.

Comparative Advantage and Education Subsidization

De-Xing Guan[1]

December 2005

Abstract

Education is usually subsidized by the government because human capital accumulation would have positive external effects from a social viewpoint. But it is *average* human capital that generates externalities in most endogenous growth models. We argue that different *distributions* or *compositions* of human capital in a society will have different mechanisms to propagate the external effects of education, and would have very different implications for long-run growth rates of a country. Some industries would depend on a few distinguished workers, but others would rely mostly on the team work. The allocation of education subsidies would therefore depend on the *comparative advantage* of a country in producing human capital through education. Using most of the government funds to subsidize a handful of industrial or academic *superstars*, as done by many developing (and some developed) countries, is not always the best way to foster economic growth.

[1]Department of Economics, National Taipei University. This draft is very preliminary and incomplete. Comments are welcome and can be sent to guan@mail.ntpu.edu.tw.

1. Introduction

Education is usually subsidized by the government in almost all countries in the world. The foundation of education subsidization rests mostly on the positive external effects that education might generate through human capital accumulation. Everyone would agree that we should subsidize education, but the question is *how* to allocate government funds to education? Should primary and secondary schools get more funds, or should tertiary education share more of them? Should the money go to more ordinary universities and professors, or just to a handful of *superstars* (Rosen (1981)): instituions or persons that have outstanding publications? Should the government subsidize schools or directly subsidize the students themselves?

To answer these questions we must first *identify* what are the externalities generated by education, and how to measure the effects of them. At least back to Gary Becker's (1993) treatise on *Human Capital*, economists have paid much attention to the measurement of the external effects of human capital accumulation. Though the returns to private investment in education are significant, the social returns to education remain ambiguous and inconclusive. For examples, Becker (p. 249) found that "Only a limited amount could be said about the social gains from education because of ignorance about the external effects. This ignorance is closely connected with ignorance about the "residual" in calculations of the contribution of various factors to growth."

Mincer (1974) tried to measure the returns of education (as well as experience) by estimating the so-called *Mincerian wage equation*, but his result has upward bias towards the positive externalities of human capital accumulation, as argued recently by Ciccone and Peri (2005). Acemoglu and Angrist (2001) claimed that the private returns to education are about seven percent, but the social returns are less than one percent, and not significantly different from zero. Krueger and Lindahl (2001, p. 1130) also found that "From the micro evidence, however, it is unclear whether the social return to schooling exceeds the private return,...The macro-economic evidence of externalities in terms of technological progress from investments in higher education seems to us more fragile,..."

In spite of the above ambiguity about the social returns of education, most people still believe that education is good for the society as a whole, because "A stable and democratic society is impossible without a minimum degree of literacy and knowledge on the part of most citizens and without widespread acceptance of some common set of values. Education can contribute to both...The education of my child contributes to your welfare by promoting a stable and democratic society (Friedman (2002, p. 86))."[2] Of coure it is hard to measure empirically how large this *cultural* externality could be, but it is no doubt that there are positive externalities from education in a broader sense.

Now if the government tries to subsidize education to promote long-run economic growth, ignoring any cultural effects, how should it do to achieve this goal? Most endogenous growth

[2] Friedman also suggested that it is better to subsidize students rather than schools by giving *education vouchers* to parents to guarantee a minimum education level.

theories use *average* human capital as a proxy to represent the external effects of education. Lucas (1988) was a standard example. Though this assumption is useful in some respects, it is of limited use to study the *distributional* or *compositional* effects from human capital accumulation.

For example, in a recent article Aghion and Howitt (2005), using the idea of Acemoglu, Aghion and Zilibotti (2006), argued that if there are two strategies to adopt new technologies: innovation and imitation, then countries with different composition of skilled and unskilled workers would have different methods to promote growth. For countries with relatively abundant skilled workers, they have less *distance* to the technological frontier, and hence should devote more resources to innovation. On the other hand, countries with more unskilled labor, and more distant from the technology frontier, should focus on imitation. So the composition of human capital in a country matters for economic growth.

On the other hand, it is usually thought that advanced countries should do more research work, and the developing countries ought to be followers to imitate technologies discovered by the leaders. This is the standard image when people talk about comparative advantage among countries. Is this image always right? Take a recent example in the cold Baltic country Estonia. Two young men, Jaan Tallinn and Sten Tamkivi, have developed a software *Skype* for free calls over the Internet. According to the usual image, people in Estonia should be imitators, and hardly be innovators. But the case of Skype tremendously changes the viewpoint of ordinary people that innovation is usually the comparative *disadvantage* for small developing countries.

In many countries the government allocates more, or even most, education subsidies to some outstanding universities or persons. Peking University, the most outstanding university in China, has been receiving a seven-year, 3.6 billion (yuan) funding from the Chinese government to subsidize research and to recruit more distinguished scholars from abroad. The government of Taiwan has also granted a five-year, 50 billion (NT dollars) subsidy to some selected universities around the island. A more recent example came from the South Korea's superstar in stem cells research, Hwang Woo Suk. Professor Hwang is famous for cloning a dog called "Snuppy" and has been heavily subsidized by the Korean government (65 million US dollars since 1998), but resigned in late December from the Seoul National University for his scandal in faking data in a paper published in the journal *Science*.

