US-Europe Differences in Technology Adoption and Growth

The Role of Education and Other Policies*

Dirk Krueger       Krishna B. Kumar

March 2003

Abstract

European economic growth has been weak, compared to the US, since the 80s. In previous work (Krueger and Kumar (2003)), we argued that the European focus on specialized, vocational education might have been effective during the 60s and 70s, but resulted in a growth gap relative to the US during the subsequent information age, when new technologies emerged more rapidly. In this paper, we extend this framework to assess the importance of education policy, when compared to labor market rigidity and product market regulation, which have also been suggested as reasons for US-Europe differences.

Households decide between acquiring general education, which allows them to work in high-tech firms, and less costly skill-specific education, which is of value only to low-tech firms that use established production methods. High-tech firms draw a workforce-specific productivity for the new technology, and decide on whether to proceed with production and pay a portion of profits toward regulation costs, or fire the workers at a cost, and redraw a new productivity-workforce combination. Analytical characterization of the balanced growth equilibrium shows that lower firing or regulation cost increases expected growth, but causes the adopting firm to set a higher productivity threshold before it proceeds with production. A higher subsidy for general education can increase expected growth, but reduces the productivity threshold. An increased rate of technology availability can increase the gap in the growth rates of economies that differ in their policies. A “decomposition” exercise using a calibrated version of our model assigns a major role to education policy in explaining US-Europe growth differences.

Keywords: Technology adoption, Education policy, Eurosclerosis.

JEL Classification: O40, O30, I21

*Prepared for the Carnegie-Rochester Conference, April 2003. Krueger: Department of Economics, Stanford University, Stanford, CA 94305 (dkrueger@stanford.edu). Kumar: Department of Finance and Business Economics, USC Marshall School of Business, Los Angeles, CA 90089 (kbkumar@usc.edu). Charlotte Yeo provided excellent research assistance. Krueger acknowledges support from NSF grant SES-0004376 and Kumar from the USC Marshall School’s research fund.
1 Introduction

European economic growth has been weak, compared to the US, since the 80s. This is true for the growth rate of per capita GDP as well as for labor productivity in the manufacturing sector. During this period, Europe has lagged in technology adoption, and a “technology gap” with the US has emerged. In previous work (Krueger and Kumar (2003)), we developed a model of education and technology adoption to argue that the European focus on specialized, vocational education might have worked well during the 60s and 70s, but not during the subsequent information age when new technologies emerged at a more rapid pace.

In that model, two-period lived agents acquire skill-specific (vocational) or general education in the first period. Acquiring general education causes disutility, which is decreasing in ability, but allows the agent to work for a higher wage in firms that adopt newer technologies. Technology adoption by firms is costly and is limited by the exogenously available frontier. Higher profits accrue to adopting firms, but only in the current period, as the latest production practice spills over to the entire economy in the next. Higher subsidies to general education can result in higher growth, because a higher fraction of generally-educated workers, which lowers their relative wages, makes it conducive for high-tech firms to adopt new technologies at higher rates. More importantly, when the rate of available technologies increases, two countries with different general education subsidies that grew at the same, maximal rate earlier can now diverge. Firms in the country with the lower subsidy no longer find it optimal to adopt technologies at the new maximal rate, given the higher wages that prevail for labor with general education.

How does this channel compare with other explanations suggested for the same phenomenon? How important is each channel quantitatively? To answer these questions, in the present model we extend the high-tech firm’s optimization problem to incorporate labor and product market regulations. First, an adopting firm has to pay a fraction of its profits to the government; this is intended to capture the costs of bureaucracy and product market regulations. Second, the adopting firm’s problem is given a dynamic dimension within each model period. There are several ways to implement new technologies, and the firm has to experiment until it finds a suitable match between the technology and the task-specific productivity of its workers. The firm can either retain its current productivity draw for production, or fire its workers and search for a new productivity. Firing labor is costly; this cost serves as a proxy for labor market distortions. In this way, we account for the three facets of policy we want to study in explaining US-Europe growth differences - education subsidy, product market regulation, and labor market rigidity.

The model is simple enough to allow analytical characterization of the impact of these policies on the balanced growth equilibrium. A higher firing cost increases the cost of re-drawing a new productivity match, causing firms to accept inferior productivity draws; this lowers wages in the adopting sector, and therefore decreases the incentive to acquire general
education. The expected growth rate decreases. A higher regulation cost has a similar effect, but the low productivity threshold results from lower benefit of redrawing a new productivity rather than higher cost. A higher education subsidy increases general education attainment by lowering its relative costs, lowers relative wages for workers with general education and thus makes the firm more willing to accept lower productivity matches. An increased rate of technology adoption can increase the gap in growth rates of economies that focus on different types of education. The wage rate in the adopting sector that is consistent with technology adoption at the maximal available rate decreases (that is, the labor requirement to support maximal growth increases), which puts an economy geared towards skill-specific education at a disadvantage.

We calibrate our stylized model to match empirical policy parameters in the US and Europe – Germany and Italy in particular – and investigate counterfactual policies for Europe that would bridge its growth gap with the US. While a decrease in the firing costs and, to a lesser degree, regulation costs will help reduce this gap, the difference in education policies explains the majority of the gap.

Lawrence and Schultze (1987) is an early collection of papers that propose explanations for US-Europe growth differences. While the European regulatory environment and labor market frictions receive extensive attention here and in subsequent research, there has been little academic and policy debate on another reason suggested in that volume: “Workers must have general training to adapt to new tasks, and European education, which has encouraged apprenticeships that provide specific skills, must adapt.” Our work, which focuses on this neglected aspect, is related to the following strands of literature.

Depreciation of skills on account of job losses or technological change has received recent attention in Ljungqvist and Sargent (1998), Gould, Moav, and Weinberg (2001), and Violante (2002). The diminished role of skill-specific education during times of rapid technological progress is also a feature of Galor and Tsiddon (1997) and Galor and Moav (2000). Unlike these papers, we jointly model the technology adoption and education acquisition decision to focus on the effect of education policy on economic performance. Wasmer (2003) argues that differences in labor mobility and labor market performance between the US and Europe can be attributed to different investment strategies of workers; European workers invest more in specific human capital and US workers in general human capital.

There is extensive use of firing costs to capture labor market rigidities in matching models in the macro-labor literature, with the focus typically on unemployment. In the context

---

1 These two countries have education systems that are most different from the US, and are thus ideal candidates for a quantitative investigation of our hypothesis.

2 See Krueger and Kumar (2003) for a more detailed discussion of this literature and how it relates to our theoretical model.

3 For instance, Hopenhayn and Rogerson (1993) argue that firing costs reduce employment by decreasing hiring as well as firm entry.
of US versus Europe, the recent study by Pries and Rogerson (2001) argues that higher dismissal costs in Europe increase the duration of unemployment through tougher hiring standards rather than through the usual channel of vacancy creation. More directly relevant is Nagypál (2002), who argues that high dismissal costs in Europe hinder learning about match quality ("experimentation" in the economy), and reduce average productivity by a quantitatively significant amount. Our modeling of product market regulations as a cost is motivated by Nicoletti, Scarpetta, and Boylaud (1999), who present a detailed taxonomy of such regulations, and identify barriers to entrepreneurship (administrative burdens and regulatory and administrative opacity) as an important component.

We model labor and product market rigidities in a fairly reduced-form fashion that does not do justice to the nuances of policy detailed in the above studies. Our focus is not the labor market or output market per se, but the effect of their regulation on technology adoption and economic growth. Consequently these policies are modeled to directly impact firms in the "high tech" sector, and then propagate through the rest of the economy via their general equilibrium effects on relative wages. Given this caveat and the slant toward education policy, an appropriate interpretation of our quantitative results is not that education policies matter exclusively, but rather that the quantitative significance of education policies is robust to the inclusion of concomitant explanations.

The rest of the paper proceeds as follows. In Section 2, we present a brief discussion of the stylized facts that motivate our study; detailed discussions are deferred to Section 5. The economic environment is presented in Section 3, and the Balanced Growth Path (BGP) is characterized in Section 4. Section 5 discusses the details of our calibration strategy and empirical targets, and the quantitative results are presented in Section 6. Section 7 concludes. Analytical derivations are relegated to the appendix.

2 Stylized Facts

The three stylized facts that motivate us are: i) a growth gap between the US and Europe since the 80s, ii) a gap in the use of Information and Communication Technology (ICT) between the US and Europe, and iii) a stronger European focus on vocational education. These facts are surveyed in detail in Krueger and Kumar (2003) and in Section 5; we present only a brief synopsis here.

Per capita growth in Europe, which was higher than in the US in the 70s, was lower in the 80s and further dropped in the 90s. More relevant for our study is the data on labor productivity growth in the manufacturing sector taken from the European Competitiveness Report 2001, and presented in Table 1A.

Scarpetta et. al. (2000) report that the manufacturing productivity growth rates for Germany and Italy was equal to or higher than those of the US during the earlier period of 1970-79. They also argue that the manufacturing sector has played a more important
role than services in terms of aggregate productivity growth. Table 1A shows that US labor productivity has outpaced that of Germany from as early as the mid-80s. While Italy did better than the US in the initial period, during the latter period its labor productivity growth was only half of US productivity growth. The EU average is 3.0% for the early period and 3.1% for the later period. The productivity growth gap is even more pronounced in technology-driven industries such as office machineries and computers, pharmaceuticals and aircraft; 8.3% for the US in the 90s and 3.5% for the European Union. Scarpetta et. al. (2000) capture the increasing gap between the US and Europe by looking at manufacturing productivity levels with the US level normalized to 100; this is shown in Table 1B. While the productivity level for Germany, as in several other European countries, was converging toward the US level until 1980, the gap has widened since then.

Table 1A

Manufacturing Labor Productivity Growth

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>US</td>
<td>2.3%</td>
<td>4.3%</td>
</tr>
<tr>
<td>Germany</td>
<td>2.0%</td>
<td>3.8%</td>
</tr>
<tr>
<td>Italy</td>
<td>3.8%</td>
<td>2.2%</td>
</tr>
</tbody>
</table>

Table 1B

Manufacturing Productivity Level: GDP Per Person

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>US</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>Germany</td>
<td>33.6</td>
<td>63.0</td>
<td>79.0</td>
<td>87.1</td>
<td>73.1</td>
<td>68.2</td>
</tr>
</tbody>
</table>

There is also evidence that, with the exception of Sweden, Finland, and the Netherlands, Europe lags behind the US in technology usage. The gap between the US and Europe, whether measured by the share of ICT equipment investment as a share of total investment or by the contribution of ICT capital to output growth has increased since the 80s. The contribution of ICT capital to output growth has been increasing for all countries, but the gap between the US and European countries has also been increasing. Schreyer (2000) reports that during 1990-1996, 0.42 percentage points of per-capita output growth in the US can be attributed to ITC capital-deepening, up from 0.28 points in the second half of the 80s. In Germany, these numbers increased from 0.12 to only 0.18.\(^4\)

Finally, the European focus on skill-specific, vocational education is readily evident from statistics reported in OECD (1997, 2001). More than 70% of students at the upper secondary

\(^4\)Italy experienced a similar trend. Stiroh (2002) shows that IT-intensive industries experienced significantly larger labor productivity gains than other industries; he also finds a strong correlation between IT capital accumulation and labor productivity.
level in Germany and Italy are enrolled in vocational or apprenticeship programs. Vocational education in the US is typically imparted in two-year community colleges; in 1994 only about 10% of enrollees were working toward a vocational degree. The net entry rate into universities, where general education is primarily imparted, is more than twice the rates of most European countries. Higher enrollment has led to higher university attainment in the US labor force. Finally, while the percentage of GDP devoted to primary and secondary education was about the same for the US and Germany in 1997, the percentage devoted to tertiary education was significantly higher in the US. Consequently, the expenditure per student relative to GDP per capita on tertiary education was higher in the US than in Germany, whereas for post-secondary non-tertiary (vocational) education the situation is reversed. We conclude that the US has a stronger focus on general education and Europe on skill-specific education.

These stylized facts motivate the construction of our model featuring endogenous choice of households between general and vocational education and technology adoption by firms, in order to study US-Europe growth and technology gaps. The details of this model are presented next.

3 The Environment

The economy is populated by a continuum of households and two continua of identical firms. Firms in one sector, the adopting or high-tech sector, potentially adopt new technologies in every period, and firms in the nonadopting or low-tech sector do not. There is a single nonstorable consumption good in each time period and households supply labor to the firms. In this section we describe the maximization problems of a typical firm in each sector and of a typical household, and finally define equilibrium and a balanced growth path. We will cast our environment in recursive form from the beginning, therefore omitting the sequential representation.

3.1 Firms and Technology Adoption

All firms in our economy are owned by infinitely lived entrepreneurs. Profits are immediately consumed since entrepreneurs, like workers, do not have access to an intertemporal storage technology or long-lived assets.

At each point of time there is a single final consumption good in the economy, which serves

---

5 We abstract from important industrial organization issues such as firm entry and exit decisions to keep our model analytically and computationally tractable. This allows us to focus on the education, labor market and product market channels as sources of growth differences between the US and Europe.

6 Krueger and Kumar (2003) provide a sequential representation for an economy that has similar features to the ones present here.
as the numeraire. The aggregate technology that is freely available for usage in production by all firms in the current period is denoted by $A$. It is the result of past technology adoption decisions (which are described below) and taken as given by all firms in the current period. A firm in the sector that does not adopt new technologies in the current period produces output according to

$$Y_n = A\beta (H_n)^\theta,$$

where $Y_n$ is output of a typical nonadopting firm and $H_n$ are the effective units of labor employed; $\beta > 0$ is a parameter that governs the relative productivity of the low-tech and the high-tech sector, and $\theta < \frac{1}{2}$ is an intensity parameter. As will be described below, each worker in the low-tech sector has productivity $h > 1$. Therefore, if the typical nonadopting firm hires $n_n$ workers from the competitive labor market, the above production function can be written as $Y_n = A\beta (n_n h)^\theta$. Every nonadopting firm takes the wage $W_n$ per efficiency unit in the sector as given and maximizes profits.

Firms in the technology-adopting (“high-tech”) sector can choose the level of technology $a'$ with which to produce today, subject to the constraint

$$a' \leq A_f,$$

where $A_f$ is the productivity level of the frontier technology. This frontier technology is assumed to grow at a constant exogenous rate $\lambda$; next period’s frontier technology is given by $A'_f = \lambda A_f$. An adopting firm incurs a technology adoption cost $\frac{1}{2}(a' - A)^2$ if it decides to use technology $a'$ in the current period and the freely available “common practice” is $A$. It is proportional to the complexity of the common practice $A$ and strictly convex in the distance of the technological “leap”.7

The output of a typical firm in the adopting sector $Y_a$ is given by

$$Y_a = a' (H_a)^\theta,$$

where $H_a$ are the effective units of labor employed. As will be described below, $E_h$ is the stochastic productivity per worker in this firm. Therefore, if the typical adopting firm hires $n_a$ workers, the above production function can be written as $Y_a = a' (n_a E_h)^\theta$. Adopting firms take wages $W_a$ per efficiency unit as given.

