Literacy traps: Society-wide education and individual skill premia

Vidya Atal,* Kaushik Basu,[†] John Gray[‡] and Travis Lee[§]

Using a model of an O-ring production function, the present paper demonstrates how certain communities can get caught in a low-literacy trap in which each individual finds that it is not worthwhile investing in higher skills because others are not high-skilled. The model sheds light on educational policy. It is shown that policy for promoting human capital has to take the form of a mechanism for solving the coordination failure in people's choice of educational strategy.

Key words   education, literacy, O-ring, skill formation, traps

JEL classification   D20, I28, J31

Accepted 13 June 2009

1 Introduction

The idea that an economy or a group of people can get caught in a low-level trap from which it is, in principle, possible to escape but no *individual* has it within his or her power to break out of is an old one in economics, but its importance has remained undiminished. There has been a lot of important research in theoretical development economics over the past half century that is rooted in this broad idea and in trying to explain why certain economies get caught at low levels of per capita incomes. Among Tapan Mitra's many fields of enquiry in economic theory, poverty traps has been a significant one. In 1995, in a joint paper with Mukul Majumdar, for instance, he explores the relation between increasing returns and poverty traps and how an economy can be caught in poverty, although once it is wrenched out of the trap it can grow unassisted (Majumdar and Mitra 1995; see also Majumdar and Mitra 1982; Dechert and Nishimura 1983). This work is a natural extension of the idea of the vicious circle of poverty found in Nurske (1953) and Rosenstein-Rodan (1943) and also the idea that there is a close connection between underdevelopment and multiple equilibria (Basu 1997).

* Department of Economics, Cornell University, Ithaca, USA.
[†] Department of Economics, Cornell University, Ithaca, USA. Email: kb40@cornell.edu.
[‡] Department of Economics, Cornell University, Ithaca, USA.
[§] Department of Economics, Cornell University, Ithaca, USA.
We are grateful to Talia Bar and an anonymous referee for several useful suggestions.

Although the dominant discussion of low-level traps has occurred in the context of a nation's or a collectivity's income, it is possible to carry the broad idea over to other indicators of a nation's well-being (see Hoff and Stiglitz 2001). In the present paper, we try to show that something similar may happen regarding literacy. A nation can get caught in a low-literacy or low-education trap. Once caught in this situation, it is not in the interest of any individual to incur cost and acquire a lot of skills. It is the skill-lessness of others that makes it not worthwhile for each individual to acquire many skills, and, therefore, they are all trapped in a vicious circle.

Our analysis has important policy implications. A nation caught in a low-literacy trap cannot break out of it just by providing schools for it is not the supply of schools that is a bottleneck but the demand for higher education. Hence, the rather abstract model that we are about to construct can shed light on significant policy questions such as when do we need to make education compulsory and when will the simple act of making schools available take care of the problem of underinvestment in human capital. The model sheds interesting light on how, in a certain class of equilibria, giving a subsidy to education may have no effect on promoting education. In the process, we obtain some insights into the design of policy that will be effective.

The core of our model is based on the idea of an O-ring production function, introduced in the literature by Kremer (1993) (see also Kremer and Maskin 1996). The idea is this. Because so much of today's work takes the form of the assembly line, either literally, as in the manufacture of cars, or, in effect, as in software services, where small groups are engaged in doing different parts of a large job, a malfunction in one part can undo the benefits of the other tasks that are done well. The metaphor is that of the space shuttle Challenger disaster in 1986, which was caused by the malfunctioning of a tiny component of the spaceship, the O-ring. The idea that there will be this kind of spillover effect of education among workers seems natural enough and there has been a lot of empirical and theoretical work on this (Rauch 1993; Benabou 1993; Redding 1996; Acemoglu and Angrist 2000; Kremer, Miguel, and Thornton 2004; Moretti 2004).

