

The Economics of Inefficient Technology Use

Paul Beaudry and Patrick Francois

February 5, 2009

Abstract: The adoption and diffusion of technological knowledge is generally regarded as a key element in a country's economic success. In this paper we examine whether asymmetric information regarding who knows how to run a new technology efficiently can explain a set of observations regarding within and cross-country patterns of technology diffusion. In particular, we show how the dynamics of adverse selection in the market for technological knowhow can explain (1) why inefficient technology use may take over a market even when better practice is available, (2) why widespread inefficient use may persist unless a critical mass of firms switch to best practice, (3) why efficient adoption of new technologies is more likely to occur where the existing technology is already productive, where wages are already relatively high, and where the new technology is not too great an advance over the old one, and (4) why the international mobility of knowledgeable individuals does not guarantee the diffusion of best practice technology across countries.

JEL No. O33

Paul Beaudry¹

Department of Economics, University of British Columbia, and NBER
#997 – 1873 East Mall, Vancouver, B.C., Canada, V6T 1Z1
paulbe@interchange.ubc.ca

Patrick Francois

Department of Economics, University of British Columbia, CEPR and CIAR
#997 – 1873 East Mall, Vancouver, B.C., Canada, V6T 1Z1
francois@interchange.ubc.ca

¹We thank Ashok Kotwal, seminar participants at the UBC macroeconomic workshop and the Canadian Institute for Advanced Research's IOG group for helpful comments on preliminary results from this project.

1 Introduction

It is widely accepted that differences in technological efficiency play an important role in explaining cross country differences in economic performance. However, it is not the case that all firms in poor countries are technologically backward in comparison to firms in rich countries. Instead, poor countries are generally characterized by a very skewed distribution of productivity across firms, with most having very low productivity, and a very small subset following what may be called best practice.² In contrast, rich countries are characterized by a different distribution, though there is still considerable variation in productivity there, a higher proportion of firms operate closer to best practice. If, as seems likely, technological efficiency depends partly on knowing how to use technologies well, then it seems reasonable to draw two conclusions from these facts: 1) knowledge does not flow freely between firms anywhere but, 2) it does seem to flow better between firms in rich countries.

Embodied knowledge, like skills, are a natural place to start when thinking about facts like these because such knowledge can never move entirely freely across firms. This contrasts with disembodied knowledge, like blueprints, that can be transmitted almost everywhere instantaneously. In this paper, we examine the process of embodied knowledge diffusion. We are particularly interested in the knowledge of how to run a technology according to best practice. This is something that is not typically learned in schools or universities, though some background from these places is often essential. Instead, such embodied knowledge transmission tends to predominately occur through person to person interaction in the workplace.³ In its simplest form, this occurs when people without the knowledge work alongside those with it, and thereby learn how best to efficiently implement a technology. The knowledge is thus transmitted as a by product of the production process.

²This well established fact is a central organizing feature of Banerjee and Duflo's (2005) handbook survey. It is also evident in the cross country data presented in Hsieh and Klenow (2007) for example.

³Thoenig and Verdier (2004) provide extensive discussion of the evidence in support of this assertion. Most of it is drawn from the business economics literature and is best summarized by a quote they include from a highly influential article in the Harvard Business Review. "*Evidence from research conducted since the mid-1960s shows that [...] managers get two thirds of their information to make decisions from face-to-face or telephone conversations; they acquire the remaining third from documents, most of which come from outside the organization.*" (Davenport, 1994)

Since such knowledge is embodied, and access to it excludable, it is reasonable to conjecture that markets could work well in transmitting it. This is confirmed when studying the distribution of steady state productivities by Chari and Hopenhayn (1991) and has been shown along decentralized dynamic trajectories in Beaudry and Francois (2006) as well. The conclusion is that in decentralized frictionless markets, prices work well in correctly signalling the value of knowledge and coordinating such person to person diffusion activities across the economy. Knowledge diffuses efficiently via such a transmission process so that, at any point in time, the distribution of productivities in an economy optimally reflects the social valuation of the knowledge and the opportunity costs of learning.

Although the frictionless competitive setting is a nice benchmark, it ignores an important real-world impediment to knowledge transmission that may help in explaining cross-country productivity differences and productivity distributions within rich and poor countries. Such a process of diffusion will generally be hampered by the fact that trade in knowledge is subject to adverse selection. As uninformed individuals do not know best practice, it is hard for them to identify which of those claiming to know and use best practice can actually do so. In particular, uninformed individuals may be willing to accept lower wages to work alongside an expert, but such behavior will incite non-experts to falsely claim expertise in order to hire them at these low wages.

Here we study the resulting dynamic adverse selection problem created by such knowledge asymmetries, to see whether it can shed light on aspects of knowledge transmission in reality.⁴ In particular, we examine whether adverse selection can explain why best practice may not diffuse across firms within a country, leaving inefficient technology use as the dominant mode of production. We also examine whether asymmetric information about identity of

⁴There may be other impediments to such diffusion, for example, liquidity constraints on the part of workers purchasing the knowledge may play a role. However, the widespread use of apprenticeships, and apprenticeship like contracts in many LDCs, suggests that liquidity constraints, though often severe, can be overcome. Despite this, the evidence regarding on-the-job skills transmission in LDCs, which we briefly consider after the main results, suggests it is generally poorly done. We will argue that this is not because the problem of knowledge diffusion is insurmountably difficult, but instead because it is an equilibrium outcome for agents to not invest in learning how to use technologies efficiently. These choices then lead to the widespread propagation of inefficient production methods.

knowledgeable individuals can explain why the diffusion of best practice is achieved rapidly in some countries and not others, and what features determine this. Finally, we explore why international migration of workers with knowledge of best practice does not always guarantee convergence across countries.⁵

The information transmission in our model regarding best practice is a form of “learning by seeing”. It differs from learning by doing in that the latter implies increasing productivity in the same task by repetition.⁶ Learning by seeing, in contrast, involves working as an unskilled, and perhaps subordinate, worker to a skilled one, so that the skilled one’s practical knowledge is transmitted.⁷ We study economies where this practical knowledge required to use a technology efficiently initially exists in some individuals, and where the blueprint is freely available, so that there are no impediments to knowledge flows into an economy.⁸

We will demonstrate that adverse selection can cause the continued misuse of technology to prevail as a steady state.⁹ In that case, person to person knowledge diffusion breaks down almost entirely. It will be seen that escaping from such a situation may require a critical mass of firms to use best practice. We will also show that efficient adoption of new

⁵This does not suggest that institutional impediments play no role, but instead that the relevant institutional shortcomings seem to affect the diffusion of knowledge within economies’ borders rather than impediments to the knowledge crossing borders in the first place. An alternative possibility not explored here, but the focus of Thoenig and Verdier (2004), is that firms may take actions to limit the spread of useful knowledge beyond their boundaries. They explore the implications of this for the form of optimal incentive contracts for agents implementing new technologies. They argue that changes in this contractual form interacted with changes in the degree of competition through time are consistent with patterns of long-run development.

⁶Chari and Hopenhayn (1991) provide a formal analysis of this type of knowledge transmission. Their analysis is primarily concerned with the steady state outcome of such a process, and in marked contrast to the present, explore an environment free from imperfections in transmission. Beaudry and Francois (2006) focus more heavily on the dynamics of this process, but again do so in an environment free from imperfections.

⁷As Fluitman (1992) notes, apprentices learn primarily by watching their master and being “corrected” when they err. We think that this is an important form of learning when it comes to the diffusion of new technologies, as suggested by micro case studies of technological laggards in LDCs such as McKinsey Global Institute (2001).

⁸Considerable empirical literature has documented that local firms and their managers often get their start as employees of multinational firms (Katz 1987, Hobday 1995, Hall and Khan 2003). We are silent in this paper about the source of the knowledge as our focus is within country transmission. In our earlier paper we presented evidence of such cross-country diffusion through worker training and relocation.

⁹There are already theories that can explain why firms with differing, exogenously given, productivity levels can coexist – Restuccia and Rogerson (2007), Hsieh and Klenow (2007), Bartelsman, Haltiwanger and Scarpetta (2006). Our focus is on a more primitive feature as we treat a firm’s productivity as its own choice.

technologies is more likely to occur in countries where the existing technology is already productive, where wages are already relatively high, and where the new technology does not represent too great an advance over the old one. We thus unearth an effect which makes poor and low wage countries – i.e., those most in need of efficiently implementing productivity improving technologies – less able to do so than the rich, where new technologies represent only incremental improvements on current best practice.

The main body of the paper examines the dynamics of information transfer in a model of learning-by-seeing where asymmetric information is present. After the paper's main theoretical results are established, we discuss evidence from developing countries on person to person knowledge transfer within firms and we argue that this evidence suggests that the asymmetries of information we focus upon here are potentially of first order importance in constraining knowledge flows. The model, combined with this evidence, thereby provides a novel explanation for the persistent low productivity levels known to exist in many countries. We also discuss how the model can explain why so few skilled individuals wish to migrate from rich countries where their skills are abundant and hence in low demand, to poor countries, where their skills are scarce. The main mechanism involves the absence of a premium to knowledge in poor countries. We show that in countries that are in a low-skill steady state, where skill transmission is not advancing, the unskilled are unwilling to pay a high premium for the opportunity to acquire skills since they expect the resulting learning to have little value. Consequently, the genuinely skilled, who would be able to effectively teach are better off remaining in their skill abundant homeland.

The paper proceeds as follows. The next section develops the basic model, and analyzes both the steady states and the transitional dynamics associated with a situation where firms can either be run efficiently by experts, or inefficiently by having non-experts fill key positions. Section 3 extends the model by allowing for a pre-existing inferior technology in which expertise is not needed, and which can also be chosen as a means of production. The productivity of the pre-existing technology is meant to capture the initial state of development of an economy, and this extension of the model allows us to examine how initial conditions

affect the diffusion of best practice. Section 4 considers the migration incentives induced by this process of skill formation, and Section 5 concludes.

2 Preferences and Technology

Consider an economy where the population (or labor force) is of constant size and normalized to 1. At each instant a new cohort of size δ enters and a measure δ of workers dies; with the property that $\int_{-\infty}^t \delta e^{-\delta(t-s)} ds = 1$. Each individual is risk neutral and controls one unit of labor per instant. Let $\rho > 0$ be the instantaneous discount rate. Discounted expected lifetime utility for an individual is:

$$U = \int_0^{\infty} c(t) e^{-(\delta+\rho)t} dt,$$

where $c(t)$ denotes consumption. Time is continuous.

There is one good in the economy, and it can be produced in a technologically efficient manner, or an inefficient manner. Whether the good is produced in an efficient manner depends on whether the manager is an expert (i.e. knows how to run the technology efficiently) or a non-expert. If an expert (efficient) manager hires L workers, he produces $F(1, L)$ units of the good. If a non-expert manager hires L workers, he produces $F(1 - \theta, L)$ units of the good.

Individuals in the economy can be of only two types, experts or non-experts. The experts are individuals that can manage efficiently, and the mass of such individuals in the economy is denoted by S_t . Non-experts are individuals that can either supply their labor to the market in order to be hired as workers, or they can alternatively become inefficient managers. Let us denote by ϕ_t the fraction of non-experts that choose to be inefficient managers. This implies that the quantity of labor supplied to the market as workers is $(1 - \phi_t)(1 - S_t)$.

In the above, we are assuming that all experts choose to work as managers. It would be reasonable to allow experts to potentially supply their labor as workers if they found it profitable. However, given the restrictions we will later introduce, allowing such a choice would not change any of our results since experts would always choose to be managers.

2.1 Learning to become an expert

We will assume that at time 0, there are a set $S_0 > 0$ of experts in the economy (which could be extremely small) and that all subsequent born individuals are born as non-experts. This set S_0 can be thought of a set of individuals that have learned to use the technology efficiently abroad, or who have learned to use the technology efficiently by investing in R&D. Initially, we do not want to focus on the determination of S_0 , but instead we want to ask whether with $S_0 > 0$ the number of experts will grow and always push out inefficient production when non-experts can learn from experts. In particular, we will assume that non-experts can become experts by being hired and working alongside an expert manager. The learning process is stochastic with the instantaneous rate of becoming an expert when working with an expert being denoted Ω . Knowledge acquisition is assumed binary, that is, a non-expert working alongside an expert for an instant dt either acquires proficiency with probability Ωdt , or remains non-expert with probability $1 - \Omega dt$. In contrast, we assume that an individual working as a non-expert manager, or an individual working alongside a non-expert manager, has a zero instantaneous probability of becoming an expert: one must work alongside an expert to learn.