### 3.2 Labor and Product Market Regulations

As mentioned in the introduction, we model product market regulations and labor market rigidities in a highly stylized way. We discuss these model features while describing the sequence of events within a given period for adopting firms.

---

7 If $A$ is viewed as the number of machines, the cost of retooling each machine will be constant on a balanced growth path.
1. A representative adopting firm hires $n_a$ number of workers.

2. The firm draws a worker productivity level $E_h \sim F(.)$, with associated continuous probability density function $f$. We assume that $E_h$ is perfectly correlated across workers and that draws are independent over time. The productivity draw can be viewed as specific to an adopting firm and the workforce it hires; it captures how effectively the firm can produce the final good if it proceeds with the current implementation of the new technology.

3. Given the above productivity draw for its workforce, the firm decides whether to proceed with the implementation of a new technology and produce the final good accordingly, using the current workforce. If the firm does not implement the technology, it pays a firing cost $A * f n_a E_h$, instantly hires new workers $n_a$ and draws again. Here, $f$ is the resource cost of firing a unit of effective labor; these resources are lost and don’t benefit anyone.\(^8\) If the firm decides to implement the technology after this “sampling” phase, it chooses the extent of adoption, $a'$, optimally.\(^9\) In order to enter the output (“product”) market to sell its product, the firm has to pay a fraction $C$ of its profits from production. This can be viewed as the resource cost of administrative burden and regulatory compliance arising from government policy; again these resources are lost. The firm then pays workers wages $W_a$. Output net of wages, regulation, firing and adoption costs is consumed by the adopting entrepreneur as profits.

Several elements of our formulation deserve further discussion. First, the interpretation of the third stage is that there are several ways of implementing a new technology, and some work well given the labor force and some do not. If a draw is viewed as building a prototype, or conducting a trial run of production, firing workers is akin to shutting the plant down when the trial run fails, and attempting to re-implement the technology in a different way, hoping for a better match productivity between the technology and the workers. Since the labor productivity $E_h$ is workforce-specific, firing workers is necessary in order to sample the aggregate labor market for a different set of workers.\(^{10}\)

\(^8\)Qualitatively, little hinges on whether the firing costs are paid per effective unit or raw unit of labor; indeed in the calibration it is convenient to consider this as the cost per worker. Our assumption that these costs are lost resources that do not benefit anyone is in line with the “real resource costs which include the costs associated with following whatever procedures are necessary in order to dismiss a worker,” assumed by Pries and Rogerson (2001).

\(^9\)For instance, a firm decides first on whether it should computerize and what computers work best, before it decides on how much of its operations will be computerized.

\(^{10}\)We are implicitly assuming that the output produced when the workers are fired is zero; therefore, a failed attempt need not only be viewed as a technology match that did not produce results, but also as a production attempt that resulted in very low output.
Second, given that $E_h$-draws are independent over time, no firm has an incentive to revise its choice $n_a$ after an unsuccessful draw. Third, we assume that new draws can occur instantly, which allows us to abstract from discounting within a model period and does not introduce other issues of timing within a period. Fourth, the cost parameters $f$ and $C$ are intended to capture labor market regulations and product market frictions, respectively. These costs are proportional to the current technology; this assumption yields the empirically plausible implication that equilibrium costs are a constant fraction of GDP in a growing economy.

### 3.3 The Firms’ Problems

Given the common practice $A$, the static problem of a typical nonadopting firm reads as

$$\max_{n_n} A \beta (n_n h)^\theta - W_n n_n h,$$

with necessary and sufficient first order condition

$$W_n = \theta A \beta (n_n h)^{\theta - 1}. \quad (1)$$

Adopting firms have three choices to make in any given period, namely how many workers $n_a$ to hire, and then, conditional on a productivity draw, whether to go ahead with production or whether to redraw, and finally, conditional on producing, the extent of adoption, $a'$.

Let’s first consider the third stage of the problem. Conditional on hiring $n_a$ workers and a productivity draw, $E_h$, the firm solves

$$\Pi(E_h, n_a) = \max_{a' \leq A_f} a' (n_a E_h)^\theta - W_a n_a E_h - \frac{A}{2} \left( \frac{a'}{A} - 1 \right)^2. \quad (2)$$

In our framework, the workers are relatively passive and the firm does all the experimenting. There is no bargaining and division of surplus as every worker, by assumption, will be paid her marginal product in a competitive labor market. While the extensive literature on individual worker-firm matching has produced several useful insights, modeling such micro-mechanisms would lead us too far astray from the main purpose of our paper of analyzing labor market friction as a concomitant cause for US-Europe growth effects. Therefore, it seems least ad-hoc to assume that the firm pays workers their marginal product as wages. We can envision competitive employment agencies that hire the adoption labor force and place them with an adoption firm, with the common understanding that workers would be placed in a different firm in case they are not paid competitively. However, this agency can guarantee wages only in expectation, since the productivity $E_h$ is random and specific to a firm. We, therefore, have

$$W_a(E_h, n_a) = \theta a' (n_a E_h)^{\theta - 1}. \quad (3)$$

\[11\] However, this will abstract from the cost of waiting for a better productivity draw.
When choosing the level of technology, $a'$, the firm takes these wages as given and beyond its control. Below we discuss how technology adoption decisions of firms impact next period’s common practice.

In the second stage, the firm decides whether to go ahead with production, conditional on a productivity draw $E_h$. Let $V(E_h, n_a)$ denote the period value of a firm with draw $E_h$ and number of workers, $n_a$. The function $V$ solves the functional equation

$$V(E_h, n_a) = \max_{\text{adopt, not adopt}} \left\{ (1 - C) \Pi(E_h, n_a), -f n_a E_h * A + \int_1^{E_h} V(E_h', n_a) dF(E_h') \right\}, (4)$$

where the choice is between adopting and producing, or dismissing the workforce and hoping for a more suitable workforce-productivity combination. We use “proceed” and “adopt” interchangeably, as we do “not adopt” and “fire” or “dismiss”. Let $a'(E_h, n_a)$ denote the optimal technology choice of a firm with productivity draw $E_h$ and workers $n_a$, and by $E_h(n_a) \in [1, h]$ the threshold productivity draw above which a firm produces rather than redraws. A unique such threshold will exist since the first part of the Bellman equation is strictly increasing in $E_h$ (as profits are) and the second part is strictly decreasing in $E_h$.

Finally, in the first stage of the firms’ problem each firm chooses

$$n_a \in \arg \max_n \int V(E_h, n) dF(E_h).$$

### 3.4 Households and Education Decisions

There is a measure 1 of households that live for two periods. Households in our economy make only one economic decision, namely which type of education to obtain in the first period of their lives. A household can opt to obtain vocational or general education. There is a utility cost, $e(a)$, of obtaining general education, which depends on an agent’s innate ability $a \in [0, 1]$ for higher education. We assume that $a$ is uniformly distributed across the population; therefore, the cumulative distribution function is $F(a) = a$. We also assume that $e(a)$ is strictly decreasing in $a$.

An agent who acquires vocational education earns a wage $W_{nh}$ in the second period of life. This agent can work in nonadopting firms with the technologies that were adopted last period, and have become established practice this period. This agent’s task-specific productivity will be at its highest possible value, $h$, because the agent has received training for that specific technology.

The benefit that an agent obtains from general education is qualification to work in the adopting sector with newer technologies. While the technology per se will be more

---

12Treating $f, n_a > 0$ as parametric, one can show that this Bellman equation (normalized by $A$) satisfies Blackwell’s sufficient conditions, and is indeed a contraction with modulus $F(E_h(n_a)) < 1$, where $E_h(n_a)$ solves $\int_1^{E_h} E_h dF(E_h) = (1 - C) \int_1^{E_h} \pi(E_h, n_a) dF(E_h)$. In other words, even though there is no discounting, the existence of positive firing costs yields a well-behaved value function. As will be apparent below, $n_a = 0$ cannot be an equilibrium.
productive, the agent’s task-specific productivity in this risky environment, $E_h$, is stochastic and lies anywhere in $[1, h]$. This agent earns the random wage $W_a E_h$. Thus, when making the education decision, households confront the trade-off of a higher cost of obtaining general education, net of any subsidy, and a potentially higher, although stochastic, wage in the second period of life.

We assume that households do not have access to insurance contracts against bad realizations of individual labor productivity and thus consume their wages in the second period of life. Agents have preferences representable by the utility function

$$U(a) = I_g \{E \log (W_a E_h) - \log(e(a)) + \log(S_g)\} + (1 - I_g) \{\log (W_n h) + \log(S_v)\}$$

(5)

where $I_g = 1$ if the household chooses general education and $I_g = 0$ if the household chooses vocational education. The expectation $E$ is taken with respect to the stochastic productivity $E_h$. Here $S_g$ and $S_v$ denote government subsidies for general and vocational education, respectively. Given that all elements apart from $e(a)$ in the utility function are independent of $a$, it follows that there exists a threshold ability level $a^*$ above which households choose to obtain general education and below which they obtain vocational education.

### 3.5 Government Education Policy

We assume that the government spends a total amount of $A \ast G$ on education. This formulation is consistent with education expenditures that are a constant fraction of GDP along a balanced growth path. Given the threshold $a^*$ defined above, the governments’ budget constraint, assumed to hold period by period, is

$$a^* S_v + (1 - a^*) S_g = A \ast G.$$ 

(6)

---

13 If we assume that there are only two skill levels, 1 and $h$, and that $T$ is the probability that a generally educated agent is unsuited for the new technology and ends up with the low productivity, the expected agent productivity can be written as $E_h = T \cdot 1 + (1 - T) \cdot h$. Hence the notation $E_h$. It is easier to deal with $E_h$ directly, instead of the underlying $T$. In Krueger and Kumar (2003) we study conditions under which it is optimal for vocationally educated agents to work only in the nonadopting sector, and generally educated agents to work only in the adopting sector; in this paper, we directly assume it.

14 A breakup of wage profiles by vocationally versus generally educated workers is hard to find. CERGE-EI (1997, VII.2) reports that the age-wage profile for agents with vocational education in Czechoslovakia is very flat, while the profile for those with university education starts around the same wage, but shows a steep growth with age. This lends some credibility to our assumption that vocationally educated agents work in the nonadopting sector and do not experience the steep wage growth one would expect in an environment where technology changes rapidly.

15 Discounting across periods serves little purpose in our setup, and we have abstracted from it.
3.6 Equilibrium

Let $\eta_g$ denote the fraction of young agents that obtained general education in the last period and thus are available for work in the high-tech sector. At any point in time the aggregate state of the economy is described by the vector $z = (\eta_g, A, A_f)$. In order to define a recursive equilibrium we have to describe the law of motion for the state vector. Next periods’ education allocation, $\eta'_g$, is determined by the decision of young households in the current period; that is, $\eta'_g = 1 - a^*$. The frontier technology evolves exogenously according to

$$A'_f = \lambda A_f.$$  \hspace{1cm} (7)

The commonly available technology next period, $A'$, is assumed to be the average level of technology adopted by high-tech firms this period; that is\(^{16}\)

$$A' = \left[1 - F(E_h(n_a))\right]^{-1} \int_{E_h \geq E_h(n_a)} a'(E_h, n_a) dF(E_h).$$  \hspace{1cm} (8)

Individual adopting firms perceive themselves as small and therefore unable to affect tomorrow’s aggregate state $A'$ through their technology adoption decision today. It is the aggregate technology $A'$ rather than the individual technology $a'$ that determines tomorrow’s starting technology for an adopting firm; this makes its adoption problem, conditional on $E_h$, essentially static. Adoption yields a technological advantage only in the current period, via higher current profits, before the competition catches up.

We can view the newly available technology as filled with “bugs”, and the process of adoption as discovering these bugs and working around them in order to realize the potential gains of the technology. In the subsequent period, when these bugs become known and fixed, a more robust technology is available to all. In this paradigm, (8) can be interpreted as the expected number of bugs that are discovered this period and will be fixed by the next period.

We can now define a recursive equilibrium:

**Definition 1** A Recursive Equilibrium is a policy function for workers, $I_g(a; z)$, a policy function for nonadopting firms, $n_a(z)$, a value function $V(E_h, n_a; z)$ and policy functions for the adopting firms, $n_a(z)$, $E_h(n_a, z)$, $a'(E_h, n_a; z)$, wage functions $W_a(z)$ and $W_a(E_h, n_a, z)$, an education cut-off function $a^*(z)$, government policy functions $S_g(z), S_v(z)$ and an aggregate law of motion $\Phi$ mapping today’s aggregate state into tomorrow’s aggregate state such that

1. Given wage functions, policy functions by firms, government policy, and the aggregate law of motion, $I_g(a; z)$ maximizes (5) for all $a \in [0, 1]$. The education cut-off $a^*(z)$

\(^{16}\)This average is a cross-sectional average over the continuum of firms. We are implicitly assuming a law of large numbers, which allows us to use the distribution function for productivity draws of a single firm, $F$, as the population distribution.
satisfies
\[ I_g(a, z) = 1 \text{ for all } a \geq a^*(z) \text{ and } I_g(a, z) = 0 \text{ for all } a \leq a^*(z). \] (9)

2. \( W_n(z) \) and \( n_n(z) \) satisfy (1). \( W_a(E_h, n_a, z) \) and \( n_a(z) \) satisfy (3).

3. The value function \( V(E_h, n_a; z) \) solves Bellman equation (4) and \( \bar{E}_h(n_a; z) \) is the associated optimal productivity cut-off.

4. Given \( W_a(E_h, n_a, z) \), the technology adoption decision solves (2).

5. The government policy satisfies (6), given \( a^*(z) \).

6. The labor market clears: \( \eta_g = n_a(z) \) and \( 1 - \eta_g = n_n(z) \).

7. The aggregate law of motion \( \Phi \) is induced by the optimal policy functions \( I_g(a; z) \) and \( a'(E_h, n_a; z) \).

The last part of the definition can be explicitly stated in the following way: \( A'_f \) is given by (7), \( A' \) is given by (8) and \( \eta'_g = 1 - a^*(z) \), where \( a^*(z) \) is in turn defined by the optimal education acquisition policy \( I_g(a; z) \) via (9).

### 3.7 Balanced Growth Path

In our analysis we will restrict ourselves to balanced growth equilibria. We first normalize all growing variables by dividing by the level of the current technology, \( A \). We let \( x(E_h, n_a) = \frac{a'(E_h, n_a)}{A} \) denote the growth rate chosen by a generic adopting firm and by
\[ E(x) = \left[ 1 - F(\bar{E}_h(n_a)) \right]^{-1} \int_{E_h \geq E_h} x(E_h, n_a) dF(E_h) \] the average growth rate of technology of adopting firms, which by (8) equals the constant aggregate growth rate of the economy along a balanced growth path.