In Kremer's O-ring model, the skill that workers bring to their task is innate to the worker. If, however, we introduce education in the model, whereby each worker has the choice of incurring some cost (in terms of both time and money) and improving their skills and ability to do their jobs better, then interesting equilibria arise, including the possibility that workers will get caught in a low-education trap. This is the central idea that is pursued in the present paper, and while poverty traps are a pervasive topic in economics, low-literacy traps seem to have received much less attention. The work most related to our paper is that by Jones (2008). He constructs a random matching model in which there is endogenous human capital accumulation. Each individual faces the choice to be trained as a generalist or a specialist, with the value of being a specialist increasing as the density of specialists in the population rises. In Jones' model, for certain parameter values, there is the possibility of multiple equilibria because the economy could be one of specialists or generalists. Another related exercise (Basu and Weibull 2003; Horowitz 2008) studies the punctuality traits of a collectivity, where each person benefits from other people's punctuality and also the marginal return to increased individual punctuality rises with the level of other people's punctuality. This strategic complementarity easily leads to multiple equilibria, whereby

two societies of a priori identical individuals can get caught in, respectively, a tardy and a punctual equilibrium.

2 Model

2.1 A primer on O-rings

It is useful to begin by briefly summarizing the O-ring model, while at the same time adapting it a little to our present need. There is one consumer good in the economy. Its production takes place in factory units or, simply, factories. Each firm can own one or more factories. In each factory n tasks ($n \geq 2$) are done, each task being done by one worker. Denote a worker's skill by q, where $0 \leq q \leq 1$. We can interpret q as the probability that the worker finishes his or her task successfully. Let q_i be the skill level that goes into task i; that is, the worker employed on task i has a skill level q_i and let B be the output produced *per worker* in a factory when all tasks are performed successfully. The "production function" in which x denotes the expected output is as follows:

$$x = q_1 \ldots q_n n B \equiv \prod_{i=1}^{n} q_i n B. \tag{1}$$

It is easy to see that if all tasks are performed at skill level 1, then total output from the factory will be nB and so the per worker output is B.

To start with, let us take the skill levels of workers to be exogenously given. The decision-making by the firms can be modeled in two different ways. The traditional route is to assume that there are many price-taking firms and free entry. Because there is a continuum of worker types, there is a continuum of wages, one for each type of worker. Let $w(q)$ be the market wage schedule exogenously given to the firms. We will throughout take the price of the product to be one. In this case the firm's problem is to choose n workers for operating a factory so as to maximize its profit. Hence, the firm's problem is the following:

$$\max_{\{q_i\}} \left[\prod_{i=1}^{n} q_i n B - \sum_{i=1}^{n} w(q_i) \right].$$

This gives us the following first-order condition for each task i:

$$w'(q_i^*) = \prod_{j \neq i} q_j^* n B. \tag{2}$$

In addition, Kremer (1993) proves that it is always in the firm's interest to have all its tasks done by workers of the same skill level. This is called the "skill-clustering theorem" in Basu (1997), where a short proof is available.

Theorem 1 (Skill-clustering) *If $(q_1^*, \ldots, q_n^*)$ maximize a firm's profit, then (in addition to (2)) $q_1^* = \cdots = q_n^*$.*

In light of the skill-clustering theorem, (2) can be written as

$$w'(q) = q^{n-1} nB,$$

where q is the skill of labor chosen for each task by a firm.

Because in equilibrium each firm earns zero profit, a firm employing workers of skill q must satisfy

$$q^n nB - nw(q) = 0$$
$$\text{or, } w(q) = q^n B. \tag{3}$$

Hence, we know that in equilibrium the wage schedule for different worker qualities will be given by this equation.

An alternative approach, which, however, will not be pursued here, is to assume that there are two or more firms and these are Bertrand oligopsonists. Each firm announces the wage it is willing to pay for each type of worker, and workers go to the firm offering the highest wage, ties being broken arbitrarily. The "equilibrium" of this oligopsony is simply the Nash equilibrium of the normal-form game among the firms. As we know from standard oligopsony, in equilibrium each firm will earn zero profit. The logic of this is obvious. If there is a firm that earns positive profit, then another firm could offer its workers a slightly higher wage and woo them away. So the initial outcome would not have been a Nash equilibrium. If, in addition, we assume away the integer problem, that is, assume that, for each wage announced by the firm, either no worker will agree to work or any number of workers will, then the wage schedule in equilibrium will be exactly as shown by (3).

However, we shall go with the traditional approach of taking this to be a model of perfectly competitive firms with free entry.

