to be employed full-time, while the lower ability individuals become non-employed and also drop out of school.

5 Conclusion

There has been much debate about the contribution of the falling minimum wage to the widening wage inequality in the US. The real value of the minimum wage eroded over the 1980s, losing 30 percent of its initial value. At the same time - in the early 1980s - there was an unprecedented surge in inequality. The wage gap widened between any two points in the wage distribution, and the college premium increased sharply. However, to my knowledge, there are no attempts in the literature to assess the quantitative significance of falling minimum wages for wage inequality in the context of a general equilibrium model.

In this paper I propose a general equilibrium model to analyze the effects of a permanent decrease

[33] This mechanism does not appear in the model presented here, but it could be easily incorporated if firms would need to buy their machines one period in advance.

[34] All the studies discussed above find that minimum wages have positive effects on the low-skilled wages. These effects would also lead to changes in the skill premium and in the educational attainment, thus would qualitatively not alter the main message of this paper.

in the value of the minimum wage on inequality. This model incorporates minimum wages, endogenous educational choices and endogenous technological progress. All these components are relevant in their own right: minimum wages affect the educational decisions of individuals through their effect on job and earning opportunities; educational decisions shape the skill composition of the labor force and the ability composition of different skill groups; the supply of high- and low-skilled labor affects the direction of technological change and the direction of technological change affects the educational decision of individuals.

The analysis in general equilibrium reveals that a reduction in the minimum wage affects overall inequality through three channels. First, a reduction in the minimum wage widens the range of abilities present on the labor market, thereby increasing the difference between any two percentiles in the distribution. Second, it differentially affects the shares of high- and low-skilled workers at every percentile in the wage distribution, thus increasing overall inequality. A third channel is the reduction in the skill premium per efficiency unit, which reduces inequality. Therefore, a reduction in the minimum wage affects inequality at the top end of the wage distribution, even if only to a smaller extent.

References

ÁBRAHÁM, Á. (2008): "Earnings inequality and skill-biased technological change with endogenous choice of education," *Journal of the European Economic Association*, 6(2-3), 695–704.

ACEMOGLU, D. (1998): "Why do new technologies complement skills? Directed technical change and wage inequality," *Quarterly Journal of Economics*, 113(4), 1055–1089.

——— (2007): "Equilibrium bias of technology," *Econometrica*, 75(5), pp. 1371–1409.

AGELL, J., AND K. E. LOMMERUD (1997): "Minimum wages and the incentives for skill formation," *Journal of Public Economics*, 64(1), 25 – 40.

AGHION, P., AND P. HOWITT (1992): "A model of growth through creative destruction," *Econometrica*, 60(2), 323–351.

ALESINA, A., AND J. ZEIRA (2006): "Technology and labor regulations," NBER Working Papers 12581, National Bureau of Economic Research, Inc.

AUTOR, D. H., L. F. KATZ, AND M. S. KEARNEY (2005): "Rising wage inequality: the role of composition and prices," Harvard Institute of Economic Research Working Papers 2096, Harvard - Institute of Economic Research.

——— (2008): "Trends in U.S. wage inequality: revising the revisionists," *Review of Economics and Statistics*, 90(2), 300–323.

AUTOR, D. H., L. F. KATZ, AND A. B. KRUEGER (1998): "Computing inequality: have computers changed the labor market?," *Quarterly Journal of Economics*, 113(4), 1169–1213.

AUTOR, D. H., A. MANNING, AND C. L. SMITH (2009): "The role of the minimum wage in the evolution of U.S. wage inequality over three decades: a modest re-assessment," *MIT Working Paper*.

BAKER, M., D. BENJAMIN, AND S. STANGER (1999): "The highs and lows of the minimum wage effect: a time-series cross-section study of the Canadian law," *Journal of Labor Economics*, 17(2), 318–350.

BÁRÁNY, Z. L. (2011): "Increasing Skill Premium and Skill Supply - Steady State Effects or Transition?," mimeo.

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

——— (2005): "Changes in U.S. Wages, 1976-2000: Ongoing Skill Bias or Major Technological Change?," *Journal of Labor Economics*, 23(3), 609–648.

BERMAN, E., J. BOUND, AND Z. GRILICHES (1994): "Changes in the demand for skilled labor within U.S. manufacturing: evidence from the annual survey of manufactures," *Quarterly Journal of Economics*, 109(2), 367–97.

BLANCHARD, O. J. (1985): "Debt, Deficits and Finite Horizons," *Journal of Political Economy*, 93(2), 223–247.

BOWLUS, A. J., AND J.-M. ROBIN (2004): "Twenty Years of Rising Inequality in U.S. Lifetime Labour Income Values," *The Review of Economic Studies*, 71(3), 709–742.

BROWN, C. (1999): "Minimum wages, employment, and the distribution of income," vol. 3, Part B of *Handbook of Labor Economics*, pp. 2101 – 2163. Elsevier.

BROWN, C., C. GILROY, AND A. KOHEN (1982): "The Effect of the Minimum Wage on Employment and Unemployment," *Journal of Economic Literature*, 20(2), pp. 487–528.

CAHUC, P., AND P. MICHEL (1996): "Minimum Wage Unemployment and Growth," *European Economic Review*, 40(7), 1463–1482.