Is subsidizing a handful of superstars always the best way to maximize economic growth? [3] One Nobel Laureate plus ninty-nine high school graduates is not productively equivalent to one hundred workers who all have college degrees. Which one is more productive depends on the *comparative advantage* of the production and international trade structure in a country. But what are the key factors to determining a country's comparative advantage?

Grossman (2004) studied this issue in an environment of imperfect labor contract. If contracts are imperfect, then "...that national differences in the distribution of talent can be an independent source of comparative advantage..." (p.211). He used Japan and the United

[3] Indeed there are no strong evidences that education would cause economic growth. Some authors even claimed that the causality is the other way around: it is economic growth which causes more schooling (Bils and Klenow (2000)).

States as an example to illustrate how the distribution of talent matters for comparative advantage in these two countries. Because on average Japan has a more homogeneous labor force than the United States, if there are two sectors, say automobiles and software, then Japan has the comparative advantage in producing automobiles and the United states in making software. This is because automobile production is a large-scale manufacturing which has the disadvantage in attracting the most talented individuals as compared to software production. In other words, automobiles need more team work than software does.

Another kind of theory about comparative advantage came from the differences in total factor productivity (TFP) or technology frontier. Using David Ricardo's trade theory, Eaton and Kortum (2002), which was further extended to a general equilibrium framework by Alvarez and Lucas (2005), proposed an interesting model where the comparative advantage of a country was identified with the variance of an exponentially (or extreme value) distributed technology frontier. Unlike Acemoglu, Aghion and Zilibotti (2006) and Aghion and Howitt (2005), where their technology frontiers are deterministic and are represented by the maximum level of all technologies available, both Eaton and Kortum (2002) and Alvarez and Lucas (2005) adopted a stochastic version of technology frontier or TFP.

As is well known and empirically confirmed, the long-run aggregate production function is approximately Cobb-Douglas[4], and it was shown by Jones (2005) that Cobb-Douglas function in the balanced growth path (BGP) is the limit of two possible distributions of ideas. If there is an infinite number of discovery of ideas then the assumption that the ideas follow a Pareto distribution will lead to a steady state Cobb-Douglas aggregate production function. Or, if the number of ideas is finite, then it is Poisson distribution for ideas to converge to Cobb-Douglas in the long run. In either case the stochastic version of TFP has more solid foundation than the deterministic one (Acemoglu, Aghion and Zilibotti (2006) and Aghion and Howitt (2005)) would have because it is more empirically oriented.

In this paper we propose a stochastic version of the distance to technology frontier, where education subsidization could be discussed based on the comparative advantage of a country. As argued by Acemoglu, Aghion and Zilibotti (2006) and Aghion and Howitt (2005), when industries in a country are approaching the technology frontier, it is more efficient to innovate the new technologies or ideas by themselves, but when they are far away from the frontier, the better strategy is to imitate the ideas discovered in the frontier by other industries in more advanced countries.

One shortcoming of their setup is that they used the distance to the highest level of the technologies available, that is, the world technology frontier, and this is an one-parameter (that is, the maximum value) representation of the distribution of new ideas. It is convenient but obviously it could not capture the variance of ideas distribution, which has been emphasized by Eaton and Kortum (2002) and Alvarez and Lucas (2005) in representing comparative advantage of countries. And comparative advantage is the key to understanding how government should allocate its subsidies to various institutions and persons more efficiently to promote long-run economic growth.

[4] This is the so-called *Steady-State Growth Theorem*: if a neoclassical growth model exhibits steady state growth with constant and positive factor shares, then either the production function is Cobb-Douglas or technical change is labor-augmenting. See Jones (2005, p. 525).

If we want to discuss the effects of the whole distribution of human capital (or ideas) on technological changes, we should use all the moments it has, not just one of them. For examples, Lucas (1988) used only the first moment (average value) of the human capital to represent human capital externality, and Acemoglu, Aghion and Zilibotti (2006) and Aghion and Howitt (2005) used the maximun value of world technology to represent the technology frontier, the ratio of it and the current technology level is the parameter they have called the *distance to the frontier*. Both examples ignored the role played by the second moment, that is, variances.

In this paper we borrow from the statistics literature (for example, Oller (1987)) to construct a theoretical model, where a distance between probability distributions can be used to measure the distance between the distributions of human capital and ideas in different sectors or countries. In such a model both mean and variance could have effects on the distance to technology frontier as well as the comparative advantage of countries. We hope that this could give us a better framework to discuss the relations between comparative advantage and education subsidization.

