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

Zsófia L. Bárány*

This version: September 20, 2012

Abstract: The last 30 years witnessed a large increase in wage inequality accompanied by a fall in minimum wages. By introducing skill choice into a two-sector production economy, this paper links these patterns theoretically. First, a lower minimum wage increases inequality through its effects on educational attainment. Second, it reduces inequality by altering the direction of technological change. The calibrated model suggests that the first effect dominates; the fall of the real minimum wage during the 1980s in the US explains 50(10) percent of the increase in the lower-tail(upper-tail) wage gap and 15 percent of the increase in the skill premium.

Keywords: Minimum wage, education, technology, wage inequality


* I am grateful to Francesco Caselli and to Tom Cunningham, Mirko Draca, Radha Iyengar, 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 for many discussions, comments and suggestions that have improved this paper.

Economics Department, Sciences Po, 28 rue des Saints-Pères, 75015, Paris, France. email: zsophia.barany@sciences-po.org
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. 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.

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. The paper predicts that lower minimum wages significantly increase inequality. This overall increase is the result of two opposing forces. On the one hand, the educational and ability composition of the labor force changes, leading

to an increase in inequality. On the other hand, the relative supply of high-skilled labor decreases, which reduces the skill-bias of technology, hence compressing inequality.

Moreover, this paper contributes to the debate about the driving forces behind the changes in the structure of wages, by showing that in a general equilibrium setting the correlation between minimum wages and upper tail inequality, contrary to the claim of many, is not spurious: I provide a theoretical channel through which changes in minimum wages can affect inequality along the entire wage distribution. I find that even though minimum wages affect the bottom end of the wage distribution more, their impact on the top end is significant as well.

In most of 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 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 a change in the minimum wage alters the optimal education decisions, which implies a change in the skill composition of the labor force at the aggregate level. 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 minimum wages, education, and technological change, the quantitative general equilibrium effects of changes in the minimum wage on inequality could be quite different from what simple partial equilibrium reasoning may suggest.

I build a two sector growth model, where the labor supply and the available technology 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.

Figure 2: The dynamic contribution of minimum wages to the increase in inequality
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.

A decrease in the minimum wage increases the observed skill premium and widens the gaps between different percentiles of the wage distribution. Figure 2 shows the contribution of the 15 percent decline in the real value of the minimum wage to the change in these measures of inequality over time. In the case of the observed skill premium, the explanatory power significantly declines over time; initially the fall in the minimum wage accounts for almost 70 percent, whereas by 2006 it only accounts for 15 percent of the observed increase. The contribution of the fall in the minimum wage to the widening of the wage gaps is more stable: it explains on average one fifth of the observed increase in the 90/10 wage differential, about half of the increase in the 50/10 wage gap, and it explains on average a 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. Second, the skill composition of the employed changes gradually. The inflow 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, and increases at the top end. 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. 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  

---

2The effects of minimum wages on unemployment 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.

3The empirical evidence on the effects of minimum wages on educational attainment is mixed. Neumark and Wascher (2003) and Neumark and Nizalova (2007) find that higher minimum wages reduce educational attainment among the young. Sutch (2010) finds that minimum wages induce more human capital formation. I will confront the predictions of my paper with the empirical findings in Section 4.

4Lemieux (2006) and Autor et al. (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 change the evolution of inequality due to composition effects (Chandra (2003), Neal (2004), Mulligan and Rubinstein (2008), Olivetti and Petrongolo (2008)). In Section 4 I discuss their relation to the findings of this paper.
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 et al. (1993), Krueger (1993), Berman et al. (1994), Autor et al. (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 et al. (1996), Lee (1999)). 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 et al. (2008), Autor et al. (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 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 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 et al. (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 et al. (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 analyses the effects of minimum wages on educational attainment. Cahuc and Michel (1996) show that if min-

---

\(^5\)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.

\(^6\)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)).
imum 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 encourage 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 1 describes the model, section 2 the equilibrium and the steady state, section 3 details the calibration, section 4 provides the quantitative analysis, section 5 decomposes the effects, and finally section 6 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 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

The unique final good is produced in perfect competition by combining the two intermediate goods:

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

where $Y_l$ and $Y_h$ are the the intermediate goods produced by the low- and high-skilled workers. The elasticity of substitution between the two intermediates is $1/(1-\rho)$, with $\rho \leq 1$. Perfect competition implies that the relative price of the two intermediate goods is:

$$p \equiv \frac{p_h}{p_l} = \gamma \left( \frac{Y_l}{Y_h} \right)^{1-\rho}. \quad (1)$$

Intermediate good production is also perfectly competitive in both sectors $s \in \{l, h\}$. I simplify notation by allowing a representative firm:

$$Y_s = A_s N_s^\beta \quad \text{for} \quad s = \{l, h\}, \quad (2)$$
where $\beta \in (0, 1)$, $N_s$ is the amount of effective labor employed.\(^7\) $A_s$ is the technology level in sector $s$, which is the result of firms’ optimizing behavior. Firms decide the quantity, $x^j_s$, of a machine with quality $q^j_s$ to use, for each machine line $j \in [0, 1]$ available in sector $s$. The productivity in sector $s$ is then given by:

$$A_s = \frac{1}{1-\beta} \int_0^1 q^j_s(x^j_s)^{1-\beta} \, dj \quad \text{for} \quad s \in \{l, h\}.$$  

The productivity of high- and low-skilled workers is different, because they use a different set of machines. Producers of intermediate goods choose the quantity of machines ($x^j_s$) depending on the price and on the supply of effective labor it complements ($N_s$).

Since intermediate good production is perfectly competitive, industry demand for machine line $j$ of quality $q^j_s$ and price $\chi^j_s$ is:

$$X^j_s = \left( \frac{p_s q^j_s}{\chi^j_s} \right)^\frac{1}{\beta} N_s \quad \text{for} \quad s \in \{l, h\} \quad \text{and} \quad j \in [0, 1].$$ \hfill (3)

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

R&D firms with a successful invention have perpetual monopoly rights over the machine they patented. I assume that quality improvements are sufficiently large ($\bar{q} > (1 - \beta)^{-\frac{1-\beta}{\beta}}$), so that even if the second highest quality machine were sold at marginal cost, firms would prefer to buy the best quality machine, the leading vintage at the monopoly price. Therefore the profit-maximizing price of the leading vintage with

\(^7\)The exact definition of $N_s$ is specified later, in the labor supply section.
quality $q$ is $\chi(q) = \frac{q}{1-\beta}$.

The value of owning the leading vintage of quality $q$ in line $j$ and 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 $t+1$:

$$V_{js}^j(t, q) = \beta(1-\beta)\left(\frac{1}{\pi_s(t)} \right) \frac{(1-\beta)}{1+r} \pi_s(t) \beta N_s + \frac{1}{1+r}(e^{-\eta z_{js}^j(t,q)})V_{js}^j(t+1, q) \text{ for } s = \{l, h\} \text{ and } j \in [0,1].$$

(4)

Where $\pi_s(t)$ is the period profit in sector $s$ per unit of quality, given monopoly pricing and industry demand (3). If total R&D spending is $z_{js}^j(t,q)$, the probability that quality $q$ remains the leading vintage in line $j$ in period $t+1$ is $e^{-\eta z_{js}^j(t,q)}$.

Free entry into the R&D sector implies that profit opportunities have to be exhausted: the expected return from R&D investment (the left hand side) has to equal its cost (the right hand side):

$$E(t)\left(V_{js}^j(t+1,q_{js}^j(t)) \frac{z}{\pi_s(t, q_{js}^j(t))} \right) = Bq_{js}^j(t)z \text{ for } s = \{l, h\} \text{ and } j \in [0,1].$$

(5)

Notice that since for any firm $k$ both the expected return and the costs are proportional to its R&D investment, in equilibrium, only the total amount of R&D spending targeted at improving line $j$ in sector $s$ is determined.

### 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 $\lambda$. 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, $c \in [\underline{c}, \overline{c}] \in [0,1]$. These assumptions imply that both the size of the population, and the distribution of costs and abilities are constant over time.

---

8With probability $1-e^{-\eta z_{js}^j(t,q)}$ a higher quality machine is invented, and this quality level becomes obsolete.

9Here I assume that each R&D firm is sufficiently small so that $\pi_s(t)$ does not depend on $z_{js,k}^j(t)$. Assuming larger R&D firms would yield a similar result, with a more complicated expression.