Before moving on, it should be clarified that in this section we assumed, as in Kremer's original model, that individuals have exogenously-given skill levels. There is no presumption that everybody has the same level of skill. Hence, what will be observed on the market is a scatter of skill levels and a scatter of wages, determined by the above equation.[1] Each firm will be a cluster of skill types.

2.2 Endogenizing level of education

Let us now allow for the possibility that individuals do not come with an immutable skill level but can acquire skill through education. To make it possible to conduct a formal analysis, we have to take a slightly novel route in developing the idea of an equilibrium. We shall assume that there are two periods. In the first, workers choose their level of education and in the second period, with education as given, firms make their decisions as in a standard competitive model with free entry: in other words, exactly as described in Subsection 2.1. In the first period, the workers essentially do a Nash-type calculation. That

---

[1] The one technical assumption used here is that, for each skill level, the number of persons with that level of skill is divisible by n. There are ways of getting around this assumption but no great conceptual insight would occur from such an exercise.

is, each worker calculates what would happen if he or she deviated and chose some other level of education. If he or she could not do better by any such deviation, then the existing choice of education for all workers is an "equilibrium."

Formally, in the first period, each worker chooses to obtain a certain level of skill q through education. We will assume that the cost of education that provides the level of skill q is given by $c(q)$ with $c'(q) \geq 0$. In the second period, the firms take their decisions about what kinds of workers to hire for the different tasks, with wages being treated as exogenous by each firm. The second period equilibrium is reached when we find a wage schedule (i.e. a wage for each level of skill) such that each firm maximizes profits and earns zero. In other words, the second period works as described in Subsection 2.1. After the second period, firms earn their payoffs (we already know this will be zero in equilibrium) and each worker receives his or her payoff, which is equal to the wage earned by the worker minus the cost of education.

In defining the equilibrium formally in this two-period model, let us focus on a re-finement of what was described informally above. The refinement is an outcome in which all workers voluntarily choose the same level of skill. We shall call this the "symmetric equilibrium," with the frequent indulgence of dropping the epithet "symmetric," because we are not going to talk about a non-symmetric equilibrium in the present paper.

A skill level q and a wage equal to $q^n B$ for each of these workers is a (*symmetric*) equilibrium if and only if

(i) $q^n B \geq c(q)$ and
(ii) for all $\hat{q}$, the wage that a worker who individually deviates to $\hat{q}$ earns is such that wage minus the cost of that education, namely, $c(\hat{q})$, is less than or equal to $q^n B - c(q)$.

In other words, all workers earn enough to cover their education cost and no worker by unilaterally deviating to some other level of education can do better.

To formalize condition (ii), we need to describe what wage a worker, who unilaterally deviates to $\hat{q}$ when everybody else has chosen q, will earn. With a slight abuse of notation, denote this wage by $w(\hat{q}; q)$ and denote the profit of the firm hiring this person by $\pi(\hat{q}; q)$. Clearly,

$$
\begin{aligned}
\pi(\hat{q}; q) &= \hat{q} q^{n-1} nB - (n-1) w(q) - w(\hat{q}; q) \\
&= \hat{q} q^{n-1} nB - (n-1) q^n B - w(\hat{q}; q).
\end{aligned}
$$

The firm will hire this worker if and only if

$$
\pi(\hat{q}; q) \geq \pi(q) = 0.
$$

Therefore, $w(\hat{q}; q)$ is the maximum possible wage the worker can get while ensuring that the above inequality is satisfied. Otherwise, the firm will refuse to employ this worker. It is now easy to derive that when all other workers have skill q, the wage of a worker with skill $\hat{q}$ will be given by:

$$
w(\hat{q}; q) = \hat{q} q^{n-1} nB - (n-1) q^n B. \tag{4}
$$

What is interesting and makes our analysis easy to conduct is a property of $w(\hat{q}; q)$. The property is the following. The graph of $w(\hat{q}; q)$ as $\hat{q}$ changes is always given by the straight line that is tangent to $w(q)(= q^n B)$ at q.

It is worth emphasizing that in this paper we are confining our attention to unilateral deviations in the spirit of Nash. There may indeed be opportunities for groups of people to form coalitions and deviate from an equilibrium. Inquiry into notions of coalition proof equilibria in models of this kind can be interesting but lie beyond the scope of the present paper.