The paper is organized as follows: in Section 2, we briefly review the setups of Acemoglu, Aghion and Zilibotti (2006) and Aghion and Howitt (2005), and also Eaton and Kortum (2002) and Alvarez and Lucas (2005), which act as the starting point of our approach. Section 3 presents our model. We use a cost-minimization approach instead of the profit-maximization one adopted by Aghion and Howitt (2005) because the former can better capture the decision problems firms would face in their choice of innovation and imitation strategies. Section 4 summarizes.

2. Distance to Technology Frontier and Comparative Advantage

Assume that a typical country has two kinds of labor: skilled and unskilled. The numbers of which can be represented by S and U, as used by Aghion and Howitt (2005). Aghion and Howitt asked the question: whether European countries should pay more attention and also more funding to higher education in order to catch up with the United States? Obviously the education strategy and subsidization policy are different in these two areas. So the authors claimed that they would have different strategies in promoting economic growth, if we take into account the comparative advantage they have in both areas. Aghion and Howitt assumed the following TFP (or, indirectly, human capital) accumulation function for a country:

$$\dot{A} = u_m^{\sigma} s_m^{1-\sigma} \bar{A} + \rho u_n^{\phi} s_n^{1-\phi} A \qquad (1)$$

where $\rho > 0$, $0 < \sigma < 1, 0 < \phi < 1, A$ is the average technology level (or TFP) this country has, $\bar{A}$ is world technology frontier, $\dot{A}$ is the time derivative of A, $u_m(u_n)$ is the number of workers with primary/secondary education used in imitation (innovation), and $s_m(s_n)$ is the number of workers with tertiary education used in imitation (innovation).

Aghion and Howitt have implicitly assumed that workers with lower (primary/secondary) education would be better at producing goods through imitation, and workers with higher (tertiary) education are better at producing goods by innovating a new technology themselves. To capture this idea they further assume $\sigma > \phi$. In other words, the productivity would be higher if skilled workers (with higher education) are allocated to the production of goods using innovation, and if unskilled workers (with lower education) go to produce goods by imitation. The representative firm in this model tries to maximize profit:

$$\delta(u_m^{\sigma} s_m^{1-\sigma} + \rho u_n^{\phi} s_n^{1-\phi} a) - w_u(u_m + u_n) - w_s(s_m + s_n) \tag{2}$$

subject to $u_m + u_n = U$, $s_m + s_n = S$, and $a = A/\bar{A}$, the inverse of the distance to technology frontier. δ is the price of goods produced by imitation, $\delta\rho a$ is the price of goods produced by innovation and w_u, w_s are the wage rates of unskilled and skilled workers, respectively. They found that "...a marginal increase in the fraction of workers with higher education enhances productivity growth all the more the closer the country is to the world technology frontier." (p. 21) They interpreted this result as a version of the famous *Rybczynski Theorem* in international trade: "...a marginal increase in the supply S of highly educated workers leads to an even greater number of skilled workers being employed in innovation." (p. 21)

As mentioned in Section 1, Aghion and Howitt (2005) treated the technology frontier as the maximum of available technologies, which is deterministic and has exogenously given growth rates. Therefore it could not adequately address the questions of comparative advantage and the aggregate technology frontier at the BGP, which is usually characterized by the Cobb-Douglas production function.

On the other hand, Eaton and Kortum (2002) and Alvarez and Lucas (2005) have used a different framework to discuss the relationship between comparative advantage and technical changes. They assumed that in each period the technology or TFP is drawn from an exponential (Gumbel, or Type-I extreme value) distribution. Specifically, Eaton and Kortum assumed the TFP has a Frechet (or Type-II extreme value, which can be transformed into exponential distribution through change of variables) distribution as follows[5]:

$$F_i(z) = \exp(-T_i z^{-\theta}) \tag{3}$$

where $F_i(z) = \Pr(Z_i < z)$, Z_i is country i's TFP, $T_i > 0$, and $\theta > 1$ are two parameters governing the distribution of TFP. In terms of Eaton and Kortum, T_i represents the absolute advantage of country i, θ determines the comparative advantage among countries, and "...a lower value of θ, generating more heterogeneity, means that comparative advantage exerts a stronger force for trade..." (p. 1747) Alvarez and Lucas (2005) developed a general equilibrium version of the Eaton-Kortum model, where the comparative advantage was also governed by θ.

To sum up, Aghion and Howitt (2005) and Acemoglu, Aghion and Zilibotti (2006) have used the idea of distance to the technology frontier to capture the comparative advantage between skilled and unskilled workers. Because this kind of comparative advantage is deterministic and exogenously given (because of their assumption that $\sigma > \phi$), it could not adequately address some of the empirical findings at the BGP, as suggested by Jones (2005). In contrast, Eaton and Kortum (2002) and Alvarez and Lucas (2005) adopted a stochastic version of

[5]For a brief but useful introduction to the extreme value distributions, see Billingsley (1995, pp. 195-197).

technical changes. They found that the variance of exponential distribution, which TFP progress was assumed to follow, determined the comparative advantage of a country, and this result has been more consistent with the theoretical results of Jones (2005) and many others.