10The choice of heterogeneous time cost and independent cost and ability distributions is explained in the Calibration section.
Each individual has to decide whether to acquire education in the first period of his life. Only those born in period $t$ can enroll to study in period $t$. 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.\footnote{In the calibration exercise I set the length of a period to be five years.} The time cost of education is idiosyncratic and is determined at birth. An individual who completes education becomes high-skilled and has the option of working in the high-skilled sector for life. High-skilled workers with ability $a$ earn wage $w_{h,a}(t)$ in period $t$. Those who choose not to acquire education, remain low-skilled and can start working in the period they are born as low-skilled. The wage in period $t$ for a low-skilled worker with ability $a$ is $w_{l,a}(t)$.

I model innate ability as a factor that increases individual productivity. Each worker supplies one unit of raw labor inelastically, which translates to $a$ units of efficiency labor for someone with ability $a$. The effective supply of high(low)-skilled labor is defined as the total amount of efficiency units of high(low)-skilled labor available in the economy.

Using monopoly pricing and the implied demand for machines, the wage can be expressed as:

$$w_{s,a}(t) = a \beta (1 - \beta)^{\frac{1-2\beta}{\beta}} p_s(t) \frac{Q_s(t)}{w_s(t)} \quad \text{for} \quad s = \{l, h\}. \quad (6)$$

Where $Q_s(t) = \int_0^1 q_s(j) dj$ denotes the average quality of the leading vintages in sector $s$. Since ability is equivalent to efficiency units of labor, it can be separated from other factors determining the 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.

The government imposes a minimum wage $w(t)$ in every period. Workers cannot be paid less than the minimum wage, hence those with marginal productivity below the minimum wage in period $t$ are unemployed in period $t$.

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:

\[
a_s(t) \equiv \frac{w(t)}{\beta(1-\beta)^{\frac{1-2\alpha}{\beta}}p_s(t)^\frac{1}{\beta}Q_s(t)} \quad \text{for } s = \{l, h\}. 
\]

(7)

Workers with skill level \(s\) and innate ability \(a \geq a_s(t)\) work in sector \(s\) in period \(t\).\(^{12}\)

This implies that whether an individual is unemployed in a given period only depends on his ability and education level. Let \(1(a \geq a_s(t))\) 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:

\[
E(t) \sum_{j=0}^{\infty} \left( \frac{\lambda}{1+r} \right)^j u(t+j),
\]

where \(u(t+j)\) is their consumption of the final good in period \(t+j\), \(\lambda\) 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 lifetime earnings of an educated individual using (6) can be expressed as:

\[
W_{h,a,c}(t) = a\beta(1-\beta)^{\frac{1-2\alpha}{\beta}} \left( \sum_{s=l}^{\infty} \left( \frac{\lambda}{1+r} \right)^{s-t} Q_h(s)p_h(s)^\frac{1}{\beta} \mathbb{1}(a \geq a_h(s)) - cQ_h(t)p_h(t)^\frac{1}{\beta} \mathbb{1}(a \geq a_h(t)) \right)
\]

(8)

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.

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

\[
W_{l,a,c}(t) = a\beta(1-\beta)^{\frac{1-2\alpha}{\beta}} \sum_{s=l}^{\infty} \left( \frac{\lambda}{1+r} \right)^{s-t} Q_l(s)p_l(s)^\frac{1}{\beta} \mathbb{1}(a \geq a_l(s)).
\]

(9)

\(^{12}\)If \(w_h(t) < w_l(t)\), then high-skilled individuals with ability \(a \in [a_l(t), a_h(t)]\) could work in the low-skilled sector rather than be unemployed. However, I later show that in equilibrium \(w_h(t) > w_l(t)\) for all \(t\).
Notice that the lifetime earnings of a low-skilled worker do not depend on \( c \), while the earnings of a high-skilled worker is decreasing in \( c \). 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 \( a \geq a_h(t) \), then this cutoff time cost is:

\[
    c_a(t) = \frac{\sum_{s=t}^{\infty} \left( \frac{1}{1+r} \right)^{s-t} \left( Q_h(s)p_h(s)\frac{1}{\beta}1(a \geq a_h(s)) - Q_l(s)p_l(s)\frac{1}{\beta}1(a \geq a_l(s)) \right)}{Q_h(t)p_h(t)\frac{1}{\beta}}
\]

If \( a < a_h(t) \), then

\[
    c_a(t) = \begin{cases} 
    \infty & \text{if } \sum_{s=t}^{\infty} \left( \frac{1}{1+r} \right)^{s-t} \left( Q_h(s)p_h(s)\frac{1}{\beta}1(a \geq a_h(s)) - Q_l(s)p_l(s)\frac{1}{\beta}1(a \geq a_l(s)) \right) \geq 0 \\
    c & \text{if } \sum_{s=t}^{\infty} \left( \frac{1}{1+r} \right)^{s-t} \left( Q_h(s)p_h(s)\frac{1}{\beta}1(a \geq a_h(s)) - Q_l(s)p_l(s)\frac{1}{\beta}1(a \geq a_l(s)) \right) < 0
    \end{cases}
\]

*Proof.* See Appendix. \( \square \)

Education is worth the investment for an individual with ability \( a \) and cost \( c \) if \( W_{h,a,c}(t) > W_{l,a,c}(t) \). 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 offers greater protection against unemployment. The second case arises when \( a \) is such that \( aw_l(t) < w_l(t) < aw_h(t) \), 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: \( w_l(s) < w_h(s) \).

### 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.
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 \( \{a_h(t), a_l(t)\}_{t=1}^{\infty} \), 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 \( \{q^j_h(t), q^j_l(t)\}_{t=1}^{\infty} \) for \( j \in [0, 1] \), average qualities \( \{Q_h(t), Q_l(t)\}_{t=1}^{\infty} \), investments into R&D \( \{\bar{\pi}^j_h(t), \bar{\pi}^j_l(t)\}_{t=1}^{\infty} \), and values of owning the leading vintage \( \{V^j_h(t, q), V^j_l(t, q)\}_{t=1}^{\infty} \) for \( j \in [0, 1] \) and \( q > 0 \), where \( c_a(0) \) for all \( a \in [0, \infty) \), \( \{a_h(0), a_l(0), N_h(0), N_l(0)\} \), and \( \{w(t), Q_h(t), Q_l(t)\}_{t=1}^{\infty} \) are given, 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 unemployed whose marginal productivity is below the minimum wage: \( \{q_h(t), q_l(t)\}_{t=0}^{\infty} \) satisfy (7),

3. the effective labor supplies are compatible with individual education decisions, cutoff abilities for unemployment, and \( c_a(0), a_h(0), a_l(0) \):

\[
N_l(t) = (1 - \lambda) \sum_{j=0}^{\infty} \lambda^j \int_{a_l(t)}^{\infty} \int_0^{\pi} a f(a) g(c) 1(a \geq a_l(t)) dcda, \tag{10}
\]

\[
N_h(t) = (1 - \lambda) \int_{a_h(t)}^{\infty} \int_0^{\pi} a f(a) g(c) 1(a \geq a_h(t)) dcda + (1 - \lambda) \sum_{j=1}^{\infty} \lambda^j \int_{a_h(t)}^{\infty} \int_0^{\pi} a f(a) g(c) 1(a \geq a_h(t)) dcda. \tag{11}
\]

4. the intermediate good prices are compatible with the profit-maximizing behavior of competitive firms, and the final good price normalized to one:

\[
p_l(t) = \left( 1 + \gamma^{1+ \frac{\beta}{(1-\rho)p}} \left( \frac{Q_h(t)}{Q_l(t)} \frac{N_h(t)}{N_l(t)} \right) \right)^{\frac{1-\rho}{\rho}}, \tag{12}
\]

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

5. \( \{V^j_h(t, q)\}_{t=0}^{\infty} \) is the expected discounted present value of profits from a machine of quality
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 for all \( j \in [0, 1] \).

7. the average quality in sector \( s \) is compatible with optimal R&D spending on all lines \( j \) in \( s \):

\[
Q_s(t+1) = \int_0^1 E(t)(q_j^s(t+1))dj = \int_0^1 q_j^s(t) \left( (1 - e^{-\eta \tau_i^l(t,q_j^l(t))})\bar{q} + \left( e^{-\eta \tau_i^l(t,q_j^l(t))} \right) \right) dj \quad \text{for } s = \{l, h\}.
\]

(14)

and the quality of the leading vintage in line \( j \) in sector \( s \), \( q_j^l(t+1) \), is a realization of the stochastic process of innovation.