What we are now ready to demonstrate is that this model can have multiple symmetric equilibria. In other words, it is possible to have a very low level of education that is an equilibrium in the sense that if everybody chooses it, nobody can do better by deviating, and there is also a possibility of a very high level of education in equilibrium. A society can simply get caught in a low literacy trap. Between two societies, one highly skilled and another with rudimentary skills, there may be no fundamental difference. They can be mere victims of their history. Using the property of $w(\hat{q}; q)$ mentioned above, these results are easy to prove. This can be done with a few simple examples; and that is what we do presently. The last section of the paper goes into the large policy implications for what the government could do to promote education and the acquisition of human capital and skills.

3 Linear cost function and literacy trap

Consider a linear cost function for education as follows. Individuals are born with some level of skill, say z. Alternatively, this is a level of skill that comes to us costlessly. Most human beings can perform basic tasks without having to undergo any formal training. To acquire skill beyond z, a worker has to incur a cost, which increases linearly with the level of skill. To sum up, the cost of education for attaining skill q is given by:

$$c(q) = \begin{cases} 0, & \text{for } 0 \leq q \leq z \\ a\,(q - z), & \text{for } z < q \leq 1, \end{cases} \tag{5}$$

where a is such that

$$z^{n-1} < \frac{a}{nB} < 1 \text{ and } \left(\frac{nz}{n-1}\right)^{n-1} \geq \frac{a}{nB}. \tag{6}$$

The first assumption guarantees that the cost of education is neither very high so that no one chooses to get more skill than z, nor very low so that everyone chooses to become an expert. The second assumption guarantees that $c(q) \leq w(q)$ for all $q \in [0, 1]$.
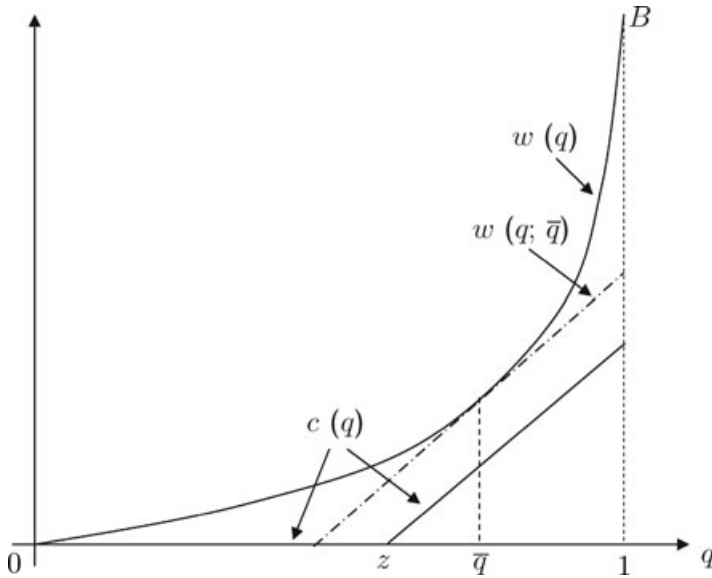
Define $\bar{q}$ such that

$$w'(\bar{q}) = a$$

Figure 1 Linear cost function.

or

$$\bar{q} = \left(\frac{a}{nB} \right)^{\frac{1}{n-1}}.$$

$\bar{q}$ is illustrated in Figure 1.

Claim 1 *Every worker acquiring skill $\bar{q}$ and earning a wage of $\bar{q}^n B$ is an equilibrium.*

PROOF: Suppose all workers have chosen $\bar{q}$. We know that perfect competition among firms with free entry will drive wages to $\bar{q}^n B$. It has already been seen earlier that firms offer the wage $w(q; \bar{q})$ given by (4) to any worker who unilaterally deviates to a different skill level q.