In what follows we would like to propose a new definition of the distance to technology fontier, where it can accommodate the original definition of the distance by Aghion and Howitt (2005) and Acemoglu, Aghion and Zilibotti (2006) to the stochastic environment of Eaton and Kortum (2002) and Alvarez and Lucas (2005).

3. A Simple Model of TFP

In this section we relate human capital accumulation with TFP and its growth. As more and more empirical evidences have shown, the quality of institutions of a country is very important for long-run economic growth. Usually it is not easy to distinguish between TFP and institutions because both variables measure the productive efficiency from inputs to outputs. We do not propose to do this in this paper either. Instead we use educational externalities, stemming from human capital accumulation, to approximate TFP or institutions of a country. This is similar to the setup of Lucas (1988), among others, but the main difference between Lucas and this paper is that we use the whole human capital distribution to generate externalities, rather than the average human capital level, as done by Lucas. Of course, we are not the first one to adopt such assumption. Ciccone and Peri (2005, Appendix A.7) was another example.

The foundation of introducing other moments of human capital distribution into the model to represent human capital externalities rests at least on two observations. First, as mentioned above, the variance of technology distribution, and hence variance of human capital distribution (their relationship will be discussed later), would be an important factor in determining comparative advantage of a country. Second, the diversity of talents in a society, as emphasized by Grossman (2004), would contribute to the comparative advantage in producing tradeable goods, even between advanced countries such as automobile industries in Japan and the software industries in the United States. Some industries or countries are better at working as a team, but others might rely on some outstanding individuals or superstars. No matter based on which foundation, a stochastic version of TFP (and human capital) distribution is needed.

Denote the probability distribution of TFP in a typical country by $f(A)$. In the frameworks of Aghion and Howitt (2005) and Acemoglu, Aghion, and Zilibotti (2006) it is $\bar{A}$, the mean value of the world technology frontier, which would enter the production function, not the whole distribution of TFP. But we would like to use moments other than the mean value of human capital and hence TFP distribution to capture the external effects which could be lost by only using the first moment of the distribution. This will distinguish the present model from that of others. Let

$$f(A) = g[f(h)] = [f(h)]^{\gamma} \tag{4}$$

where the parameter $\gamma > 0$, representing the positive externalities generated by human capital. There are three special cases of equation (4).

Case1.

In a deterministic framework, if $f(h) = (\int_0^\infty hN(h)dh)/(\int_0^\infty N(h)dh)$, where the "number" of workers with human capital level h is $N(h)$, then this reduces to the case of Lucas (1988), where it is the average level of human capital that would have external effects.

Case 2.

If a deterministic function $f(h) \in \arg\max\{g[f(h)]\}$, then this is the case of Aghion and Howitt (2005) and Acemoglu, Aghion and Zilibotti (2006), where it is the mean value of TFP drawn from the world technology frontier (that is, $\bar{A}$) that matters for external effects of human capital.

Case 3.

This is the case we will use in the present paper. Let $f(h)$ be the probability density that will generate human capital externalities, and the strength is governed by γ. If we further assume that $f(h) = \lambda\gamma h^{\gamma-1}\exp(-\lambda h^{\gamma})$, $\lambda > 0$, which is a Weibull (or Type-III extreme value) distribution, then $f(A) = g[f(h)] = [f(h)]^{\gamma}$ would follow an exponential distribution [6], as assumed by Eaton and Kortum (2002) and Alvarez and Lucas (2005).

Assume that there are many representative firms in a typical country, the final good of this country is produced by two kinds of workers, skilled and unskilled, and the supply of these workers are fixed at S and U respectively, the same as assumed in Aghion and Howitt (2005). There is no physical capital, and the only capital good is human capital. All markets are supposed to be perfectly competitive. In each period each firm faces the problem: to innovate a new technology by itself, or to imitate the highest level of technology provided by other firms in this country or in some other countries in the world.

There is a tradeoff between these two strategies of technology adoption. The benefit of innovation is that the firm could capture most of the rent created by the innovation in terms of, say, patents. But there are innovation costs (R&D costs and other sunk costs) which might make firms hesitating about whether they should spend the money to innovate a new technology by themselves. On the other hand, imitation has its advantage because firms can more cheaply get the technology they want without investing money to do their own research. The costs of imitation are twofold. First, there would be so many firms in this pool because being a leader is more costly than being followers, and this means that the profits would be lower for the market is more competitive. In this paper, we call competing in the market of innovation the *competition at the extensive margin*, and competing in the market of imitation the *competition at the intensive margin*. Firms would like to choose the margin

[6] Johnson, Kotz, and Balakrishnan (1994, p. 551).

in which they want to compete to minimize their costs, or equivalently, to maximize their profits. [7]

The second disadvantage of imitation is the so-called *political trap* emphasized by Acemoglu, Aghion and Zilibotti (2006). They argued that when firms in a country is used to imitating others, the government would not have incentives to create a better infrastructure to attract more advanced technologies, and the country in question would more and more rely on the ideas or technologies discovered by other countries. This would make this country more dependent on foreigners and hence more difficult to catch up with the more advanced countries.