2.2 Information Structure

The identity of an expert is assumed to be private information in the sense that an individual that becomes an expert recognizes this, but others cannot observe it. In principle, non-experts would be willing to accept lower wages to work alongside experts versus working for non-experts managers, however the asymmetry of information will make such separate treatment impossible and therefore cause the wage paid to workers, denoted w , to be independent of worker type.^{10 11} This represents the main friction in the model, that is, we

¹⁰We are also assuming that a worker does not observe other information about a manager, such as his revenues or the number of employees, since such information could allow the worker to infer the managers type. However, it is worth noting that the essence of our results is robust to allowing workers to observe the number of employees hired by a manager. What is crucial to maintain pooling is that workers cannot observe revenue of managers.

¹¹This pooling seems likely to be true in many situations. Since workers do not know the exact form of training they need to receive to become experts, they do not know whether they are receiving training that

assume that non-experts cannot readily identify experts because they do not know what characterizes an expert. This leads to an adverse selection problem when deciding whether to supply labor to a manager: both expert and non-expert managers will want to claim that they are experts.

2.3 Market Structure

The output and labor markets are assumed to be perfectly competitive, and there is free entry of managers. That is, any individual can decide at any moment to become a manager and hire workers on the labor market at wage w_t . The net income received by an expert manager that hires an optimal amount of workers will be denoted R_t^E and the net income obtained by a non-expert manager will be denoted R_t^I .

2.4 Managerial Decisions

The manager's problem is a simple static problem corresponding to choosing the number of workers to hire so as to maximize revenue, taking the price of labor as given. The problem of an expert manager can be stated as follows, where L_t^E represents the number of workers hired by an expert manager,

$$R_t^E = \max_{L^E} F(1, L^E) - w_t L^E$$

The optimization problem for the non-expert manager is identical except for the fact that the technology under his control is less productive and hence his revenue R_t^I will be given by

$$R_t^I = (1 - \theta)R_t^E \quad (3)$$

and L^I , which denotes the number of workers hired by a non-expert manager, will satisfy¹²

is of value, and which can therefore transform them into an expert manager, or whether they are learning skills that do not correspond to best practice. If firms were able to commit to a training schedule and to prove they had the knowledge required to train workers, then this would not be the case, and certainly this is possible in some markets, perhaps through firm reputations. Nonetheless as we discuss after the main results, there is considerable micro-level evidence from LDCs suggesting that this asymmetric problem is ubiquitous in the labor market.

¹²If we assumed that workers could observe the number of workers hired by the firm, this would require the non-expert manager set $L^I = L^E$ if he desires to mimic an efficient manager. We have solved the model under this assumption as well, and nothing qualitatively changes.

$$L_t^I = (1 - \theta)L_t^E. \quad (4)$$

2.5 The decision problem of a non-expert

A non-expert must decide each period whether to simply offer his labor time to the market, or whether to become a manager and hire workers. If he decides to be a manager, he makes a flow revenue of R_t^I since he is a non-expert, and he remains a non-expert. In contrast, if he decides to be a worker, he receives a wage w_t and he has a chance to learn to become an efficient manager. The instantaneous rate at which he becomes an expert is denoted $\tilde{\Omega}$, and this rate is determined in equilibrium depending on the likelihood that he is hired by an expert manager or a non-expert. Since non-experts may be indifferent between the two strategies, it is useful to allow them to follow mixed strategies. To this end, let ϕ_t ($0 \leq \phi \leq 1$) be the probability that a non-expert chooses to be a manager at time t . Here we use the same notation to represent the probability that a non-expert chooses to become a manager and the fraction of non-experts that become managers, since in equilibrium both will be the same. To solve this decision problem, we use Bellman's Principle of optimality. With V_t^N representing the expected discounted utility associated with currently being a non-expert, and V_t^E representing the expected discounted utility associated with being an expert, Bellman's principle implies that

$$(\delta + \rho)V_t^N = \max_{\phi_t, 0 \leq \phi_t \leq 1} \phi_t R_t^I + (1 - \phi_t)w_t + (1 - \phi_t)\tilde{\Omega}_t(V_t^E - V_t^N) + \dot{V}_t^N$$

The associated FOC is

$$(R_t^I - w_t - \tilde{\Omega}_t(V_t^E - V_t^N)) = 0 \quad \text{if } 0 < \phi_t < 1 \quad (5a)$$

$$(R_t^I - w_t - \tilde{\Omega}_t(V_t^E - V_t^N)) \leq 0 \quad \text{if } \phi_t = 0 \quad (5b)$$

$$(R_t^I - w_t - \tilde{\Omega}_t(V_t^E - V_t^N)) \geq 0 \quad \text{if } \phi_t = 1 \quad (5c)$$

Which implies that when $0 < \phi < 1$, V^N can be rewritten as

$$(\delta + \rho)V_t^N = w_t + \tilde{\Omega}_t(V_t^E - V_t^N) + \dot{V}_t^N$$

Again using Bellman's principle, V_t^E can be written as

$$(\delta + \rho)V_t^E = R_t^E + \dot{V}_t^E$$

Taking the difference between these two last equations, and denoting by λ_t the difference between V^E and V^N ($\lambda_t = V_t^E - V_t^N$), we get

$$(\delta + \rho)\lambda_t = R_t^E - w_t - \tilde{\Omega}_t\lambda_t + \dot{\lambda}_t$$

More generally, for ϕ either in the interior or at the boundaries, λ_t will need to satisfy

$$(\delta + \rho)\lambda_t = R_t^E - \phi_t R_t^I - (1 - \phi_t)w_t - (1 - \phi_t)\tilde{\Omega}_t\lambda_t + \dot{\lambda}_t \quad (6)$$

where ϕ_t now represents the optimal values of ϕ . In addition to (6), λ will need to satisfy the transversality condition (7).

$$\lim_{t \rightarrow \infty} \exp^{-\rho t} \lambda_t = 0 \quad (7)$$

In the case where there is an interior solution ($0 < \phi < 1$), the optimality condition (5) has a simple interpretation. It states that the current loss between being a laborer versus a manager must be compensated by the expected gain associated with potentially working alongside an expert, since in this case the gain to becoming an expert is λ_t and the probability is $\tilde{\Omega}$. Equation (6) also has a simple interpretation, especially in a steady state situation with $\dot{\lambda} = 0$. For example, when $\phi < 1$ in the steady state (which will always be the case), $\lambda = \frac{R^E - w}{\delta + \rho + \tilde{\Omega}}$, which implies that the value of being an expert is simply the actualized difference between receiving R^E and w , where the discount factor incorporates the fact that a non-expert may become an expert.

2.6 Learning and accumulation dynamics

The dynamic equation governing evolution of experts depends on (1) the fraction of individuals that supply their work time to the labor market $(1 - \phi_t)(1 - S_t)$, (2) the probability that a laborer is hired by an expert manager $(\frac{S_t L_t^E}{S_t L_t^E + (1 - S_t)\phi_t L_t^I} = \frac{S_t}{S_t + (1 - S_t)\phi_t(1 - \theta)})$, and (3) the rate at which one learns conditional on being hired to work alongside an expert (Ω). Hence, the dynamics of S is given by

$$\dot{S} = \tilde{\Omega}_t(1 - \phi_t)(1 - S_t) - \delta S_t \quad (8)$$

where $\tilde{\Omega}_t$ is given by

$$\tilde{\Omega}_t = \Omega \frac{S_t L_t^E}{S_t L_t^E + (1 - S_t)\phi_t L_t^I} = \Omega \frac{S_t}{S_t + (1 - S_t)\phi_t(1 - \theta)} \quad (9)$$

2.7 Equilibrium

An Equilibrium for this economy is composed of a sequence of wages and managerial returns $\{w_t, R_t^I, R_t^E\}$, a sequence of allocations $\{L_t^E, L_t^I, \phi_t, S_t\}$ and a learning rate $\tilde{\Omega}_t$, such that:

- 1) Given $\{w_t, R_t^I, R_t^E\}$, and $\tilde{\Omega}_t, \phi_t$ solves the non-experts decision problem with respect to occupational choice.
- 2) Given $\{w_t\}$, L_t^E and L_t^I solve the managers' problems, with R_t^E and R_t^I being the resulting revenues for the expert and non-expert manager respectively.
- 3) S_t and $\tilde{\Omega}_t$ obey equations (8) and (9).
- 4) The market for workers clears, that is,

$$(1 - \phi_t)(1 - S_t) = S_t L_t^E + (1 - S_t)\phi_t L_t^I. \quad (10)$$

An equilibrium can be obtained by finding a solution to Equations (1) through (10). Since this is a rather large system of equations, it is helpful to reduce its dimension. To this end, it is helpful to recognize that in equilibrium w_t and R_t^E satisfy the following marginal

product conditions:¹³

$$w_t = F_2(S_t + (1 - S_t)\phi_t(1 - \theta), (1 - S_t)(1 - \phi_t)) \quad (11)$$

$$R_t^E = F_1(S_t + (1 - S_t)\phi_t(1 - \theta), (1 - S_t)(1 - \phi_t)) \quad (12)$$

Using (11), (12), we can rewrite Equations (5) and (6) as follows:

$$\begin{aligned} ((1 - \theta)F_1(S_t + (1 - S_t)\phi_t(1 - \theta), (1 - S_t)(1 - \phi_t)) &- F_2(S_t + (1 - S_t)\phi_t(1 - \theta), (1 - S_t)(1 - \phi_t)) \\ &- \tilde{\Omega}_t\lambda_t) = 0 \quad \text{if } 0 < \phi_t < 1 \end{aligned} \quad (13a)$$

$$\begin{aligned} ((1 - \theta)F_1(S_t + (1 - S_t)\phi_t(1 - \theta), (1 - S_t)(1 - \phi_t)) &- F_2(S_t + (1 - S_t)\phi_t(1 - \theta), (1 - S_t)(1 - \phi_t)) \\ &- \tilde{\Omega}_t\lambda_t) \leq 0 \quad \text{if } \phi_t = 0 \end{aligned} \quad (13b)$$

$$\begin{aligned} ((1 - \theta)F_1(S_t + (1 - S_t)\phi_t(1 - \theta), (1 - S_t)(1 - \phi_t)) &- F_2(S_t + (1 - S_t)\phi_t(1 - \theta), (1 - S_t)(1 - \phi_t)) \\ &- \tilde{\Omega}_t\lambda_t) \geq 0 \quad \text{if } \phi_t = 1 \end{aligned} \quad (13c)$$

$$\begin{aligned} (\delta + \rho)\lambda_t &= F_1(S_t + (1 - S_t)\phi_t(1 - \theta), (1 - S_t)(1 - \phi_t)) (1 - \phi_t(1 - \theta)) \\ &- (1 - \phi_t)F_2(S_t + (1 - S_t)\phi_t(1 - \theta), (1 - S_t)(1 - \phi_t)) \\ &- (1 - \phi_t)\tilde{\Omega}_t\lambda_t + \dot{\lambda}_t \end{aligned} \quad (14)$$

The problem of finding an equilibrium can now be reduced to finding paths for $\{\lambda_t, S_t, \phi_t\}$ that satisfy equations (13), (14) and (8), when (9) is used to replace $\tilde{\Omega}_t$. To be part of an

¹³To see for example that $w_t = F_2(S_t + (1 - S_t)\phi_t(1 - \theta), (1 - S_t)(1 - \phi_t))$, start with the fact that $w_t = F_2(1 - \theta, L_t^I)$. Then use the market clearing condition (10), and equation (4), to solve L_t^I as a function ϕ_t and S_t . Finally exploit the homogeneity of degree zero of the marginal production condition to write w_t as given in (11). Similar steps can be used to derive (12).

equilibrium, this solution must in addition satisfy the transversality condition (7) and be such that $0 \leq \phi_t \leq 1$. Since Equation (8) is an intra-temporal condition, it is helpful to think of this equation as determining ϕ_t as a function of λ_t and S_t , and thereby noting that the problem can be thought as a system of two first order differential equations in λ_t and S_t . This observation will allow us to depict the equilibrium using a phase diagram in the λ - S space, where this space can be divide into two regions: a first region where $0 < \phi < 1$, and a second region where $\phi = 0$.¹⁴ The equation that delimits the two regions in the λ - S space is given by

$$\lambda = \frac{(1 - \theta)F_1(S, (1 - S)) - F_2(S, (1 - S))}{\Omega} \quad (15)$$

The region where $\lambda < \frac{(1-\theta)F_1(S,(1-S))-F_2(S,(1-S))}{\Omega}$ represents points where $0 < \phi < 1$, while the region where $\lambda \geq \frac{(1-\theta)F_1(S,(1-S))-F_2(S,(1-S))}{\Omega}$ represents points where $\phi = 1$.¹⁵

Before characterizing the dynamics of this system, let us begin by briefly examining what happens if $\theta = 1$. In this case, no one will choose to be an inefficient manager ($\phi_t = 0$), since it would produce no output, and therefore the dynamics for S are governed by $\dot{S} = \Omega(1 - S_t) - \delta S_t$. This will result is S converging to $\frac{\Omega}{\Omega + \delta}$ regardless of the starting value for $S > 0$. Since we want experts to prefer to be managers than to be workers along such a path, we need to have $R_t^E > w_t$. This is achieved by the following assumption:

Assumption (1) : $F_1(\frac{\Omega}{\Omega + \delta}, \frac{\delta}{\Omega + \delta}) > F_2(\frac{\Omega}{\Omega + \delta}, \frac{\delta}{\Omega + \delta})$

The dynamics for this case can be easily represented in the $\lambda - S$ space since the pair of equations representing equilibrium dynamics reduces to

$$\dot{\lambda}_t = F_1(S_t, (1 - S_t)) - F_2(S_t, (1 - S_t)) - (\Omega + \delta + \rho)\lambda_t$$

$$\dot{S} = \Omega(1 - S_t) - \delta S_t$$

¹⁴In equilibrium, ϕ is never equal to 1 since this is incompatible with market clearing.