Let us now check how a worker, who unilaterally deviates from $\bar{q}$ will do. Note that by assumption (6), $w(\bar{q}) \geq c(\bar{q})$ and $z < \bar{q} < 1$. Now, from (4), we have

$$\frac{\partial w(q; \bar{q})}{\partial q} = \bar{q}^{n-1} nB = w'(\bar{q}) = a.$$

This implies that $w(q; \bar{q})$ is parallel to $c(q)$ for all $q \in [z, 1]$. This in turn means that

$$w(\bar{q}; \bar{q}) - c(\bar{q}) = w(q; \bar{q}) - c(q) \text{ for all } q \in [z, 1].$$

However, $w(\bar{q}; \bar{q}) = w(\bar{q})$. Therefore,

$$w(\bar{q}) - c(\bar{q}) \geq w(q; \bar{q}) - c(q) \text{ for all } q \in [0, 1].$$

Hence, $\bar{q}$ is an equilibrium. It is interesting to note that deviations to the interval $[z, 1]$ leave the deviating worker exactly as well off as before, and all other deviations make the worker worse off.   □

Claim 2   *In this example, with the linear cost function given by* (5), *there are two other equilibria, one in which all workers choose* $q_* = z$ *and another in which all workers choose* $q^* = 1$.

PROOF: First, suppose all workers choose $q^* = 1$ and one of them contemplates deviating from this common choice to a lower level of quality $\hat{q} < 1$. Again, the tangent line to the graph of $w(q)$ at $q^* = 1$ gives the wage for obtaining quality $\hat{q}$. Now, for all $\hat{q} < q^*$, the following is true:

$$\frac{\partial w(\hat{q}; q^*)}{\partial \hat{q}} = w'(q^*) > w'(\bar{q}) = a.$$

The inequality holds because $w(q)$ is a convex function and $q^* = 1 > \bar{q}$. Therefore, the loss in wages due to deviation to $\hat{q}$ from q^* more than offsets the cost savings because $w(\hat{q}; q^*)$ is more steeply sloped than the cost function. Hence, deviation would lead to a lower payoff.

Similarly, if all workers choose $q_* = z$, then deviation to $\hat{q} > z$ is not advantageous because $q_* = z < \bar{q}$ implies that

$$\frac{\partial w(\hat{q}; q_*)}{\partial \hat{q}} = w'(q_*) < w'(\bar{q}) = a.$$

Finally, deviation to $\hat{q} < z$ lowers wages without reducing costs, so workers will not do that.   □

It is easy to see that there does not exist any other equilibrium apart from the three described above in this model with the linear cost function given by (5). It is worth cautioning the reader that in the present paper we are focusing entirely on symmetric equilibria. It is possible to describe economies in which there will exist some non-symmetric equilibria; that is, ones in which different clusters of individuals acquire different levels of skills. However, such equilibria will typically be non-generic and, therefore, are ignored in the present paper.

3.1 Literacy trap and big push

Note that, while $\bar{q}$ is an equilibrium, it is not "stable" in the following sense. In a society where all workers are skilled up to level $\bar{q}$, if it is possible to increase everyone's skill a little bit, then each of the workers will deviate further *away* from $\bar{q}$. That is, they will increase their skill; and note that this dynamic will continue until the equilibrium q^* is reached or gradually approached. However, if everybody's skill was lowered a little, then a downward dynamic would start up and society could go all the way to the equilibrium q_*.

Finally, we have a big "literacy trap" at the education-level that provides skill $q_* = z$. A "big push" that drives the entire economy beyond the skill level $\bar{q}$ can start up innate
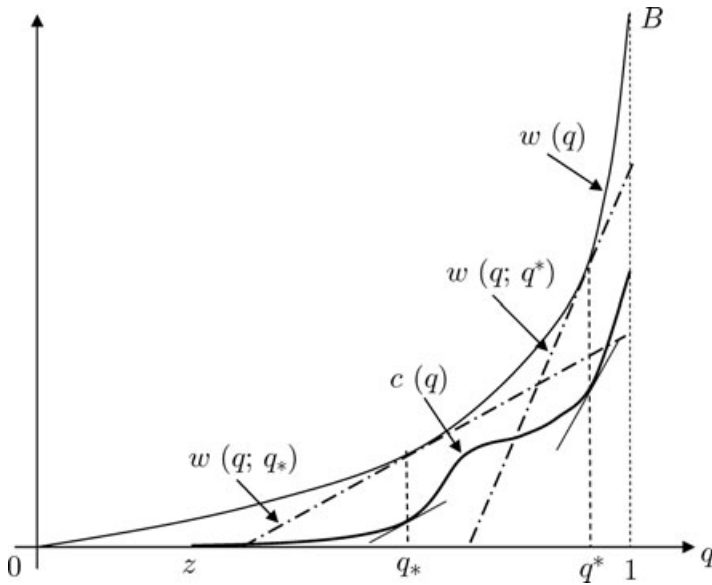
Figure 2 Nonlinear cost function with multiple equilibria.