For simplicity, consider there are two decision stages in each period, and firms only minimize one-period costs or maximize one-period profits. The only intertemporal choice of firms is to make decisions on technology adoptions. Human capital will promote economic growth through its role in generating external effects on technology or TFP accumulation. Besides, the only factor of production is labor. Population or the supply of labor is a constant. Because all markets are assumed to be competitive, in long-run equilibrium the profits would be zero, and firms would operate at the minimum value of the long-run average cost curve (LAC).

When choosing innovation strategy, or competition at the extensive margin, the total cost of a representative firm is assumed to be

$$C_E = c_n(d[f(\bar{A}), f(A)], ...)Y_n, \tag{5}$$

and when choosing imitation strategy, or competition at the intensive margin, the total cost of a representative firm is assumed to be

$$C_I = c_m(d[f(\bar{A}), f(A)], ...)Y_m, \tag{6}$$

where $c_n(d[f(\bar{A}), f(A)], ...)$ is the long-run average cost if firms adopt innovation strategy in the first decision stage, $c_m(d[f(\bar{A}), f(A)], ...)$ is the long-run average cost if firms adopt imitation strategy, Y_n is output of innovating firms, and Y_m is output of imitating firms. The costs of using innovation or imitation strategies are assumed to be functions of the distance between two probability distributions $f(\bar{A})$ and $f(A)$: $d[f(\bar{A}), f(A)] \geq 0$, where $f(\bar{A}) = g[f(\bar{h})] = [f(\bar{h})]^{\gamma}$ is the distribution from which world technology is drawn and $f(A) = g[f(h)] = [f(h)]^{\gamma}$ is the distribution from which the technology of firms can be drawn. It is assumed that the technology can be represented by the external effects of human capital, where $f(\bar{h})$ is the distribution of world human capital, $f(h)$ is the distribution of this country's human capital, and $\gamma > 0$ indicates the positive external effects of human capital.

The distance between two TFP distributions need more comments. The LACs are usually functions of factor prices and output levels. The novelty here is that we introduce the

[7] This observation is not unique in economics. A similar concept could be found in business literature. For example, in a recent book Kim and Mauborgne (2005) used the term *blue ocean strategy* to indicate the idea of creating uncontested markets to make competition irrelevant. This is similar to our concept of the competition at the extensive margin. And competition in the intensive margin is similar to the original *red ocean strategy.*

concept of distance between distributions to capture the degree of relative ease in adopting technologies. If the technology level of firms is far away from that of the world frontier, then the distance between the distributions $f(\bar{A})$ and $f(A)$ becomes larger, and it is more costly for firms to innovate new technologies for producing goods, because it takes them more funds and efforts to do research work. Therefore we assume that $\partial c_n/\partial d[f(\bar{A}), f(A)] > 0$. On the other hand, if the technology level of firms is closer to that of the world frontier, then the distance between the distributions $f(\bar{A})$ and $f(A)$ becomes smaller, and it is more costly for firms to imitate new technologies for producing goods, because otherwise the firm will lose the opportunity to earn the potential monopoly rent created by the new technology it innovates. So we would assume $\partial c_m/\partial d[f(\bar{A}), f(A)] < 0$. For simplicity, let both LACs be linear functions of the distance. Specifically,

$$c_n(d[f(\bar{A}), f(A)], ...) = \mu_n\{1+d[f(\bar{A}), f(A)]\} \tag{7}$$

$$c_m(d[f(\bar{A}), f(A)], ...) = \mu_m\{\alpha-\{1+d[f(\bar{A}), f(A)]\}\} \tag{8}$$

where $\mu_n > 0, \mu_m > 0$ are both positive coefficients, and assume that $\alpha > 1 + d[f(\bar{A}), f(A)]$. In order to be consistent with competitive equilibrium and also consistent with the long-run Cobb-Douglas result for aggregate production functions, as suggested by Jones (2005), we would assume that Y_m and Y_n are both Cobb-Douglas:

$$Y_m = \bar{A}u_m^{\sigma}s_m^{1-\sigma} \tag{9}$$

$$Y_n = Au_n^{\phi}s_n^{1-\phi} \tag{10}$$

where as in Section 2, $u_m(u_n)$ is the number of workers with primary/secondary education used in imitation (innovation), $s_m(s_n)$ is the number of workers with tertiary education used in imitation (innovation). But the point here is that we do not need the condition that $\sigma > \phi$, as required by Aghion and Howitt (2005). This is important for thinking about policies of education subsidization because, as argued by Grossman (2004), a country with more homogeneous workers would have less volatile human capital, and would have the comparative advantage in producing goods which need more team work, such as the automobile industries in Japan. The variability of work force could be captured by the second moment (variance) of human capital distribution, so if we ignore the effects of variance we would probably get the wrong answer about how to subsidize different industries or persons.