¹⁵In the $\lambda - S$ region where $\lambda < \frac{(1-\theta)F_1(S,(1-S))-F_2(S,(1-S))}{\Omega}$, equation (8) uniquely defines ϕ as a function of S and λ .

The phase diagram for this system is depicted in Figure 1. In this figure, we have represented the $\dot{\lambda} = 0$ locus (which is downward sloping) and the $\dot{S} = 0$ locus (which is vertical). As can be seen, the equilibrium path is a saddle that converges to the tuple $\left\{ \frac{F_1(S_t, (1-S_t)) - F_2(S_t, (1-S_t))}{(\Omega + \delta + \rho)}, \frac{\Omega}{\Omega + \delta} \right\}$.

The question we want to address is when, if ever, will the equilibrium of our economy converge to this tuple when $\theta < 1$, that is, will the equilibrium outcome converge to $\phi = 0$ when running the technology using a non-expert manager produces positive output. If the economy converges to this outcome, we will say that it converges to a situation of efficient technology use. It should not be surprising that our economy with $\theta < 1$ will not always converge to a situation of efficient technology use, since it is possible that using the technology inefficiently is actually a first best outcome when θ is sufficiently small. In order to best delimit the effect of adverse selection on equilibrium outcomes, it is appropriate to focus our attention on the set of θ s for which it would be socially optimal to converge to the efficient use of technology. This set is given under Assumption 2.¹⁶

Assumption 2: $\theta > \theta^*$, where θ^* is defined by

$$\frac{\Omega F_1\left(\frac{\Omega}{\Omega + \delta}, \frac{\delta}{\Omega + \delta}\right)}{\delta + \rho + \Omega} + \frac{(\delta + \rho) F_2\left(\frac{\Omega}{\Omega + \delta}, \frac{\delta}{\Omega + \delta}\right)}{\delta + \rho + \Omega} = \max_{S^*} F(S^*(1 - \theta^*), 1 - S^*).$$

Assumption 2 shall be imposed throughout the remainder of the paper as it makes sense to restrict analysis to situations where it would be optimal to converge to the efficient use of technology in the absence of adverse selection.

The first question we address is whether, with $\theta^* < \theta < 1$, our economy will ever converge to the efficient use of technology or whether adverse selection will always prevent the economy from reaching this outcome. Proposition 1 provides a partial answer to this question.

Proposition 1: *There exists a $\bar{\theta} > \theta^*$, such that if $1 > \theta > \bar{\theta}$, then the equilibrium*

¹⁶The first best outcome for this economy is defined as the sequence $\{c_t, S_t, \phi_t\}$ that maximizes discounted consumption subject to the resource constraint $c_t \leq F(S_t + (1 - S_t)\phi_t(1 - \theta), (1 - S_t)(1 - \phi_t))$ and the accumulation equation for S . In Beaudry and Francois (2006), it is shown that Assumption 2 is necessary and sufficient for the first best to converge to the efficient use of technology.

outcome of the economy converges to the efficient use of technology. $\bar{\theta}$ is defined by $\bar{\theta} = 1 - \frac{F_2(\frac{\Omega}{\delta}, 1)}{F_1(\frac{\Omega}{\delta}, 1)}$.

Proposition 1 indicates that if a non-expert is sufficiently inefficient when acting as a manager (θ sufficiently large), then the economy will converge to the efficient use of technology even if workers cannot distinguish an expert manager from whom they can learn, from an inefficient one from whom they have nothing to learn. A typical example of the equilibrium dynamics for λ and S in this case is depicted in Figure 2. In this Figure, we have superimposed the locus of points which delimit the region where $\phi = 0$ and where $\phi > 0$, this line is denoted by $\phi = 0$. Points below this line are points where $\phi_t > 0$, while points above the line are ones where $\phi_t = 0$. An important element in this Figure, in contrast to later cases we will cover, is that the $\phi = 0$ line crosses the X axis at a value of $S < \frac{\Omega}{\Omega + \delta}$. It is worth noting that once the system enters into the region where $\phi_t = 0$, it exhibits dynamics identical to that shown in Figure 1.

Proposition 1 indicates that the presence of private information regarding the identity of managers is not fatal with respect to an economy progressively eliminating inefficient production techniques with efficient production techniques. The main force driving this result arises from the fact that when θ is sufficiently large, the incentive for a non-expert to become a manager is quite low and therefore most workers are being hired by experts, which allows them to learn and displace inefficient managers. An interesting aspect that can be seen in Figure 2 is that, for low values of S , the value of becoming an expert generally increases with the greater supply of experts. This reflects the fact that the adverse selection problem is weakening as S increases and thereby allows experts to make higher revenues as the supply of S increases.¹⁷

In light of Proposition 1, the natural follow up question is whether the absence of public information on the quality of managers can stop the economy from converging to the efficient use of technology, or whether such an informational asymmetry only affects transitional

¹⁷If the quality of managers were public information, the value of being an expert (λ) would never increase as the result of an increase in the supply of experts.

dynamics. Proposition 2 provides a partial answer.

Proposition 2: *There exists a $\underline{\theta}$, $\theta^* < \underline{\theta} < \bar{\theta}$, such that if $\theta^* < \theta < \underline{\theta}$, then the unique steady state equilibrium outcome is characterized by $S = 0$. Moreover, this steady state is locally stable. With $\underline{\theta}$ being defined as $\underline{\theta} = \bar{\theta}(\frac{\rho+\delta}{\Omega+\rho+\delta})$.*

Proposition 2 indicates that the absence of symmetric information regarding the quality of managers can stop the economy from using a technology inefficiently in the long run even if it would be socially optimal to converge to efficient use. Proposition 2 even goes further and indicates that the outcome in this case will be for best practice to be displaced completely by inefficient use of the technology. This is a case of the “bad driving out the good”. The reason for this result is that with θ not too large, there will be many non-experts deciding to be managers and hence most workers will be hired by such managers, learn nothing, and not help in the diffusion of knowledge required for efficient use. Instead some of these workers will go on to become inefficient managers themselves and perpetuate inefficient technology use. At the same time, the experts will only be hiring a small group of workers and this group is not large enough to allow growth in S . The dynamics for λ and S in this case are illustrated in Figure 3. As we can see from the figure, even if the economy starts with a large set of experts, the adverse selection problem will cause this set to dwindle over time and converge to zero.

Notice in Figure 3 that the $\phi = 0$ line crosses the x-axis at a value for S that is greater than $\frac{\Omega}{\delta+\Omega}$. This leads to a different shape of the $\dot{S} = 0$ curve, with a downward portion appearing in the region where $\phi > 0$. Moreover, at $S = \frac{\Omega}{\delta+\Omega}$, the value of λ implied by the $\phi = 0$ line is above $\frac{F_1(\frac{\Omega}{\delta},1)-F_2(\frac{\Omega}{\delta},1)}{\delta+\rho+\Omega}$. It is this last characteristic which makes a steady state with $S = \frac{\Omega}{\delta+\Omega}$ impossible.

Propositions 1 and 2, cover cases where either θ is quite high or quite low, and it is because of these extremes that we obtain very different results. The case that remains to be studied is when θ is between these two extremes, that is, $\underline{\theta} < \theta < \bar{\theta}$. Proposition 3 provides a first step in this direction.

Proposition 3: For $\underline{\theta} < \theta < \bar{\theta}$, there are three equilibrium steady states and these correspond to cases where either $S = 0$, $S = \frac{\Omega}{\Omega + \delta}$ or $S = S^H$, with $S^H = \left[\frac{1-\mu}{\mu} \left(\frac{\delta}{\Omega} + \frac{1}{1-\theta} \right) + 1 + \frac{\delta}{\Omega} \right]^{-1}$, and $\mu = \frac{(\delta + \rho)}{\theta \Omega} \left(1 - \theta - \frac{F_2\left(\frac{\Omega}{\delta}, 1\right)}{F_1\left(\frac{\Omega}{\delta}, 1\right)} \right)$. Moreover, if S_0 is close to zero, then there is an equilibrium that converges to $S = 0$; and if S_0 is close to $\frac{\Omega}{\Omega + \delta}$, then there is an equilibrium that converges to $S = \frac{\Omega}{\Omega + \delta}$.

Figure 4 illustrates the three steady states mentioned in the Proposition.¹⁸ As can be seen in the Figure, in addition to the steady state equilibrium with efficient use of technology, and the steady state equilibrium with no experts, there is a third steady state in which both efficient and inefficient use co-exist. We will refer to this equilibrium as the heterogeneous steady state equilibrium. In this third steady state, experts manage to hire a sufficient number of workers so as to ensure their replacement, but they do not hire a sufficient number to allow the number of experts to keep growing and take over the market. The Proposition further indicates that the behavior of the economy in this case is related to initial conditions, in that both the efficient use steady state and the inefficient use steady state are locally (saddle path) stable. However, the proposition is silent on the nature of dynamics for S close to the heterogeneous steady state equilibrium. The common configuration, whereby the two stable steady states would be interspersed by an unstable one, does not necessarily hold here. A necessary condition for the heterogeneous steady state equilibrium to be unstable is provided in the appendix. Via numerical simulations, we have not been able to find any cases where this equilibrium is stable. Moreover, the following proposition presents a simple sufficient condition, which is far stronger than necessary, under which the heterogeneous equilibrium is unstable. We will persist with it from hereon.

Proposition 4: If $\rho < \Omega$, then for $\underline{\theta} < \theta < \bar{\theta}$, the steady state where $S = S^H$ (the heterogeneous steady state) is locally unstable, and this equilibrium exhibits hysteresis.

Figure 4 depicts the equilibrium dynamics for the case where $\rho < \Omega$ and $\underline{\theta} < \theta < \bar{\theta}$. The

¹⁸In comparison to Figure 3, note that the value of λ implied by the $\phi = 0$ line at $S = \frac{\Omega}{\Omega + \delta}$ is positive but smaller than $\frac{F_1\left(\frac{\Omega}{\delta}, 1\right) - F_1\left(\frac{\Omega}{\delta}, 1\right)}{\Omega + \delta + \rho}$.

configuration presented in this figure has a rather standard structure for a system with three steady states. In this case, the long run properties of the economy depend on the starting values. If the economy starts with $S < S^H$, then it converges to the steady state where the technology is used inefficiently everywhere (no-experts). In contrast, if the economy starts with $S > S^H$, then it converges to a situation where there is efficient use of the technology. In other words, the economy exhibits hysteresis, with the outcome at $S = \frac{\Omega}{\Omega+\delta}$ being Pareto superior to the case where $S = 0$.¹⁹ The economic forces causing hysteresis follow quite naturally from our previous discussion. For low levels of S_0 , many non-experts choose to become managers and produce inefficiently. For this reason workers are not willing to accept low wages for working alongside a (randomly assigned) manager since they know that they are unlikely to become experts. Hence, experts do not hire enough workers to ensure their replacement and the expert managers gradually disappear. In contrast, when S_0 is high, non-experts find it attractive to supply their time to the market, as opposed to becoming inefficient managers, as they expect a substantial possibility of becoming expert. This allows experts to hire sufficient laborers to favor replacement and growth of experts. Obviously, this dependency of the long run equilibrium outcome on the initial level of S points to the importance of understanding the determinants of S_0 , and this will be discussed in a later section. If we think of S_0 as being a quantity that could be manipulable by policy, the results of Proposition 4 suggests that favoring a small amount S_0 may not be sufficient to generate the efficient use of a technology, it may instead be necessary to favor (subsidize) a substantial entry of S_0 into an economy if one wants to guarantee efficient use in the long run.