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, \( \tau_j^s = \tau_s^* \) for \( j \in [0, 1] \) and \( \tau_s^* \) is given by:

\[
\beta(1 - \beta)^{-\beta} \frac{1}{\tau_s^*} p_s^* N_s^* = B \tau_s^* \frac{1 + r - e^{-\eta \tau_s^*}}{1 - e^{-\eta \tau_s^*}} \quad \text{for } s = \{l, h\}.
\]

(15)

The above equation shows that R&D effort in a sector is increasing in the period profit from machine sales.\(^{13}\) 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: \( \tau_h^* = \tau_l^* = \tau^* \), and the growth rate of the economy is: \( g^* = 1 + (\bar{q} - 1)(1 - e^{-\eta \tau^*}) \). From (15) R&D spending in the two sectors

---

\(^{13}\)To see this, take the derivative:

\[
\frac{\partial \tau^*}{\partial \tau^*} \left( 1 + \frac{r}{1 - e^{-\eta \tau^*}} \right) = 1 + \frac{r}{1 - e^{-\eta \tau^*}} \left( 1 - \frac{\eta \tau^* e^{-\eta \tau^*}}{1 - e^{-\eta \tau^*}} \right).
\]

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

\[
1 \geq e^{-\eta \tau^*} (1 + \eta \tau^*).
\]

For \( \tau^* = 0 \) this holds with equality, while the right hand side is decreasing in \( \tau^* \). QED
is equal if:

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

Combining the relative price (1), (16) with the production of intermediate goods gives:

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

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 (6) with (16) and (17):

\[ \frac{w^*_h(t)}{w^*_l(t)} = \left( \frac{p_h^*}{p_l^*} \right)^{\frac{1}{1-\rho}} \left( \frac{Q_h(t)}{Q_l(t)} \right)^{\frac{1}{1-\rho}} = \gamma^{\frac{1}{1-\rho}} \left( \frac{N_h^*}{N_l^*} \right)^{\frac{\beta \rho}{1-\rho} - 1}. \quad (18) \]

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.

The relative price of the two intermediates depends negatively on the relative supply of high-skilled workers. If there are more high-skilled workers, high-skilled intermediate production is greater, other things being equal. The technology effect reinforces this, since more R&D is directed towards the larger sector (from (17)), implying a higher relative quality, \( Q^* \). Intuitively, having more high-skilled workers and better high-skilled technologies, leads to more high-skilled intermediate production, and lowers the relative price of the intermediate.

Since the relative quality depends positively, while the relative price depends negatively on the relative supply of skilled workers, the net effect depends on which influences the wages more. This ultimately depends on the elasticity of substitution between the two intermediates. If \( \rho > 1/(1+\beta) \), then the two intermediates are highly substitutable, and the skill premium per efficiency unit of labor is an increasing function of the relative supply of skills, as relative technology responds sufficiently to offset the negative effect of relative price. If \( \rho < 1/(1+\beta) \), then the two intermediates are less substitutable, and the skill premium per efficiency unit of labor is decreasing in
the relative supply, as relative technology does not respond sufficiently to offset the
effect of relative price.\footnote{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).}

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:

\[
\frac{\bar{w}_h^*(t)}{\bar{w}_l^*(t)} = \frac{w_h^*(t)}{w_l^*(t)} \frac{\bar{a}_h}{\bar{a}_l},
\]

where \(\bar{a}_h^*\) is the average ability among the high-skilled and \(\bar{a}_l^*\) 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.

From Remark 1, the skill premium has to be greater than one in at least one period. As the skill premium per efficiency unit is constant (from (18)), \(w_h^*(t) > w_l^*(t)\) has to hold for all \(t \geq 0\).

The threshold ability of unemployment for the low-skilled is defined in (7), combining this with steady state wages yields:

\[
\bar{a}_l^* = \frac{\bar{w}(t)}{a_l^* \beta (1 - \beta) \frac{1 - 2 \beta}{\gamma} (p_l^*)^{\frac{\rho}{1 - \rho}} Q_l^*(t)}.
\] (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, \(g^*\).\footnote{If the minimum wage was growing at a slower(faster) rate than the low skilled wage per efficiency unit, then \(a_l^*\) would be falling(rising) over time, which would lead to a changing supply of high- and low-skilled labor.} Since the growth in average quality is driving wage growth, let \(\tilde{w} \equiv \frac{w(t)}{Q_l(t)}\) denote the normalized minimum wage, which has to be constant in the steady state. Using (12) and (16), the low-skilled cutoff ability for employment is given by:

\[
\bar{a}_l^* = \frac{\tilde{w}}{\beta (1 - \beta) \frac{1 - 2 \beta}{\gamma} \left(1 + \gamma \left(\frac{N_h^*}{N_l^*}\right)^{\frac{\beta}{1 - \rho}}\right)^{-\frac{1 - \rho}{\rho}}}.
\] (20)
Given $a^*_l$, and using (7) and (18) the cutoff ability for the high-skilled is given by:

$$a^*_h = a^*_l \gamma^{\frac{1}{1-\rho}} \left( \frac{N_h^*}{N_l^*} \right)^{1-\frac{\beta}{\gamma}}.$$  \hfill (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: $a^*_h < a^*_l$. 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: $a < a^*_h$, $a \in [a^*_h, a^*_l)$ or $a \geq a^*_l$.

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

Now consider an individual with ability $a \in [a^*_h, a^*_l)$. 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 $a \geq a^*_l$, who is always employed regardless of his skill level. Such an individual acquires education if the present value of his earnings as high-skilled (8) exceed his present value earnings as low-skilled (9).

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

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

Proof. Building on Lemma 1 and using that in equilibrium \( 1(a \geq a^*_s) = 1 \) for all \( k \geq 0, \)
for \( s = l, h, \) and \( a \geq a^*_s, \) 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. A higher 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 3** 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 3 depicts educational choices in the steady state. Individuals with ability
lower than \( a^*_h \) are unemployed and do not acquire education \((U)\). Between the two

---

The growth rate depends on the equilibrium amount of R&D investment \( \pi^* \), which is an increasing
function of \( (p_s^*)^\frac{1-\sigma}{2\sigma} \)

\[
N_s^* = \left( (N_s^*)^{\frac{1}{1+\rho}} + (N_h^*)^{\frac{1}{1+\rho}} \right)^{\frac{1-\rho}{1+\rho}}.
\]
thresholds, \( a^*_h \leq a < a^*_l \), everyone acquires education and becomes high-skilled to avoid unemployment. Finally individuals with ability above \( a^*_l \) acquire skills if their time cost is below \( c^* \). The steady state supply of high- and low-skilled labor is:

\[
N^*_h = \left( 1 - \frac{\lambda}{g(c^*)} \right) \int_{a^*_h}^{a^*_l} af(a)da + \left( 1 - \frac{\lambda}{g(c^*)} \right) \int_{a^*_l}^\infty af(a)da,
\]

\[
N^*_l = \left( 1 - G(c^*) \right) \int_{g^*_l}^\infty af(a)da.
\]

This completes the description of the steady state: the three cutoff values, \( a^*_h, a^*_l \) 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.

The paper’s main goal is to analyze the qualitative and quantitative effects of minimum wages on inequality. In the steady state, these effects are fully described by the changes in the three cut-off values: \( a^*_l, a^*_h \) and \( c^* \). It is useful to define the relative share of high-skilled workers as \( N^* \equiv N^*_h/N^*_l \), which can be expressed as:

\[
N^* = \left( \frac{1 - \frac{\lambda}{g(c^*)}}{1 - G(c^*)} \right) \int_{a^*_h}^{a^*_l} af(a)da + \left( \frac{1 - \frac{\lambda}{g(c^*)}}{1 - G(c^*)} \right) \int_{a^*_l}^\infty af(a)da.
\]

The first term is the relative supply of efficiency units of high-skilled labor, by those who acquire education in order to avoid unemployment, while the second is by those who acquire education in order to achieve higher wages. Notice that the second term only depends on the cutoff time cost of education, while the first term depends on all three cutoff values.