At the first decision stage each firm must decide whether to innovate or to imitate technologies. After choosing technology adoption strategy at the extensive/intensive margins, firms begin to produce final goods according to the Cobb-Douglas technology available to them at the second decision stage. We assume that these two decisions are accomplished within the same period. We therefore have the problem for firms: given w_u, w_s, firms try to minimize total costs in choosing technology adoption strategies at either extensive or intensive margins, subject to equations (1), (4), (7), (8), (9), and (10). The competitive equilibrium in this model can be stated more clearly as follows:

Definition

Given $f(h), f(\bar{h}), d[f(\bar{A}), f(A)]$, a competitive equilibrium is a sequence of numbers of workers and wage rates: $\{u_m, u_n, s_m, s_n, w_u, w_s\} \in \arg\min\{C_E, C_I\}$, such that both unskilled and skilled labor markets clear: $u_m + u_n = U$, $s_m + s_n = S$.

This equilibrium is generally not Pareto optimal because both workers and firms could not capture all the benefits they generate through positive human capital externalities. Government should subsidize workers whose efforts will maximize outputs or minimize LACs. The solution to firms' problem is standard: to equalize the total costs of producing at extensive and intensive margins such that there would be no further arbitrage opportunities for workers to switch between sectors. When the costs of choosing technology at the extensive margin are lower, workers will switch to firms adopting innovation strategy. This will raise the labor costs of firms at extensive margin such that their comparative advantage will eventually disappear. The same argument applies equally to the case of firms choosing technology adoption strategy at the intensive margin. Therefore, in the long-run equilibrium, cost minimization requires that $C_E = C_I$, or

$$\mu_n\{1+d[f(\bar{A}), f(A)]\}Au_n^{\phi}s_n^{1-\phi} = \mu_m\{\{\alpha-\{1+d[f(\bar{A}), f(A)]\}\}\bar{A}u_m^{\sigma}s_m^{1-\sigma} \tag{11}$$

Because all markets are competitive, free entry and exit would imply that the wage rates of unskilled workers are equalized when they produce goods for firms choosing technology adoption strategies at either extensive or intensive margins. The same argument applies to the skilled workers, too. This means that the marginal product of labor (MPL) for unskilled workers (as well as skilled workers) would be the same no matter they work for firms at extensive or intensive margins. These are mathematically equivalent to

$$\phi A u_n^{\phi-1}s_n^{1-\phi} = \sigma\bar{A}u_m^{\sigma-1}s_m^{1-\sigma} \tag{12}$$

$$(1-\phi)Au_n^{\phi}s_n^{-\phi} = (1-\sigma)\bar{A}u_m^{\sigma}s_m^{-\sigma} \tag{13}$$

Using equations (11)-(13), together with equation (1) and labor market clearing conditions: $u_m + u_n = U$, $s_m + s_n = S$, we can determine six endogenous variables: $u_m, u_n, s_m, s_n, w_u, w_s$. When we get the equilibrium values of these variables, the TFP growth rates at the BGP, say $g = \dot{A}/A$, could be determined by equations (1), (9), and (10):

$$g = (u_m^*)^{\sigma}(s_m^*)^{1-\sigma}\tfrac{\bar{A}}{A}\{1+\rho(\tfrac{\mu_m}{\mu_n})(\tfrac{\alpha-\{1+d[f(\bar{A}),f(A)]\}}{1+d[f(\bar{A}),f(A)]})\} \tag{14}$$

where

$$u_m^* = \{U-S[(\tfrac{1-\sigma}{\sigma})(\tfrac{\phi}{1-\phi})]\{(\tfrac{\phi A}{\sigma\bar{A}})[(\tfrac{1-\sigma}{\sigma})(\tfrac{\phi}{1-\phi})]^{\phi-1}\}^{1/(\sigma-\phi)}\}\{1-[(\tfrac{1-\sigma}{\sigma})(\tfrac{\phi}{1-\phi})]\}^{-1} \tag{15}$$

$$s_m^* = u_m^*\{\{(\tfrac{\phi A}{\sigma\bar{A}})[(\tfrac{1-\sigma}{\sigma})(\tfrac{\phi}{1-\phi})]^{\phi-1}\}^{1/(\sigma-\phi)}\}^{-1} \tag{16}$$

As argued by Aghion and Howitt (2005, p. 20), there is a Rybczynski effect in their model provided that $\sigma > \phi$. Mathematically, this requires that $\partial^2 g/\partial a\partial S > 0$ if Rybczynski theorem is satisfied. This means that if the supply of skilled labor (S) increases, it will increase the marginal effects of the inverse of the distance from technology frontier on TFP growth rate at the BGP (that is, $\partial g/\partial a$). A larger a indicates a smaller $1/a$, which is equivalent to a smaller $\bar{A}/A$, and the TFP level of the country in question would be closer to the world technology frontier. In this circumstance, the skilled workers would have comparative advantage in producing final goods. And the increase in their supply would enhance this advantage in terms of Rybczynski.