Propositions 1 to 4 describe the equilibrium outcomes that arise in the presence of adverse selection, while under Assumption 2, the efficient outcome is to have the economy converge to the efficient use of technology in all these cases. A question that arises in such a situation is whether the equilibrium outcomes described in these Propositions are second best in the sense that a social planner facing the same asymmetry of information as individuals would

¹⁹The Pareto superiority follows from Assumption 2.

choose the same allocations. The answer to this question is no: the equilibrium outcomes are not constrained socially efficient.²⁰ In general, a social planner would want to favor worker training by subsidizing employment in jobs paying wage w_t . This would have the effect of reducing the number of inefficient managers and thereby mitigating the adverse selection problem. As a result the economy would move more quickly toward the efficient use of technology.

3 The propagation of efficient technology use

In the previous section we analyzed a situation where an economy's allocation problem was restricted to a choice between running a technology in an efficient or inefficient manner. In this setup, individuals did not have any alternative use of their time besides working with this technology. In this section we extend our analysis to the case where individuals have an additional choice in terms of an outside option \bar{w} . We want to examine how such an outside option interacts with the decision to run the technology in an efficient or inefficient manner. There are at least two closely related reasons to look at this issue. First, we would like to know whether being poor (or initially technologically behind), in terms of having a lower outside option, favors the diffusion of efficient technology use or whether it hinders it. A priori, it is not clear whether a low outside option would favor or hinder the spread of efficient technology use, since a lower outside option makes the gains to adopting an efficient technology greater, but it could also exacerbate the adverse selection problem. In this case the outside option can be thought as representing the wage level associated with an older technology in use in the economy, and we can think of the technology $F(\cdot)$ as being a newly available technology which can be run in an inefficient or efficient manner.

Answering how the level of an outside option affects technology use will offer insight regarding whether there is a complementarity aspect to efficiently using a technology, that is, whether the spread of efficient technology use in one generation of technology, through

²⁰Given the asymmetry of information, a social planner could still attain the first best by promising to pay all individuals the same amount. In such a case, individuals would have no incentive to lie about their identity thereby allowing the best to be implemented.

its effect on the reservation value of time, could help the spread of efficient use in another generation. Alternatively, instead of different generations of technology, we can think of \bar{w} as representing the value of time in another industry. By examining how changes in \bar{w} affects the diffusion of technological knowledge, we can potentially shed light on the issue of positive or negative spillovers between sectors, whereby the efficient use of technology in one sector could feedback on the diffusion process in another sector though the effect on reservation wages.

To examine how the presence of an outside option affects the allocation process between efficiently and inefficiently run firms, we need to add to the household's problem an additional decision margin. Let ϕ_{1t} represent the fraction of non-experts who decide to become inefficient managers, and let ϕ_{2t} represent the fraction of non-experts who choose the outside option \bar{w} . This leaves $(1 - \phi_{1t} - \phi_{2t})$ of non experts who supply their work time to managers of the technology $F(\cdot)$ as workers. As before, the fraction ϕ_1 and ϕ_2 can also be interpreted as the probabilities that govern a non-expert decision problem. Bellman's principle of optimality implies that the value of being a non-expert at time t will satisfy:

$$(\delta + \rho)V_t^N = \max_{\phi_{1t}, \phi_{2t}} \phi_{1t}R_t^I + (1 - \phi_{1t} - \phi_{2t})w_t + \phi_{2t}\bar{w} + (1 - \phi_{1t} - \phi_{2t})\tilde{\Omega}_t(V_t^E - V_t^N) + \dot{V}_t^N$$

The associated FOCs are

$$(R_t^I - w_t - \tilde{\Omega}_t\lambda_t) = 0 \quad \text{if } 0 < \phi_t < 1 \quad (16a)$$

$$(R_t^I - w_t - \tilde{\Omega}_t\lambda_t) \leq 0 \quad \text{if } \phi_t = 0 \quad (16b)$$

$$(R_t^I - w_t - \tilde{\Omega}_t\lambda_t) \geq 0 \quad \text{if } \phi_t = 1 \quad (16c)$$

and

$$(\bar{w} - w_t - \tilde{\Omega}_t\lambda_t)(1 - \phi_{2t})\phi_{2t} = 0 \quad (17)$$

Using the fact that the value function for an expert satisfies $(\delta + \rho)V_t^E = R_t^E + \dot{V}_t^E$, λ will need to satisfy

$$(\delta + \rho)\lambda_t = R_t^E - \phi_{1t}R_t^I - (1 - \phi_{1t} - \phi_{2t})w_t - \phi_{2t}\bar{w} - (1 - \phi_{1t} - \phi_{2t})\tilde{\Omega}_t\lambda_t + \dot{\lambda}_t \quad (18)$$

where ϕ_{1t} and ϕ_{2t} are optimal values. Given that in equilibrium $R_t^E = F_1(S_t + (1 - S_t)\phi_{1t}(1 - \theta), (1 - S_t)(1 - \phi_{1t} - \phi_{2t}), 1)$, $w_t = F_2(S_t + (1 - S_t)\phi_{1t}(1 - \theta), (1 - S_t)(1 - \phi_{1t} - \phi_{2t}))$ and that $R_t^E = R_t^I(1 - \theta)$, finding an equilibrium for the economy with an outside option can be reduced to finding a tuple $\{S_t, \phi_{1t}, \phi_{2t}, \lambda_t\}$, with $1 < \phi_{it} \leq 1$, that satisfy equations (16), (17), (18), the transversality condition (7), the accumulation equation for S_t given by

$$\dot{S}_t = \tilde{\Omega}_t(1 - \phi_{1t} - \phi_{2t})(1 - S_t) - \delta S_t \quad (19)$$

where $\tilde{\Omega}_t$ satisfies $\tilde{\Omega}_t = \Omega \frac{S_t}{S_t + (1 - S_t)\phi_{1t}(1 - \theta)}$

In analyzing the case with an outside option, we again want to focus on a situation where it would be socially optimal to converge toward using the technology $F(\cdot)$ efficiently, with workers abandoning the outside option in favor of the new technology. In addition to Assumption 2, the condition that guarantees that it would be socially desirable to converge toward using $F(\cdot)$ in a efficient manner is given by Assumption 4.

Assumption 4:

$$\bar{w} < \frac{\Omega F_1\left(\frac{\Omega}{\Omega + \delta}, \frac{\delta}{\Omega + \delta}\right)}{\delta + \rho + \Omega} + \frac{(\delta + \rho) F_2\left(\frac{\Omega}{\Omega + \delta}, \frac{\delta}{\Omega + \delta}\right)}{\delta + \rho + \Omega}$$

As in the case without an outside option, it can be shown that if $\theta \geq \bar{\theta}$, the economy converges to the efficient use of the technology, and if $\theta \leq \underline{\theta}$ then the economy converges to a situation where $S = 0$. Note that in this case, $\bar{\theta}$ and $\underline{\theta}$ take on the same values as in the case without the outside option. The situation of interest is when $\underline{\theta} < \theta < \bar{\theta}$, that is, how does the outside option affect outcomes when there are three steady states. Proposition 5 highlights the role played by the outside option in favoring convergence to the efficient use of technology.

Proposition 5: *If the outside option $\bar{w} > \hat{w} = (1 - \theta)F_1(\frac{\Omega}{\delta}, 1)$, then even if $\underline{\theta} < \theta < \bar{\theta}$, the economy will converge to the steady state equilibrium characterized by $S = \frac{\Omega}{\Omega + \delta}$, $\phi_{1t} = \phi_{2t} = 0$, that is, the economy converges to the efficient use of the technology.*

The interesting aspect of Proposition 5 is that it indicates that a higher outside option favors convergence to the efficient use of technology. In particular, if $\underline{\theta} < \theta < \bar{\theta}$, we saw in Proposition 4 that for S_0 not too large, the economy would converge to $S = 0$ if $\bar{w} = 0$. In contrast, Proposition 5 indicates that this will not happen if \bar{w} is sufficiently high. The reason that the outside option matters is that it affects ϕ_{1t} . In particular, at the point where a non-expert is indifferent between becoming a manager, exercising his outside option, or supplying his time as a laborer, an increase in the outside option leads to a greater exit of non-experts from management than from the supply of laborers, and hence this reduces the adverse selection problem and allows experts to increase their hiring, thereby favoring growth in efficient technology use. Mechanically, what the outside option is doing is that it is eliminating the middle steady state depicted in Figure 4. A typical equilibrium configuration for a case where $\bar{w} \geq \hat{w}$ is illustrated in Figure 5.

Proposition 6 offers a counterpart to Proposition 5 by showing that if the outside option is not sufficiently high, then the economy will not converge to the efficient use of the technology.

Proposition 6: *If $\bar{w} < \hat{w}$, $\underline{\theta} < \theta < \bar{\theta}$, and $S_0 < S^H$, then the economy will not converge to the efficient use of the technology.*

Propositions 5 and 6 demonstrate that the diffusion of efficient technology use can depend on the level of productivity in a backup or incumbent technology. In other words, these proposition suggest a type of complementarity in efficient technology use. If a society uses one technology efficiently, this is likely to favor the efficient use of another technology, either in another sector or in the same sector later in time. This suggests that high wages favor high productivity.

When $\theta < \bar{\theta}$, we now know that the economy may converge to a situation with $S = 0$ (that is, a situation with no experts), either when $\bar{w} < \hat{w}$ or when $\theta < \underline{\theta}$. If this is the case, it is of

interest to know whether the final outcome involves non-experts working inefficiently with the new technology $F(\cdot)$, or if instead everyone maintains the old technology captured by the outside option \bar{w} . The answer to this question depends on whether the outside option \bar{w} is greater or less than $\max_{\phi} F((1-\theta)\phi, (1-\phi))$, that is, on whether the inefficiently managed new technology is more or less productive than the outside option. If $\bar{w} > \max_{\phi} F((1-\theta)\phi, (1-\phi))$, then in the absence of experts it is optimal to remain with the old technology and this will be the equilibrium outcome. Conversely when $\bar{w} < \max_{\phi} F((1-\theta)\phi, (1-\phi))$, then with $S = 0$ it is optimal to drop the outside option and produce inefficiently with the new technology, and again this will be the equilibrium outcome. Hence the model implies that adverse selection in the diffusion of management skills may be why an economy may not adopt a new technology or may adopt it, but use it in an inefficient manner.

3.1 Discussion

Are our findings consistent with micro level accounts of the technology diffusion process in LDCs? We argue here that such studies suggest that on-the-job skills acquisition plays a major role in improving worker know-how, but is poorly done. This is particularly true in the economies with lowest existing productivity, i.e., those of sub-Saharan Africa, as is consistent with our findings above. Moreover, there is some evidence that precisely the impediments arising from information asymmetries that are the focus of our theory seem to play a key role in limiting diffusion.

Numerous large scale overviews attest to the important role played by skills that are learned on the job: In an African context, Banji (2004); in East-Asian countries, (De Ferranti et. Al. p.119); in Latin American countries (Batra 2002). More detailed micro-level studies corroborate these overviews. For example, Ciceon (2001) found, in a cross-industry sample of Mexican firms, that about two thirds of training was provided in-house, which De Ferranti et. al. argue is in line with numbers reported elsewhere in Latin America. Biggs Shah and Srivastava (1995), in their comprehensive survey of modern enterprises in three sub-Saharan African economies, found skills shortages to be a key impediment to productivity

– with difficulties in workplace transmission playing a prominent role. Stolovich and Ngoa-Nguele (2001) document both the importance of on-the-job learning, and the numerous shortcomings in the way such learning is implemented, especially in African countries. They report that: Instructors are poorly trained; or not experienced. Funds and training materials are not contributed; Training developers and managers rarely possess the competencies to fulfill their roles; Needs are rarely systematically assessed; Training practices are loosely structured, or entirely unstructured, and limited, short-term goals drive training activities.