forces that will then take the economy all the way to the good equilibrium. All smaller efforts will keep pulling workers back to $q_* = z$. This has one heartening implication. In an economy with widespread illiteracy, the cost of raising human capital might not be as much as appears at first sight. This is because the funding needed to promote education will not have to be sustained endlessly through time. As soon as a threshold is crossed, the accumulation of human capital and skills can be left to natural forces and will continue unabated.

4 More general cost function and literacy trap

Although we illustrated our main results with the linear case, there is no need to confine the analysis to such cases. Virtually all results carry over to the more general nonlinear cases. Consider the nonlinear cost function as shown in Figure 2. As in the case of a linear cost function, suppose primary education that provides skill z is free. Then the cost of education increases at an increasing rate. After some point, the behavior of the cost function changes and as someone becomes more educated, the lower is his or her marginal cost of education. After a very high level of education is attained, the marginal cost starts increasing again.

It is clear from the largely self-explanatory figure above that even with nonlinear cost of education we may have multiple equilibria. In this particular example, we have two symmetric and strict Nash equilibria. As shown in Figure 2, q_* and q^* with firms offering wages $w(.)$ given by (3) are the two Nash equilibria and both are stable. Therefore, a literacy trap might occur in a society where all the workers optimally choose to attain the skill q_*.

It is easy to see that the non-convexity of the cost function is by no means necessary for our claim to be valid. We make such an assumption purely for visual ease. If the cost function is linear or gently convex between the two equilibrium points in Figure 2, so as to render the entire cost function convex, we would still have the same equilibria.

5 Policy interventions

Increasing literacy and the advancement of human capital has been a major focus of policy-making certainly in developing countries but also in developed nations. Evidently, there are two sides to this policy. There has to be a demand for education on the part of parents making decisions for their children and young adults making decisions for their own education. In addition, there has to be a supply of schools so that parents who wish to educate their children *can* do so. In popular discourse, it is often said that poor parents do not want to educate their children. This has met, rightly, with strong criticism (see PROBE Team 1999). However, this must not blind us to the fact that the intensity of demand for education can vary and this can make a difference to the literacy outcome of a nation (see PROBE Team 1999; Drèze and Kingdon 2001). It is believed that the rewards from education (the so-called "skill premium") have been rising in the developing world; and there is now some hard evidence on this (see Arbache, Dickerson, and Green 2004; Azam 2009). When this happens, it is not surprising to find that the demand for obtaining education will also become stronger. It is now said in India, given that missionary schools had historically played a major role in the country, that all you have to do is to think of a good English name, like John or Thomas or Mary and add the prefix "Saint" and suffix "School" to it, and you will be in the education business with students flocking to you.

In our model, it is easy to see that the same country where the demand for education is low because the skill premium is low can change to an equilibrium with high premium and high demand for education. In Figure 1, if we start from a case where the country is caught at a low literacy trap at z, it is not worthwhile for any individual to seek more education. The skill premium is just not high enough to make this worthwhile. If, in contrast, education rises and goes past $\bar{q}$ for everybody, then people will invest even more in education and the nation will come to rest at a very high level of education for all.

Our model allows us to separate out the demand and supply aspects quite neatly and so enables us to take a more sophisticated view on policy. We can think of government-subsidized education as an intervention that lowers the cost of education. This can have interesting effects depending on how it is done. Suppose government gives a small flat subsidy s for all levels of education. Contrary to what is presumed, this might have no effect on education. This will be true for all the equilibria depicted in Figures 1 and 2, excepting at z in Figure 1. To boost education, government has to vary the subsidy with respect to the level of education. In other words, the government needs to have a non-constant function $s(q)$. The total cost of education is then given by $[c(q) - s(q)]$. By suitably altering the slope of $s(q)$, the state can boost education. Indeed, the net expenditure on education could be very small if the subsidy function is chosen artfully. If the economy is caught at equilibrium

z in Figure 1, then a constant small subsidy will have a small positive effect on education. Beyond a critical point, it will have a huge effect, pushing skill acquisition all the way to the maximum value 1, with no need for any subsidy in the new equilibrium.