The direct effect of a lower minimum wage is that it excludes fewer people from the labor market, thus reducing \( a^*_h \) and \( a^*_l \). A lower \( a^*_l \) unambiguously reduces the first term in \( N^* \), while a lower \( a^*_h \) increases it. The only way the overall affect on the first term in \( N^* \) can be positive, is if the distance between \( a^*_h \) and \( a^*_l \) increases so much as to compensate for the higher denominator. This requires a substantial increase in
the skill premium, which did not occur in any of the simulations. The change in the minimum wage has indirect effects as well: as both the high- and the low-skilled labor supply expands, the total supply of labor increases, which increases the growth rate of the economy. This in turn leads to higher expected gains from acquiring education, thereby pushing $c^*$ upwards. The increase in $c^*$ increases both the first and the second term in $N^*$. Finally, as $N^*$ changes, the skill premium changes as a result of the price effect and the directed technology effect, of which the latter, depending on $\rho$, reinforces or mitigates the initial impact of $N^*$ on the skill premium. A change in $N^*$ affects $c^*$ and the distance between $a^h_k$ and $a^l_j$ through the skill premium. In general, the overall affect on $c^*$, on $N^*$, and on the skill premium is ambiguous. However, numerical solutions reinforce intuition: the direct effect is larger, and even though $c^*$ increases, the supply of low-skilled labor expands more, thus overall $N^*$ decreases as the minimum wage decreases.

4 Calibration

Table 1 summarizes the calibrated parameter values. The first three parameters are taken from the literature; $\beta$, 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 $1.05^{5/1} - 1$. The probability of survival, $\lambda$, is chosen such that individuals in expectation spend 45 years working and studying.\(^{17}\)

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.

\(^{17}\)The expected lifespan of someone who has a per period survival probability of $\lambda$ is

$$E(j) = \sum_{j=1}^{\infty} j \lambda^{j-1}(1 - \lambda) = \frac{1}{1 - \lambda}$$

Solving for $E(j) = 9$ gives $\lambda = 1 - \frac{1}{9}$.
Table 1: Calibrated parameters

<table>
<thead>
<tr>
<th></th>
<th>β</th>
<th>r</th>
<th>λ</th>
<th>τ</th>
<th>σ</th>
<th>ρ</th>
<th>γ</th>
<th>η</th>
<th>˜q</th>
<th>B</th>
<th>w</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2/3</td>
<td>1.05^5</td>
<td>-1</td>
<td>8/9</td>
<td>0.82</td>
<td>0.73</td>
<td>0.9</td>
<td>1.15</td>
<td>0.25</td>
<td>2.08</td>
<td>0.15</td>
</tr>
</tbody>
</table>

Figure 4, which represents the hourly wages of high- and low-skilled individuals, offers a good starting point for identifying the distribution of abilities and costs of education. 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.

![Figure 4: Hourly wages of the high- and low-skilled in 1981](image)

Wages are calculated from the CPS MORG supplements. Wages are the exponent of the residuals from regressing log hourly wage on age, age square, sex and race. Those who attended college are high-skilled, everyone else is low-skilled. The lines represent the kernel density estimate produced by Stata.

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.\(^{18}\) However, if the costs were purely consumption costs,

---

\(^{18}\)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.
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.\textsuperscript{19}

I assume a uniform time cost distribution on $[0, \tau]$, with $\tau \leq 1$, allowing a maximum of five years for studies if $\tau = 1$. I assume that ability is lognormally distributed.\textsuperscript{20} Since all variables of interest in the steady state calibration and in the quantitative assessment of the transition are invariant to the mean of the ability distribution, I normalize this mean to be one.\textsuperscript{21}

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 $\bar{w}_s = \bar{a}_s w_s$, where $\bar{a}_s$ 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:

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

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

If the distribution of time costs and abilities is known, cutoff values for unemployment, $a^*_h, a^*_l$ and time cost $c^*$ can be found by matching the fractions of unemployed, low- and high-skilled workers. The thresholds $a^*_h, a^*_l$ and $c^*$, and the parameters of the ability and cost distributions are sufficient to calculate the average ability in both education groups, $\bar{a}_h, \bar{a}_l$ (see Figure 3 and the Appendix).

Multiplying the relative ability of a person by the average ability in his education

\textsuperscript{19}In the Appendix I discuss of a homogeneous cost, a distribution of consumption and time costs.

\textsuperscript{20}The results are robust to alternative distributions.

\textsuperscript{21}This normalization is equivalent to:

$$E(a) = e^{\mu + \frac{1}{2} \sigma^2} = 1 \Leftrightarrow \mu = -\frac{1}{2} \sigma^2$$

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.
Table 2: Moments

<table>
<thead>
<tr>
<th>Moment</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$L_u$</td>
<td>0.0693</td>
<td>0.1023</td>
</tr>
<tr>
<td>$L_I$</td>
<td>0.5338</td>
<td>0.4923</td>
</tr>
<tr>
<td>$L_h$</td>
<td>0.3554</td>
<td>0.3964</td>
</tr>
<tr>
<td>$g$</td>
<td>0.0800</td>
<td>0.0798</td>
</tr>
<tr>
<td>$\bar{w}_h/\bar{w}_l$</td>
<td>1.3344</td>
<td>1.0518</td>
</tr>
<tr>
<td>$\bar{w}/w_{50}$</td>
<td>1.1072</td>
<td>1.2942</td>
</tr>
<tr>
<td>$w_{90}/w_{50}$</td>
<td>1.7060</td>
<td>2.4252</td>
</tr>
<tr>
<td>$w_{50}/w_{10}$</td>
<td>1.7006</td>
<td>2.0778</td>
</tr>
<tr>
<td>$\bar{w}_h/\bar{w}$</td>
<td>1.1796</td>
<td>1.0280</td>
</tr>
</tbody>
</table>

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 $\sigma$ and $\overline{\sigma}$.  

I calibrate the remaining parameters to minimize the weighted distance between moments of the initial steady state and the same moments from the data.  

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. The second type is also crucial, since I analyze the effects of minimum wages on inequality. Finally, matching the growth rate, which is governed by the R&D process, determines the responsiveness of technology. The moments and the fit of the model are summarized in Table 2.  

The results presented in the following sections are all qualitatively unchanged for different parameter values, while the quantitative results are also fairly

\[ a_i = \frac{\overline{a}_i}{\overline{a}_s} = \frac{w_s(a_i)}{\overline{w}_s \overline{a}_s}. \]

---

22The details of the maximization can be found in the Appendix.

23The weight of the $i$th moment is the estimated standard deviation of the $i$th moment in the data.

24I run a grid search over the set of parameter values and find the set that globally minimizes the distance from the moments.
robust (see Appendix for additional graphs).

5 Quantitative Analysis

To analyze the role of the drop in minimum wages during the 1980s in the increasing inequality since then, I consider an unanticipated permanent 20 per cent decrease in the normalized minimum wage.\textsuperscript{25} 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.\textsuperscript{26}

As the two top left panels in Figure 5 show, the lower minimum wage excludes fewer people from the labor market by lowering the unemployment threshold for both the high- and the low-skilled. At the moment of the announcement, both $a^*_h$ and $a^*_l$ drop almost immediately to their new steady state value.\textsuperscript{27} The path of the cutoff time cost for acquiring education is shown in the first row of Figure 5. This threshold $c^*$ initially overshoots and then decreases monotonically towards its new steady state value, which is higher than the original. This pattern can be understood by looking at the path of the skill premium (first row, right panel) and the path of the growth rates (bottom right panel). The initial jump in the skill premium, which is due to the large initial inflow from unemployment into the low-skilled labor force, drives the overshooting of $c^*$, then as the skill premium decreases, so does $c^*$. Counteracting

\textsuperscript{25}Figure 1 shows that the real value of the minimum wage decreased by about 30 percent until the late 1980s, while the minimum wage compared to the average high- and low-skilled wage decreased by about 20 percent. In the transitional dynamics I mimic this pattern by a one-time 20 percent drop in the value of the normalized minimum wage. The change in the normalized minimum wage is not necessarily the same as the change in the minimum wage compared to the average wage, but the transition shows that it is sufficiently close. Using the normalized minimum wage implies:

\[ \tilde{w}_1 \equiv \frac{w(t)}{Q(t)} = \mu(t)\beta(1-\beta)^{\frac{1-2\alpha}{\alpha}}(p_t(t))^{\frac{1}{\alpha}}, \]

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

\[ \tilde{w}_2 \equiv \frac{w(t)}{\bar{w}_1(t)} = \mu(t)\bar{w}(t). \]

These clearly do not imply the same dynamics for $\mu(t)$, but since the magnitude of the change in both $p_t(t)$ and $\bar{w}_1(t)$ is small, their effect will be dominated by the drop in $\tilde{w}$ throughout the transition.