In many countries apprenticeships correspond well with the implicit on-the-job learning agreement we have studied here. In Africa, these are the dominant form of learning in small and medium firms (Banji 2004). This relatively simple contract (training, in return for reduced wages, and perhaps a fee) has the ability to overcome the liquidity constraints workers may face in paying for skills. Apprenticeships are ubiquitous in small and medium sized enterprises in much of Africa. For example, Adams and Johanson (2004 p.131) note that in West Africa it is not uncommon to find more apprentices than regular wage employees in informal firms. They have also increased in importance. Haan and Serriere (2002, 50,57 133) note that, in Senegal, apprenticeships grew from 40% of the informal sector workforce in 1980 to 70% in 1995. In Benin, the number of traditional apprenticeships increased by more than 10% per year between 1979 and 1992. In Cameroon, apprenticeships have enrollments of over 200,000 compared with total public training of 14,000. In Ghana 80 to 90% of all basic skills training comes from traditional or informal apprenticeships (Atchoarena and Delluc 2001 p.225).

Almost all studies of workplace training in African countries emphasize the poor performance of firms on the in-house training dimension; see Stolovich and Ngoa-Nguele (2001, p.461). What are the causes of this? For the smaller enterprises, which are the largest employers, the problems seem to arise due to the informal and unstructured nature of the traditional learning relationship. Adams and Johanson (p.135) suggest that the informal sector masters themselves lack awareness of the shortcomings in their own skills. Even where they understand that in order to transmit better skills to their apprentices, they would have

to upgrade their own, they do not. The authors suggest that information symmetries play a key role:

“Entrepreneurs (master craftsmen) who have upgraded their skills often are not able to increase their prices in order to reflect their improved product. The incentives for new training amongst suppliers is weak given the lack of demand and the risks involved in pioneering new training.” p.135

This, in words, seems to correspond with the consequences of the adverse selection problem we have studied. Since there is no way for the training provider to credibly signal the quality of the skills he will impart, it is impossible to charge for it. Adams and Johanson note that even when contracts can be written to ensure some training is provided, problems arise with respect to quality:

“Even when such contracts can be devised, the informal manner in which some skills are imparted on the job will make it difficult to monitor training and enforce contracts, leaving the risk of underinvestment”

Since future wages will capitalize the value of the skills, the firms should be able to obtain worker effort cheaply. But again, the impossibility of contracting seems to thwart these efforts.

“.. difficulties in monitoring the training commitments in the enterprises and the potential variance between wages paid in current and in future employment can produce underinvestments in general skills training.” p.29.

Adams and Johanson explicitly blame this state of affairs on the lack of objective quantifiable characteristics in the training agreement:

“One of the major shortcomings of traditional apprenticeship training is the lack of quality assurance, through either monitoring the process or applying objective end of training assessments.” p.146.

It is interesting to note that where governments have attempted to improve this process of on-the-job related knowledge diffusion they have focused on overcoming the asymmetric information problem that we study. Some governments have attempted to set up external (to the firm) accreditation agencies. The National Vocation Training Institute in Ghana introduced competency based testing of apprentices. South Africa subsidized “Learnerships”, structured workplace learning programs offering a qualification upon completion. Informal sector associations have also attempted upgrading by certifying skills and sometimes providing supplementary training and common examinations in Ghana, Kenya, Cameroon, Tanzania and Zimbabwe; see Adams and Johanson (p. 138)., and Haan and Serriere (p.151) for similar initiatives in Senegal and Benin.

By means of subsidies to training, overseeing bodies have also attempted to improve the content and form of training. Providers receive transfers for training in return for oversight bodies establishing some degree of influence over content. Payroll levies of this form have been used to reimburse employers for the costs of on-the-job training in many European countries, and also in Singapore, Malaysia, Morocco, Turkey and a host of Latin American countries. Others have used matching grants schemes towards similar ends: Japan, Korea, Singapore, Germany, Holland and Scandinavia; see De Ferranti et. al. p.129 for details.

In summary, this literature suggests that on-the-job skills are an important accompaniment to the use of productive technologies, and to increasing productivity. But even though such skills are valuable, the fact that they must be imparted in-house, in largely unstructured and informal workplace settings, gives rise to transactions problems. As we have emphasized, a key problem seems to be that individuals willing to “buy” the skills are unsure as to what they are buying, and cannot easily monitor the training transaction. Consequently, individuals willing to “sell” the skills can have very weak incentives to deliver training of quality. As our theoretical model has shown, this can lead economies to either implement new technologies inefficiently, or to maintain in use old technologies that are less efficient than freely available new ones. This is particularly evident in the poorest economies of sub-Saharan Africa where traditional means of workplace learning that correspond roughly to

apprenticeships, have been seen to entrench inefficient production methods.

4 Migration incentives and the determinants of S_0

Up to now, we have treated S_0 as exogenously given. In this section we briefly discuss the implications of allowing $S_0 = 0$, and consider the incentives for foreign experts to migrate, and thereby potentially ignite a process of diffusion of technological knowhow. In particular, consider the situation where the foreign expert is paid $F_1(\frac{\Omega}{\delta}, 1)$ in his home country, that is, he is assumed to come from an economy which has converged to the efficient use of the technology. In this case, we will say that the foreign expert has an incentive to migrate if the flow revenue he could make by temporarily reallocating is greater than $F_1(\frac{\Omega}{\delta}, 1)$. In an economy with $S_0 = 0$, the allocation of non-experts between management and labor will be such that $R^l = w$. In other terms, the share of non experts, denoted $\phi_{(S_0=0)}$, working as managers when $S_0 = 0$ will be determined by $(1 - \theta)F_1(\phi_{(S_0=0)}(1 - \theta), 1 - \phi_{(S_0=0)}) = F_2(\phi_{(S_0=0)}, 1 - \phi_{(S_0=0)})$. So a foreign expert will want to migrate if $F_1(\phi_{(S_0=0)}(1 - \theta), 1 - \phi_{(S_0=0)}) > F_1(\frac{\Omega}{\delta}, 1)$. Proposition 7 provides a characterization of when foreign experts would want to migrate to an economy with no experts.

Proposition 7: *An economy with $S_0 = 0$ will be attractive to migration by foreign experts if $\theta > \bar{\theta}$, or if $\underline{\theta} < \theta \leq \bar{\theta}$ and $w > \hat{w}$. Otherwise, foreign experts would not have an incentive to migrate.*

The first aspect to note from Proposition 7 is that migration incentives arise precisely when the economy will converge to efficient use when $S_0 > 0$. Hence, for situations where $\theta > \bar{\theta}$ or where $\underline{\theta} < \theta \leq \bar{\theta}$ and $w > \hat{w}$, it is reasonable to expect the efficient use of technology to prevail in the presence of unhindered international migration.²¹ In contrast, if neither of these conditions are met, the inefficient use of the technology captured in previous propositions

²¹Instead of focusing on incentives for foreign experts to migrate, we could have focused on incentives for domestic residents to go abroad to learn the relevant skills. However, in such a case we would still want to ensure that such newly trained individuals have an incentive to return to their country of origin, which corresponds to having incentives for foreign experts to migrate.

appears as a robust equilibrium outcome. Accordingly, an economy stuck in a situation with no experts would generally need some form of intervention if efficient technology use is to diffuse. However, such an intervention would not necessarily need be in the form of subsidies, but instead could take the form of a policy aimed primarily at coordinating the inflow of a critical mass of experts. For example, if $\underline{\theta} < \theta \leq \bar{\theta}$, then for it to be attractive to foreign experts to migrate, it must simply be the case that $S > S_H$; for in this situation $R^E > F_1(\frac{\Omega}{\delta}, 1)$.²² If foreigners expect that there would be an inflow of a critical mass of experts greater than S^H , this would be sufficient to create the incentive for the inflow.

5 Conclusion

When skills that are key to successfully implementing technologies must be learned on-the-job, and when such learning cannot be fully contracted over, firms can profit by providing poor training to their workers. The evidence suggests that, in many developing economies, such training is indeed performed quite poorly, and this process of skill diffusion is weak. We believe that this has the potential to explain some part of the persistently low productivity observed in many less developed economies.

The model here suggests that, in the presence of such problems, economies may not converge to full-scale implementation of more productive technologies, even where it is efficient to do so. These economies may be subject to hysteresis, where small scale implementation of efficient use will fail whereas larger scale use would have been successful. Technologies will be more likely to be used efficiently where they do not represent too great a departure over existing production methods. Paradoxically, economies with weak existing technologies, where opportunity costs are low, will be those most likely to fail in implementing widespread efficient technology use.

The model also provides an explanation for the relatively small observed flows of skilled individuals from economies where their skills are abundant (rich countries) to economies

²²From the steady state conditions defining the heterogeneous steady state with $S = S_H$, $R^E = F_1(\frac{\Omega}{\delta}, 1)$ and hence at $S = S^H$ foreign experts would be exactly indifferent between migrating or not.

where their skills are scarce and needed (the poor). This works counter to the standard neo-classical effect favoring scarcity. Here, skilled individuals reap equilibrium returns from their role in diffusing knowledge to the unskilled. When the poor countries are also the ones in which this process of knowledge diffusion is performed badly due to the presence of firms who provide substandard training, the incentives for the skilled to migrate will be weak.

These considerations suggest remedial policies. Firstly, the level of on-the-job learning undertaken in a decentralized economy will not, generally, correspond with the optimal second best (information constrained) choices of a planner. There is a case for subsidizing such on-the-job learning. Secondly, though it is an equilibrium for economies to perform such training poorly, we show that it need not be an inherent characteristic of the country. If proper training in modern technologies is widespread enough, or if such training can be effectively subsidized (even for only a limited period) then an economy can move out of an inefficient equilibrium to one in which it uses technologies at a par with world best practice. The paper suggests a set of key institutions and policies that facilitate this process. These would be those that help in increasing the diffusion of expertise relative to non-expert and poorly structured training. These include organizations monitoring the content of instruction on-the-job, subsidies to training in return for curriculum compliance, adjudication agencies designed to test the competency of trainees, and organizations that monitor and accredit those providing the training. The developing world sees many examples of such organizations and institutions, and the present paper suggests that their relative successes may play a role in understanding some part of cross-country productivity differences.

Appendix: Proofs.

The following 4 lemmas will be useful to prove Propositions 1-4. These lemmas will describe properties of the steady states for the dynamic system defined by equations (8), (13),(14) and the transversality condition (7). A steady state for this system is a triplet λ, S, ϕ that satisfies the following equations.

$$0 = \tilde{\Omega}(1 - \phi)(1 - S) - \delta S \quad (A1)$$

$$\begin{aligned} (\delta + \rho)\lambda &= F_1\left(\frac{S + (1 - S)\phi(1 - \theta)}{(1 - S)(1 - \phi)}, 1\right) (1 - \phi(1 - \theta)) - (1 - \phi)F_2\left(\frac{S + (1 - S)\phi(1 - \theta)}{(1 - S)(1 - \phi)}, 1\right) \\ &\quad - (1 - \phi)\tilde{\Omega}\lambda \end{aligned} \quad (A2)$$

$$(1 - \theta)F_1\left(\frac{S + (1 - S)\phi(1 - \theta)}{(1 - S)(1 - \phi)}, 1\right) - F_2\left(\frac{S + (1 - S)\phi(1 - \theta)}{(1 - S)(1 - \phi)}, 1\right) = \tilde{\Omega}\lambda \quad \text{if } 0 < \phi < 1 \quad (A3)$$

$$(1 - \theta)F_1\left(\frac{S + (1 - S)\phi(1 - \theta)}{(1 - S)(1 - \phi)}, 1\right) - F_2\left(\frac{S + (1 - S)\phi(1 - \theta)}{(1 - S)(1 - \phi)}, 1\right) \leq \tilde{\Omega}\lambda \quad \text{if } \phi = 0 \quad (A4)$$

where $\tilde{\Omega}$ is given by

$$\tilde{\Omega} = \Omega \frac{S}{S + (1 - S)\phi(1 - \theta)} \quad (A5)$$

$$0 \leq \phi \leq 1 \quad (A6)$$

In the above, the case with $\phi = 1$ is disregarded since it is never part of an equilibrium.

Lemma 1: The system of equation A1-A6 admits the solution $\phi = 0$, $S = \frac{\Omega}{\delta + \Omega}$ and $\lambda = \frac{F_1(\frac{\Omega}{\delta}, 1) - F_2(\frac{\Omega}{\delta}, 1)}{\delta + \Omega + \rho}$ iff $\theta \geq \underline{\theta}$. This is the unique solution to the system with $\phi = 0$. Moreover, this solution is a locally (saddle path) stable steady state of the dynamic system defined by Equations 8, 13, 14.