The Rybczynski effect still holds in the present model since, from equation (14), we have both $\partial^2 g/\partial a \partial S > 0$ and $\partial^2 g/\partial d[f(\bar{A}), f(A)]\partial S > 0$. Therefore with either definition of the distance to technology frontier, the more abundant the skilled labor is, the higher productive efficiency firms that operate at the extensive margin would have. But the economic interpretation is quite different here. In Aghion and Howitt (2005) and Acemoglu, Aghion, and Zilibotti (2006) the distance to technology frontier was measured by the distance between the *mean values* of two kinds of TFP, that is, between $\bar{A}$ and A. In our framework, the distance is measured by the distance between the *distributions* of TFP. A larger difference between $\bar{A}$ and A does not necessarily mean a larger difference between $f(\bar{A})$ and $f(A)$, or $d[f(\bar{A}), f(A)]$.

If it is postulated that human capital for both typical country and the whole world follow a Weibull distribution, as in Case 3 above, $f(h) = \lambda\gamma h^{\gamma-1}\exp(-\lambda h^{\gamma})$ and $f(\bar{h}) = \lambda\gamma\bar{h}^{\gamma-1}\exp(-\lambda\bar{h}^{\gamma})$, $\lambda > 0$, then $f(A) = g[f(h)] = [f(h)]^{\gamma}$ and $f(\bar{A}) = g[f(\bar{h})] = [f(\bar{h})]^{\gamma}$ would follow an exponential distribution.Then both the mean values and variances of TFP distribution would influence $d[f(\bar{A}), f(A)]$, and henceforth would have effects on the long-run productivity growth rate g. Because $d[f(\bar{A}), f(A)]$ depends in a quite complicated way on the mean and variance of TFP distribution, we could not reach any monotonic relationship between $d[f(\bar{A}), f(A)]$ and these two moments. Nevertheless it is obvious that variance does have effects on long-run technology growth, which was absent in many endogenous growth models such as Lucas (1988) and Aghion and Howitt (2005). A concrete example will be useful to show the difference between this paper and Aghion and Howitt (2005), among others.

Example

In this example we take the externality parameter to be unity ($\gamma = 1$) such that the solution to the TFP growth rate at the BGP would be much simpler and more tractable. When $\gamma = 1$, $f(A) = f(h)$. Following the assumption of Eaton and Kortum (2002), let

$$F(\bar{A}) = \exp(-T_1\bar{A}^{-\theta_1}) \tag{17}$$

$$F(A) = \exp(-T_2 A^{-\theta_2}) \tag{18}$$

be the cumulative density functions for $\bar{A}$ and A respectively. The economic interpretation of parameters $T_1, T_2, \theta_1, \theta_2$ are the same as those in equation (3). Then the distance $d[f(\bar{A}), f(A)]$ would be (Oller(1987, p. 20)):

$$d[f(\bar{A}), f(A)] = \tfrac{\pi}{\sqrt{6}} \log \tfrac{1+\delta}{1-\delta} \tag{19}$$

where

$$\delta = \left\{ \frac{[\log(T_1/T_2)+(1-\beta)(\theta_2-\theta_1)/\theta_1\theta_2]^2+(\pi/\sqrt{6})^2(\theta_2-\theta_1)^2/\theta_1^2\theta_2^2}{[\log(T_1/T_2)+(1-\beta)(\theta_2-\theta_1)/\theta_1\theta_2]^2+(\pi/\sqrt{6})^2(\theta_2+\theta_1)^2/\theta_1^2\theta_2^2} \right\}^{\frac{1}{2}} \tag{20}$$

β is the Euler constant (= 0.577...), and π is 3.14159...

As argued in Section 2 and by Eaton and Kortum (2002) and Alvarez and Lucas (2005), the parameters θ_1 and θ_2 govern the comparative advantage of a country. A smaller θ_2 indicates, for example, that the typical country has comparative advantage in producing final goods. But here this advantage should be compared with that in the distribution of world technology frontier. In other words, θ_1 should also be taken into account, not only θ_2, as in previous studies.

A simple calculation would show that a smaller θ_1, or a larger θ_2, induces a larger δ, and therefore a greater $d[f(\bar{A}), f(A)]$, and this in turn results in a lower TFP growth rate g of the typical country. It is the *relative* magnitudes of the variance parameters that matter. And even different countries have the same absolute advantage in producing goods ($T_1 = T_2$), it is comparative advantage that determines which country would produce what kind of goods.