CARD, D. (1992a): "Do Minimum Wages Reduce Employment? A Case Study of California, 1987-89," *Industrial and Labor Relations Review*, 46(1), pp. 38–54.

——— (1992b): "Using Regional Variation in Wages to Measure the Effects of the Federal Minimum Wage," *Industrial and Labor Relations Review*, 46(1), pp. 22–37.

CARD, D., AND J. E. DINARDO (2002): "Skill-Biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles," *Journal of Labor Economics*, 20(4), 733–783.

CARD, D., L. F. KATZ, AND A. B. KRUEGER (1994): "Comment on David Neumark and William Wascher, "Employment Effects of Minimum and Subminimum Wages: Panel Data on State Minimum Wage Laws"," *Industrial and Labor Relations Review*, 47(3), pp. 487–497.

CARD, D., AND A. B. KRUEGER (1994): "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania," *American Economic Review*, 84(4), 772–93.

——— (1995): *Myth and Measurement: The New Economics of the Minimum Wage*. Princeton University Press, Princeton, NJ.

——— (2000): "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania: Reply," *American Economic Review*, 90(5), 1397–1420.

CARNEIRO, P., AND S. LEE (2011): "Trends in Quality-Adjusted Skill Premia in the United States, 1960-2000," *American Economic Review*, 101(6), 2309–49.

CASELLI, F. (1999): "Technological Revolutions," *American Economic Review*, 89(1), 78–102.

CHANDRA, A. (2003): "Is the Convergence of the Racial Wage Gap Illusory?," Working Paper 9476, National Bureau of Economic Research.

Cunningham, J. (1996): "The Impact of Minimum Wages on Youth Employment, Hours of Work, and School Attendance: Cross-sectional Evidence from the 1960 and 1970 Censuses," in *The Economics of Legal Minimum Wages*, pp. 25–63. Washington, DC: American Enterprise Institute.

Currie, J., and B. C. Fallick (1996): "The Minimum Wage and the Employment of Youth Evidence from the NLSY," *Journal of Human Resources*, 31(2), 404–428.

Deere, D., K. M. Murphy, and F. Welch (1995): "Employment and the 1990-1991 Minimum-Wage Hike," *The American Economic Review*, 85(2), pp. 232–237.

DiNardo, J., N. M. Fortin, and T. Lemieux (1996): "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," *Econometrica*, 64(5), 1001–1044.

Eckstein, Z., and É. Nagypál (2004): "The Evolution of U.S. Earnings Inequality: 1961-2002," *Quarterly Review*, 28(2), 10–29.

Ehrenberg, R. G., and A. J. Marcus (1982): "Minimum Wages and Teenagers' Enrollment-Employment Outcomes: A Multinomial Logit Model," *The Journal of Human Resources*, 17(1), pp. 39–58.

Flinn, C. J. (2002): "Labour Market Structure and Inequality: A Comparison of Italy and the U.S.," *The Review of Economic Studies*, 69(3), 611–645.

Galor, O., and O. Moav (2000): "Ability-Biased Technological Transition, Wage Inequality, and Economic Growth," *Quarterly Journal of Economics*, 115(2), 469–497.

Goldin, C., and L. F. Katz (2007): "Long-Run Changes in the Wage Structure: Narrowing, Widening, Polarizing," *Brooking Papers on Economic Activity*, 2, 135–165.

——— (2008): *The Race Between Education and Technology*. The Belknap Press of Harvard University Press, Cambridge, MA.

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

Juhn, C., K. M. Murphy, and B. Pierce (1993): "Wage Inequality and the Rise in Returns to Skill," *Journal of Political Economy*, 101(3), 410–442.

Katz, L. F., and A. B. Krueger (1992): "The Effect of the Minimum Wage on the Fast-Food Industry," *Industrial and Labor Relations Review*, 46(1), pp. 6–21.

Katz, L. F., and K. M. Murphy (1992): "Changes in the Relative Wages, 1963-1987: Supply and Demand Factors," *Quarterly Journal of Economics*, 107(1), 35–78.

KILEY, M. T. (1999): "The Supply of Skilled Labour and Skill-Biased Technological Progress," *Economic Journal*, 109(458), 708–24.

KOENIGER, W., AND M. LEONARDI (2007): "Capital Deepening and Wage Differentials: Germany versus US," *Economic Policy*, 22, 71–116.

KRUEGER, A. B. (1993): "How Computers Have Changed the Wage Structure: Evidence from Microdata, 1984-1989," *Quarterly Journal of Economics*, 108(1), 33–60.

KRUSELL, P., L. E. OHANIAN, J.-V. RÍOS-RULL, AND G. L. VIOLANTE (2000): "Capital-Skill Complementarity and Inequality: A Macroeconomic Analysis," *Econometrica*, 68(5), 1029–1054.

LEE, D. S. (1999): "Wage Inequality in the United States During the 1980s: Rising Dispersion or Falling Minimum Wage?," *Quarterly Journal of Economics*, 114(3), 977–1023.

LEMIEUX, T. (2006): "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?," *American Economic Review*, 96(3), 461–498.

MACHIN, S., AND A. MANNING (1994): "The Effects of Minimum Wages on Wage Dispersion and Employment: Evidence from the U.K. Wages Councils," *Industrial and Labor Relations Review*, 47(2), pp. 319–329.