Proof: With $\phi = 0$, $S = \frac{\Omega}{\delta + \rho}$ and $\lambda = \frac{F_1(\frac{\Omega}{\delta}, 1) - F_2(\frac{\Omega}{\delta}, 1)}{\delta + \Omega + \rho}$ is clearly the unique solution to A1 and A2. Therefore, what needs to be checked is that A4 is satisfied at this solution. Evaluating A4 at this solution gives a restriction on θ which corresponds exactly to the condition $\theta \geq \underline{\theta}$.

The linearized dynamic equations for S and λ evaluated at this solution takes the form:

$$\dot{S} = -(\Omega + \delta)(S_t - S^{ss})$$

$$\dot{\lambda} = -\left(\frac{F_1(\frac{\Omega}{\delta}, 1) - F_2(\frac{\Omega}{\delta}, 1)}{(1 - S^{ss})^2}\right)(S_t - S^{ss}) + (\delta + \rho + \Omega)(\lambda - \lambda^{ss})$$

where S^{ss} and λ^{ss} represent the steady state values.

This system has one root greater than zero and one root smaller than zero, and hence it is locally (saddle path) stable.

Lemma 2: The system of equations A1-A6 admits only one solution with $S = 0$, moreover this solution is a steady state of the dynamic system given by equations 8, 13, 14. If $\theta > \bar{\theta}$, this steady state is unstable, if $\theta < \bar{\theta}$ this steady state is locally (saddle path) stable.

Proof: Let z be defined by $(1 - \theta)F_1(z, 1) = F_2(z, 1)$. The $S = 0$, $\phi = \frac{z}{z+1-\theta}$, $\lambda = \frac{\theta F_1(\frac{\phi(1-\theta)}{1-\phi}, 1)}{\delta + \rho}$ is the only solution of A1-A6 with $S=0$. The linearized dynamic equation for S and λ , evaluated at this steady state, takes the form:

$$\dot{S} = \left(\frac{\Omega(1 - \phi^{ss})}{\phi^{ss}(1 - \theta)} - \delta\right)(S_t - S^{ss})$$

$$\dot{\lambda} = -\left(\frac{\theta F_{11}(\frac{\phi}{1-\phi}, 1)}{(1 - \phi^{ss})^2}\right)(S_t - S^{ss}) + (\delta + \rho)(\lambda - \lambda^{ss})$$

If $\theta > \bar{\theta}$, then this system has two roots greater than zero and therefore is unstable. If $\theta < \bar{\theta}$, then this system has one root greater than zero and one less than zero and hence the dynamic system is locally (saddle path) stable.

Lemma 3: The system of equations A1, A2 A3 admits only one solution with $S \neq 0$. This solution will have $0 < \phi < 1$ and $0 < S < \frac{\Omega}{\Omega + \delta}$ iff $\underline{\theta} < \theta < \bar{\theta}$.

Proof: With $S \neq 0$, the system of equations A1-A3 can be solved explicitly for the unique solution. This solution is given by

$$S = S^H = \frac{1}{\frac{1-\mu}{\mu} \left(\frac{\frac{\delta}{\Omega}(1-\theta)+1}{(1-\theta)} \right) + \left(1 + \frac{\delta}{\Omega} \right)}$$

$$\phi = \frac{1}{1 + \frac{\delta}{\Omega}(1-\theta) \left(1 + \frac{\mu}{1-\mu} \right)}$$

$$\lambda = \frac{\theta F_1\left(\frac{\Omega}{\delta}, 1\right)}{\delta + \rho}$$

where

$$\mu = \left(\frac{\delta + \rho}{\theta \Omega} \right) \left(1 - \theta - \frac{F_2\left(\frac{\Omega}{\delta}, 1\right)}{F_1\left(\frac{\Omega}{\delta}, 1\right)} \right)$$

For this solution to be such that $0 < \mu < 1$ and $0 < S < \frac{\Omega}{\delta + \Omega}$, it must be the case that $0 < \mu < 1$. The condition $0 < \mu < 1$ is equivalent to the condition $\underline{\theta} < \theta < \bar{\theta}$.

Lemma 4: $\underline{\theta} > \theta^*$

Proof: The definition of $\underline{\theta}$ combined with the definition of θ^* given in Assumption 2 implies

$$(1 - \underline{\theta})F_1\left(\frac{\Omega}{\delta}, 1\right) = \max_{\phi} F(\phi(1 - \theta^*), 1 - \phi)$$

At this maximum denoted by ϕ^* necessarily: $(1 - \phi^*)F_1(\phi^*(1 - \theta^*), 1 - \phi^*) = F_2(\phi^*(1 - \theta^*), 1 - \phi^*)$. Also from the definition of $\underline{\theta}$, we know that

$$(1 - \underline{\theta})F_1\left(\frac{\Omega}{\delta}, 1\right) > F_2\left(\frac{\Omega}{\delta}, 1\right)$$

It then follows from the above that:

$$F_2\left(\frac{\Omega}{\delta}, 1\right) < F_2(\phi^*(1 - \theta^*), 1 - \phi^*)$$

and the converse is necessarily true for the first argument:

$$F_1\left(\frac{\Omega}{\delta}, 1\right) > F_1(\phi^*(1 - \theta^*), 1 - \phi^*)$$

Multiplying both sides of this inequality by $(1 - \theta^*)$ and using the first equality above then implies that

$$\max_{\phi} F(\phi(1 - \theta^*), 1 - \phi) < (1 - \theta^*)F_1\left(\frac{\Omega}{\delta}, 1\right)$$

Combining the first and last equation, we get

$$\underline{\theta} > \theta^*.$$

Proof of Proposition 1: Lemma 1 implies that there is one steady state with $S = \frac{\Omega}{\Omega + \delta}$, and that this steady state is (saddle path) stable. Lemma 2 implies that there is one steady state with $S = 0$, and this steady state is unstable. Finally, Lemma 3 implies that there are no other steady states. Hence, Lemmas 1 2 and 3 imply that efficient technology use is the only stable steady state outcome when $\theta > \bar{\theta}$, and finally Lemma 4 implies that $\bar{\theta} > \theta^*$ since $\bar{\theta} > \underline{\theta}$.

Proof of Proposition 2: Lemmas 1 2 and 3 imply that $S = 0$ is the only steady state when $\theta < \underline{\theta}$. Moreover, Lemma 2 implies that this steady state is locally stable. Finally Lemma 4 implies that $\underline{\theta} > \theta^*$.

Proof of Proposition 3: Lemmas 1 2 and 3 imply that there are three steady states to the dynamic system when when $\underline{\theta} < \theta < \bar{\theta}$. Moreover, Lemma 1 and Lemma 3 imply that both the steady state with $S = 0$ and the one with $S = \frac{\Omega}{\Omega + \delta}$ are locally stable.

Proof of Proposition 4: Lemma 3 implies that the dynamic system defined by equations 8, 13a and 14 has a steady state with $0 < S < \frac{\Omega}{\Omega + \delta}$ denoted S^H , and $0 < \phi < 1$. The dynamic system is:

$$\begin{aligned} \dot{\lambda} &= (\delta + \rho) \lambda_t - \theta F_1(\psi_t, 1) \\ \dot{S} &= \Omega \frac{S}{\psi} - \delta S, \end{aligned}$$

where $\psi = \frac{S+(1-S)\phi(1-\theta)}{(1-\phi)(1-S)}$, and where ϕ is implicitly defined as a function of S and λ by

$$(1-\theta)F_1(\psi_t, 1) - F_2(\psi_t, 1) = \frac{\Omega\lambda S}{(1-\theta)(1-S)\psi} = \frac{\Omega\lambda\left(1 + \frac{(1-\theta)}{\psi}\right)}{(1-\theta)\frac{(1-S)}{S} + 1}. \quad A7$$

Furthermore, from Lemma 3, the steady heterogeneous steady state has $\psi = \frac{\Omega}{\delta}$, $S = 1/\left(\frac{1-\mu}{\mu}\left(\frac{\delta}{\Omega} + \frac{1}{1-\theta}\right) + 1 + \frac{\delta}{\Omega}\right)$, where

$$\mu = \left(\frac{\delta + \rho}{\theta\Omega}\right) \left(1 - \theta - \frac{F_2\left(\frac{\Omega}{\delta}, 1\right)}{F_1\left(\frac{\Omega}{\delta}, 1\right)}\right) \quad A8$$

Linearizing this system yields:

$$\begin{bmatrix} \dot{\lambda} \\ \dot{S} \end{bmatrix} = \begin{bmatrix} (\delta + \rho) - \theta F_{11}(\psi, 1)\psi_\lambda & -\theta F_{11}\psi_S \\ -\frac{S}{\psi^2}\psi_\lambda\Omega & -\frac{\Omega S}{\psi^2}\psi_S \end{bmatrix} \begin{bmatrix} \lambda \\ S \end{bmatrix}.$$

The eigenvalues of this system are the γ solving:

$$\gamma^2 - \left((\delta + \rho) - \theta F_{11}\psi_\lambda - \frac{\Omega S}{\psi^2}\psi_S\right)\gamma - (\delta + \rho)\frac{\Omega S}{\psi^2}\psi_S = 0.$$

For this dynamic system to be unstable it is necessary and sufficient that the γ solving this equation are both positive. This will be the case if the following two conditions hold:

$$\begin{aligned} \psi_S &< 0 \\ (\delta + \rho) - \theta F_{11}\psi_\lambda - \frac{\Omega S}{\psi^2}\psi_S &> 0 \end{aligned}$$

The terms ψ_λ , ψ_S above are determined by totally differentiating (A7):

$$\begin{aligned} &\left[(1-\theta)F_{11}(\psi, 1) - F_{12}(\psi, 1) + \frac{\lambda\Omega(1-\theta)}{\psi^2\left(1 + \frac{(1-S)}{S}(1-\theta)\right)} \right] d\psi \\ &= \left[\frac{\Omega\left(1 + \frac{(1-\theta)}{\psi}\right)}{1 + \frac{(1-S)}{S}(1-\theta)} \right] d\lambda + \left[\frac{\lambda\Omega\left(1 + \frac{1-\theta}{\psi}\right)(1-\theta)}{\left(1 + \frac{(1-S)}{S}(1-\theta)\right)^2 S^2} \right] dS. \quad A9 \end{aligned}$$

Also, in this steady state:

$$\begin{aligned} \frac{(1-S)}{S}(1-\theta) + 1 &= \frac{1}{\mu}\left(\frac{1-\theta}{\psi} + 1\right) \\ \Rightarrow \frac{\mu}{S} &= \frac{\delta}{\Omega} + \frac{1}{1-\theta} - \frac{\mu\theta}{1-\theta}. \quad A10 \end{aligned}$$

Part A. $\psi_S < 0$

It is necessary to show that $(1 - \theta) F_{11}(\psi, 1) - F_{12}(\psi, 1) < -\frac{\lambda\Omega(1-\theta)}{\psi^2(1+\frac{(1-S)}{S}(1-\theta))}$ at $\psi = \frac{\Omega}{\delta}$.

This is equivalent to showing that $\frac{d}{d\psi} [(1 - \theta) F_1(\psi, 1) - F_2(\psi, 1)] < \frac{d}{d\psi} \left[\frac{\lambda\Omega(1+\frac{(1-\theta)}{\psi})}{1+\frac{(1-S)}{S}(1-\theta)} \right]$ at $\psi = \frac{\Omega}{\delta}$.

We know that, at $\psi = \frac{S}{1-S} < \frac{\Omega}{\delta}$, it must be the case that $(1 - \theta) F_1(\psi) - F_2(\psi) > \lambda\Omega$. This is necessary from $0 < \phi < 1$. We also know from (A7) that, at $\psi = \frac{\Omega}{\delta} (1 - \theta) F_1(\psi, 1) - F_2(\psi, 1) = \frac{\Omega\lambda(1+\frac{(1-\theta)}{\psi})}{(1-\theta)\frac{(1-S)}{S}+1}$, and from Lemma 3 that this is the only point at which these two curves cross. It necessarily follows then that the curve $(1 - \theta) F_1(\psi, 1) - F_2(\psi, 1)$ crosses $\frac{\Omega\lambda(1+\frac{(1-\theta)}{\psi})}{(1-\theta)\frac{(1-S)}{S}+1}$ from above at $\psi = \frac{\Omega}{\delta}$. That is, since they are both negatively sloped, that $\frac{d}{d\psi} [(1 - \theta) F_1(\psi, 1) - F_2(\psi, 1)] < \frac{d}{d\psi} \left[\frac{\lambda\Omega(1+\frac{(1-\theta)}{\psi})}{1+\frac{(1-S)}{S}(1-\theta)} \right]$, which implies that $\psi_S < 0$.