#### 3.7.1 Firms

Dividing the optimality condition for the nonadopting firms by \( A \) yields
\[ w_n = \theta \beta (n_n h)^{\theta - 1}, \] (10)
where \( w_n = \frac{W_n}{A} \) is the normalized wage in the nonadopting sector.

In the adopting sector, conditional on hiring \( n_a \) workers and drawing a productivity shock \( E_h \), the firm solves
\[ \pi(n_a, E_h) = \max_{x \leq \lambda} x (n_a E_h)^{\theta} - w_a n_a E_h - \frac{1}{2} (x - 1)^2, \] (11)
where the de-trended wage \( w_a \), taken as given by the firm, equals the marginal product:
\[ w_a = \theta x (n_a E_h)^{\theta - 1}. \] (12)
The dynamic programming problem of the firm becomes

\[ v(E_h, n_a) = \max_{\text{adopt, not adopt}} \left\{ (1 - C) \pi(E_h, n_a), -fn_aE_h + \int_1^h v(E_h', n_a) dF(E_h') \right\} \]  \tag{13} \]

where \( v(E_h, n_a) = \frac{V(E_h, n_a)}{A} \) is the normalized value function and \( n_a \in \arg \max_n \int v(E_h, n) dF(E_h) \) is the optimal labor demand of the firm.

### 3.7.2 Households

Household preferences can be rewritten as (ignoring constants irrelevant for maximization)

\[ u(a) = I_g \{ E \log (w_a E_h) - \log(e(a)) + \log(s_g) \} + (1 - I_g) \{ \log (w_n h) + \log(s_v) \} \]  \tag{14} \]

where the subsidy levels \( s_v = \frac{S_v}{A} \) and \( s_g = \frac{S_g}{A} \) of the government have to obey its budget constraint

\[ a^* s_v + (1 - a^*) s_g = G \]  \tag{15} \]

### 3.7.3 Equilibrium

**Definition 2** A balanced growth equilibrium is an optimal education decision for workers, \( I_g(a) \), labor demand for nonadopting firms, \( n_n \), a value function \( v(E_h, n_a) \) and policies for the adopting firms, \( n_a, \tilde{E}_h(n_a), x(E_h, n_a) \), wages \( w_n \) and \( w_a(E_h, n_a) \), an education cut-off function \( a^* \), government policies \( s_g, s_v \) and an aggregate growth rate \( E(x) \) and education allocation \( \eta_g \) such that

1. Given wages, \( I_g(a) \) maximizes (14) for all \( a \in [0, 1] \). The education cut-off \( a^* \) satisfies
   \[ I_g(a) = 1 \text{ for all } a \geq a^* \text{ and } I_g(a) = 0 \text{ for all } a \leq a^*. \]  \tag{16} \]
2. Wages satisfy (10) and (12).
3. The value function \( v(E_h, n_a) \) solves Bellman equation (13) and \( \tilde{E}_h(n_a) \) is the associated optimal productivity cut-off.
4. \( x(E_h, n_a) \) solves (11).
5. The government policy satisfies (15), given \( a^* \).
6. The labor market clears: \( \eta_g = n_a = 1 - a^*; 1 - \eta_g = n_n \).
7. The aggregate growth rate is given by
   \[ E(x) = \left[ 1 - F(\tilde{E}_h) \right]^{-1} \int_{E_h \geq \tilde{E}_h(n_a)} x(E_h, \eta_g) dF(E_h). \]

We will use \( s = \frac{s_v}{s_e} \) as our educational policy variable; given \( s \) and \( G \), the individual subsidies \((s_g, s_v)\) can be derived from the definition of \( s \) and equation (15).
4 Analytic Characterization of the Balanced Growth Path

We now analytically characterize the balanced growth path of this economy. Given the motivation of the paper, we are particularly interested in the comparative statics with respect to the education policy parameter, $s$, parameters intended to capture product and labor market frictions, $C, f$, and the speed of technological innovation $\lambda$.

Our strategy is to first solve the different stages of the adopting firm’s problem for a schedule that relates the optimal productivity threshold $\bar{E}_h$, beyond which the firm chooses the adoption option, for a given measure $\eta_g$ of households available for work in the high-tech sector. This can be viewed as a human capital “demand” schedule. We then derive a similar schedule from the households’ education problem – the productivity threshold $\bar{E}_h$, with its implied wage differential, that is necessary to induce a given measure $\eta_g$ of households to work in the high-tech sector. A high $\bar{E}_h$ is indicative of the high wages households expect before they are willing to supply a given $\eta_g$. For this reason, the household condition, denoted by $\bar{E}_{HH}$, can be viewed as a human capital “supply” schedule. Finally we combine these two schedules to solve for the balanced growth path education allocation and productivity threshold. All other BGP equilibrium values follow.

4.1 Firms

We first discuss the solution of the optimal technology adoption problem, conditional on an acceptable $E_h$ draw and then analyze the determination of the optimal threshold $\bar{E}_h$.

4.1.1 The Technology Adoption Decision

Recall that $H_a = n_a E_h$ is the effective labor used for production. The adoption problem in (11) is

$$\max_{x \leq \lambda} x H_a^\theta - w_a H_a - \frac{1}{2} (x - 1)^2,$$

with associated first order condition

$$(H_a)^\theta \begin{cases} 
= x - 1 & \text{if } x < \lambda \text{ (interior growth)} \\
\geq \lambda - 1 & \text{if } x = \lambda \text{ (maximal growth)}. 
\end{cases}$$

The labor market equilibrium implies $n_a = \eta_g$ and therefore $H_a = \eta_g E_h$.

Motivated by the above first-order condition, it is convenient to define, for an arbitrary $\eta_g$, the $E_h$ beyond which $x = \lambda$ (i.e. adoption is maximal):

$$\bar{E}_h (\eta_g) \equiv \min \left\{ \max \left\{ 1, \frac{(\lambda - 1)\frac{1}{\eta_g}}{} \right\}, h \right\}.$$

In general, there exist cutoffs $\eta_g^1$ and $\eta_g^3$, such that, for $\eta_g \in [1, \eta_g^1]$ only interior growth occurs, for $\eta_g \in [\eta_g^1, \eta_g^3]$ either interior (when $E_h < \bar{E}_h$) or maximal growth (when $E_h \geq \bar{E}_h$)
is possible, and for $\eta_g \in [\eta_g^3, 1]$ only maximal growth occurs. These cutoffs are defined by:

$$\eta_g^1 \equiv \frac{(\lambda - 1)^{\frac{1}{\theta}}}{h} < (\lambda - 1)^{\frac{1}{\theta}} \equiv \eta_g^3.$$  

(17)

Figure 1 depicts the maximal-growth cutoff function $\tilde{E}_h(\eta_g)$. The $\tilde{E}_h$ function and the $\eta_g$ cutoffs depend only on the model parameters.

![Figure 1: Cutoff $\tilde{E}_h(\eta_g)$ for maximal growth](image)

Using the adopting firm’s first order conditions we obtain, when $E_h < \tilde{E}_h$ (interior growth)

$$x = 1 + (H_a)^{\theta},$$
$$w_a = \theta x (H_a)^{\theta-1} = \theta (H_a)^{\theta-1} + \theta (H_a)^{2\theta-1},$$
$$\pi_i(E_h, \eta_g) = (1 - \theta) (H_a)^{\theta} + \left(\frac{1}{2} - \theta\right) (H_a)^{2\theta}.$$  

(18)

When $E_h \geq \tilde{E}_h$ (maximal growth) we obtain,

$$x = \lambda,$$
$$w_a = \theta \lambda (H_a)^{\theta-1},$$
$$\pi_m(E_h, \eta_g) = (1 - \theta) \lambda (H_a)^{\theta} - \frac{1}{2} (\lambda - 1)^2.$$  

(19)

17It can be seen that $\tilde{E}_h(0) = h$; that is, as $\eta_g \to 0$, only interior growth is possible. A sufficient condition for only interior growth to occur, for all $\eta_g$ is $(\lambda - 1)^{\frac{1}{\theta}} \geq h$. When $\eta_g \to 1$, if $(\lambda - 1)^{\frac{1}{\theta}} \leq 1$, $\tilde{E}_h(\eta_g) = 1$, and only maximal growth occurs; if $1 < (\lambda - 1)^{\frac{1}{\theta}} < h$, then $\tilde{E}_h$ is interior and growth could either be interior (whenever $E_h < \tilde{E}_h$) or maximal (whenever $E_h \geq \tilde{E}_h$).
In all these expressions, \( H_a = \eta_g E_h \), is a random variable.

The wage rate in the adoption sector, \( w_a \), as well as the actual wage for a worker, \( w_a E_h \), is decreasing in the availability of adoption labor, \( \eta_g \); the wage is increasing in the productivity draw, \( E_h \). The profit function, \( \pi(E_h, \eta_g) \), is an increasing function of \( E_h \) in both segments – it is subscripted with \( i \) for the interior region and \( m \) for the maximal region – with a kink at \( \tilde{E}_h(\eta_g) \).\(^\text{18}\) Profits are also increasing in \( \eta_g \), which is relevant for the discussion of firm behavior below. For low \( \eta_g \), the firm relies on a high enough productivity draw to recoup the cost of adoption and to earn profits; for high \( \eta_g \), the firm can afford to proceed with production for a lower productivity draw.

The optimality condition for the non-adopting firm in (10), and the non-adoption labor market clearing condition implies that \( w_n = \theta \beta \left( (1 - \eta_g) h \right)^{\theta-1} \).

### 4.1.2 The Dynamic Problem

The high-tech firms’ decision \( x(E_h, \eta_g) = A^X \) is a purely static one; given the complete spillover assumed, the adoption decision does not affect the technology that this firm has access to in the next period. However, its second stage problem of deciding which implementation route to follow to increase productivity is dynamic across sub-periods within a model period and is specified by (13).

Profits \( \pi \) are an increasing function of \( E_h \); the “dismiss” option is decreasing in \( E_h \). We therefore anticipate a threshold \( \tilde{E}_h \in [1, h] \) such that for all \( E_h < \tilde{E}_h \) the firm fires and searches for a better workforce-technology match and for all \( E_h \geq \tilde{E}_h \) it proceeds with production using the current workforce. In this sub-section, we characterize the firm’s BGP condition \( \tilde{E}_h(\eta_g) \).

Based on the productivity required for maximal growth, \( \tilde{E}_h(\eta_g) \), one of two cases arises. In the first case, depicted in Figure 2, the threshold productivity \( \tilde{E}_h \) is less that the productivity needed for maximal growth. Therefore, some or all of the \( E_h \) draws that the firm decides to proceed with will not result in maximal growth rate and thus the expected aggregate growth rate \( E(x) \) is less than \( \lambda \). Two sub-cases are possible. In case 1a, if \( \eta_g < \eta_g^1 \), the measure of generally-educated agents is low enough to cause \( \tilde{E}_h = h \), which implies that every draw of \( E_h \) will result in non-maximal growth. For \( \eta_g \) greater than \( \eta_g^1 \), but less than a yet-to-be determined valued \( \eta_g^2 \), some draws of \( E_h \) will result in interior growth and some will result in maximal growth for that firm.

In the second case, depicted in Figure 3, \( \tilde{E}_h \) is greater than or equal to the productivity needed for maximal growth. Every draw of \( E_h \) for which the “proceed” option is chosen by an individual firm will result in maximal growth; therefore, the expected aggregate growth rate \( E(x) \) equals \( \lambda \). Again, there are two sub-cases. If \( \eta_g \) is greater than or equal to \( \eta_g^2 \), but

\(^{18}\)One can show, by evaluating both \( \pi \)’s at \( E_h = \tilde{E}_h(\eta_g) \), that \( \pi(E_h, \eta_g) \) is continuous in \( E_h \), and that the second part (when growth is maximal) is steeper than the first (when growth is interior).
less than the $\eta_g^3$, some draws of $E_h$ will not be compatible with maximal growth, but the firm will not choose to proceed with these (case 2a). In the second sub-case, for large enough $\eta_g$, that is, for $\eta_g \geq \eta_g^3$, only draws that are compatible with maximal growth will even occur (case 2b). The expression that determines the cutoff $\eta_g^2$ is derived as a by-product of the derivation of the firm’s $E_h(\eta_g)$ curve.

Case 1a: $1 < \hat{E}_h < \hat{\hat{E}}_h(\eta_g) = h$
(Only $x < \lambda$ possible; $Ex < \lambda$; $0 < \eta_g < \eta_g^3$)

Case 1b: $1 < \hat{E}_h < \hat{\hat{E}}_h(\eta_g) < h$
($x < \lambda$ or $x = \lambda$ possible; $Ex < \lambda$; $\eta_g \in [\eta_g^3,1]$)

Case 2a: $1 < \hat{E}_h(\eta_g) < \hat{\hat{E}}_h < h$
($x < \lambda$ or $x = \lambda$, possible, but $Ex = \lambda$)

Case 2b: $1 = \hat{E}_h(\eta_g) < \hat{\hat{E}}_h < h$
(Only $x = \lambda$ possible; $Ex = \lambda$; $\eta_g \in [\eta_g^3,1]$)

Figure 2: Firm’s Bellman Equation when $\overline{E}_h < \hat{E}_h$

Figure 3: Firm’s Bellman Equation when $\overline{E}_h \geq \hat{E}_h$
For ease of exposition, the above figures have been drawn assuming that the threshold $\bar{E}_h$ is interior for all $\eta_g$. We will consider the corner cases later. If the “proceed” option lies everywhere below the “fire” option, $\bar{E}_h = 1$, and if it lies everywhere above the “fire” option, $\bar{E}_h = h$.

4.1.3 The Firm Condition $\bar{E}_h(\eta_g)$

The solution for the Bellman equation depends on whether case 1 or case 2 is assumed. We first assume that case 1 applies and derive the $\bar{E}_h(\eta_g)$ function. In the appendix (section A.1), we show that the $\bar{E}_h(\eta_g)$-schedule is implicitly defined by

$$f \eta_g \left\{ (1 - F(\bar{E}_h)) \bar{E}_h + \int_{1}^{\bar{E}_h} E_h dF(E_h) \right\} > \eta_g (1 - C) \left[ \int_{\bar{E}_h}^{\bar{E}_h} \pi_i(E_h, \eta_g) dF(E_h) + \int_{\bar{E}_h}^{h} \pi_m(E_h, \eta_g) dF(E_h) - (1 - F(\bar{E}_h)) \pi_i(E_h, \eta_g) \right]. \tag{20}$$

The left hand side is the marginal cost of firing the workforce and waiting for a better draw and the right hand side is the marginal benefit of such a draw in the form of extra expected net profits. If (20) holds with $> \text{ at } \bar{E}_h = 1$, then the cost of firing is so high that the firm accepts any productivity draw, and $\bar{E}_h(\eta_g) = 1$. On the other hand, if (20) holds with $< \text{ at } \bar{E}_h = h$, the benefit of hiring is so high relative to the cost that the firm is willing to wait for the highest draw, and $\bar{E}_h(\eta_g) = h$. Otherwise an interior productivity threshold $\bar{E}_h(\eta_g)$ is uniquely determined, with (20) holding as an equality.