MATTILA, J. P. (1996): "The Impact of Minimum Wages on Teenage Schooling and on the Part-Time/Full-Time Employment of Youths," in *The Economics of Legal Minimum Wages*, pp. 25–63. Washington, DC: American Enterprise Institute.

MULLIGAN, C. B., AND Y. RUBINSTEIN (2008): "Selection, Investment, and Women's Relative Wages Over Time," *The Quarterly Journal of Economics*, 123(3), 1061–1110.

NEAL, D. (2004): "The Measured Black-White Wage Gap among Women Is Too Small," *Journal of Political Economy*, 112(S1), pp. S1–S28.

NEUMARK, D., AND W. WASCHER (1992): "Employment Effects of Minimum and Subminimum Wages: Panel Data on State Minimum Wage Laws," *Industrial and Labor Relations Review*, 46(1), pp. 55–81.

——— (1994): "Employment Effects of Minimum and Subminimum Wages: Reply to Card, Katz, and Krueger," *Industrial and Labor Relations Review*, 47(3), pp. 497–512.

——— (1995): "Minimum-Wage Effects on School and Work Transitions of Teenagers," *American Economic Review*, 85(2), 244–49.

——— (1996): "The effects of minimum wages on teenage employment and enrollment: Evidence from matched CPS surveys," in *Research in Labor Economics, Vol. 15*, pp. 25–63. Greenwich, Conn.: JAI Press.

——— (2000): "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania: Comment," *American Economic Review*, 90(5), pp. 1362–1396.

——— (2003): "Minimum Wages and Skill Acquisition: Another Look at Schooling Effects," *Economics of Education Review*, 22, 1–10.

——— (2006): "Minimum Wages and Employment: A Review of Evidence from the New Minimum Wage Research," Working Paper 12663, National Bureau of Economic Research.

OLIVETTI, C., AND B. PETRONGOLO (2008): "Unequal Pay or Unequal Employment? A Cross-Country Analysis of Gender Gaps," *Journal of Labor Economics*, 26(4), pp. 621–654.

RAVN, M. O., AND J. R. SORENSEN (1999): "Schooling, Training, Growth and Minimum Wages," *Scandinavian Journal of Economics*, 101(3), 441–57.

TURNER, M. D., AND B. DEMIRALP (2001): "Do higher minimum wages harm minority and inner city teens," *Review of Black Political Economy*, 28(4), 95–121.

WELLINGTON, A. J. (1991): "Effects of the Minimum Wage on the Employment Status of Youths: An Update," *Journal of Human Resources*, 26(1), 27–46.

Appendix – For online publication

Labor Supply

Proof of Lemma 1

Proof. First consider the case where $\mathbf{1}(a \geq \underline{a}_h(t)) = 1$. The individual for who $c(0) = c_a(t)$ is indifferent between acquiring education and not: $\mathbf{W}_{h,a,c(0)}(t) = \mathbf{W}_{l,a,c(0)}(t)$. Consider someone whose cost of acquiring education is $c(1) < c_a(t)$, for this individual $\mathbf{W}_{h,a,c(1)}(t) > \mathbf{W}_{h,a,c(0)}(t)$, while $\mathbf{W}_{l,a,c(1)}(t) = \mathbf{W}_{l,a,c(0)}(t)$, therefore due to transitivity $\mathbf{W}_{h,a,c(1)}(t) > \mathbf{W}_{l,a,c(1)}(t)$. The opposite holds for someone who has $c_2 > c_a(t)$.

Now consider the case where $\mathbf{1}(a \geq \underline{a}_h(t)) = 0$. Then, as income in the first period as high-skilled is zero, the optimal educational decision is independent of the individual's cost of acquiring education. It is therefore sufficient to compare the lifetime earnings of high- and low-skilled as if acquiring education would not take any time. If the income as high-skilled is higher, it is optimal to acquire education for all individuals with such a, i.e. $c_a(t) = 0$, whereas if the income as low-skilled is higher, then not acquiring education is optimal: $c_a(t) = \infty$. □

Steady State

R&D spending

Using that the steady state profits in sector s are constant:

Lemma 2. *The total R&D spending on any line for a given quality is constant along the BGP:* $\overline{z}^{j,s*}(t)(q) = \overline{z}^{j,s*}_{t+T}(q) = \overline{z}^{j,s*}(q)$ *for all* $t, T \geq 0$.

Proof. The R&D spending on each line has to be either constant or growing at a constant rate along the balanced growth path. This implies that the equilibrium total R&D spending on line j in sector s can be written as: $\overline{z}^{j,s*}_{t+T}(q) = \gamma^T \overline{z}^{j,s*}(t)(q)$. Where $\gamma > 0$ is the growth rate of the R&D spending on line j in sector s for a given quality q. In what follows I denote $\overline{z}^{j,s*}(t)(q)$ by $z(t)$. Conditional on quality q, the per period profit is constant, $\pi^*_s q$, since both N^*_s and p^*_s are constant along the BGP. Iterating forward (4), the value of owning the leading vintage on line j with quality q at time $t+T$ can be written as:

$$V_{t+T}(q) = q\pi^*_s \sum_{\tau=0}^{\infty} \frac{e^{-\eta z(t)\gamma^T \frac{\gamma^\tau - 1}{\gamma - 1}}}{(1+r)^\tau}.$$