Part B $(\delta + \rho) - \theta F_{11}\psi_\lambda - \frac{\Omega S}{\psi^2}\psi_S > 0$

Using (A9) we re-write the relevant condition as:

$$(\delta + \rho) \left[\frac{1}{\mu} \left(\frac{1 - \theta}{\psi} + 1 \right) ((1 - \theta) F_{11} - F_{21}) + \frac{\lambda\Omega(1 - \theta)}{\psi^2} \right] < \theta F_{11}\Omega \left(1 + \frac{1 - \theta}{\psi} \right) + \frac{\Omega\mu}{S} \frac{\lambda\Omega(1 - \theta)}{\psi^2}.$$

A sufficient condition for this is: $(\delta + \rho) \frac{1}{\mu} (1 - \theta) F_{11} - \theta F_{11}\Omega < 0$, and $\frac{\Omega\mu}{S} > \delta + \rho$.

Since $F_{11} < 0$, the first condition becomes: $(\delta + \rho) \frac{1}{\mu} (1 - \theta) - \theta\Omega > 0$, which follows immediately from the definition of μ in (A8).

Using (A10) the second condition rearranges to: $\mu\theta < 1 - \frac{\rho}{\Omega} (1 - \theta)$. Since $0 < \mu < 1$ in this steady state, then a sufficient condition for this to hold is that $\rho < \Omega$.

Proofs for Section 3

The following 3 lemmas, which will mirror Lemmas 1-3, will be useful to prove Propositions 5 and 6. These lemmas will describe properties of the steady states for the dynamic system defined by equations (16), (17),(18) and (19). A steady state for this system is a tuple $\lambda, S, \phi_1, \phi_2$ that satisfies the following equations.

$$0 = \tilde{\Omega}(1 - \phi_1 - \phi_2)(1 - S) - \delta S \quad (A11)$$

$$(\delta + \rho)\lambda = F_1\left(\frac{S + (1 - S)\phi_1(1 - \theta)}{(1 - S)(1 - \phi_1 - \phi_2)}, 1\right)(1 - \phi_1(1 - \theta)) - (1 - \phi_1 - \phi_2)F_2\left(\frac{S + (1 - S)\phi_1(1 - \theta)}{(1 - S)(1 - \phi_1 - \phi_2)}, 1\right)$$

$$- \phi_2 \bar{w} - (1 - \phi_1 \phi_2) \tilde{\Omega} \lambda \quad (A12)$$

$$(1-\theta)F_1\left(\frac{S+(1-S)\phi_1(1-\theta)}{(1-S)(1-\phi_1-\phi_2)}, 1\right) - F_2\left(\frac{S+(1-S)\phi_1(1-\theta)}{(1-S)(1-\phi_1-\phi_2)}, 1\right) = \tilde{\Omega} \lambda \quad \text{if } 0 < \phi_1 < 1 \quad (A13)$$

$$(1-\theta)F_1\left(\frac{S+(1-S)\phi_1(1-\theta)}{(1-S)(1-\phi_1-\phi_2)}, 1\right) - F_2\left(\frac{S+(1-S)\phi_1(1-\theta)}{(1-S)(1-\phi_1-\phi_2)}, 1\right) \leq \tilde{\Omega} \lambda \quad \text{if } \phi_1 = 0 \quad (A14)$$

$$\bar{w} - F_2\left(\frac{S+(1-S)\phi_1(1-\theta)}{(1-S)(1-\phi_1-\phi_2)}, 1\right) = \tilde{\Omega} \lambda \quad \text{if } 0 < \phi_2 \leq 1 \quad (A15)$$

$$\bar{w} - F_2\left(\frac{S+(1-S)\phi_1(1-\theta)}{(1-S)(1-\phi_1-\phi_2)}, 1\right) \leq \tilde{\Omega} \lambda \quad \text{if } \phi_2 = 0 \quad (A16)$$

$$0 \leq \phi_1 \leq 1, \quad 0 \leq \phi_2 \leq 1, \quad (A17)$$

$$0 \leq \phi_1 + \phi_2 \leq 1 \quad (A18)$$

where $\tilde{\Omega}$ is given by

$$\tilde{\Omega} = \Omega \frac{S}{S+(1-S)\phi_1(1-\theta)} \quad (A19)$$

In the above, the case with $\phi_1 = 1$ is disregarded since it is never part of an equilibrium. Also, when $\phi_2 = 1$, the definition of a the equilibrium implies that $\bar{w} - F_2\left(\frac{S+(1-S)\phi_1(1-\theta)}{(1-S)(1-\phi_1-\phi_2)}, 1\right) = \tilde{\Omega} \lambda$

Lemma 5: The system of equation A11-A19 admits the solution $\phi_1 = 0$, $\phi_2 = 0$, $S = \frac{\Omega}{\delta+\rho}$ and $\lambda = \frac{F_1(\frac{\Omega}{\delta}, 1) - F_2(\frac{\Omega}{\delta}, 1)}{\delta + \Omega + \rho}$ iff $\theta \geq \underline{\theta}$. This is the unique solution to the system with $\phi_1 = \phi_2 = 0$. Moreover, this solution is a locally (saddle path) stable steady state of the dynamic system defined by Equations (16),(17),(18) and (19).

Proof: Most of Lemma 5 follows directly from Lemma 1, except that it must be verified that Equation (A16) is satisfied, that is,

$$\bar{w} - F_2\left(\frac{\Omega}{\delta}, 1\right) \leq \Omega \frac{F_1\left(\frac{\Omega}{\delta}, 1\right) - F_2\left(\frac{\Omega}{\delta}, 1\right)}{\delta + \Omega + \rho}$$

This condition corresponds to Assumption 4.

Lemma 6A: If $\bar{w} < F_1(z, 1)(1 - \theta)$, where z defined by $F_1(z, 1)(1 - \theta) = F_2(z, 1)$, then the system of equations (A11)-(A19) admits a unique solution with $S = 0$, and this solution takes the form $\phi_1 = \frac{z}{1+z-\theta}$, $\phi_2 = 0$, $\lambda = \frac{\theta F_1(z, 1)}{\delta + \rho}$. Moreover, this solution is a locally (saddle path) stable steady state of the dynamic system defined by Equations (16)-(19) if $\theta < \bar{\theta}$, and it is an unstable steady state if $\theta > \bar{\theta}$.

Proof: From Lemma 2, we know that this solution is the unique solution that satisfies (A11)-(14), (A17) and (A19). The solution also satisfies (A16) since $\bar{w} < F_1(z, 1)(1 - \theta) = F_2(z, 1)$. Hence it is the unique solution with $S = 0$ when $\bar{w} < F_1(z, 1)(1 - \theta)$. The fact that this solution is a stable steady state of the dynamic system when $\theta < \bar{\theta}$, and is an unstable steady state when $\theta > \bar{\theta}$, also follows from Lemma 2.

Lemma 6B: If $\bar{w} > F_1(z, 1)(1 - \theta)$, then the system of equations (A11)-(A19) admits a unique solution with $S = 0$, and this solution takes the form $\phi_1 = 0$, $\phi_2 = 1$, $\lambda = \frac{\theta \bar{w}}{(1-\theta)(\delta+\rho)}$.²³ Moreover, this solution is a (saddle path) stable steady state of the dynamic system defined by Equations (16)-(19) if $\bar{w} < F_1\left(\frac{\Omega}{\delta}, 1\right)(1 - \theta)$, otherwise it is unstable.

Proof: With $S = 0$ and $\bar{w} \neq F_1(z, 1)(1 - \theta)$, then (A13)-(A17) imply that either $\phi_2 = 0$ and $0 < \phi_1$ or the converse $\phi_1 = 0$ and $0 < \phi_2$. If $\phi_2 = 0$, the the only potential solution is that given in Lemma 2. But this solution only satisfies (A16) if $\bar{w} < F_1(z, 1)(1 - \theta)$, hence the solution must have $\phi_1 = 0$ and $\phi_2 = 1$. $\lambda = \frac{\theta \bar{w}}{(1-\theta)(\delta+\rho)}$ is equivalent to stating that $F_1\left(\frac{\phi_1(1-\theta)}{(1-\phi_1-\phi_2)} = \bar{w}\right)$, hence this solution satisfies (A11)-(A19). Note that in this solution $\tilde{\Omega} \neq 0$.

²³The solution is unique in terms on of ϕ_1 and ϕ_2 , but λ can vary between 0 and $\frac{\theta \bar{w}}{(1-\theta)(\delta+\rho)}$ and still satisfy (A11)-(A19) by having (A14) hold with strict inequality. However these additional steady states implies prices that do not satisfy the definition of qn equilibrium given in the text.

The linearized dynamics around this steady state, takes the form:

$$\dot{S} = \left(\frac{\Omega}{\tilde{Z}}\right) - \delta(S_t - S^{ss})$$

$$\dot{\lambda} = (\delta + \rho)(\lambda - \lambda^{ss})$$

where \tilde{Z} is defined by $(1 - \theta)F_1(\tilde{Z}, 1) = \bar{w}$.

When $\theta < \bar{\theta}$ and $\bar{w} < F_1(\frac{\Omega}{\delta}, 1)(1 - \theta)$, then $\tilde{Z} > \frac{\Omega}{\delta}$, and the system has one root smaller than zero and one root greater than zero. Otherwise, $\tilde{Z} < \frac{\Omega}{\delta}$, and both roots are positive.

Lemma 7: The system of Equations (A11)-(A19) admits a solution with $0 < S < \frac{\Omega}{\Omega + \delta}$ iff $\underline{\theta} < \theta < \bar{\theta}$. Moreover this solution is unique, and it is an unstable steady state of the dynamic system given by Equations (16)-(19).

Proof: Let us begin by showing that when $\bar{w} \neq (1 - \theta)F_1(\frac{\Omega}{\delta}, 1)$ then the only potential configuration for a solution with $0 < S < \frac{\Omega}{\Omega + \delta}$ is to have $0 < \phi_1 < 1$, and $\phi_2 = 0$. With $0 < S < \frac{\Omega}{\Omega + \delta}$, Equation (A11) and (A18) imply that $\frac{S + (1-S)\phi_1(1-\theta)}{(1-S)(1-\phi_1-\phi_2)} = \frac{\Omega}{\delta}$ and that $\phi_2 < 1$. If $\phi_1 = 0$. Then from Equation (A12) and (A15), we know that $(\Omega + \delta + \rho)\lambda = F_1(\frac{\Omega}{\delta}, 1) - F_2(\frac{\Omega}{\delta}, 1)$ and that $\bar{w}(\Omega + \delta + \rho) = \Omega F_1(\frac{\Omega}{\delta}, 1) + (\delta + \rho)F_2(\frac{\Omega}{\delta}, 1)$, but this last condition is ruled out by Assumption 4. Finally, from Equations (A13) and (A15), we see that $0 < \phi_1 < 1$ and $0 < \phi_2 < 1$ is only possible when $\bar{w} = (1 - \theta)F_1(\frac{\Omega}{\delta}, 1)$, hence the only potential configuration is $\phi_2 = 0$ and $0 < \phi_1 < 1$. Then note that the only potential solution (A11)-(A13) with $\phi_2 = 0$ and $0 < S < \frac{\Omega}{\delta + \Omega}$ is the one described in Lemma 3. From Lemma 3, we know that this solution exists, is unique and will satisfy (A16), only in the case where $\bar{w} < (1 - \theta)F_1(\frac{\Omega}{\delta}, 1)$.

For the non-generic case where $\bar{w} = (1 - \theta)F_1(\frac{\Omega}{\delta}, 1)$, there is an indeterminacy in the determination of ϕ_1 and ϕ_2 . Any combination of ϕ_1 and ϕ_2 that satisfy $\frac{S + (1-S)\phi_1(1-\theta)}{(1-S)(1-\phi_1-\phi_2)} = \frac{\Omega}{\delta}$ can be part of a solution where S and λ remain defined as in Lemma 3.

The stability properties of this steady state follow from Proposition 4.

Proof of Proposition 5: Lemmas 5, 6 and 7 imply that efficient technology use is the only stable steady state equilibrium when $\underline{\theta} < \theta < \bar{\theta}$, and $\bar{w} > (1 - \theta)F_1(\frac{\Omega}{\delta}, 1)$.