In the appendix (section A.2), we characterize the $\bar{E}_h(\eta_g)$ condition in detail; here we confine ourselves to a brief discussion. It is convenient to divide the above condition by $\eta_g$ throughout; we will argue that as $\eta_g \to 0$, $\bar{E}_h \to h$, separately. In the following discussion, when we refer to condition (20), it is this modified condition we refer to.

For a given set of cost parameters, $f$, $C$, and labor availability, $\eta_g$, the left hand side of (20) is an increasing function of $\bar{E}_h$; waiting for a higher $E_h$’s entails more rounds of firings and higher firing costs. The right hand side is decreasing in $\bar{E}_h$. While a higher $\bar{E}_h$ implies a higher profit for the adopting firm, it reduces the range of $E_h$ over which actual production takes place for those profits to be realized. The decreasing right hand side evidently captures the diminishing returns inherent in the process of experimenting with $E_h$. Since the left hand side is increasing and the right hand side is decreasing in $\bar{E}_h$, an intersection in $[1, h]$ occurs. However, where the intersection occurs depends on $\eta_g$; we turn to this next.

The left hand side of (20) is independent of $\eta_g$. As argued in the appendix, one can find a sufficient condition to ensure that the right hand side is decreasing in $\eta_g$. Given these properties of (20) with respect to $\bar{E}_h$ and $\eta_g$, the intersection of the two sides, which gives the threshold $\bar{E}_h$, is decreasing in $\eta_g$. The concavity of the production and thus the profit functions in $\eta_g$ imply that the benefit of $\eta_g$ increases slower than the firing cost; indeed

---

19 In the range of $\eta_g$ in which $\bar{E}_h = h$ (case 1a), the $\pi_m$ term is absent.
\( \theta \), which controls this concavity, plays an important role in the above-mentioned sufficient condition.

We show that the \( \bar{E}_h(\eta_g) \) schedule starts at \( h \) (and potentially stays at \( h \) for a neighborhood of zero); for very low supply of adoption labor, the firing costs are negligible relative to the profits associated with the best possible draw so that the firm is willing to wait for it. It then decreases monotonically. The firm condition is depicted in Figure 4.\(^{20}\)

At some point the decreasing function \( \bar{E}_h(\eta_g) \) crosses the \( \bar{E}_h(\eta_g) \) function, depicted as a dotted line in Figure 4. Beyond that point, case 2 applies. The \( \eta_g \) that corresponds to this switchover, which we have labeled \( \eta_g^2 \), is implicitly defined by setting \( \bar{E}_h = \bar{E}_h = (\lambda - 1) \frac{\theta}{\eta_g} \) in (20). The condition, depicted in Figure 4, should therefore be interpreted as a composite of the two cases.

In the appendix (section A.3), we show that the \( \bar{E}_h(\eta_g) \) condition for case 2 is implicitly

\[^{20}\text{As shown in the appendix, if } f < (1 - C)(1 - \theta) \lambda \left[ \int_{1}^{h} (E_h)^{\theta} dF(E_h) - 1 \right], \text{ we obtain } \bar{E}_h(1) > 1.\]
given by:  
\[ f\eta_g \left\{ \left(1 - F(\bar{E}_h)\right) \bar{E}_h + \int_1^{\bar{E}_h} E_h dF(E_h) \right\} = (1 - C) \left[ \int_{\bar{E}_h}^h \pi_m(E_h, \eta_g) dF(E_h) - \left(1 - F(\bar{E}_h)\right) \pi_m(E_h, \eta_g) \right]. \]  
(21)

The characterization of \( \bar{E}_h(\eta_g) \) for this case parallels the one for the previous case. In the appendix (section A.4), we argue that the \( \bar{E}_h(\eta_g) \) schedule is monotonically decreasing in \( \eta_g \). It is depicted as the \([\eta_g^2, 1]\) portion of the \( \bar{E}_h(\eta_g) \) curve in Figure 4. We also provide a condition on \( f \) and other parameters of the model to ensure \( \bar{E}_h(1) > 1 \). Intuitively, the firing cost cannot be so high as to cause the firm to proceed with any \( E_h \)-draw.

Taken together, the decreasing \( \bar{E}_h(\eta_g) \) curve is interpreted as follows. As \( \eta_g \) increases, the firing cost increases linearly since it is paid for every person hired. Profits increase, but less than linearly, given the diminishing returns to \( H_a \). Therefore the firm “experiments” less in adoption – alternately, profits are high on account of high \( \eta_g \), and therefore the firm does not have to set a high productivity threshold in order to realize high profits and meet adoption costs.

4.1.4 Dependence of \( \bar{E}_h(\eta_g) \) on \( f \)

For a given \( \eta_g \), an increase in the unit firing cost of labor, \( f \), increases the marginal cost of drawing another \( E_h \), and the adopting firm experiments less and accepts lower productivity matches for any given \( \eta_g \); that is, \( \bar{E}_h(\eta_g) \) decreases. We graph the equilibrium implication of this shift in the firm condition once we have characterized the household condition below.

4.1.5 Dependence of \( \bar{E}_h(\eta_g) \) on \( C \)

Unlike an increase in the firing cost, an increase in the entry or regulation costs decreases the benefit of redrawning instead of the cost. But the outcome is similar; the adopting firm experiments less and accepts lower productivity matches for any given \( \eta_g \). We graph the equilibrium implication of this shift in the firm condition later. To study how the growth gap between the US and Europe changes with \( \lambda \), we next examine the dependence of the adopting firm’s condition on \( \lambda \).

\[ \text{We make assumptions in the appendix to guarantee that this condition holds as an equality.} \]

\[ \text{It is easy to see from (20) and (21), by setting } \pi_i = \pi_m \text{ at } \bar{E}_h = \bar{E}_h, \text{ that the schedule } \bar{E}_h(\eta_g) \text{ is continuous at } \eta_g^2. \]

\[ \text{In the appendix (section A.5), we provide analytical details on the dependence of } \bar{E}_h(\eta_g) \text{ on } f, C, \text{ and } \lambda. \]

20
4.1.6 Dependence of $\bar{E}_h(\eta_g)$ on $\lambda$

There are two effects of an increase in the growth rate of the available technology, $\lambda$. The minimum adoption labor supply, $\eta_g$, needed to induce the firm to adopt at the maximum possible rate increases. Since the fixed cost of adopting at the maximum rate, $\frac{1}{2} (\lambda - 1)^2$, increases, the wage in the adoption sector has to be low enough to keep profits high, which is ensured by higher $\eta_g$. This effect is captured by the rightward shift of $\bar{E}_h(\eta_g)$.

When $\lambda$ increases, the expected profit from a higher draw increases, unless only interior draws occur. This increase in the marginal benefit of a redraw causes the firm to set a higher adoption threshold; the firm can afford the firing cost for lower productivity draws. This effect causes $\bar{E}_h(\eta_g)$ to increase. We graph this condition later.

4.1.7 Dependence of $\bar{E}_h(\eta_g)$ on parameters: Summary

In summary, the dependence of the firm condition on $\eta_g$ and the various parameters is given by:

$$\bar{E}_h = \bar{E}_h(\eta_g; f, C, \lambda)$$

$$-; -, -, +$$

4.2 Households

Workers take as given the adopting firm behavior, as characterized by $\bar{E}_h(\eta_g)$, and decide on their education, which will result in a BGP education allocation $\eta_g$.

In order to calculate the expected utilities in (14), denote the value of a vocationally educated agent on the BGP by $v_v$, and that of a generally educated agent by $v_g$. Conditional on being kept by the firm, the expected utility of a latter agent is

$$v_g = (1 - F(\bar{E}_h)) \left[ \frac{1}{(1 - F(\bar{E}_h))} \int_{\bar{E}_h}^h \log (w_a E_h) dF(E_h) \right] + F(\bar{E}_h) v_g,$$

which implies

$$v_g = \frac{1}{(1 - F(\bar{E}_h))} \int_{\bar{E}_h}^h \log (w_a E_h) dF(E_h).$$

The condition that pins down the threshold ability for obtaining general education in (14) is given by

$$v_g - v_v \equiv \frac{1}{(1 - F(\bar{E}_h))} \int_{\bar{E}_h}^h \log (w_a E_h) dF(E_h) - \log (w_n h) = \log (e (a^*)) - \log \left( \frac{\gamma_g}{\gamma_v} \right).$$
Assuming a parametric form for the cost function, \( c(a) = \frac{1}{a} \), noting \( a^* = 1 - \eta_g \), and recalling the definition of our education policy variable, \( s = \frac{s_g}{s_e} \), the threshold condition can be written as

\[
\frac{1}{1 - F(E_h)} \int_{E_h}^{h} \log (w_a E_h) dF(E_h) - \log (w_n h) = \log \left( \frac{1}{1 - \eta_g} \right) - \log (s). \tag{22}
\]

From the optimality condition (10) and the market clearing condition in the low-tech sector it follows that

\[
w_n h = \theta \beta h^{\theta} / (1 - \eta_g)^{1 - \theta}. \tag{23}
\]

From (18) and (19) we observe that wages in the high-tech sector depend on whether growth is interior maximal. In particular

\[
w_a E_h = \begin{cases} 
\theta (\eta_g)^{\theta - 1} (E_h) + \theta (\eta_g)^{2\theta - 1} (E_h)^{2\theta} & \text{when } x < \lambda \\
\theta \lambda (\eta_g)^{\theta - 1} (E_h)^{\theta} & \text{when } x = \lambda 
\end{cases} \tag{24}
\]

The equilibrium household condition depends on which of these two cases occurs.

### 4.2.1 Household Condition \( \bar{E}_{hH}^H (\eta_g) \)

In the appendix (section A.6), we derive the expression that implicitly defines the household condition \( \bar{E}_{hH}^H (\eta_g) \) by using (23) and (24) in (22). For the case of \( x < \lambda \) this yields

\[
\frac{1}{1 - F(E_h)} \left\{ \int_{E_h}^{E_h} \log \left[ \theta (\eta_g)^{-\theta} (E_h)^{\theta} + \theta (E_h)^{2\theta} \right] dF(E_h) + \int_{E_h}^{h} \log \left[ \theta \lambda (\eta_g)^{-\theta} (E_h)^{\theta} \right] dF(E_h) \right\}
\]

\[
\begin{aligned}
\equiv & \log \left[ \theta \beta h^{\theta} \right] + (1 - 2\theta) \log (\eta_g) - (2 - \theta) \log (1 - \eta_g) - \log (s). \tag{25}
\end{aligned}
\]

The left hand side of this expression can be viewed as the benefit of general education in the form of higher expected wages, and the right hand side as the cost net of subsidy, which includes the foregone wages in the non-adoption sector in addition to the disutility of general education. If the left hand side is greater at \( \bar{E}_h = 1 \) or the right hand side greater at \( \bar{E}_h = h \), a corner solution obtains.

In the appendix (section A.7), we characterize the \( \bar{E}_{hH}^H (\eta_g) \) schedule in detail; here we limit ourselves to a brief discussion. The right hand side of (25) is independent of \( \bar{E}_h \); the benefit of general education is increasing in expected wages, which is in turn increasing in the productivity threshold. An intersection in \([1, h]\) occurs. However, as in the firm condition, where the intersection occurs depends on \( \eta_g \).

It can directly be seen that the right hand side of (25) is increasing in \( \eta_g \); this results from an increase in the foregone wages in the labor-scarce non-adoption sector, as well as the increased disutility of inducing inframarginal agents to acquire general education. It can be shown that the left hand side is decreasing in \( \eta_g \); adoption wages decrease with the labor supplied.
Given these properties of (25) with respect to $\bar{E}_h$ and $\eta_g$, the intersection of the two sides, which gives the threshold $\bar{E}_h^{HH}(\eta_g)$, is increasing in $\eta_g$. In the appendix we show that when $\eta_g \to 0$, the wage premium is so high that even an $\bar{E}_h = 1$ is enough to attract entry into general education. The $\bar{E}_h^{HH}(\eta_g)$ curve starts at 1, (potentially) stays at 1 for an interval, and increases after that. This condition is depicted in Figure 5. As with the $\bar{E}_h(\eta_g)$ schedule case 1 applies only up to the intersection with the $\bar{E}_h(\eta_g)$ curve; case 2 applies beyond that, and is discussed below. We denote by $\eta_g^{HH}$ the point where this intersection occurs.

In the appendix (section A.8), we show that the expression that implicitly defines the household condition $\bar{E}_h^{HH}(\eta_g)$ for the case that $x = \lambda$ is:

\[
\frac{1}{(1 - F(\bar{E}_h))} \left\{ \int_{\bar{E}_h}^{h} \log \left[ \theta \lambda (E_h)^{\theta} \right] dF(E_h) \right\} \\
\leq \log [\theta^\gamma h^\theta] + (1 - \theta) \log (\eta_g) - (2 - \theta) \log (1 - \eta_g) - \log (s).
\]  

(26)

The analysis of this condition parallels that of the first case, and the details are given in the appendix (section A.9). The left hand side of (26) is strictly increasing in $\bar{E}_h$ and the right hand side is independent of $\bar{E}_h$; the left hand side is independent of $\eta_g$ and the right hand side is strictly increasing with $\eta_g$. This indicates that the intersection of the two sides, which gives the threshold $\bar{E}_h$ necessary for any given labor supply $\eta_g$, is increasing in $\eta_g$. We can show that as $\eta_g \to 1$, $\bar{E}_h \to h$. When $\eta_g$ is very high, even the lowest ability agents obtain general education. The disutility cost is so prohibitive that the $\bar{E}_h$, and thus

![Figure 5: The Household Condition $\bar{E}_h^{HH}(\eta_g)$](image)
the expected wage premium, has to be very high to induce entry. Therefore, the $E_{HH}^{H}(\eta_g)$ curve increases up to $h$, and (potentially) stays at $h$ for an interval. This interval is depicted as the $[\eta_g^{HH},1]$ portion of the $E_{HH}^{H}(\eta_g)$ curve in Figure 5.

Taken together, the increasing $E_{HH}^{H}(\eta_g)$ curve is interpreted as follows. As $\eta_g$ increases, a downward pressure on adoption wages arises, and the disutility of the marginal general education enrollee increases. To counter these, the productivity levels $E_h > \bar{E}_h$ with which the firm produces and pays wages have to be high.

### 4.2.2 Dependence of $E_{HH}^{H}(\eta_g)$ on $s$

From both (25) and (26), we see that the right hand sides decrease with $s$. Given the increasing left hand sides, this means that the point of intersection, $\bar{E}_h$, if interior, shifts down for a given $\eta_g$. As the subsidy for general education increases, households are willing to supply a given $\eta_g$ for lower thresholds $\bar{E}_h$ and thus lower wage premia. We will graph the equilibrium implication of this shift in the household condition induced by a change in $s$ below.