\textsuperscript{26}I approximate the equations that have to hold throughout the transition up to second order, as in Schmitt-Grohe and Uribe (2004), see the Appendix for details.

\textsuperscript{27}See the Appendix for discussion.
Figure 5: Transition of the main variables
The horizontal axis denotes the year, with the drop in the normalized minimum wage occurring in 1981.

the decline in the skill premium is the increase in growth rates. 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.

Taking the path of the three cutoffs $a^*_h$, $a^*_l$, and $c^*$ as given, the paths of the effective supply of high- and low-skilled labor (depicted in row 2 of Figure 5) can be understood. Figure 6 plots the effect of changes in the cutoffs on the high- and low-skilled effective labor supply and on the employment status of individuals. The initial steady state thresholds are denoted by $a^*_{l0}$, $a^*_{h0}$, $c^*_{0}$, while the new steady state values are denoted by $a^*_{l1}$, $a^*_{h1}$, $c^*_{1}$. The maximum value of $c^*$, which is reached in the period of the announcement is denoted by $c^*_{01}$.

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 of the announcement, and in subsequent generations. This is because the option of acquiring education is only available at birth.\footnote{Allowing individuals to retrain themselves in later periods would not change the steady state, it would only speed up the transition process.}
Figure 6: Change in the optimal education and employment status
Ability is on the horizontal axis, the time cost is on the vertical axis. The initial steady state cutoffs are: $a^*_0, a^*_0, c^*_0$, while the new ones are: $a^*_1, a^*_1, c^*_1$. I denote the maximum threshold time cost that is reached in the period of the announcement by $c^{*01}$.

In Figure 6 the green color represents individuals who would have either acquired education and worked as high-skilled before the change (for $a \geq a^*_0$) or would have been low-skilled and unemployed (for $a < a^*_0$), while now they start to work in the low-skilled sector when they are born. On the other hand, the grey color represents individuals who would have either worked as low-skilled before the change (for $a \geq a^*_1$), or would have been low-skilled and unemployed (for $a < a^*_1$), whereas now they find it optimal to acquire education and work as high-skilled. Since $c^*(t)$ does not jump immediately to its new steady state level, there are ability-cost pairs that imply different educational decisions throughout the transition. The lighter colors represent those ability-cost pairs, for which early generations acquire, while later generations opt out from education, as $c^*(t)$ decreases towards its new steady state after the initial overshooting. The light gray area is an initial gain for the high-skilled group, while the lighter green is a later gain for the low-skilled group compared to the initial steady state.

However, most of the initial impact on inequality comes from the inflow of individuals from older generations of unemployed into low-skilled employment. Those whose ability is between $a^*_1$ and $a^*_0$ are low-skilled and unemployed before the change,
but in the period of the announcement they can immediately start working as low-skilled workers. Their entry into the workforce instantaneously increases the supply of low-skilled workers, which is reflected by the jump in $N_l$. This effect outweighs the others on impact; since this change involves individuals from several generations, while the others take effect slowly as new generations are born.

The bottom left panel of Figure 5 shows the overall effect of these changes on the relative supply of skills, $N_h/N_l$: the relative supply of skills decreases on impact. This is largely driven by the mass entry by older generations from unemployment into the low-skilled labor force at the time of the announcement. The effect of this can be seen as $N_l$ jumps up. As time passes the effect of the initial increase in the supply of low-skilled workers is diminished, the relative supply of high-skilled workers starts increasing more, and growth in the supply of low-skilled workers decreases. Figure 5 shows that both supplies increase gradually, and both measures rise above their initial level in the long run.

Figure 7: Average skill premium, relative raw labor supply and wage gaps
The vertical dashed line represents 2006, the year to which I am comparing the results to. The top left panel represents the change in the observed skill premium compared to its initial value, the bottom left panel shows the path of the relative supply of raw high-skilled labor, the right panel shows the wage gaps.

The variables with empirically observable counterparts are the supply of high- and low-skilled raw labor, $L_h$ and $L_l$, the average skill premium, $\bar{w}_h/\bar{w}_l$, and the wage
The relative raw labor supply is shown in the bottom left panel of Figure 7. Its path is very similar to that of the effective labor supply, but the magnitude of change is quite different. This difference in magnitude 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 much larger than the measure of those joining the high-skilled workforce, reflected in the significant 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 of those joining the low-skilled. This is demonstrated by the only slight 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 it decreases relatively more among the low-skilled.

The top left panel in Figure 7 represents the change in the observed skill premium compared to its initial value. 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:

\[
\frac{\bar{w}_h^*(t)}{\bar{w}_l^*(t)} = \frac{w_h^*(t) \bar{a}_h^*(t)}{w_l^*(t) \bar{a}_l^*(t)},
\]

where \(\bar{a}_h^*(t)\) is the average ability among the high-skilled and \(\bar{a}_l^*(t)\) is the average ability among the low-skilled. Therefore the observed 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 observed skill premium increases on impact and then decreases gradually, as does the skill premium per efficiency unit of labor. However, unlike the skill premium per efficiency unit, the average skill premium converges to a value higher than its initial value in the long-run. This is due to compositional effects: since the average ability in the low-skilled labor force decreases more than in the high-skilled labor force, the average skill premium increases relative to its initial value.

\[29\] I only look at cross-sectional inequality measures, rather than lifetime inequality. This is because in this model, since there are no lifetime income dynamics and the employment status of individuals does not fluctuate, there isn’t a significant difference between the two.
Comparing the model’s prediction to the data, it is clear that the contribution of the declining minimum wage to the change in the observed skill premium declined significantly over time (from 70 percent to 15 percent). This is due to the fact that in the data between 1981 and 2006 the average skill premium increased continuously, (see Figure 1), while in the model, the observed skill premium only increases initially, which is followed by a continuous decline. The fact that the sharpest increase in the skill premium occurred in the 1980s, when the model predicts an increase in the observed skill premium suggests that the sharp increase was partly due to the drop in minimum wages.\textsuperscript{30}

The widening wage inequality is well captured by the increasing gap between the wages of workers in the 90th, 50th and 10th percentile. The right panel in Figure 7 shows the monotonic increase in these measures during the transition. These wage gaps increase due to two factors: changes in the skill premium per efficiency unit, and compositional effects.

Changes in the skill premium only increase inequality in the period of the announcement; from the third period onwards these changes compress the wage distribution (see Figure 5 second row right panel).

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 drops, the threshold abilities for unemployment 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. The second 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.

All three wage gaps increase the most in the period of the announcement, since the skill premium and the compositional effects both put an upward pressure on them in

\textsuperscript{30}Note also that the raw relative skill supply has been increasing continuously in the US throughout this period. However, the growth rate of the relative supply was significantly reduced in the 1980s. In the model, this is the period, where the raw relative skill supply drops. These observations could be suggestive that there is another reason for which the skill supply is constantly increasing together with the skill premium. Perhaps in 1981 the economy was not in the steady state yet.

29
this period. After the first period, the wage gaps widen further, but at a slower rate. 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.

The path of the wage gaps in the model and in the data are very similar: all increase continuously, and all increase the most steeply during the 1980s. From the 1990s both the 50/10 wage differential flattened out, while the 90/50 gap kept expanding but at a slower rate. The models predictions are in line with this: the contribution of the drop in the minimum wage is stable for the 50/10 wage gap at around 45 percent, while it is slowly declining for the 90/50 wage gap, on average at around 10 percent. The contribution of falling minimum wages to the increase in the 50/10 wage gap is larger, which is quite natural, as it is this group which it affects directly. Nonetheless, the fall in minimum wages also explains a significant fraction of the increase in the 90/50 wage differential.

5.1 Relation to empirical studies

The two most important channels in the model are the permanent change in employment status and the change in equilibrium educational decisions as a reaction to the lower minimum wage. Both of these channels have been extensively studied in the empirical labor literature.