The model suggests that the fiscal burden of boosting education may not be too high, because all we need is a short period boost, after which the natural incentives in the system kick in and little further outside intervention is needed. For this same reason, it might be worthwhile for a country caught in a low-literacy equilibrium to have a policy of compulsory education, which forces parents to educate their children. If this can be sustained for a while, the need for force will vanish because the high education of the rest of the population will raise the education premium for each individual.

References

Acemoglu, D., and J. Angrist (2000), "How large are human capital externalities? Evidence from compulsory schooling laws," *NBER Macroeconomics Annual* **2000**, 9–59.

Arbache, J. S., A. Dickerson, and F. Green (2004), "Trade liberalisation and wages in developing countries," *The Economic Journal* **114**, F73–F96.

Azam, M. (2009), "India's increasing skill premium: Role of demand and supply," Institute for the Study of Labor Discussion Paper No. 3968, Bonn, Germany.

Basu, K. (1997), *Analytical Development Economics: The Less Developed Economy Revisited*, Cambridge, MA: MIT Press.

Basu, K., and J. Weibull (2003), "Punctuality: A cultural trait as equilibrium," R. Arnott, R. Kanbur, B. Greenwald, and B. Nalebuff, eds, *Economics for an Imperfect World: Essays in Honor of Joseph Stiglitz*, 163–82, Cambridge, MA: MIT Press.

Benabou, R. (1993), "Workings of a city: Location, education, and production," *The Quarterly Journal of Economics* **108**, 619–53.

Dechert, D., and K. Nishimura (1983), "A complete characterization of optimal growth paths in an aggregated model with a non-concave production function," *Journal of Economic Theory* **31**, 332–54.

Drèze, J., and G. Kingdon (2001), "School participation in rural India," *Review of Development Economics* **5**, 1–24.

Hoff, K., and J. Stiglitz (2001), "Modern economic theory and development," G. Meier, and J. Stiglitz, eds, *Frontiers of Development Economics*, 389–459, Oxford: Oxford University Press.

Horowitz, A. (2008), "The punctuality puzzle: Can current punctuality initiatives in low-income countries succeed?" Working Paper, Sam W. Walton College of Business, University of Arkansas, Fayetteville, AR.

Jones, B. (2008), "The knowledge trap: Human capital and development reconsidered," NBER Working Paper No. 14138, NBER, Cambridge, MA.

Kremer, M. (1993), "The O-ring theory of economic development," *The Quarterly Journal of Economics* **108**, 551–75.

Kremer, M., and E. Maskin (1996), "Wage inequality and segregation by skill," NBER Working Paper No. 5718, NBER, Cambridge, MA.

Kremer, M., E. Miguel, and R. Thornton (2004), "Incentives to learn," NBER Working Paper No. W10971, Cambridge, MA.

Majumdar, M., and T. Mitra (1982), "Intertemporal allocation with a non-convex technology: The aggregative framework," *Journal of Economic Theory* **27**, 101–36.

Majumdar, M., and T. Mitra (1995), "Patterns of trade and growth under increasing returns: Escape from the poverty trap," *The Japanese Economic Review* **46**, 207–25.

Moretti, E. (2004), "Workers' education, spillovers, and productivity: Evidence from plant-level production functions," *American Economic Review* **94**, 656–90.

Nurske, R. (1953), *Problems of Capital Formation in Underdeveloped Areas*, Oxford: Oxford Press.

PROBE Team (1999), *The Public Report on Basic Education in India*, New Delhi: Oxford University Press.

Rauch, G. E. (1993), "Productivity gains from geographic concentration of human capital: Evidence from the cities," *Journal of Urban Economics* **34**, 380–400.

Redding, S. (1996), "The low-skill, low-quality trap: Strategic complementarities between human capital and R&D," *The Economic Journal* **106**, 458–70.

Rosenstein-Rodan, P. N. (1943), "Problems of industrialisation of Eastern and South-Eastern Europe," *The Economic Journal* **53**, 202–11.