Given $V_{t+T}(q)$ the equilibrium level of R&D spending is z_{t+T} if (5) is satisfied:

$$\frac{1}{1+r}\frac{(V_{t+T}(q))}{z_{t+T}}(1 - e^{-\eta z_{t+T}}) = q.$$

This has to hold for all $T > 0$, implying that

$$\begin{aligned}
&\textstyle\sum_{k=0}^{\infty} \frac{e^{-\eta z(t)\frac{\gamma^k - 1}{\gamma - 1}}}{(1+r)^k}(1 - e^{-\eta z(t)}) = \\
&\textstyle\frac{1}{\gamma}\sum_{k=0}^{\infty} \frac{e^{-\eta z(t)\gamma\frac{\gamma^k - 1}{\gamma - 1}}}{(1+r)^k}(1 - e^{-\eta z(t)\gamma}) = \ldots \\
&\textstyle\frac{1}{\gamma^T}\sum_{k=0}^{\infty} \frac{e^{-\eta z(t)\gamma^T\frac{\gamma^k - 1}{\gamma - 1}}}{(1+r)^k}(1 - e^{-\eta z(t)\gamma^T}).
\end{aligned}$$

To simplify notation denote $a_k \equiv \frac{\gamma^k - 1}{\gamma - 1}$ and $\eta z(t) \equiv b$. Since the above should hold for any $T > 0$, this implies that the difference between two consecutive terms should be zero. Taking logarithm and derivative with respect to T yields the following condition:

$$0 = \ln\gamma\left(-1 + \left(\frac{b\gamma^T e^{-b\gamma^T}}{1 - e^{-b\gamma^T}} - \frac{b\gamma^T\sum_{k=0}^{\infty}\frac{a_k e^{-b\gamma^T a_k}}{(1+r)^k}}{\sum_{k=0}^{\infty}\frac{e^{-b\gamma^T a_k}}{(1+r)^k}}\right)\right). \tag{23}$$

This has to hold for all $T > 0$, even as $T \to \infty$. There are three cases: $\gamma > 1$, $\gamma < 1$ and $\gamma = 1$. For $\gamma = 1$ the above trivially holds for all $T > 0$.
For $\gamma > 1$ taking the limes yields:

$$\begin{aligned}
&\lim_{T\to\infty}\left(\frac{b\gamma^T e^{-b\gamma^T}}{1 - e^{-b\gamma^T}} - \frac{b\gamma^T\sum_{k=0}^{\infty}\frac{a_k e^{-b\gamma^T a_k}}{(1+r)^k}}{\sum_{k=0}^{\infty}\frac{e^{-b\gamma^T a_k}}{(1+r)^k}}\right) = \\
&0 - \lim_{T\to\infty}\frac{b\gamma^T\sum_{k=0}^{\infty}\frac{a_k e^{-b\gamma^T a_k}}{(1+r)^k}}{\sum_{k=0}^{\infty}\frac{e^{-b\gamma^T a_k}}{(1+r)^k}} < 0
\end{aligned}$$

Where the second term is non-negative, implying a negative value as T grows very large. Hence, for $\gamma > 1$ (23) does not hold for all $T > 0$.
For $\gamma < 1$ I will show that the second term in the brackets is strictly smaller than 1, except in the limit. Denote $x \equiv b\gamma^T$, then as $T \to \infty$, $x \to 0$. The first term is smaller than 1 for any $x > 0$:

$$\frac{xe^{-x}}{1 - e^{-x}} < 1 \Leftrightarrow e^{-x}(1 + x) < 1$$

For $x = 0$, $e^{-x}(1 + x) = 1$. The derivative of the left hand side is $-e^{-x}x$, which is negative for all $x > 0$, implying that for any $x > 0$ the above inequality strictly holds.
The second term in the brackets is strictly positive for all $T > 0$ and finite. This implies that the term in the brackets is strictly smaller than 1 for any finite T. Hence (23) does not hold for any $T > 0$. □

Therefore in the steady state $\overline{z}^{j,s*}$ is constant for a line with quality q. This also implies that the value of owning the leading vintage with quality q in line j and sector s is constant in the steady state. Its value can be expressed from iterating (4) forward and using the above lemma as:

$$V^{j,s*}(t)(q) = \frac{q\beta(1-\beta)^{\frac{1-\beta}{\beta}}(p_s^*)^{\frac{1}{\beta}}N_s^*}{1 - \frac{e^{-\eta\overline{z}^{j,s*}(q)}}{1+r}}.$$

Note that the value of owning a leading vintage is proportional to its quality level. This observation leads to the following corollary:

Corollary 1. *In the steady state the total R&D spending on each line within a sector is constant and equal:* $\overline{z}^{j,s*}(t) = \overline{z}^{k,s*}_{t+v} = \overline{z}^*_s$ *for all* $j, k \in s$ *and all* $v \geq 0$*.*

Proof. Using (5) and the steady state value of owning a leading vintage, the total amount of R&D spending on line j in sector s with quality q is implicitly defined by:

$$\beta(1-\beta)^{\frac{1-\beta}{\beta}}(p^*_s)^{\frac{1}{\beta}}N^*_s = B\overline{z}^{j,s*}(q)\frac{(1+r-e^{-\eta\overline{z}^{j,s*}(q)})}{1-e^{-\eta\overline{z}^{j,s*}(q)}}.$$

The left hand side only depends on sector specific variables, hence the total amount of R&D spending on improving line j in sector s is independent of the current highest quality, q on that line. Since it is only the quality level that distinguishes the lines from each other within a sector the corollary follows. $\square$

From Corollary 1 the total amount of R&D spending on each line within a sector is equal and constant over time. This equilibrium R&D spending is given by (15). In the steady state $\overline{z}^*_h = \overline{z}^*_l = \overline{z}^*$ and the growth rate is $g^* = 1 + (\overline{q}-1)(1-e^{-\eta\overline{z}^*})$.