The first channel in the model relies on the negative employment effect of minimum wages. Most of the literature up to the late 1980s supported the conventional view that minimum wages reduce employment among teenagers (Brown et al. (1982)). In the early 1990s several papers relying on natural experiments challenged this view. They found that minimum wages have no, or small positive effects on employment (Wellington (1991), Card (1992a), (1992b), Katz and Krueger (1992), Card and Krueger
(1994), Machin and Manning (1994), Card and Krueger (1995)). Other papers confirmed the conventional view that minimum wages reduce employment (Neumark and Wascher (1992), Deere et al. (1995), Currie and Fallick (1996), Baker et al. (1999)). However, these latter studies are controversial, as shown by the exchange in Card et al. (1994), Card and Krueger (2000) and Neumark and Wascher (1994), (2000). It seems that the difference in the results of these studies stem from the differences of the statistical models used. Most importantly 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. The studies that find negative employment effects also incorporate lagged minimum wages, thus look at minimum wage effects in a longer time horizon. This suggests, as Baker et al. (1999) and Neumark and Wascher (1994) point out, that employment responds to long run, permanent changes in the minimum wage, as is true in the model presented here. Even though the general empirical results suggest that the minimum wage has longterm effects, none of the studies considered incorporating general equilibrium effects in their specifications.

The second channel works through the changing educational decisions of individuals as a response to a decrease in the minimum wage. In the model, the response of individuals to a decrease in the minimum wage depends on their ability level. On the one hand, the range of abilities for which individuals acquire education in order to avoid unemployment, and meet the productivity requirement given by the minimum wage shifts down. On the other hand, as the skill premium increases, a larger fraction of the population with high enough ability acquires education.

The model predicts that the overall effect in the population is increased education attainment. Nevertheless, within the employed population education attainment decreases. These predictions provide a reasonable explanation regarding the mixed empirical findings on the effects of minimum wages on aggregate educational attainment.\footnote{Neumark and Wascher (1995b), (1995a), (1996), (2003), Chaplin et al. (2003), Neumark and Nizalova (2007) find negative effects, Mattila (1996) and Sutch (2010) find positive effects, while Warren and Hamrock (2012) find no effects of minimum wages on high school enrollment/completion rates.}

Another prediction of the model is that the effect of minimum wages on educational attainment is heterogeneous along the ability dimension. As ability is unobservable,
this prediction cannot be easily tested. However, there are several empirical papers
that condition on various observable individual characteristics, which which are likely
be correlated with unobservable ability. These papers find that the effect of mini-
mum wages is heterogeneous depending on the individual characteristics, such as
ethnicity, race and family characteristics (Cunningham (1996), Ehrenberg and Marcus
(1982), Turner and Demiralp (2001), Chaplin et al. (2003)). The main mechanisms of
the model are echoed by the empirical studies. Higher minimum wages have a dis-
placement effect, leading to an increase in the number of not enrolled/not employed
teenagers (Mattila (1996), Cunningham (1996), Turner and Demiralp (2001)). As min-
imum wages increase employers substitute towards higher skilled individuals, who
might drop out of school to be employed full-time (Neumark and Wascher (1995b),
(1995a), (1996), (2003))), while other individuals need to acquire more education to be
productive enough (Sutch (2010)).

6 Decomposition

I consider three simplified versions of the model, in order to better understand the
contributions of changing technology and education to the effects of minimum wages
on the patterns of wage inequality. I shut down the adjustment first of both technology
and educational attainment, then of educational attainment and finally of technologi-
cal adjustment.

First consider a model, where the production side is as before, but technology and
education are fixed. In such an economy lowering the minimum wage affects the
wage distribution only through an expansion of low-skilled employment. A lower
minimum wage allows people who have been previously unemployed, and are hence
low-skilled, to enter the low-skilled labor market. With constant technology, this de-
creases the wage per unit of efficiency for the low-skilled, thereby increasing the skill
premium. However, since education is fixed, this does not translate into an increase
in the supply of high-skilled labor. The average ability in the low-skilled sector de-
creases, hence the observed skill premium increases more than the skill premium per
efficiency unit. In this setup there are no transitional dynamics, as low-skilled em-
ployment expands in the period of the announcement, the skill-premium responds, and there are no further adjustments.

Second, consider an economy where education is fixed, but technology changes endogenously. In such a setup, a lower minimum wage increases the supply of low-skilled labor, thus increasing the skill premium initially. This does not lead to an increase in the supply of skills, as educational choices are fixed. Transition takes time, as technology needs to adapt to the new relative labor supplies. In the long-run, technology becomes less skill-biased and the skill premium per efficiency unit falls below its original value.

Finally, consider a model where educational choices are made optimally, but technology is fixed. The key difference is that since growth is exogenous, there is no feedback from the effective labor supplies to the direction and rate of technological improvements. Therefore, the relative supply of skills only affects the skill premium through the price effect, as the market size effect is shut down. Hence, in this setup, the skill premium per efficiency unit is always decreasing in the relative supply of skills. As a reaction to the lower minimum wage, first the supply of low-skilled effective labor increases, which increases the skill premium, leading to an increase in the supply of high-skilled labor. The relative supply of skills decreases, leading to an increase in the skill premium per efficiency unit, unlike in the full model. The transition takes a long time, just as in the full model, since complete educational adjustment takes several generations.

Figure 8 compares the transitional dynamics in the four cases. The graphs support the claim that most of the initial effects are due to the inflow from unemployment into the low-skilled labor market. This inflow of low-skilled workers is somewhat offset by the skill acquisition of low-ability previously unemployed individuals in the endogenous education models.

In line with this, the observed skill premium increases the most in the case of exogenous technology and exogenous education (see top left panel in Figure 8). With endogenous technology, the initial impact is the same, but is diminished in the long-run as technologies become less skill-biased. When education is endogenous, but technology is not, then the observed skill premium changes less throughout the transition.
then in the full model. This is due to the fact, that the rate of technological change stays the same, while in the full model the growth rate increases, thus prompting more high ability individuals to acquire education. Thus in the case of endogenous education and exogenous technology the average ability of the low-skilled decreases more compared to the average ability of the high-skilled.

The remaining panels in Figure 8 show the patterns of widening wage gaps. In all three graphs, the biggest initial impact is in the case of exogenous education, implying that most of the initial increase is due to the inflow of previously unemployed workers into the low-skilled labor market. However, if education is exogenous, then there are no further changes in the wage gaps. On the other hand, if education is endogenous, then the expansion of the wage gaps continue throughout the transition. In the long-run, the wage gaps increase the most in the case of endogenous education, suggesting that compositional effects play a significant role in the widening dispersion of wages, and that the change in technology does not have a quantitatively big impact on overall

Figure 8: The role of education and technology in inequality
The vertical dashed line represents 2006, the final year, to which I am comparing the results.
inequality.

7 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 analyse the effects of a permanent decrease 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.

The full effects of minimum wage reductions are only realized in the long run.
Minimum wages affect the educational decisions of individuals in successive cohorts. New cohorts have to replace old ones for the new equilibrium to be reached. Through considering three simplified models, I show that the initial and highest increase in all measures of inequality is due to the inflow from unemployment in the period of the announcement. After this period, the observed skill premium contracts, while the widening of the wage distribution continues due to compositional changes in both ability and skills.
References


38


Appendix - For Online Publication

Labor Supply

Proof of Lemma 1

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

Now consider the case where \( 1(a \geq a_h(t)) = 0 \). 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:
\[ z_{j,s}^*(t) = z_{j,s}^*(q) = z_{j,s}^*(q) \text{ for all } t, T \geq 0. \]

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: \( z_{j,s}^{i,s*}(q) = \gamma^T z_{j,s}^{i,s*}(q) \). 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 \( z_{j,s}^{i,s*}(q) \) by \( z(t) \). Conditional on quality \( q \), the per period profit is constant, \( \pi(q) \), 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:

\[
V_{t+T}(q) = q\pi_s^T \sum_{\tau=0}^{\infty} \frac{e^{-\eta z(t)\gamma^T z_{t+1}^T}}{(1 + r)^\tau}.
\]

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

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

This has to hold for all \( T > 0 \), implying that

\[
\sum_{k=0}^{\infty} \frac{e^{-\eta z(t)\gamma^T k_{t+1}^T}}{(1 + r)^k} (1 - e^{-\eta z(t)}) = \frac{1}{\gamma} \sum_{k=0}^{\infty} \frac{e^{-\eta z(t)\gamma^T k_{t+1}^T}}{(1 + r)^k} (1 - e^{-\eta z(t)\gamma^T}).
\]

To simplify notation denote \( a_k \equiv \gamma^k \frac{1}{\gamma^k - 1} \) 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:

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

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 tines yields:

\[
\lim_{T \to \infty} \left( \frac{b\gamma T e^{-b\gamma T}}{1 - e^{-b\gamma T}} - \frac{b\gamma T \sum_{k=0}^{\infty} \frac{a_k e^{-b\gamma T} a_k}{(1 + r)^k}}{\sum_{k=0}^{\infty} \frac{e^{-b\gamma T} a_k}{(1 + r)^k}} \right) = 0
\]

\[
- \lim_{T \to \infty} \frac{b\gamma T \sum_{k=0}^{\infty} \frac{a_k e^{-b\gamma T} a_k}{(1 + r)^k}}{\sum_{k=0}^{\infty} \frac{e^{-b\gamma T} a_k}{(1 + r)^k}} < 0
\]

Where the second term is non-negative, implying a negative value as \( T \) grows very large. Hence, for \( \gamma > 1 \) (24) 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 \( x > 0 \):

\[
\frac{xe^{-x}}{1 - e^{-x}} < 1 \iff e^{-x}(1 + x) < 1
\]

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 (24) does not hold for any \( T > 0 \).