### 4.2.3 Dependence of $E_{HH}^{H}(\eta_g)$ on $\lambda$

As in the case of the analysis of firms, (17) implies that the $E_{H}(\eta_g)$ threshold curve shifts to the right when $\lambda$ increases, and the thresholds $\eta_g^1$ and $\eta_g^3$ increase.

In both (25) and (26), the right hand sides – the cost of acquiring general education – do not depend on $\lambda$. A mere examination of the left hand side of (26) shows that it is increasing in $\lambda$. In the appendix (section A.10), we show this is also true for (25). Given the flat right hand sides, the threshold, $\bar{E}_h$, decreases.

Expected adoption wages are increasing in $\lambda$; therefore, such an increase induces supply of a given $\eta_g$ even for lower thresholds $\bar{E}_h$.

---

24 As in the firm condition, it is easy to see from (25) and (26) by setting $E_h = \bar{E}_h$ that the schedule $E_{HH}^{H}(\eta_g)$ is continuous at $\eta_g^{HH}$.

25 The intervals for $\eta_g$ for which corner solutions arise change, however. The corner $E_h = 1$ occurs because the benefit of general education exceeds the cost for any $E_h$. Since the cost of general education decreases with $s$, the benefit of general education exceeds the cost for a larger interval of $\eta_g$. Likewise, the corner $E_h = h$ occurs because the cost of general education exceeds the benefit for any $E_h$. Since the cost of general education has decreases in $s$, the cost of general education exceeds the benefit for a smaller interval of $\eta_g$.

26 The behavior of the $\eta_g$ intervals in which $E_h = 1$ or $E_h = h$ is very similar to the one discussed above for an increase in $s$; the only difference is that an increase in $\lambda$ works by increasing the benefit while the increased $s$ works by decreasing the cost of general education.
4.2.4 Dependence of $\bar{E}_{h}^{HH}(\eta_g)$ on parameters (summary)

To summarize, the household condition is given by

$$\bar{E}_{h}^{HH} \equiv \bar{E}_{h}^{HH}(\eta_g; s \lambda).$$

4.3 Existence and Uniqueness of Balanced Growth Path Equilibrium

Given the strictly decreasing, continuous firm condition, $\bar{E}_{h}(\eta_g)$, and the strictly increasing, continuous household condition, $\bar{E}_{h}^{HH}(\eta_g)$, we obtain a unique BGP equilibrium. It is depicted in Figure 6; the intersection of the two conditions yield the equilibrium education allocation $(\eta_g)^*$ and productivity threshold $(\bar{E}_{h})^*$.

![Figure 6: The BGP Equilibrium](image)

In the example shown in Figure 6, the intersection between $\bar{E}_{h}(\eta_g)$ and $\bar{E}_{h}^{HH}(\eta_g)$ occurs in case 2a of the firm’s condition where $\bar{E}_{h} \geq \bar{E}_{h}$. Therefore all firms always choose maximal growth. As discussed in section 4.1.2, if the intersection occurs in case 1, the expected growth rate of the economy satisfies $E(x) < \lambda$. Note that once $(\eta_g)^*$ and $(\bar{E}_{h})^*$ are determined, all other variables, such as the expected growth rate, expected wages, and profits on the BGP can be determined. For instance, the average growth rate in the economy is given by

$$E(x) = \frac{1}{1 - F((\bar{E}_{h})^*)} \left\{ \int_{(\bar{E}_{h})^*}^{E(h)^*} \left[ 1 + \left[ (\eta_g)^* (E_h)^* \right]^\theta \right] dF(E_h) + \lambda \left( 1 - F((\bar{E}_{h})^*) \right) \right\}.$$
If \((\bar{E}_h)^* \geq \left(\bar{E}_h\right)^*\), this condition reduces to \(E(x) = \lambda\).\(^{27}\) The expected growth rate \(E(x)\) is increasing in the equilibrium values of both \(\eta_g\) and \(\bar{E}_h\).

### 4.4 Comparative Statics

The previous analysis of how the firm and household conditions depend on the parameters \((f, C, s, \lambda)\) makes it straightforward now to characterize their influence on the BGP equilibrium.

#### 4.4.1 Increase in Firing Cost, \(f\)

An increase in the firing cost parameter \(f\) (our proxy for labor market regulations) affects only the firm condition by shifting it down. As a result, the equilibrium threshold as well as the general education attainment and the expected growth rate decreases. Firms are willing to produce with lower \(E_h\)-draws rather than fire employees; this translates into lower wages in the adoption sector and a lower incentive to obtaining general education. Figure 7 depicts this situation.

![Figure 7: Effect of an increase in \(f, C\)](image)

Increase in ‘\(f\)’ or ‘\(C\)’ moves intersection from case 2a \((Ex = \lambda)\) to case 1b \((Ex < \lambda)\).

\(^{27}\) \((\bar{E}_h)^* = \frac{(2-1)^{\frac{1}{\eta_g}}}{(\eta_g)}\)
4.4.2 Increase in Entry Cost, $C$

An increase of the entry cost parameter $C$ (our proxy for product market regulations and bureaucratic costs) also only shifts the firm’s condition downward, again decreasing the equilibrium threshold, general education attainment, and expected growth $E(x)$. Firms are willing to accept lower $E_h$-draws as the benefits of higher draws are diminished, which results in lower wages in the adoption sector and a lower incentive to obtain general education.

4.4.3 Increase in General Education Subsidy, $s$

The subsidy parameter $s$ (our proxy for general education focus) only impacts the household condition; an increase in $s$ shifts the household’s condition down, reducing the equilibrium productivity threshold $E_h$, but increasing general education attainment $\eta_g$. If the latter effect dominates, expected growth $E(x)$ increases. When the effective cost of general education decreases, households require a lower threshold $E_h$ and wage premium to choose general education. While a decrease in $f$ or $C$ or an increase in $s$ will all result in higher general education attainment, the subsidy increase reduces $E_h$ while the others increase it. This implies that the average number of draws before successful production, given by $1/(1 - F(E_h))$, declines with $s$ (and increases with $f, C$). 28

\[ \bar{E}_h \]

![Diagram](image)

**Figure 8: Effect of an increase in $s$**

28 The mean number of attempts is given by $\sum_{j=1}^{\infty} j [F(E_h)]^{j-1} (1 - F(E_h)) = \frac{1}{1 - F(E_h)}$. We have chosen to abstract from discounting in our setup for sake of simplicity; the advantage of the subsidy scheme is even more readily apparent if discounting is taken into account.
4.4.4 Dependence of BGP on $\lambda$

Analyzing the dependence of the BGP equilibrium on the speed of technology availability, $\lambda$, is more complicated because all three schedules – $\tilde{E}_h(\eta_g)$ (shifts right), $\bar{E}_h(\eta_g)$ (shifts up), and $\bar{E}_h^{HH}(\eta_g)$ (shifts right) – are affected. The effect on the equilibrium productivity threshold $\bar{E}_h$ is ambiguous, whereas the equilibrium general education attainment increases unambiguously. The higher productivity threshold set by the firm, as well as the direct increase in the maximal growth rate, increase the expected wage in the adoption sector. Therefore the incentive to acquire general education increases, and the resulting $\eta_g$ is higher. Figure 9 depicts this situation.

![Figure 9: Effect of an increase in $\lambda$](image)

What happens to the US-Europe (expected) growth gap when $\lambda$ increases? The answer to this question is again complicated by the shifts of all three curves. In Figure 10, we hold the household curves at their old positions to achieve graphical tractability. The accompanying box explains how a growth gap, initially zero, can expand with an increase in $\lambda$. The behavior

---

29 In the particular example shown in Figure 9, case 1b obtains both before and after the increase. However, if the shifting of the $\tilde{E}_h$ curve is more muted, which will happen if $\theta$ is not too low, it is possible that the new intersection is in case 2a, and maximal growth obtains.

For ease of drawing, the entire firm condition is shown as shifting upward. In the interval $[1, \eta_g^1]$, when $\tilde{E}_h = 1$, and only interior growth is possible, a change in $\lambda$ has no effect. The argument made in section 4.1.6 is for a marginal change in $\lambda$ that leaves the relevant case for $\tilde{E}_h$ unchanged. As long as the intersection of the household and firm conditions does not correspond to case 1a (only interior growth draws), the figure will be an accurate representation of an increase in $\lambda$. 

28
of the maximal growth cutoff function, $\tilde{E}_h(\eta_g)$, is important for this possibility. While the shifts of the firm and household conditions increase the equilibrium general education attainment for Europe, this increased attainment may still fall short of the new cutoff, $\eta_{g1}$, needed for maximum growth. Alternately, the wages in the adopting sector have to be low enough to induce the firm to adopt at the higher maximum speed. In the US, where general education subsidy and attainment are higher, $\eta_g$ can exceed even the increased cutoff $\eta_{g1}$. It may therefore continue to grow at the higher potential rate.

![Graph showing the dependence of the US-Europe growth gap on $\lambda$.](image)

**Figure 10: Dependence of the US-Europe growth gap on $\lambda$**

Our analytical characterization has revealed the qualitative impact of the policy variables $(f, C, s)$ on the BGP equilibrium. However, to quantify the relative importance of each policy’s contribution to the US-Europe growth gap requires experimenting with counterfactuals using a calibrated model. We proceed to do this next.

5 **Calibration**

We choose a model period to represent roughly $T = 20$ years. We have assumed that a newly adopted technology becomes “common practice” within one model period. Atkeson and Kehoe (2001) discuss the diffusion of electricity in U.S. manufacturing establishments, and report that the time required for the technology to diffuse from 5% to 50% – a measure of the speed of diffusion that is less sensitive to the choice of starting date – “occurred over the 20 years from 1899 to 1919.” Therefore, a choice of 20 years for a model period seems to
be a reasonable start.\footnote{The technology of the integrated circuit was invented in 1959 and found large scale use in the personal computer in 1981, a bit more than 20 years later. And the full potential of the personal computer in turn was unleashed with the pervasive use of the internet nearly another 20 years later. These observations lend further credibility to our choice of the model period. Note that our choice of the period length has a bearing on the interpretation of the magnitudes of some of the parameters in the model, in particular the gross growth rate of the frontier technology $\lambda$ and the firing cost $f$.}

Parameters describing the production technology, $(\lambda, \theta, \beta)$, maximum task-specific productivity, $h$, its distribution $F$, and government policy $(C, f, s)$ need to be chosen. We first pick a subset of parameters motivated by independent empirical evidence. The remaining parameters are then chosen so that our model broadly matches selected empirical growth and education allocation statistics.

### 5.1 Predetermined Parameters

The share parameter $\theta$ in the production function is set to 0.35. Recall that the production function is of the form $AH^\theta$. Since $H$ is the effective units of labor, one can view $\theta$ as the intensity on raw labor as well as on human capital. For instance, the production function in the adopting sector is: $x(\eta_g)^\theta (E_h)^\theta$; with our choice of $\theta = 0.35$, the labor share in this sector equals the commonly assumed value of 0.7. Klenow and Rodriguez (1997) explicitly write the production function as $Y = K^\alpha H^\gamma (AL)^{1-\alpha-\gamma}$, and use values of $\gamma = 0.28$ and $\alpha = 0.3$, which implies that $1 - \alpha - \gamma = 0.42$. In choosing $\theta = 0.35$, we slightly overstate the human capital intensity and understate the labor intensity, compared to their study.

We choose the distribution over productivity draws, $F$, to be uniform over its support $[1, h]$, thus avoiding the introduction of further free parameters.

Finally, in order to calibrate the relative education subsidies $s$ for both the US and Europe we directly use expenditure per student data. The OECD (2001, Table B1.1), reports the figures presented in Table 2.

**Table 2**

<table>
<thead>
<tr>
<th>Country</th>
<th>Upper sec.</th>
<th>Univ. tertiary</th>
</tr>
</thead>
<tbody>
<tr>
<td>US</td>
<td>7,764</td>
<td>19,802</td>
</tr>
<tr>
<td>Germany</td>
<td>9,519</td>
<td>10,139</td>
</tr>
<tr>
<td>Italy</td>
<td>6,340</td>
<td>6,295</td>
</tr>
</tbody>
</table>

We use upper secondary expenditure as a proxy for $s_v$, and university tertiary expenditure as a proxy for $s_g$. Thus $s_{US}^{US} = \frac{19802}{7764} = 2.55$, $s_{GER}^{GER} = \frac{10139}{9519} = 1.07$, $s_{IT}^{IT} = \frac{6295}{6340} = 0.99$. We
synthesize this information by choosing $s^{US} = 2.55$ and $s^{EUR} = 1$ as our education policy parameters.31 The predetermined parameters are summarized in 3.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$s^{US}$</th>
<th>$s^{EUR}$</th>
<th>$F(E_h)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>0.35</td>
<td>2.55</td>
<td>1</td>
</tr>
</tbody>
</table>

### 5.2 Parameters Chosen to Match Observations

Given the manufacturing labor productivity growth reported in Table 1 of roughly 2.5% in the 70s-80s and 4% in the 90s and beyond for the US, and assuming that the US grows at the potential rate in both periods, we choose $\lambda = (1.025)^{20}$ for the rate of technological change that prevailed until 1980, and the increased rate $\lambda' = (1.04)^{20}$ for the period beyond. Note that our model will attain its empirical targets only if, given the other parameters chosen, it yields maximal growth for both US and Europe with $\lambda$, and maximal growth for the US with $\lambda'$ but sub-maximal growth for Europe. In other words, our choice of $\lambda$ (which is not varied across regions) enables, but by no means guarantees quantitative success of our model in matching the stylized productivity growth facts.

How do the above choices for $\lambda$ compare with the independent evidence available on the change in the growth rate of available technologies? Cummings and Violante (2002) compute an aggregate index of investment-specific (embodied) technical change, and report that it grew at a rate of 3% a year until 1975 and reached an average annual rate in excess of 6% in the 90s. If we assume that this change, besides augmenting the quality of (unmodeled) capital, also augments labor as in the Klenow and Rodriguez (1997) production function, our variable $A$ in the model would have to grow at rates close to 2% and 4% in the two periods.32

In order to calibrate the remaining parameters $\beta, h, f, C$, both for the US and Europe we attempt to match the following statistics (with $\lambda = \lambda'$ and all other parameters set to their

31 Germany and Italy also have data on per pupil expenditure in Tertiary type B programs that “tend to focus on occupationally-specific skills intended for direct labor market entry.” (p. 129) and on post-secondary non-tertiary programs, such as “vocational certificates in ... the United States,” or “vocational training in the dual system for holders of general upper secondary qualifications in Germany.” For Germany, these figures are $5,422$ and $10,924$, and for Italy, $6283$ ans $6,458$ respectively. For the US, there is no data for these categories. For Germany, there is a big disparity between these two figures. The upper secondary value which we use appears to be a reasonable via media. For Italy, it does not matter much which figure is used.