The price of the intermediates can be expressed from substituting the steady state relative price (16) into the intermediate good prices (13):

$$p^*_l = \left(1+\gamma\left(\frac{N^*_h}{N^*_l}\right)^{\frac{\beta\rho}{1-\rho}}\right)^{\frac{1-\rho}{\rho}} \tag{24}$$

$$p^*_h = \left(\left(\frac{N^*_h}{N^*_l}\right)^{-\frac{\beta\rho}{1-\rho}}+\gamma\right)^{\frac{1-\rho}{\rho}} \tag{25}$$

Using the steady state relative price and the steady state R&D investment:

$$B\overline{z}^*\frac{(1+r-e^{-\eta\overline{z}^*})}{1-e^{-\eta\overline{z}^*}} = \beta(1-\beta)^{\frac{1-\beta}{\beta}}\left(\gamma N^{*\frac{\beta\rho}{1-\rho}}_h + N^{*\frac{\beta\rho}{1-\rho}}_l\right)^{\frac{1-\rho}{\beta\rho}} \tag{26}$$

The right hand side is the steady state per period profit from owning the leading vintage normalized by the quality of the vintage. This profit is increasing in both N^*_h and N^*_l. If the labor supply increases, then any unit of investment into R&D has a higher expected return, since there are more people who are able to use it. This implies that the steady state R&D spending and the steady state growth rate is increasing in the effective labor supplies.

Calibration

Cost of education

For sake of brevity in the discussion of the various cases I only consider the decision of those individuals, who acquire education for higher wages and not to avoid unemployment. In all cases, there would be a range of abilities at the very bottom end of the ability distribution, where some people would acquire education to avoid unemployment, while the rest would be unemployed.

First, consider the case with a homogeneous consumption cost of acquiring education. In this case, the returns to education are increasing in ability, while the cost is fixed. In equilibrium there is a cutoff ability above which people acquire education, and below which they do not. Since both ability and wage per efficiency unit are higher for high-skilled individuals, equilibrium choices imply higher wages for high-skilled individuals. Wage distributions in this setup would not overlap, contradicting the empirically observed pattern.[35]

Second, assuming a distribution of consumption costs does not fit the empirical pattern of overlapping wage distributions either. A distribution of consumption costs implies a cutoff cost for every ability level in equilibrium. Given the cutoff for an ability level, those with the respective ability and lower cost of education acquire education, while those with cost higher than the cutoff do not. The equilibrium cutoff cost is increasing in ability: people with higher ability, have higher returns from education and are willing to pay a higher consumption cost for education. This implies that the fraction of high-skilled is increasing in the ability level, implying a higher average ability among the high-skilled. As in the previous case, high-skilled individuals have higher wages due to a higher unit wage and higher average abilities, contradicting the overlapping wage distribution pattern.[36]

Third, assuming instead, that the cost of education is a time cost, the equilibrium cutoff cost for acquiring education is independent of ability. If the ability and cost distributions are independent, then the high-skilled have higher wages only because of higher unit wages, since the average ability in the two sectors are equal. The distribution of wages for the high-skilled is a shifted and compressed version of the distribution of wages for the low-skilled. Hence, in this case predictions on the distribution of wages in the high- and low-skilled sector match well with the pattern observed in Figure 3. Therefore in the calibration and in the numerical results I assume that the cost of education is purely an idiosyncratic time cost.

Ability and Cost Distribution

Given the assumptions on the distribution of a and c, and the thresholds $\underline{a}_l^*, \underline{a}_h^*$ and c^* the high- and low-skilled effective labor supplies are:

$$\begin{aligned} N_h^* = \quad & \left((1-\lambda)\int_0^{\overline{c}}(1-c)g(c)dc + \lambda\right)\int_{\underline{a}_h^*}^{\underline{a}_l^*} af(a)da + \\ & \left((1-\lambda)\int_0^{c^*}(1-c)g(c)dc + \lambda G(c^*)\right)\int_{\underline{a}_l^*}^{\infty} af(a)da \end{aligned} \tag{27}$$

[35]If the homogeneous cost was a time cost, everyone would need to be indifferent between acquiring education or not. Since both the cost and the returns to education are linearly increasing in ability, if people were not indifferent then either everyone would acquire education or nobody would. An equilibrium based on indifference cannot be estimated from the data, since the ability, and therefore the wages of high- and low-skilled individuals are indeterminate in equilibrium.