Therefore in the steady state \( z_{j,s}^j \) 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:

\[
V_{j,s}^j(t)(q) = \frac{q\beta(1 - \beta)^{1-\beta} (p_s^*)^{\beta} N_s^*}{1 - e^{-\eta z_{j,s}^j(q)}(1 + r - e^{-\eta z_{j,s}^j(q)})^{1+r}}.
\]

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: \( z_{j,s}^j(t) = z_{k,s}^k(t) = z_{s}^s \) 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:

\[
\beta(1 - \beta)^{1-\beta} (p_s^*)^{\beta} N_s^* = Bz_{j,s}^j(q)(1 + r - e^{-\eta z_{j,s}^j(q)})^{1+r}.
\]

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.

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 \( z_h^* = z_l^* = z^* \) and the growth rate is \( g^* = 1 + (\bar{q} - 1)(1 - e^{-\eta z^*}) \).

The price of the intermediates can be expressed from substituting the steady state relative price (16) into the intermediate good prices (12):

\[
p_l^* = \left(1 + \gamma \left(\frac{N_h^*}{N_l^*} \frac{\beta \rho}{1 - \rho}\right)^{\frac{1 - \rho}{\rho}}\right)
\]

\[
p_h^* = \left(\frac{N_h^*}{N_l^*} \frac{\beta \rho}{1 - \rho} + \gamma\right)^{\frac{1 - \rho}{\rho}}
\]

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

\[
B z^* \frac{(1 + r - e^{-\eta z^*})}{1 - e^{-\eta z^*}} = \beta (1 - \beta)^{\frac{1 - \beta}{\rho}} \left(\gamma \frac{N_h^*}{N_l^*} \frac{\beta \rho}{1 - \rho} + \frac{N_l^*}{N_l^*} \frac{\beta \rho}{1 - \rho}\right)^{\frac{1 - \rho}{\rho}}
\]

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**

**Data**

I use the May and Outgoing Rotation Group supplements of the Current Population Survey for 1981. I choose 1981 as the initial steady state because from 1982 onwards, the minimum wage was not adjusted by inflation, and its real value started declining. The basic processing of the data is as in Autor et al. (2008). I divide the population into high- and low-skilled based on college education: those who attended college are high-skilled, those who did not are low-skilled. I calculate the fraction of unemployed, low-skilled and high-skilled workers using the education and the employment status.
categories. 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. The maximum likelihood yields $\sigma = 0.73$ and $\bar{c} = 0.82$, which corresponds to about four years.

**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.

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

---

32 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.

33 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.
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.\(^{34}\)

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 4. 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 \(a^*_l, a^*_h\) and \(c^*\) the high- and low-skilled effective labor supplies are:

\[
N^*_h = (1 - \lambda) \int_0^{a^*_l} (1 - c) g(c) dc + \lambda \int_{a^*_l}^{a^*_h} a f(a) da + \\
(1 - \lambda) \int_0^{c^*} (1 - c) g(c) dc + \lambda G(c^*) \int_{a^*_l}^{\infty} a f(a) da
\]

\[
N^*_l = (1 - G(c^*)) \int_{a^*_l}^{\infty} a f(a) da
\]

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,\(^{34}\)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.

\[34\]
The cutoff ability of unemployment for the low-skilled is found by matching the fraction of unemployed:

\[
U = \int (0) \underline{a}_h f(a) da \iff \underline{a}_h^* = e^{(\sigma \Phi^{-1}(U) + \mu)} \tag{30}
\]

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

\[
L_l = (1 - G(c^*)) \int_{\underline{a}_l^*}^\infty f(a) da, \tag{31}
\]

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

\[
\underline{a}_l^* = \underline{a}_h^* \frac{w_h}{w_l} = \underline{a}_h^* \frac{\overline{w}_h}{\overline{w}_l} \overline{a}_l = \underline{a}_h \underline{a}_l^* \overline{a}_h
\]

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 (28).

This way for any cost and ability distribution \(\underline{a}_h^*, \underline{a}_l^*, c^*\) 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:

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

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

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

A high-skilled individual has wage $w_h(i)$ if his ability is $a_h(i) = \frac{w_h(i)}{\bar{w}_h} a_h$, 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:

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

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:

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

(32)
Similarly, a low-skilled individual earning wage \( w_l(i) \) must have \( a_l(i) = \frac{w_l(i)}{\bar{w}_l} \bar{a}_l \), 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:

\[
P(w_l(i), l) = P(a = a_l(i)) P(c(i) \geq c^*).
\]

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

\[
\mathcal{L}(w_l(i); \sigma, \tau) = \begin{cases} 0 & \text{if } a_l(i) < a_l^* \text{ or } a_l(i) < a_h(k(i)) \\ f(a_l(i))(1 - G(c^*)) & \text{if } a_l(i) \geq a_l^* \& a_l(i) \geq a_h(k(i)) \end{cases}
\]

(33)

I calculate the likelihood of observing the sample of wage and education pairs using (32) and (33). I maximize the likelihood by choosing parameters \( \sigma \) and \( \tau \).

**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 \( \rho \), 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:

\[
\log \frac{w_h}{w_l} = \alpha(1) + \alpha_2 \log \frac{H}{L}.
\]

(34)

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 \( \hat{\alpha}_2 \) 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 \( \hat{\alpha}_2 \) is different depending on the assumptions. To
see this note that the skill premium per efficiency unit can be expressed as

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

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 \(\hat{\alpha}_2 = \beta\rho/(1 - \rho) - 1\), while along the transition it is \(\hat{\alpha}_2 = -(1 - \rho)/(1 - (1 - \beta)\rho)\). However, the estimate of \(\hat{\alpha}_2\) 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 in Katz 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 \(\hat{\alpha}_2\) for \(\rho\) is not sufficient to use these estimates in my calibration.

**Transitional Dynamics**

To use the Schmitt-Grohe and Uribe algorithm, all equations have to be defined in terms of variables that are stationary in the steady state. Let \(v^s(t)\) denote the normalized value of owning the leading vintage in sector \(s\) at time \(t\):

\[
v_h(t) = \frac{V_h(t)(q)}{q} \quad v_l(t) = \frac{V_l(t)(q)}{q}
\]

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:

\[
v^*(t+1) = B\frac{(1+r)\bar{v}^*}{1-e^{-\gamma\tau^*(t)}} \quad s = l, h
\]

\[
v^*(t) = \beta(1-\beta)\frac{\bar{p}^*(t)}{1+\rho} N^*(t) - \frac{e^{\gamma\tau^*(t)}}{1+r} v^*(t+1) \quad s = l, h
\]

\[
g^*(t+1) = 1 + (\bar{q} - 1)(1 - e^{-\gamma\tau^*(t)}) \quad s = l, h
\]

\[
p_h(t) = \left( \gamma + \gamma \frac{\beta}{(1-\beta)(1-r)} \right) \left( \frac{Q(t)}{N_h(t)} \right)^{\frac{1-\beta}{\rho}}
\]

\[
p_l(t) = \left( 1 + \gamma \frac{\beta}{(1-\beta)(1-r)} \right) \left( \frac{Q(t)}{N_l(t)} \right)^{\frac{1-\beta}{\rho}}
\]

\[
\bar{w} = a_l(t) \beta(1-\beta)\frac{\bar{p}^*(t)}{1+\rho} \left( \frac{p_h(t)}{N_h(t)} \right)^{\frac{1-\beta}{\rho}}
\]

\[
\bar{w} = a_h(t) \beta(1-\beta)\frac{\bar{p}^*(t)}{1+\rho} \left( \frac{p_l(t)}{N_l(t)} \right)^{\frac{1-\beta}{\rho}} Q(t)
\]