32 With a residual capital intensity of 0.3, and the assumed raw labor intensity of 0.35, the growth rate of $A$ in our model should be 0.65 times the growth rate of the reported technical change. Our assumption that the technical change is capital- and raw-labor-augmenting lies in between the assumptions that it is only capital augmenting and it is neutral. In another highly relevant study, Greenwood and Yorukoglu (1997) suggest rates of technical change of 3% until the 70s and 5% beyond.
values described above):

- **Observations on entry costs** are mainly used to pin down $C$, the share of profits not accruing to the entrepreneur.\(^{33}\) Using data on official cost of entry procedures from the “Entry Regulations” part of the “Doing Business” database of the World Bank we find, that, as a percentage of gross national income, that cost is 0.6% for the US, 5.8% for Germany, and 22.7% for Italy. For the US we use, as a benchmark target, a ratio of $ent^{US} = 0 = C^{US}$, and for Europe we use the German value of $ent^{EUR} = 0.058$, since the extremely high value for Italy is an outlier within the EU.\(^{34}\) Indeed, the more comprehensive index for barriers to entrepreneurship constructed by Nicoletti, Scarpetta, and Boylaud (1999) – US: 1.3, Germany: 2.1, Italy: 2.7, with 0 being least regulated and 6 most – suggests that Italy fares only slightly worse than Germany in this regard. The corresponding model statistic is

$$
ent = \frac{C}{1-F(E_h)^{\beta}} \int_{E_h}^{\bar{E}_h} \frac{\pi(E_h, \eta_g) dF(E_h)}{\beta \left[(1 - \eta_g) H\right]^\theta + \int_{E_h}^{\bar{E}_h} \frac{w_o(E_h, \eta_g) \left(\eta_g E_h\right)^{\delta} dF(E_h)}{1-F(E_h)}}
$$

where the numerator is the expected (or aggregate) total cost paid by adoption firms and the denominator is the GDP of the economy.

- **Firing costs** are used to mainly pin down $f$. We set $f = 0$ for the US; for Europe we set $f$ in such a way that firing a worker once costs 6 weeks of her average wage in the model.\(^{35}\) Nagypál (2001), suggests firing costs of six weeks’ worth of average revenues for Europe.\(^{36}\) Given that our model period is 20 years, we choose $f^{EUR}$ so that

$$
f^{EUR} = \frac{6}{20 \times 52} \times \int_{E_h}^{\bar{E}_h} \frac{w_o(E_h, \eta_g) E_h dF(E_h)}{1-F(E_h)}.
$$

\(^{33}\)It is understood that all parameters discussed below affect all equilibrium quantities, albeit to different degrees. We associate a parameter with the statistic whose equilibrium behavior it most affects.

\(^{34}\)As we will document below, this parameter does not have a large effect on equilibrium allocations.

The entry cost database is described in Djankov, La Porta, Silanes, and Shleifer (2001) and can be found in http://rru.worldbank.org/DoingBusiness/TopicReports/EntryRegulations.aspx.

\(^{35}\)As discussed in Section 3.3, $f$ cannot be set exactly to zero as the firms’ Bellman equation may not have a solution. We set it to a very small number for the US, corresponding to less than one day of average wages. Our results are not sensitive to this choice.

\(^{36}\)Pries and Rogerson (2001) consider a range of values for the real resource costs involved in dismissing a worker, the maximum of which is 10% of quarterly wages in the model (i.e. 2.5% of yearly wages). Lazear (1990) presents data on the number of months of salary given to workers upon dismissal after 10 years of service. It is zero for the US, 1 month for Germany, and 15.86 months for Italy! Nicoletti et.al. provide indices for employment protection legislation: 0.2 for the US, 2.8 for Germany, and 3.3. for Italy on a 0 to 6 scale. The World Bank entry costs database mentioned earlier also give an index measure – 0.94 for the US, 1.75 for Germany, and 1.35 for Italy. While there is variation in the exact magnitude of costs reported, the relative ranking of firing costs is preserved across the US and “Europe.”

32
The fraction of population with general education, $\eta_g$, will be used to mainly pin down $h$, which is the maximum productivity draw possible in the adopting sector and the certain productivity for agents in the nonadopting sector. Table 4 summarizes education attainment data from OECD 2001 (tables A2.1.a and C2.1). We realistically assume that workers lower secondary education or less are not able to work in the high-tech sector. In dividing the upper secondary education attainment into vocational and general, we have assumed that enrollment mirrors attainment: vocational enrollments of 6.8% for the US, and 64.6% for Germany and Italy.\footnote{The 6.8\% figure for the US is the percentage of students who completed 30\% or more of all credits in specific labor market preparation courses in 1990. See Medrich, Kagehiro, and Houser (1994).} As previously described, we associate post-secondary non-tertiary education and tertiary B education with vocational education. University tertiary education counts toward general education. Based on these figures we compute the fractions of agents with general education as, $\eta_g^{US} = 0.745$, $\eta_g^{GER} = 0.317$ and $\eta_g^{IT} = 0.196$. As calibration targets we therefore choose $\eta_g^{US} = 0.745$ and $\eta_g^{EUR} \approx 0.3$. In contrast to policy parameters, where there exist obvious differences across the US and Europe, it is less obvious that the parameter $h$ ought to vary across regions. We thus choose $h$ to roughly match the US education allocation and use the same $h$ for Europe. The (in)-ability of the model to match the European education allocations is then an important quantitative test of the model.

The wage premium is used to mainly determine $\beta$, the constant in the production function of the nonadopting sector. At the heart of our model is the endogenous education decision by households, which is driven by relative wages in the two sectors. It is therefore important to calibrate our model to be consistent with empirically measured premia for general education in both regions. The OECD (1997, Table E4.1a) provides the relative earnings of persons aged 25-64 (normalizing income of workers with upper secondary education to 100). These figures are given for tertiary and non-university tertiary (vocational) categories. For the US the numbers are 174 and 119,
for Germany 163 and 111. For the US we thus obtain a general education wage premium \( wp^{US} = \frac{174}{119} = 1.46 \). For Germany this ratio is \( wp^{GER} = 1.47 \), slightly higher.\(^{38}\) We use the German value as our target for Europe. In particular, note that this premium differs from the “college premium”. In the model, the corresponding statistic is computed by dividing expected wages in the adopting sector by nonadopting wages,

\[
wp = \frac{1}{(1-F(E_h))} \left[ \int_{E_h}^\infty w_a(E_h, \eta_g)E_h dF(E_h) \right] / w_n(\eta_g)h.
\]

Our procedure for selecting parameters yields results summarized in Table 5. In the next section we report how well the model outcomes and target statistics match, before turning to counterfactual policy experiments.\(^{39}\)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>( \lambda )</th>
<th>( \lambda' )</th>
<th>( h )</th>
<th>( C^{US} )</th>
<th>( C^{EUR} )</th>
<th>( f^{US} )</th>
<th>( f^{EUR} )</th>
<th>( \beta^{US} )</th>
<th>( \beta^{EUR} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>0.35</td>
<td>2.55</td>
<td>2.4</td>
<td>0.0</td>
<td>0.6</td>
<td>0.0</td>
<td>0.011</td>
<td>0.74</td>
<td>2.1</td>
</tr>
</tbody>
</table>

6 Quantitative Results

In this section we present our quantitative results. We will do this in two steps. First we assess whether the model can qualitatively and, more importantly, quantitatively match the key empirical targets set out in the previous section, both for the US and Europe at different points of time. Since the model outcomes broadly match the data, we use it to “decompose” the recent US-Europe productivity growth differentials by differences in the proxies for policies – the product market friction, \( C \), labor market friction, \( f \), and focus on general education, \( s \).

6.1 Quantitative Evaluation of Model

In Table 6 we summarize the quantitative predictions of our model. All variables with primes refer to the 1990s. Recall that we chose a total of 9 parameters to match 9 empirical observations.\(^{40}\) There were a total of 12 empirical statistics we reported in the previous section, so

\(^{38}\)The non-university tertiary wage is unavailable for Italy.

\(^{39}\)Note that the value for the non-adoption production function parameter, \( \beta \), which is chosen to match the wage premium of an economy, is higher for Europe than the US. This calibration outcome is reminiscent of the much-documented “wage compression” in Europe. See, for instance, Acemoglu (2002) and the references therein.

\(^{40}\)Strictly speaking, we only set out to match 7 statistics, since our choices of \( \lambda, \lambda' \) only enable, but not insure that the model matches selected productivity growth observations.
that we have three “overidentifying restrictions”. In particular, we do not explicitly attempt to reproduce European growth rates or its education allocation. First, the model attains its calibration targets almost exactly (which is not automatic, since model statistics are nonlinear functions of model parameters and a system of 9 equations in 9 variables not always has a solution). Second, the model does very well in satisfying the overidentifying restrictions as well: European growth rates in both periods as well as European education allocations are almost exactly reproduced by the model. These findings give us some confidence that our model indeed allows us to decompose the recent growth differentials between the US and Europe in a quantitatively meaningful way.

**Table 6**  
**Targets and Model Predictions**

<table>
<thead>
<tr>
<th>Variable</th>
<th>US, Target</th>
<th>US, Model</th>
<th>Europe, Target</th>
<th>Europe, Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E(x)$</td>
<td>2.5% p.a.</td>
<td>2.5% p.a.</td>
<td>2.9% p.a.</td>
<td>2.5% p.a.</td>
</tr>
<tr>
<td>$E(x)'$</td>
<td>4% p.a.</td>
<td>4% p.a.</td>
<td>3% p.a.</td>
<td>3.1% p.a.</td>
</tr>
<tr>
<td>$\eta_g$</td>
<td>74.5%</td>
<td>74%</td>
<td>30%</td>
<td>30%</td>
</tr>
<tr>
<td>$wp'$</td>
<td>1.46</td>
<td>1.46</td>
<td>1.47</td>
<td>1.47</td>
</tr>
<tr>
<td>$fir'$</td>
<td>0</td>
<td>0</td>
<td>6 weeks</td>
<td>6 weeks</td>
</tr>
<tr>
<td>$ent'$</td>
<td>0</td>
<td>0</td>
<td>0.058%</td>
<td>0.057%</td>
</tr>
</tbody>
</table>

The model also predicts how education allocations and wage premia, both in Europe as well as in the US, change in response to the increased speed at which technologies arrive. As demonstrated in the theoretical part of our paper and confirmed in Table 7, the model predicts an increase in the share of the population with general education. The increase is much more pronounced for the US, which explains why, in contrast to Europe, it continues to grow at maximal speed: the incentives to acquire general education are strong enough (because of the high differential subsidy, $s^{US} = 2.55$ vs. $s^{EUR} = 1.0$) so that a scarcity of workers with appropriate education does not slow technology adoption and economic growth.41

What is not obvious from the theoretical analysis is the response of the wage premium to a change in $\lambda$; on one hand the supply of generally skilled workers increases, on the other hand faster technology adoption fuels higher demand for those workers. The theoretical effect on the wage premium is ambiguous, but, as Table 7 shows, the demand effect dominates in both cases and leads to an increase in the wage premium in both regions; again the effect is much stronger for the US.

---

41 The European Competitiveness Report 2001 bemoans the lack of adaptability of the labor force to new technologies: “... in recent years skill shortages in important technology areas have been reported in several European countries... It appears that, unlike in previous years, when the long-term trend increase in the demand for skills was met by the supply of technology professionals from the educational system, the surge in demand for ICT-related skills in the 1990s found no corresponding supply forthcoming.”
Table 7
Ancillary Model Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>US, Model</th>
<th>Europe, Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\eta_g$</td>
<td>50%</td>
<td>28%</td>
</tr>
<tr>
<td>$\eta_g'$</td>
<td>74%</td>
<td>30%</td>
</tr>
<tr>
<td>$wp/wp$</td>
<td>1.87</td>
<td>1.07</td>
</tr>
</tbody>
</table>

6.2 Decomposition of Recent US-Europe Growth Differentials

We are now in a position to perform a model-based decomposition of the growth gap that emerged in the 1990s between Europe and the US. We start with the European situation in the 1990s (see column 5 of Table 6). We then sequentially reduce the product market friction to zero ($C = 0$), then also the labor market friction to zero ($f = 0$), and finally increase the education subsidy to US levels, thus arriving at the US situation (see column 3 of Table 6). For all experiments we adjust $\beta$ in such a way as to keep the wage premium constant at the empirically observed value (see table 6). Table 8 summarizes the results of this thought experiment.

Table 8
Decomposition of Growth Gap

<table>
<thead>
<tr>
<th>Variable/Parameter</th>
<th>Europe</th>
<th>Eur., $C = C^{US}$</th>
<th>Eur., $C = C^{US}, f = f^{US}$</th>
<th>US</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E(x)$</td>
<td>3.1%</td>
<td>3.2%</td>
<td>3.4%</td>
<td>4%</td>
</tr>
<tr>
<td>$\eta_g$</td>
<td>0.30</td>
<td>0.31</td>
<td>0.35</td>
<td>0.74</td>
</tr>
<tr>
<td>$E_h$</td>
<td>1.85</td>
<td>2.05</td>
<td>2.36</td>
<td>2.39</td>
</tr>
<tr>
<td>$\beta$</td>
<td>2.1</td>
<td>2.08</td>
<td>2.0</td>
<td>0.77</td>
</tr>
</tbody>
</table>

Our key finding is that the majority of the growth gap stems from differences in educational policy. Columns 2 and 3 of the table demonstrate that a reduction of the share of GDP lost to regulation from 6% to 0% only increases the growth rate by one-tenth of one percent and leaves the education allocation almost unchanged. Only the threshold productivity $E_h$ increases substantially, by more than 10%.

The growth gap explained by the firing cost is more sizeable, amounting to 0.2 percentage points. Lower firing costs allow firms to wait for more productive draws and yield higher expected wages in the adoption sector, increasing the fraction of the population opting for general education.

The relatively modest growth effects of labor and product market frictions implied by our model therefore leave the bulk of the growth gap, 0.6 percentage points, to be explained by the marked difference in educational focus between the US and Europe. Our quantitative results are not sensitive to the sequential order in which European labor, product market and education policies are changed to their corresponding US values: increasing $s$ from 1 to
2.55 and leaving $f,C$ at their European values again attributes more than 0.6 percentage points of the growth gap (of the 0.9 points predicted by the model) to differences in education policies.

Our results are of course specific to the model we use, but at may least suggest that Europe’s (especially Germany’s and Italy’s) focus on vocational education held these countries back in a quantitatively significant way in the ITC-age of the 90s.

7 Conclusion

The roles of labor market rigidities and excessive regulation in explaining US-Europe growth differences have been studied extensively. In this paper, we draw attention to another possible cause – the focus on skill-specific education in Europe. Our quantitative analysis with a calibrated model shows that the role of education may be significant. Educational reform, in the form of higher flexibility in educational choices made at the upper secondary level, and a greater focus on general education might be important in reducing the US-Europe growth gap that has emerged since the mid 80s.