[36]This holds even when the ability and cost distributions are independent. With a negative correlation between ability and the consumption cost of education, the two wage distributions would overlap even less.

$$N_l^* = (1 - G(c^*)) \int_{\underline{a}_l^*}^{\infty} af(a)da \tag{28}$$

Where $f(\cdot)$ is the probability density function of the ability distribution and $G(\cdot)$ is the cumulative distribution function of the cost distribution. The above expressions account for the fact that those members of the new generation who choose to acquire education only work $1 - c$ fraction of the first period of their life.

Note that the effective supply of labor is not equivalent to the measure of high- and low-skilled individuals, the difference being that the former counts the total *ability* available, while the latter counts the number of people. The measure of high-skilled, low-skilled and unemployed is given by:

$$\begin{aligned}
L_h^* = \quad & \Big((1-\lambda)\textstyle\int_0^{\overline{c}}(1-c)g(c)dc + \lambda\Big)\textstyle\int_{\underline{a}_h^*}^{\underline{a}_l^*} f(a)da + \\
& \Big((1-\lambda)\textstyle\int_0^{c^*}(1-c)g(c)dc + \lambda G(c^*)\Big)\, G(c^*)\textstyle\int_{\underline{a}_l^*}^{\infty} f(a)da, \\
L_l^* = \quad & (1 - G(c^*))\textstyle\int_{\underline{a}_l^*}^{\infty} f(a)da, \\
L^{u*} = \quad & \textstyle\int_0^{\underline{a}_l^*} f(a)da.
\end{aligned}$$

The cutoff ability of unemployment for the low-skilled is found by matching the fraction of unemployed:

$$U = \int (0)^{\underline{a}_h^*} f(a)da \quad \Leftrightarrow \quad \underline{a}_h^* = e^{(\sigma\Phi^{-1}(U)+\mu)} \tag{29}$$

The cutoff time cost is found by matching the fraction of low-skilled:

$$L_l = (1 - G(c^*)) \int_{\underline{a}_l^*}^{\infty} f(a)da, \tag{30}$$

where $\underline{a}_l^*$ satisfies (using (21)):

$$\underline{a}_l^* = \underline{a}_h^* \frac{w_h^*}{w_l^*} = \underline{a}_h^* \frac{\overline{w}_h}{\overline{w}_l} \frac{\overline{a}_l}{\overline{a}_h},$$

and $\overline{a}_h, \overline{a}_l$ are the average abilities and $\overline{w}_h, \overline{w}_l$ are the average wages in the two education groups. The average ability in a sector is the ratio of the supply of efficiency units of labor to the supply of raw labor in that sector: $\overline{a}^s = N^s/L^s$. The supply of high- and low-skilled raw labor, L_h and L_l are observed from the data, but N_h and N_l have to be calculated using (27).

This way for any cost and ability distribution $\underline{a}_l^*, \underline{a}_h^*$ and c^* is given as a function of the fraction of unemployed and low-skilled workers. Finally note that the three thresholds and the parameters of the ability and cost distribution are sufficient to calculate the average ability in both education groups.

Maximum likelihood

According to the model, if a high-skilled individual i's wage is lower than a low-skilled individual's wage, and since the skill premium is greater than one, it follows that his ability has to be lower as well.

This implies the following:

$$k(i) \equiv \arg\min_{j|w_h(i)<w_l(j)} w_l(j) \quad a_h(i) \leq a_l(k(i)).$$

Similarly, the ability of any low-skilled individual has to be higher than the ability of all high-skilled individuals with a lower wage:

$$k(i) \equiv \arg\max_{j|w_l(i)>w_h(j)} w_h(j) \quad a_l(i) \geq a_h(k(i)).$$

A high-skilled individual has wage $w_h(i)$ if his ability is $a_h(i) = \frac{w_h(i)}{\overline{w}_h}\overline{a}_h$, and he acquired education either to avoid unemployment, or because his time cost is lower than the threshold, $c(i) \leq c^*$. If he is in the first period of his life, his time cost of education must be lower than the maximum amount of time he could have spent studying. The probability of observing a high-skilled individual with wage $w_h(i)$ at age d is:

$$P(w_h(i), h, d) = \begin{cases} P(a = a_h(i)) & \text{if} \quad a_h(i) \in [\underline{a}_h^*, \underline{a}_l^*) \quad \& \quad d \geq 23 \\ P(a = a_h(i))P(c \leq \frac{d-18}{5}) & \text{if} \quad a_h(i) \in [\underline{a}_h^*, \underline{a}_l^*) \quad \& \quad d < 23 \\ P(a = a_h(i))P(c(i) < c^*) & \text{if} \quad a_h(i) \geq \underline{a}_l^* \quad \& \quad d \geq 23 \\ P(a = a_h(i))P(c(i) < \min\{c^*, \frac{d-18}{5}\}) & \text{if} \quad a_h(i) \geq \underline{a}_l^* \quad \& \quad d < 23 \end{cases}$$