Now we can compare our results with those in Aghion and Howitt (2005). In their model the comparative advantage was determined wholly by the parameter of distance to the technology frontier a. But as mentioned previously, their distance parameter only measures the distance between the mean values of A and $\bar{A}$, the plausible differences induced by other moments, especially the variance, are totally ignored. In our example it is clear that even when difference between the mean values of typical country's TFP and world frontier TFP is smaller, this does not in itself guarantee that this country should adopt the innovation strategy. If at the same time the TFP variance of this country decreases relative to that of the world frontier TFP (i.e., θ_2 increases relative to θ_1), then, as we have shown, this country would have a lower TFP growth rate at the BGP *even if* it has been closer to the world technology frontier in the sense of Aghion and Howitt (2005).

4. Summary

Because human capital would have positive externalities, this gives government the rationale to subsidize education in almost all countries in the world. But how to allocate government funds and whom should be subsidized are still open questions. The answers to these questions provided by economists are often based on models in which the externalities are generated by the average level of human capital or by mean values of the total factor productivity. In this paper we want to show that this foundation is not that solid and some modifications could be made such that we can more adequately answer the above questions.

Using the frameworks of Aghion and Howitt (2005) and Eaton and Kortum (2002), we have proposed a new definition of the distance to technology frontier to measure the costs when firms would like to compete at either the extensive or the intensive margins. The resulting competitive equilibrium is defined and the equilibrium values of work efforts are derived, such that we can calculate the stochastic distance used to measure the TFP growth rates at the BGP.

We use an example to show why Aghion and Howitt's (2005) results are not always right. If we ignore the diversity of talents a country has, as stressed by Grossman (2004), then we would reach the conclusion that the comparative advantage is totally determined by the average or mean values of our country's TFP and that of the world frontier. As the example shows, even when a country is approaching the world technology frontier *on average*, it is still not sure whether this country should innovate more technologies by itself. It also depends on the *variance* or diversity of the human capital (as well as TFP because in that example these two are the same) this country would have. The discussion of education subsidization is therefore based on the comparative advantage, and this in turn depends on the shape of human capital and TFP distributions.

References

Acemoglu, Daron, Philippe Aghion and Fabrizio Zilibotti (2006): "Distance to Frontier, Selection, and Economic Growth," *Journal of the European Economic Association* (forthcoming).

Acemoglu, Daron and Joshua Angrist (2001): "How Large Are the Social Returns to Education: Evidence from Compulsory Schooling Laws," in Ben Bernanke and Kenneth Rogoff (editors), *NBER Macroeconomics Annual 2000*, pp. 9-59.

Aghion, Philippe and Peter Howitt (2005): "Appropriate Growth Policy: A Unifying Framework," the 2005 Joseph Schumpeter Lecture to the European Economic Association. (available at: http://post.economics.harvard.edu/faculty/aghion/papers.html)

Alvarez, Fernando and Robert E. Lucas, Jr. (2005): "General Equilibrium Analysis of the Eaton-Kortum Model of International Trade," *NBER Working Paper*, No. 11764.

Becker, Gary S. (1993): *Human Capital*, 3rd edition, Chicago: University of Chicago Press.

Billingsley, Patrick (1995): *Probability and Measure*, 3rd edition, New York: John Wiley & Sons.

Bils, Mark and Peter Klenow (2000): "Does Schooling Cause Growth?" *American Economic Review*, 90, 1160-1193.

Ciccone, Antonio and Giovanni Peri (2005): "Identifying Human Capital Externalities: Theory with Applications," *Review of Economic Studies* (forthcoming).

Eaton, Jonathan and Samuel Kortum (2002): "Technology, Geography, and Trade," *Econometrica*, 70, 1741-1779.

Friedman, Milton (2002): "The Role of Government in Education," in *Capitalism and Freedom*, 40th anniversary edition, Chicago: University of Chicago Press, pp. 85-107.

Grossman, Gene M. (2004): “The Distribution of Talent and the Pattern and Consequences of International Trade,” *Journal of Political Economy*, 112, 209-239.

Johnson, Norman L., Samuel Kotz, and N. Balakrishnan (1994): *Continuous Univariate Distributions*, Vol. 1, 2nd edition, New York: John Wiley & Sons.

Jones, Charles I. (2005): “The Shape of Production Functions and the Direction of Technical Change,” *Quarterly Journal of Economics*, 120, 517-549.

Kim, W. Chan and Renee Mauborgue (2005): *Blue Ocean Strategy: How to Create Uncontested Market Space and Make Competition Irrelevant*, Cambridge, MA: Harvard Business School Press.

Krueger, Alan B. and Mikael Lindahl (2001): “Education for Growth: Why and For Whom?” *Journal of Economic Literature*, 39, 1101-1136.

Lucas, Robert E., Jr. (1988): “On the Mechanics of Economic Development,” *Journal of Monetary Economics*, 22, 3-42.

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

Oller, Jose M. (1987): “Information Metric for Extreme Value and Logistic Probability Distributions,” *Sankhya*, 49, Series A, 17-23.

Rosen, Sherwin (1981): “The Economics of Superstars,” *American Economic Review*, 71, 845-858.