As conceded above, our model may be geared towards finding a significant role for education. For instance, the education decision determines occupational choice forever. Allowing more flexibility in this regard would attribute a greater effect to labor market frictions and other impediments to occupational mobility. Indeed, it might be hard to completely disentangle the educational and labor market aspects in practice. Likewise, a model of entry and entrepreneurship is likely to find a more important role for product market and other regulations. Extending the model along these dimensions is a subject for future research.
A Appendix

A.1 Firm Condition for Case 1

Assume that case 1 is the relevant one, and equate the payoffs in (13) at the threshold to get:

\[ Ev - f\eta_g \bar{E}_h \equiv (1 - C) \left( 1 - \theta \right) \left( \eta_g \bar{E}_h \right)^\theta + \left( \frac{1}{2} - \theta \right) \left( \eta_g \bar{E}_h \right)^{2\theta} \], \quad (27)

with equality if \( \bar{E}_h \in (1, h) \). Given the distribution function \( F \) (with density \( \hat{f} \)) on \( E_h \), integrating the appropriate \( v \) over the three segments, substituting for \( Ev \) from the threshold condition and simplifying we obtain the firm condition that implicitly defines \( \bar{E}_h (\eta_g) \), in a form suitable for comparative statics:

\[
\begin{align*}
&\int \left(1 - F(E_h)\right) \bar{E}_h + \int_{1}^{\bar{E}_h} E_h dF(E_h) \\
&\quad \equiv \frac{(1 - C) (1 - \theta)}{(\eta_g)^{1 - \theta}} \left[ \int_{E_h}^{\bar{E}_h} (E_h)^{\theta} dF(E_h) - (1 - F(E_h)) (E_h)^{\theta} \right] + \frac{(1 - C) (1 - \theta)}{(\eta_g)^{1 - \theta}} \int^{\bar{E}_h} (1 - \theta) (\eta_g)^{\theta} (E_h)^{\theta} - \frac{1}{2} (\lambda - 1)^2 dF(E_h). 
\end{align*}
\] \quad (28)

A more concise version is presented in the main text as (20).

A.2 Characterizing \( \bar{E}_h (\eta_g) \) for Case 1

A.2.1 Dependence on \( \bar{E}_h \)

For a given \( \eta_g, f, C \), the LHS of (28) is an increasing function of \( \bar{E}_h \). Use Leibniz' formula for the term within curly braces to find that the derivative is \( 1 - F(E_h) - \hat{f}(E_h) \bar{E}_h + \hat{f}(E_h) \bar{E}_h > 0 \). The LHS starts out at \( f \) when \( \bar{E}_h = 1 \) and increases to \( fE(E_h) > f \) at \( \bar{E}_h = h \), independent of \( \eta_g \).

In the RHS of (28) the third term does not depend on \( \bar{E}_h \). Again use Leibniz' formula for the terms within the first set of square brackets to find the derivative as \(- (\bar{E}_h)^{\theta} \hat{f}(\bar{E}_h) - \theta (1 - F(\bar{E}_h)) (\bar{E}_h)^{\theta - 1} + (\bar{E}_h)^{\theta} \hat{f}(\bar{E}_h) < 0 \). Similarly, the terms within the second set of square brackets are also decreasing in \( \bar{E}_h \). So the RHS is decreasing in \( \bar{E}_h \). This is true, independent of whether case 1a or case 1b is relevant, since \( \bar{E}_h \) does not enter the calculation. Since the LHS is increasing and the RHS decreasing in \( \bar{E}_h \), an intersection is likely to occur. However, where the intersection occurs depends on \( \eta_g \).

A.2.2 Dependence on \( \eta_g \)

The LHS of (28) does not change with \( \eta_g \). We now argue that the RHS of (28) is decreasing in \( \eta_g \). When \( \hat{E}_h = h \) (case 1a), the third term vanishes and the first two terms are directly seen to decrease in \( \eta_g \). If \( (\lambda - 1)^p > 1 \), \( \bar{E}_h = 1 \) can be ruled out. The case of interior \( \hat{E}_h \) is more involved, as \( \bar{E}_h = \frac{(\lambda - 1)^p}{\eta_g} \) directly depends on \( \eta_g \). The derivative of the first three terms of the RHS of (28) with respect
to \( \eta_g \) can be shown to reduce to:

\[
- \frac{(1 - C)(1 - \theta)^2}{(\eta_g)^{2 - \theta}} \left[ \int_{E_h} (E_h)^\theta dF(E_h) - (1 - F(\hat{E}_h))(\hat{E}_h)^\theta \right]
\]

\[
- \frac{(1 - C)(2 - \theta)(1 - 2\theta)}{(\eta_g)^{2(1 - \theta)}} \left[ \int_{E_h} (E_h)^{2\theta} dF(E_h) - (1 - F(\hat{E}_h))(\hat{E}_h)^{2\theta} \right]
\]

\[
- \frac{(1 - C)}{(\eta_g)^2} \int_{E_h} (1 - \theta)^2 \lambda (\eta_g)^\theta (E_h)^\theta - \frac{1}{2} (\lambda - 1)^2 \right] dF(E_h).
\]

Since the terms within the square brackets are all not necessarily positive, the entire derivative cannot be readily signed. Since \( \hat{E}_h \) decreases with \( \eta_g \), thereby increasing the range over which maximal profits result, there is ambiguity regarding the dependence of the RHS of (28) on \( \eta_g \). For the above derivative to decrease in \( \eta_g \), intuitively it appears that the concavity of profits in \( \eta_g \) should be strong enough; that is, \( \theta \) should be low enough.

Using (28), one can show that the above derivative is negative if:

\[
\frac{(1 - \theta)f\eta_g}{(1 - C)} \left\{ (1 - F(\hat{E}_h)) \hat{E}_h + \int_{E_h} E_h dF(E_h) \right\} > \theta \left( \frac{1}{2} - \theta \right) \left[ \int_{E_h} (\eta_g)^{2\theta} (E_h)^{2\theta} dF(E_h) - (1 - F(\hat{E}_h))(\eta_g)^{2\theta}(\hat{E}_h)^{2\theta} \right] + \theta \left( \frac{1}{2} - \theta \right) (1 - F(\hat{E}_h))(\eta_g)^{2\theta}(\hat{E}_h)^{2\theta}.
\]

To find a sufficient condition we investigate a low value for the LHS and a high one for the RHS. Given the definition of \( \hat{E}_h \), we can write \((\lambda - 1)^2 = (\eta_g)^{2\theta}(\hat{E}_h)^{2\theta} \). Using this, evaluating the integral in the RHS at \( \hat{E}_h \), and simplifying we find one high value for the RHS is:

\[
\frac{\theta}{2}(\lambda - 1)^2 - \theta^2(\eta_g)^{2\theta}(\hat{E}_h)^{2\theta} F(\hat{E}_h)
\]

\[
- \theta \left( \frac{1}{2} - \theta \right) (\eta_g)^{2\theta}(\hat{E}_h)^{2\theta} F(\hat{E}_h) - \theta \left( \frac{1}{2} - \theta \right) (1 - F(\hat{E}_h))(\eta_g)^{2\theta}(\hat{E}_h)^{2\theta}.
\]

For case 1b that we are discussing, it is true that \(1 < \hat{E}_h < E_h < h, \) and \( \eta_g > \eta_g^1 = \frac{(\lambda - 1)^2}{h} \). Using these facts, the high value for the RHS can be made more stringent, and the sufficient condition reduces to:

\[
\frac{(1 - \theta)f(\lambda - 1)^{\frac{\theta}{h}}}{(1 - C)h} \left\{ (1 - F(\hat{E}_h)) \hat{E}_h + \int_{1}^{\hat{E}_h} E_h dF(E_h) \right\} + \frac{\theta (\lambda - 1)^2}{h^{2\theta}} \left[ \left( \frac{1}{2} - \theta \right) + \theta F(\hat{E}_h) \right] > \frac{\theta}{2}(\lambda - 1)^2.
\]

Since the LHS is increasing in \( \hat{E}_h \), setting it to its minimum value of 1 yields the sufficient condition we seek for the RHS of (28) to decrease in \( \eta_g \):

\[
\frac{(1 - \theta)f(\lambda - 1)^{\frac{\theta}{h}}}{(1 - C)h} > \theta \left[ \frac{1}{2} - \left( \frac{1}{2} - \theta \right) \right].
\]

This condition is satisfied for \( \theta \rightarrow 0; \) as conjectured earlier, enough concavity of profits in \( \eta_g \) will ensure that the sign of the derivative will be the one we seek. When \( \theta \) is at its maximum value of \( \frac{1}{2}, \) the above condition reduces to the simpler, but more stringent condition of \( f > \frac{1}{2}h. \)
The above analysis indicates that the LHS is strictly increasing and the RHS is strictly decreasing in $\hat{E}_h$; the LHS is invariant in $\eta_g$ and the RHS is strictly decreasing with $\eta_g$. These observations indicate that the intersection of the two sides, which gives the threshold $\hat{E}_h$, is decreasing in $\eta_g$.

Consider condition (28), multiplied by $\eta_g$ so that we can analyze the case of $\eta_g \to 0$; $\hat{E}_h \to h$, and maximal growth is not possible for any draw in this case. Both sides of this modified condition tend to zero as $\eta_g \to 0$, but the ratio of LHS to RHS tends to zero using L'Hospital’s rule. Since the RHS, the marginal benefit term of waiting for a draw, dominates the LHS, the marginal cost over all possible $\hat{E}_h$, we get $\hat{E}_h(0) = h$. Indeed $\hat{E}_h(\eta_g) = h$, for $\eta_g$ in a neighborhood of 0 for which the RHS of (28) exceeds $fE(\hat{E}_h)$.

Consider the following expressions for the two sides of (28) when $\eta_g \to 1$:

$$LHS(\hat{E}_h = 1; \eta_g = 1) = f; \quad LHS(\hat{E}_h = h; \eta_g = 1) = fE(\hat{E}_h) > 0,$$

$$RHS(\hat{E}_h = 1; \eta_g = 1) = (1 - C)(1 - \theta) \int_1^{(\lambda - 1)\frac{\nu}{\eta_g}} (E_h)^{\theta} dF(E_h) - 1$$

\[
(1 - C) \left( \frac{1}{2} - \theta \right) \int_1^{(\lambda - 1)\frac{\nu}{\eta_g}} (E_h)^{2\theta} dF(E_h) - 1 + (1 - C) \int_{(\lambda - 1)\frac{\nu}{\eta_g}}^{h} \left[ (1 - \theta) \lambda (E_h)^{\theta} - \frac{1}{2} (\lambda - 1)^2 \right] dF(E_h),
\]

$$RHS(\hat{E}_h = h; \eta_g = 1) = 0.$$ 

If $f < RHS(\hat{E}_h = 1; \eta_g = 1)$, $\hat{E}_h > 1$. Since the RHS is decreasing in $\eta_g$, $\hat{E}_h > 1$ at $\eta_g = 1$ guarantees the same for all $\eta_g$. The condition says that the firing cost $f$ cannot be high enough to cause the firm to make do with any $E_h$. We will assume this for now and update it when we discuss the firm condition for case 2.

Therefore $\hat{E}_h (\eta_g)$ starts at $h$ (and potentially stays at $h$ for a neighborhood of zero) and decreases monotonically to a value less than 1, if the above assumption is satisfied. At the point at which the decreasing function $\hat{E}_h (\eta_g)$ crosses $\hat{E}_h (\eta_g)$, the assumption that we are in case 1 of the firm’s Bellman equation ceases to be valid because $\hat{E}_h \geq \hat{E}_h$. Define $\eta_g^2$ implicitly by setting $E_h = \hat{E}_h = \frac{(\lambda - 1)^{\frac{\nu}{\eta_g}}}{\eta_g}$ in (28). The following condition determines $\eta_g^2$:

\[
f \left\{ \left( 1 - F \left( \frac{(\lambda - 1)^{\frac{\nu}{\eta_g}}}{\eta_g^2} \right) \right) \left( \frac{(\lambda - 1)^{\frac{\nu}{\eta_g}}}{\eta_g^2} \right) \right\} = \frac{(1 - C)}{\eta_g^2} \int_{(\lambda - 1)^{\frac{\nu}{\eta_g}}}^{\lambda - 1)\frac{\nu}{\eta_g}} \left[ (1 - \theta) \lambda (E_h)^{\theta} - \frac{1}{2} (\lambda - 1)^2 \right] dF(E_h)
\]

\[- \frac{(1 - C)(1 - \theta)}{(\eta_g^2)^{(1 - \theta)}} \left( 1 - F \left( \frac{(\lambda - 1)^{\frac{\nu}{\eta_g}}}{\eta_g^2} \right) \right) \left( \frac{(\lambda - 1)^{\frac{\nu}{\eta_g}}}{\eta_g^2} \right)^\theta - \frac{(1 - C)(\frac{\nu}{\eta_g})^{\theta}}{(\eta_g^2)^{(1 - \theta)}} \left( 1 - F \left( \frac{(\lambda - 1)^{\frac{\nu}{\eta_g}}}{\eta_g^2} \right) \right) \left( \frac{(\lambda - 1)^{\frac{\nu}{\eta_g}}}{\eta_g^2} \right)^{(1 - \theta)}.
\]

We assume that the parameters of the model are such that $\eta_g^2 < 1$; that is, when $\eta_g^2 = 1$, the LHS of (29) is lower than the RHS.

**A.3 Firm Condition for Case 2**

Given that case 2 is the relevant one, equating the payoffs in (13) at the threshold yields:

\[Ev - f\eta_g\hat{E}_h = (1 - C) \left[ (1 - \theta) \lambda (\eta_g\hat{E}_h)^{\theta} - \frac{1}{2} (\lambda - 1)^2 \right].\]
We have assumed that \( \bar{E}_h > 1 \) as discussed above; if we are in this case, \( \bar{E}_h < h \), and therefore we can write the threshold condition as an equality. Integrate \( v \) over the two relevant segments, substitute for \( E v \) from the threshold condition, then simplify to obtain the firm’s condition for this case as:

\[
\begin{align*}
\text{LHS} \left( \bar{E}_h = 1; \eta_g = 1 \right) & = f; \quad \text{LHS} \left( \bar{E}_h = h; \eta_g = 1 \right) = f E (\bar{E}_h) > 0, \\
\text{RHS} \left( \bar{E}_h = 1; \eta_g = 1 \right) & = (1 - C) (1 - \theta) \lambda \left[ \int_1^{\bar{E}_h} (E_h)^\theta dF (E_h) - 1 \right]; \quad \text{RHS} \left( \bar{E}_h = h; \eta_g = 1 \right) = 0.
\end{align*}
\]

An intersection clearly exists. To ensure, \( \bar{E}_h > 1 \), we need \( f < \text{RHS} \left( \bar{E}_h = 1; \eta_g = 1 \right) \). Since we will be in case 2, rather than case 1, for this corner to be relevant we update the earlier assumption on \( f \) with the following assumption:

\[
f < (1 - C) (1 - \theta) \lambda \left[ \int_1^{\bar{E}_h} (E_h)^\theta dF (E_h) - 1 \right].
\]

Therefore, \( \bar{E}_h (\eta_g) \) starts at the case 1 value for \( \eta_g^2 \) and decreases monotonically to a value larger than 1, if the above assumption is satisfied. The firm condition depicted in Figure 4 is a composite of the two cases discussed thus far.
A.5 Dependence of $\bar{E}_h(\eta_g)$ on $f, C, \lambda$

In both (20) and (21), an increase in $f$ shifts the left hand side upward. Given that the right hand sides of these conditions are decreasing in $\bar{E}_h$ and independent of $f$, the points of intersection shift leftward; that is, $\bar{E}_h(\eta_g)$ decreases. In the neighborhood of $\eta_g = 0$, it still is the case that $\bar{E}_h(0) = h$; the firm schedule for $\bar{E}_h(\eta_g) < h$ shifts down with an increase in $f$.