Since there is an upper bound on the ability a high-skilled individual can have, the likelihood of observing a given wage, $w_h(i)$ for a high-skilled person can be written as:

$$\mathfrak{L}(w_h(i), d; \sigma, \overline{c}) = \begin{cases} 0 & \text{if} \quad a_h(i) < \underline{a}_h^* \quad \text{or} \quad a_h(i) > a_l(k(i)) \\ f(a_h(i)) & \text{if} \quad a_h(i) \in [\underline{a}_h^*, \underline{a}_l^*) \quad \& \quad a_h(i) \leq a_l(k(i)) \quad \& \quad d \geq 23 \\ f(a_h(i))G(\frac{d-18}{5}) & \text{if} \quad a_h(i) \in [\underline{a}_h^*, \underline{a}_l^*) \quad \& \quad a_h(i) \leq a_l(k(i)) \quad \& \quad d < 23 \\ f(a_h(i))G(c^*) & \text{if} \quad a_h(i) \geq \underline{a}_l^* \quad \& \quad a_h(i) \leq a_l(k(i)) \quad \& \quad d \geq 23 \\ f(a_h(i))G(\min\{c^*, \frac{d-18}{5}\}) & \text{if} \quad a_h(i) \geq \underline{a}_l^* \quad \& \quad a_h(i) \leq a_l(k(i)) \quad \& \quad d < 23 \end{cases} \tag{31}$$

Similarly, a low-skilled individual earning wage $w_l(i)$ must have $a_l(i) = \frac{w_l(i)}{\overline{w}_l}\overline{a}_l$, and cost exceeding the cutoff time cost; $a_l(i) \geq a_h(k(i))$ must also hold. The probability of observing $w_l(i)$ is then:

$$P(w_l(i), l) = P\left(a = a_l(i)\right) P(c(i) \geq c^*).$$

The likelihood of observing wage $w_l(i)$ for a low-skilled individual is:

$$\mathfrak{L}(w_l(i); \sigma, \overline{c}) = \begin{cases} 0 & \text{if} \quad a_l(i) < \underline{a}_l^* \quad \text{or} \quad a_l(i) < a_h(k(i)) \\ f(a_l(i))(1 - G(c^*)) & \text{if} \quad a_l(i) \geq \underline{a}_l^* \quad \& \quad a_l(i) \geq a_h(k(i)) \end{cases} \tag{32}$$

I calculate the likelihood of observing the sample of wage and education pairs using (31) and (32). I maximize the likelihood by choosing parameters σ and $\bar{c}$.

Elasticity of Substitution

Note that the elasticity of substitution between the intermediate goods, $1/(1-\rho)$, is not the same as the elasticity of substitution between high- and low-skilled workers, which has been estimated by several authors. However, their estimates are not comparable to ρ, since technology is usually modeled as exogenous, while in my model it is endogenous.

The consensus value is around 1.4 based on the paper by Katz and Murphy (1992). This original estimate was based on 25 data points, and Goldin and Katz (2008) updated this estimate by including more years and found an elasticity of 1.64. The estimating equation is:

$$\log \frac{w_h}{w_l} = \alpha_1 + \alpha_2 \log \frac{H}{L}. \tag{33}$$

These estimates typically adjust for productivity differentials within a skill-group, but do not adjust for differentials between skill groups. Hence the labor aggregates H and L are between the measure of effective labor and raw labor. The parameter estimate $\widehat{\alpha}_2$ is interpreted as the inverse of the elasticity of substitution between the two types of labor. I cannot use these estimates directly for several reasons.

First of all, the interpretation of $\widehat{\alpha}_2$ is different depending on the assumptions. To see this note that the skill premium per efficiency unit can be expressed as

$$\frac{w_h}{w_l} = \gamma^{\frac{1}{1-\rho}} \frac{N_h}{N_l}^{\frac{\beta\rho}{1-\rho}-1},$$

along the balanced growth path, while it can be measured as

$$\frac{w_h}{w_l} = \gamma^{\frac{1}{1-(1-\beta)\rho}} \left(\frac{N_h}{N_l}\right)^{-\frac{1-\rho}{1-(1-\beta)\rho}} \left(\frac{Q_h}{Q_l}\right)^{\frac{\beta\rho}{1-(1-\beta)\rho}}.$$

in the transition. Thereby, the interpretation along the BGP is $\widehat{\alpha}_2 = \beta\rho/(1-\rho) - 1$, while along the transition it is $\widehat{\alpha}_2 = -(1-\rho)/(1-(1-\beta)\rho)$. However, the estimate of $\widehat{\alpha}_2$ in the transition will be biased due to the lack of a good measure of average quality in the two sectors. Second, as noted before, the measure of labor supply aggregates used inKatz and Murphy (1992) are not the effective supply of labor, which in the model determines wages. Moreover, the measure of skill premium is not the skill premium per efficiency unit w_h/w_l of the model, it is probably closer to the average skill premium. Due to these reasons, reinterpreting the implications of the value of $\widehat{\alpha}_2$ for ρ is not sufficient to use these estimates in my calibration.