\[
Q(t+1) = \frac{q_{l,h}^*(t+1)}{q_{l,h}^*(t+1)} Q(t)
\]

\[
\Delta(t) = c^*(t)(1-\beta)\frac{\bar{p}^*(t)}{1+\rho} \left( \frac{p_h(t)}{N_h(t)} \right)^{\frac{1-\beta}{\rho}} Q(t)
\]

\[
\Delta(t) = 1 - \beta(1-\beta)\frac{\bar{p}^*(t)}{1+\rho} \left( \frac{p_l(t)}{N_l(t)} \right)^{\frac{1-\beta}{\rho}} Q(t) - \left( \frac{p_h(t)}{N_h(t)} \right)^{\frac{1-\beta}{\rho}} Q(t)
\]

\[
N_h(t) = \lambda N_h(t-1)
- \lambda(1-\lambda)\frac{\tau^*}{\tau} \int_{a_h(t-1)} a f(a) da
+ \lambda(1-\lambda) \int_{a_h(t-1)} a f(a) da
- \lambda(1-\lambda) \frac{c^*(t)-c^*(t-1)^2}{\tau} \int_{a_h(t-1)} a f(a) da
+ \lambda(1-\lambda) \frac{c^*(t)-c^*(t-1)^2}{\tau} \int_{a_h(t)} a f(a) da
+ \lambda \frac{\tau^*}{\tau} \int_{a_h(t)} a f(a) da
+ \lambda \frac{\tau^*}{\tau} \int_{a_h(t)} a f(a) da
\]

\[
N_l(t) = \lambda N_l(t-1) + (1-\lambda)\frac{\tau^*}{\tau} \int_{a_l(t)} a f(a) da
\]

**Path of the unemployment thresholds**

It is not visible on the graphs, but the threshold ability for low-skilled unemployment initially stays above its steady state value and gradually falls towards it, while the threshold for high-skilled unemployment drops slightly below, then increases to its new steady state value. From equation (19) it is clear that once the normalized minimum wage does not change, it is only the price of the low-skilled intermediate that affects the path of \( \bar{w}_l \). As the bottom left panel of Figure 5 shows, the change in the
steady state price is very small, which explains the seemingly immediate jump of $\omega_t$ to its new steady state value. The movement of $\omega_h$ can be understood from (21): $\omega^h$ follows $a_tw_l(t)/w_h(t)$, therefore the initial overshooting of the skill premium (second row, right panel in Figure 5) explains the undershooting of $\omega_h$. The thresholds for unemployment do not change much after the initial drop because intermediate prices and the skill premium do not change much either.

**Robustness**

The following graphs show that the results are qualitatively unchanged for different parameter values. The explanatory power of the model is also fairly robust.

**Different $\rho$s**

*Graph 1*

*Graph 2*

**Different $\gamma$s**

*Graph 1*

*Graph 2*
Different $\sigma$s

Decomposition

I denote the initial steady state by a subscript 0 and the new steady state by a subscript 1.

Exogenous education, exogenous technology

Since the total supply of high-skilled effective and raw labor is constant $N_h^*(0) = N_l^*(1) = N_h^*$.

The equations that define the new steady state are:

$$N_l^*(1) = \int_{a_l^*(1)}^{a_h^*(0)} af(a)da + N_l^*(0)$$

Note that this adjustment only takes place if $a_l^*(1) < a_h^*(0)$, that is if the decrease in $\tilde{w}$ is large enough. When the change in the minimum wage is small, then the decline only implies that some people should not get educated, because they would be productive enough even without acquiring skills. However, since education is fixed, this would imply no adjustments in the economy.

$$a_l^*(1)(p_l^*(1))^{\frac{1}{2}} = \tilde{w}(1)$$
\[ p_t^*(1) = \left( 1 + \gamma^{1-\frac{\beta}{1-\rho+\beta \rho}} \left( \frac{N^*_h}{N^*_l(1)} \right)^{\frac{\beta \rho}{1-\rho+\beta \rho}} Q^{\frac{\beta \rho}{1-\rho+\beta \rho}} \right)^{\frac{1-\rho}{\rho}} \]

where \( Q = Q_h/Q_l \) and \( Q^* = \frac{1}{\beta} \int (0)^1 (q^{*j})^{\frac{1}{2}} (\chi^{*j})^{\frac{1-\beta}{\beta}} d\chi \). I do not explicitly model the pricing of the machines, I denote the price of a machine with quality \( q^* \) in line \( j \) by \( \chi^{*j} \). The assumption that technology is exogenous boils down to having \( Q_h \) and \( Q_l \) growing at the same constant rate. If the pricing of machines would follow monopoly pricing or competitive pricing, then this would be equivalent to a constant growth rate in the quality of each line.

Since education and technology are fixed, the new steady state is reached in the moment of the announcement. The lower bound of unemployment for the low-skilled, which implies the adjustment in the size of the low-skilled labor force. The new skill premium is:

\[ \frac{w^*_h(1)}{w^*_l(1)} = \frac{1}{\beta} \left( \frac{N^*_h}{N^*_l(1)} \right)^{\frac{1-\rho}{1-\rho+\beta \rho}} Q^{\frac{\beta \rho}{1-\rho+\beta \rho}}, \]

which is higher than before.

**Endogenous education, exogenous technology**

The unemployment cutoffs and the threshold for acquiring education are determined exactly as in the full model. The supply of high- and low-skilled workers in the new steady state are as in (28) and (29), while through the transition they are governed by the same equations as in section D of the Appendix. The threshold for low- and high-skilled unemployment are given exactly as in (21) and (20) (again the transition is as in section D of the Appendix, except for \( Q(t) = Q \) here, since technology is exogenous). The cutoff time cost for acquiring education is given by:

\[ c^* = \frac{1 - w^*_h(t)}{w^*_l(t)} \frac{1}{1 - \frac{g \lambda}{1+r}}, \]

where \( g \) is the exogenous growth rate of the economy. The skill premium is given by:

\[ \frac{w^*_h}{w^*_l} = \frac{1}{\beta} \left( \frac{N^*_h}{N^*_l} \right)^{\frac{1-\rho}{1-\rho+\beta \rho}} Q^{\frac{\beta \rho}{1-\rho+\beta \rho}}. \]
The price of intermediates is given by:

\[
p_h^*(t) = \left( \frac{N_h^*(t)}{N_l^*(t)} \right)^{1 - \frac{\beta \rho}{1 - \rho + \beta \rho}} Q^{1 - \frac{\beta \rho}{1 - \rho + \beta \rho}} + \gamma \right) \frac{1 - \rho}{\rho}
\]

\[
p_l^*(t) = \left( 1 + \gamma^{1 - \frac{\beta \rho}{1 - \rho + \beta \rho}} \left( \frac{N_h^*(t)}{N_l^*(t)} \right)^{\frac{\beta \rho}{1 - \rho + \beta \rho}} Q^{1 - \frac{\beta \rho}{1 - \rho + \beta \rho}} \right) \frac{1 - \rho}{\rho}
\]

I show that the system can be reduced to two thresholds, \(a_l^*\) and \(c^*\). The only thing left to show is that the two curves are both downward sloping, with the curve which gives \(a_l^*\) for different values of \(c\) being flatter. This curve is downward sloping as before: a higher \(c\) implies an increase in the fraction of high skilled and a decrease in the fraction of low-skilled, implying an increase in \(p_l^*(t)\). This from (20) implies a lower \(a_l^*\). The other curve, which defines the optimal \(c^*\) for any value of \(a_l\) is also downward sloping. To see this, consider an increase in \(a_l\), which increases the relative supply of skills, as \(a_l\) shifts up, the population between \(a_h\) and \(a_l\) get a bigger weight in the relative supply of skills. An increase in the relative supply decreases the skill premium, which in turn decreases \(c^*\).

The only differences are that the skill premium is always decreasing in the relative supply (see equation above) and the growth rate is exogenous and independent of the relative supply of skills.

This also implies that as in the full model, a reduction in the minimum wage reduces the unemployment threshold in both sectors, and increases the threshold cost of acquiring education.

**Exogenous education, endogenous technology**

The supply of high and low skilled workers evolves the same way as in section E.1 of the Appendix. The main difference is that the intermediate price in the new steady state is given by:

\[
p_l^*(1) = \left( 1 + \gamma \left( \frac{N_h^*}{N_l^*} \right)^{\frac{\beta \rho}{1 - \rho}} \right) \frac{1 - \rho}{\rho}
\]