For a given $\eta_g$, when the cost of entry, $C$, increases the right hand sides of both (20) and (21) shift downward. Given that the left hand sides are increasing in $\bar{E}_h$ and independent of $C$, the points of intersection shift leftward; that is, $\bar{E}_h(\eta_g)$ decreases. In the neighborhood of $\eta_g = 0$, it still is the case that $\bar{E}_h(0) = h$; the firm schedule for $\bar{E}_h(\eta_g) < h$ shifts down when there is an increase in $C$.

We can see from (17) that the $\bar{E}_h(\eta_g)$ threshold curve shifts right when $\lambda$ increases, and the thresholds $\eta_g^1$ and $\eta_g^3$ increase.

In both (20) and (21), the left hand sides – the marginal cost of firing and redrawing – do not depend on $\lambda$. Below, we provide sufficient conditions to ensure that the right hand sides are increasing in $\lambda$ for cases other than case 1a, where only interior growth rates are realized. Given the nature of the two sides discussed above, the points of intersection shift rightward; that is, $\bar{E}_h(\eta_g)$ increases. Thus $\bar{E}_h(\eta_g)$ shifts up for cases other than case 1a.

A mere examination of the RHS of (30) shows that it is increasing in $\lambda$. To see if this is true for the RHS of (28), we take the derivative with respect to $\lambda$. This derivative can be simplified to:

$$
\frac{(1-C)}{\eta_g} \int_{\bar{E}_h}^{h} \left[(1-\theta)(\eta_g)^\theta (E_h)^\theta - (\lambda - 1)\right] dF(E_h)
$$

This expression is not automatically positive; the fact that profits are positive when growth equals $\lambda$ does not imply that its derivative is positive. We need to make distributional assumptions to unambiguously sign the derivative. Anticipating the calibration, we assume a uniform distribution for productivity draws: $f(E_h) = \frac{1}{h-1}$. Carrying out the integration in the above derivative, and using the definition of $\bar{E}_h$ we can show that the expression is positive if:

$$
\frac{(1-\theta)}{(1+\theta)} (\eta_g)^\theta (h)^\theta > (\lambda - 1) - \frac{2\theta}{(1+\theta)} (\lambda - 1)^{1+\frac{1}{\theta}}.
$$

For this case (1b), $\bar{E}_h \leq h$, implies that $(\eta_g)^\theta (h)^\theta$ in the LHS cannot be smaller than $(\lambda - 1)$, and the RHS is largest at $\eta_g = 1$. Make these substitutions to get a stringent sufficient condition for the derivative to be positive as:

$$(\lambda - 1)^{\frac{1}{\theta}} > h.$$

Consistent with the sufficient condition in section (A.2), a low $\theta$ and a low $h$ are more likely to satisfy the above condition and cause the RHS of (28) to increase in $\lambda$. Given the characterization of $\bar{E}_h(\eta_g)$ discussed earlier, the firm condition increases (shifts upward) when $\lambda$ increases.
A.6 Household Condition for Case 1

Substituting the relevant expressions for the wages in (22), we find:

\[
\frac{1}{1 - F(E_h)} \left\{ \int_{E_h}^{\hat{E}_h} \log \left[ \theta (\eta_g)^{- \theta} (E_h)^{\theta} + \theta (E_h)^{2\theta} \right] dF(E_h) \right\} = \log \left[ \theta (\eta_g)^{- \theta} (E_h)^{\theta} \right] - \log (1 - \eta_g) - \log (s).
\]

Factoring out \((\eta_g)^{2\theta - 1}\) from the two integrals when \(\eta_g > 0\) – we will discuss the \(\eta_g = 0\) case separately – on the left hand side, transposing terms, and noting that \(1 - 2\theta > 0\), we obtain the expression that implicitly defines the household condition \(\hat{E}_h^{HH} (\eta_g)\) as:

\[
\frac{1}{1 - F(E_h)} \left\{ \int_{E_h}^{\hat{E}_h} \log \left[ \theta (\eta_g)^{- \theta} (E_h)^{\theta} + \theta (E_h)^{2\theta} \right] dF(E_h) \right\} = \log \left[ \theta (\eta_g)^{- \theta} (E_h)^{\theta} \right] - (1 - 2\theta) \log (1 - \eta_g) - \log (s). \tag{31}
\]

A.7 Characterizing \(\hat{E}_h^{HH} (\eta_g)\) for Case 1

A.7.1 Dependence on \(\hat{E}_h\)

The RHS of (25) is independent of \(\hat{E}_h\). Differentiate the LHS with respect to \(\hat{E}_h\) to get:

\[
\frac{\tilde{f}(E_h)}{1 - F(E_h)} \left\{ \int_{E_h}^{\hat{E}_h} \log \left[ \theta (\eta_g)^{- \theta} (E_h)^{\theta} + \theta (E_h)^{2\theta} \right] dF(E_h) \right\} = \log \left[ \theta (\eta_g)^{- \theta} (E_h)^{\theta} \right] - (1 - 2\theta) \log (1 - \eta_g) - \log (s).
\]

Since wages are higher in the maximum growth region (at \(E_h = \hat{E}_h\), interior and maximal growth wages are the same, and for higher \(E_h\), maximal growth wages are higher), the second term is larger than the first. Therefore the derivative is:

\[
\frac{\tilde{f}(E_h)}{1 - F(E_h)} \left\{ \int_{E_h}^{\hat{E}_h} \log \left[ \theta (\eta_g)^{- \theta} (E_h)^{\theta} + \theta (E_h)^{2\theta} \right] dF(E_h) \right\} = \log \left[ \theta (\eta_g)^{- \theta} (E_h)^{\theta} \right] - (1 - 2\theta) \log (1 - \eta_g) - \log (s).
\]

Therefore the LHS of (25) is increasing in \(\hat{E}_h\). Taken together with the flat RHS, this means that an intersection is likely to occur. However, where the intersection occurs depends on \(\eta_g\).

A.7.2 Dependence on \(\eta_g\)

It can directly be seen that the RHS of (25) is increasing in \(\eta_g\). To study the dependence of the LHS of (25) on \(\eta_g\), differentiate the terms within curly braces and simplify to obtain the derivative as:

\[
- \int_{E_h}^{\hat{E}_h} \frac{\theta^2 (\eta_g)^{- \theta - 1} (E_h)^{\theta}}{\theta (\eta_g)^{- \theta} (E_h)^{\theta} + \theta (E_h)^{2\theta}} dF(E_h) - \int_{E_h}^{\hat{E}_h} \theta (\eta_g)^{- 1} dF(E_h) < 0.
\]

That is, the LHS of (25) is decreasing in \(\eta_g\).
A.7.3 The Nature of $E_h^{HH} (\eta_g)$

The above analysis indicates that the LHS of (25) is strictly increasing and the RHS is independent of $\bar{E}_h$; the LHS is strictly decreasing with $\eta_g$ and the RHS is strictly increasing with $\eta_g$. These results indicate that the intersection of the two sides, which gives the threshold $\bar{E}_h$ necessary for any given labor supply $\eta_g$, is increasing in $\eta_g$.

First examine the case of $\eta_g \to 0$. Working with (31), it can be seen that both $LHS (\bar{E}_h = 1; \eta_g = 0)$ and $LHS (\bar{E}_h = h; \eta_g = 0)$ diverge to $\infty$. Examination of the RHS of (25) reveals that $RHS (\eta_g = 0) \to \log [\theta \beta h^\theta] - \log (s)$, a finite quantity. Since the benefit of general education exceeds the cost for any $\bar{E}_h$, the corner case of $\bar{E}_h = 1$ results. It is likely for a neighborhood of $\eta_g = 0$ that $LHS > RHS$ and $\bar{E}_h = 1$. When $\eta_g$ is very low, relative wages are so high that even an $\bar{E}_h = 1$ is enough to attract entry into general education.

The situation is reversed for $\eta_g \to 1$. We have that the is LHS of (25) well-defined at both endpoints of $\bar{E}_h$:

$$LHS (\bar{E}_h = 1; \eta_g = 1) = \begin{cases} \int_{\bar{E}_h}^1 \log \left[ \theta (E_h)^\theta + \theta (E_h)^{2\theta} \right] dF (E_h) + \int_{\bar{E}_h}^h \log \left[ \theta \lambda (E_h)^\theta \right] dF (E_h) \\ \log \left[ \theta (h)^\theta + \theta (h)^{2\theta} \right] \end{cases}$$

$$LHS (\bar{E}_h = h; \eta_g = 1) = \log \left[ \theta (h)^\theta + \theta (h)^{2\theta} \right].$$

When $\bar{E}_h = h$, given case 1 requires $\bar{E}_h \geq \bar{E}_h$, we need to set $\bar{E}_h = h$; L’Hospital’s rule is needed to evaluate the LHS. However, $RHS (\eta_g = 1) \to \infty$. Since the cost exceeds the benefit for any $\bar{E}_h$, the corner case of $\bar{E}_h = h$ results. It is likely for a neighborhood of $\eta_g = 1$ that $LHS < RHS$ and $\bar{E}_h = h$. When $\eta_g$ is very high, even the lowest ability agents have to obtain general education, and the disutility is so prohibitive that the $\bar{E}_h$, and thus the expected wage premium of generally educated agents, has to be very high to induce entry.

This analysis indicates that the $E_h^{HH} (\eta_g)$ curve starts at 1, (potentially) stays at 1 for an interval, increases up to $h$, and (potentially) stays at $h$ for an interval. At some $\eta_g = \eta_g^{HH}$ the $E_h^{HH} (\eta_g)$ curve will intersect with $\bar{E}_h (\eta_g)$. Define $\eta_g^{HH}$ implicitly, by setting $\bar{E}_h = \bar{E}_h = (\lambda-1)^+ / \eta_g^{HH}$ in (25):

$$= \log [\theta \beta h^\theta] + (1 - 2\theta) \log (\eta_g^{HH}) - (2 - \theta) \log (1 - \eta_g^{HH}) - \log (s).$$

Case 2, which is discussed below, becomes relevant beyond $\eta_g^{HH}$.

A.8 Household Condition for Case 2

Substituting the relevant expressions for the wages in (22), factoring out $(\eta_g)^{\theta-1}$ from the integral on the left hand side, and transposing terms we obtain the expression that implicitly defines the household condition $E_h^{HH} (\eta_g)$ in this case as:

$$\frac{1}{(1 - F (E_h))} \left\{ \int_{E_h}^h \log \left[ \theta \lambda (E_h)^\theta \right] dF (E_h) \right\} \leq \log [\theta \beta h^\theta] + (1 - \theta) \log (\eta_g) - (2 - \theta) \log (1 - \eta_g) - \log (s).$$
A.9 Characterizing $\bar{E}_{h}^{HH}(\eta_g)$ for Case 2

A.9.1 Dependence on $\bar{E}_{h}$

The RHS of (26) is independent of $\bar{E}_{h}$. The derivative of the LHS of (26) with respect to $\bar{E}_{h}$ is:

$$\frac{\bar{f}(\bar{E}_{h})}{(1 - F(E_{h}))} \left\{ \frac{1}{(1 - F(E_{h}))} \int_{E_{h}}^{h} \log \left[ \theta \lambda (E_{h})^\theta \right] dF(E_{h}) - \log \left[ \theta \lambda (\bar{E}_{h})^\theta \right] \right\} > 0.$$  

A.9.2 Dependence on $\eta_g$

The LHS of (26) is independent of $\eta_g$. Mere examination reveals that the RHS is increasing in $\eta_g$.

A.9.3 The Nature of $\bar{E}_{h}^{HH}(\eta_g)$

The above analysis indicates that the LHS of (26) is strictly increasing and the RHS is independent of $\bar{E}_{h}$; the LHS is independent of $\eta_g$ and the RHS is strictly increasing with $\eta_g$. These results indicate that the intersection of the two sides, which gives the threshold $\bar{E}_{h}$ necessary for any given labor supply $\eta_g$, is increasing in $\eta_g$.

Since both sides of the expression in (25) and (26) coincide at $\eta_g = \eta_g^{HH}$, the $\bar{E}_{h}^{HH}(\eta_g)$ schedule is continuous at $\eta_g^{HH}$. To characterize the relationship in the interval $[\eta_g^{HH}, 1]$, we compute for (26):

$$LHS (\bar{E}_{h} = \bar{E}_{h}^{HH}) = \frac{1}{(1 - F(\bar{E}_{h}^{HH}))} \left\{ \int_{E_{h}}^{\bar{E}_{h}^{HH}} \log \left[ \theta \lambda (E_{h})^\theta \right] dF(E_{h}) \right\} > \log \left[ \theta \lambda (h)^\theta \right]$$

$$LHS (\bar{E}_{h} = h) = \frac{\log \left[ \theta \lambda (h)^\theta \right] \bar{f}(h)}{-\bar{f}(h)} = \log \left[ \theta \lambda (h)^\theta \right],$$

where L'Hospital’s rule has been used for the second evaluation. The LHS is independent of $\eta_g$ and is an increasing function in $\bar{E}_{h}$, with finite endpoints.

However, the RHS ($\eta_g = 1$) diverges to $\infty$. Therefore, when $\eta_g = 1$, the cost of general education exceeds the benefit for any $\bar{E}_{h}$, and $\bar{E}_{h} = h$ results; this is likely to be true for a neighborhood of $\eta_g = 1$.

A.10 Dependence of $\bar{E}_{h}^{HH}(\eta_g)$ on $\lambda$

Take the derivative of the terms in curly braces in the LHS of (25), noting $\bar{E}_{h} = \frac{(\lambda - 1)\eta_g}{\lambda}$, with respect to $\lambda$. This derivative is $\int_{E_{h}}^{h} \frac{1}{\lambda} dF(E_{h}) > 0.$
References


47