Transitional Dynamics

To use a forward shooting algorithm all variables have to be transformed in such a way, that they are stationary in the steady state. Let $v^s(t)$ denote the normalized value of owning the leading vintage in sector s at time t:

$$v_h(t) = \tfrac{V_h(t)(q)}{q} \quad v_l(t) = \tfrac{V_l(t)(q)}{q}$$

Let $\Delta(t)$ denote the normalized present value gain per unit of effective labor from acquiring education conditional on being employed in every future period (normalized by the current quality in the low-skilled sector):

$$\Delta(t) = \sum_{j=0}^{\infty} \left(\frac{\lambda}{1+r}\right)^j \frac{w_h(t+j) - w_l(t+j)}{Q_l(t)}$$

The equations that hold throughout the transition in terms of these normalized variables are:

$$\begin{aligned}
v^s(t+1) =& B\frac{(1+r)\overline{z}^s(t)}{1-e^{-\eta\overline{z}^s(t)}} \qquad s = l, h \\
v^s(t) =& \beta(1-\beta)^{\frac{1-\beta}{\beta}}(p^s(t))^{\frac{1}{\beta}}N^s(t) + \frac{e^{-\eta\overline{z}^s(t)}}{1+r}v^s(t+1) \qquad s = l, h \\
g^s(t+1) =& 1 + (\overline{q}-1)(1-e^{-\eta\overline{z}^s(t)}) \qquad s = l, h \\
p_h(t) =& \left(\gamma + \gamma^{\frac{\beta\rho}{(1-(1-\beta)\rho)(1-\rho)}}\left(Q(t)\frac{N_h(t)}{N_l(t)}\right)^{-\frac{\beta\rho}{(1-(1-\beta)\rho)}}\right)^{\frac{1-\rho}{\rho}} \\
p_l(t) =& \left(1 + \gamma^{1-\frac{\beta\rho}{(1-(1-\beta)\rho)(1-\rho)}}\left(Q(t)\frac{N_h(t)}{N_l(t)}\right)^{\frac{\beta\rho}{(1-(1-\beta)\rho)}}\right)^{\frac{1-\rho}{\rho}} \\
\widetilde{w} =& \underline{a}_l(t)\beta(1-\beta)^{\frac{1-2\beta}{\beta}}(p_l(t))^{\frac{1}{\beta}} \\
\widetilde{w} =& \underline{a}_h(t)\beta(1-\beta)^{\frac{1-2\beta}{\beta}}(p_h(t))^{\frac{1}{\beta}}Q(t) \\
Q(t+1) =& \frac{g_h(t+1)}{g_l(t+1)}Q(t) \\
\Delta(t) =& c^*(t)\beta(1-\beta)^{\frac{1-2\beta}{\beta}}(p_h(t))^{\frac{1}{\beta}}Q(t) \\
\Delta(t) =& \beta(1-\beta)^{\frac{1-2\beta}{\beta}}\left((p_h(t))^{\frac{1}{\beta}}Q(t) - (p_l(t))^{\frac{1}{\beta}}\right) + \frac{\lambda}{1+r}g_l(t+1)\Delta(t+1) \\
N_h(t) =& \lambda^t\left(\int_{\max(\underline{a}_h(0),\underline{a}_h(t))}^{\underline{a}_l(0)} af(a)da + \frac{c^*(0)}{\overline{c}}\int_{\max(\underline{a}_l(0),\underline{a}_l(t))}^{\infty} af(a)da\right) \\
&+ \sum_{i=1}^{t-1}\lambda^{t-i}(1-\lambda)\left(\int_{\max(\underline{a}_h(i),\underline{a}_h(t))}^{\underline{a}_l(i)} af(a)da + \frac{c^*(i)}{\overline{c}}\int_{\max(\underline{a}_l(i),\underline{a}_l(t))}^{\infty} af(a)da\right) \\
&+ (1-\lambda)\left(\frac{\overline{c}-\frac{\overline{c}^2}{2}}{\overline{c}}\int_{\underline{a}_h(t)}^{\underline{a}_l(t)} af(a)da + \frac{c^*(t)-\frac{c^*(t)^2}{2}}{\overline{c}}\int_{\underline{a}_l(t)}^{\infty} af(a)da\right) \\
N_l(t) =& \lambda^t\frac{\overline{c}-c^*(0)}{\overline{c}}\int_{\max(\underline{a}_l(0),\underline{a}_l(t))}^{\infty} af(a)da + \sum_{i=1}^{t}\lambda^{t-i}(1-\lambda)\frac{\overline{c}-c^*(i)}{\overline{c}}\int_{\max(\underline{a}_l(i),\underline{a}_l(t))}^{\infty} af(a)da
\end{aligned}$$

Lifetime inequality

The following graph depicts how the inequality across expected lifetime incomes at birth is affected by the subsequent unexpected changes in the minimum wage. Lifetime inequality widens to the same extent as cross-sectional wage inequality in the long run. The graph shows that the changes in lifetime inequality are much less gradual than changes in the cross-sectional wage inequality. Lifetime inequality increases immediately rather that incrementally, due to the perpetual youth modeling: those who are born in a period live through all of the increases in cross-sectional wages.

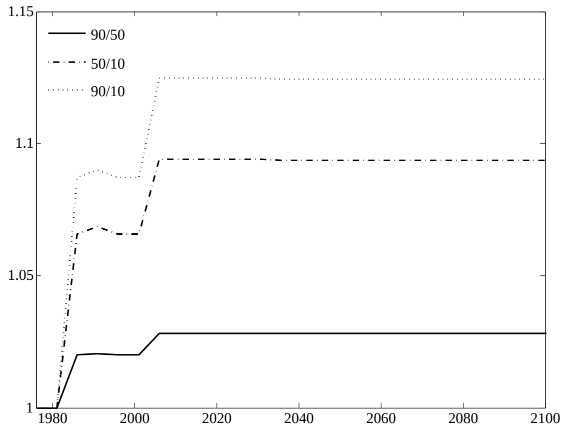

Figure 9: The widening of lifetime inequality