Proof of Proposition 6: From Lemmas 5, 6, and 7, we know that under the conditions of the proposition, there are three steady states. The steady state with $S = 0$ and $S = \frac{\Omega}{\Omega + \delta}$ are saddle path stable by Lemmas 5 and 6, while the steady state with S between these two values is unstable by Lemma 7.

Proof of Proposition 7: Migration to a country with $S = 0$ will be attractive if $F_1\left(\frac{\phi_1(1-\theta)}{1-\phi_1-\phi_2}\right) > F_1\left(\frac{\Omega}{\delta}\right)$ when evaluated at $S = 0$. This is equivalent to the condition $\frac{\phi_1(1-\theta)}{1-\phi_1-\phi_2} < \frac{\Omega}{\delta}$ when evaluated at $S = 0$. There are five cases to be considered.

Case 1) $\theta < \bar{\theta}$ and $\bar{w} < F_1(Z, 1)$.

Lemma 6A implies that $\frac{\phi_1(1-\theta)}{1-\phi_1-\phi_2}$ evaluated at $S = 0$ is equal to Z defined by $(1 - \theta)F_1(Z, 1) = F_2(Z, 1)$. From the definition of $\bar{\theta}$ we know that $F_2\left(\frac{\Omega}{\delta}, 1\right) > (1 - \theta)F_1\left(\frac{\Omega}{\delta}, 1\right)$. Together these two conditions imply that $Z > \frac{\Omega}{\delta}$ and hence that $\frac{\phi_1(1-\theta)}{1-\phi_1-\phi_2} > \frac{\Omega}{\delta}$ at $S = 0$.

Case 2) $\theta < \bar{\theta}$ and $F_1(Z, 1) \leq \bar{w} < F_1\left(\frac{\Omega}{\delta}, 1\right)$.

Lemma 6B implies that $\frac{\phi_1(1-\theta)}{1-\phi_1-\phi_2}$ evaluated at $S = 0$ is equal to \tilde{Z} defined by $(1 - \theta)F_1(\tilde{Z}, 1) = \bar{w}$. This implies that $F_1(\tilde{Z}, 1) < F_1\left(\frac{\Omega}{\delta}, 1\right)$ hence that $\frac{\phi_1(1-\theta)}{1-\phi_1-\phi_2} > \frac{\Omega}{\delta}$ at $S = 0$.

Case 3) $\theta < \bar{\theta}$ and $\bar{w} \geq F_1\left(\frac{\Omega}{\delta}, 1\right)$.

Lemma 6B implies that $\frac{\phi_1(1-\theta)}{1-\phi_1-\phi_2}$ evaluated at $S = 0$ is equal to \tilde{Z} defined by $(1 - \theta)F_1(\tilde{Z}, 1) = \bar{w}$. This implies that $F_1(\tilde{Z}, 1) \leq F_1\left(\frac{\Omega}{\delta}, 1\right)$ hence that $\frac{\phi_1(1-\theta)}{1-\phi_1-\phi_2} \leq \frac{\Omega}{\delta}$ at $S = 0$.

Case 4) $\theta \geq \bar{\theta}$ and $F_1(Z, 1) < \bar{w}$.

References:

- Adams, A.V. and R.K. Johanson (2004) Skills Development in Sub-Saharan Africa, World Bank, Washington DC.
- Atchoarena, D. and Delluc, A. (2001) Revisiting Technical and Vocational Education in Sub-Saharan Africa: An Update on Trends, Innovations, and Challenges. International Institute for Educational Planning: Paris.
- Banerjee, A. and E. Duflo (2005) Growth theory through the lens of development economics, Handbook of Development Economics, Vol 1A Amsterdam, Elsevier, pp473-552.
- Banji, O-O (2004) Learning, Knowledge and Skills: Implications for Firm Level Performance in African Industry, International Journal of Technology Management and Sustainable Development, 3, 2, 91-113.
- Bartelsman, E., J. Haltiwanger, and S. Scarpetta (2006) Reallocation and Productivity Growth: The FAQs, mimeo, University of Amsterdam.
- Batra, G. (2002) Training, Technology and Firm Level Competitiveness – Evidence from the World Business Environment Survey for Latin American and the Caribbean, World Bank Background Paper, Washington, D.C.
- Beaudry, P. and P. Francois (2006) Managerial skill acquisition and the theory of economic development, NBER working paper W11451.
- Biggs, T. M. Shah and P. Srivastava (1995) Technological Capabilities and Learning in African Enterprises, World Bank Technical Paper #288.
- Chari V.V. and H. Hopenhayn (1991) Vintage Human Capital, Growth, and the Diffusion of New Technology, Journal of Political Economy, 99, 6, 1142-1165.
- Ciceon (Centro de Investigación sobre Economía Laboral y Gestión del Conocimiento) (2001) The Incidence of Training in Mexico, World Bank Background Paper, Washington, D.C.
- Davenport T., (1994) Saving IT's Soul: Human-Centered Information Management, Harvard Business Review, March 1994.
- De Ferranti, D., G.E. Perry, I. Gill, J.L. Guasch, W.F. Maloney, Sanchez-Paramo, C. and Schady N. (2003) Closing the Gap in Education and Technology, World Bank, Washington

DC.

Fluitman, F. (1992) Traditional Apprenticeship in West Africa: Recent Evidence and Policy Options. Discussion Paper No. 34. Geneva: ILO.

Haan, H. C. and Serriere, N. (2002) Training for Work in the Informal Sector: Fresh Evidence from West Africa, ILO Training Center, Turin.

Hall, B. H., and B. Khan. (2003) Adoption of New Technology. NBER working paper 9730.

Hobday, M. (1995) Innovation in East Asia: the challenge to Japan. Aldershot, London.

Hsieh, C-T and P. Klenow (2007) Relative prices and relative prosperity, American Economic Review, June.

Katz, J. M. (1987) Technology creation in Latin American manufacturing industries. St. Martin's Press, New York.

Lucas (1990) Why Doesn't Capital Flow from Rich to Poor Countries? American Economic Review, Papers and Proceedings, Vol 80 (2) pp. 92-96.

McKinsey Global Institute (2001) India: The Growth Imperative. Report, McKinsey Global Institute.

Restuccia, D. and R. Rogerson (2007) Policy distortions and aggregate productivity with heterogeneous plants, NBER working paper #W13018.

Stolovich, H.D and D. Ngoa-Nguele (2001) Structured On-the-job Training in Developing Nations, Advances in Developing Human Resources, 3, 461-470.

Thoenig, M. and T. Verdier (2004) The Macroeconomics of Knowledge Management: Internal Hold-up Versus Technological Competition, CEPR Discussion paper no. 4710.

Figure 1

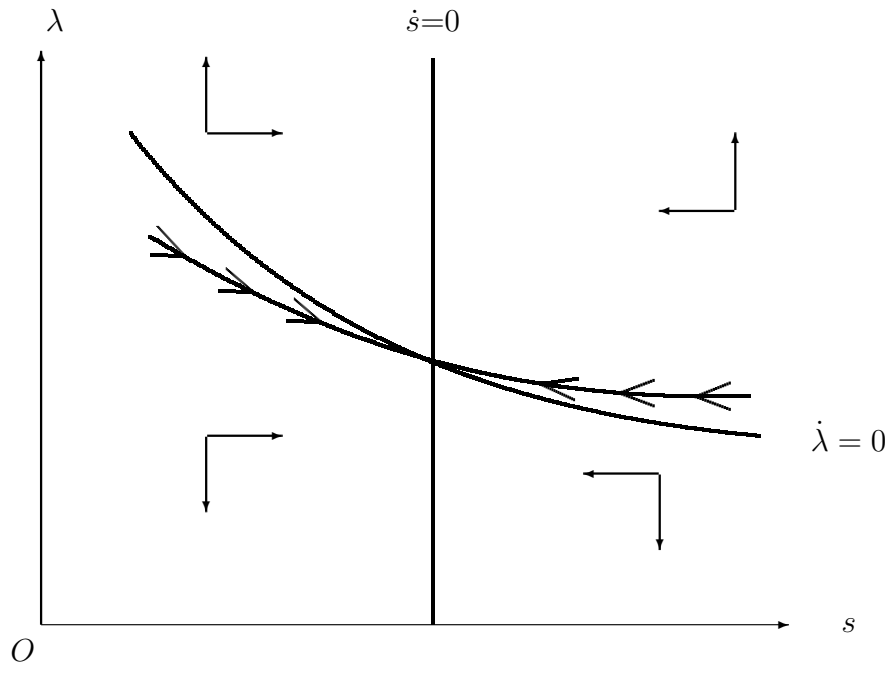


Figure 2

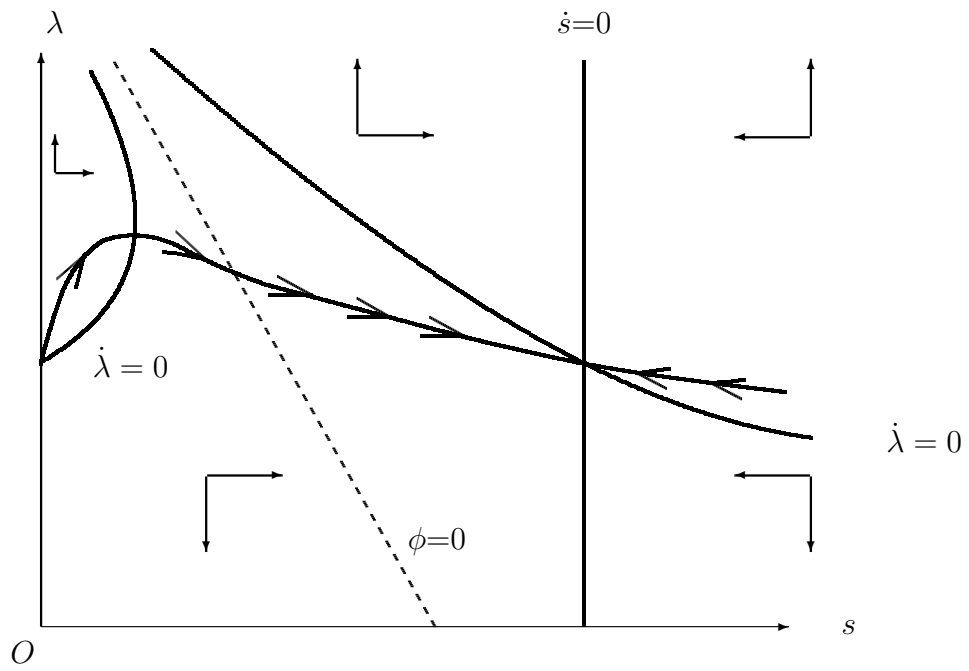


Figure 3

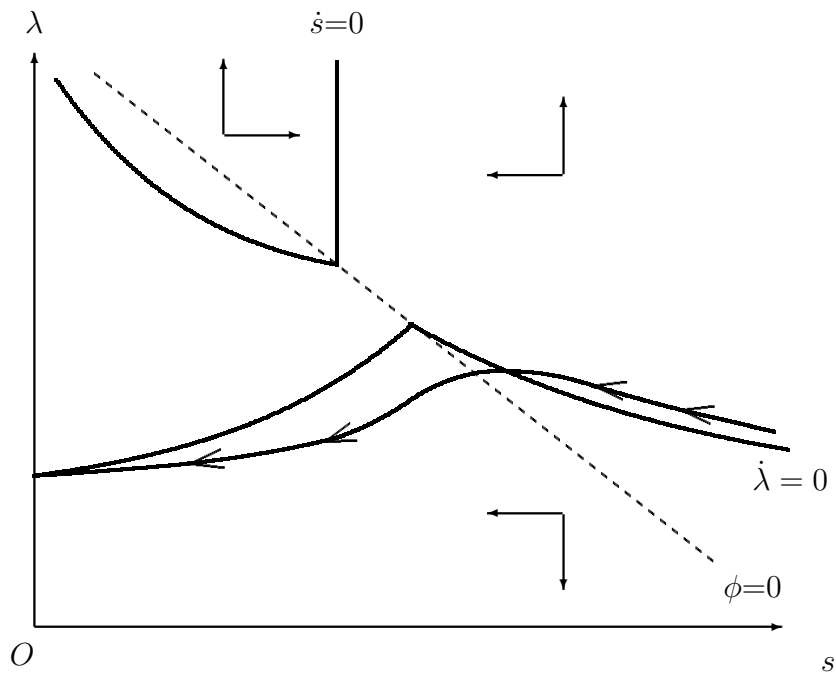


Figure 4

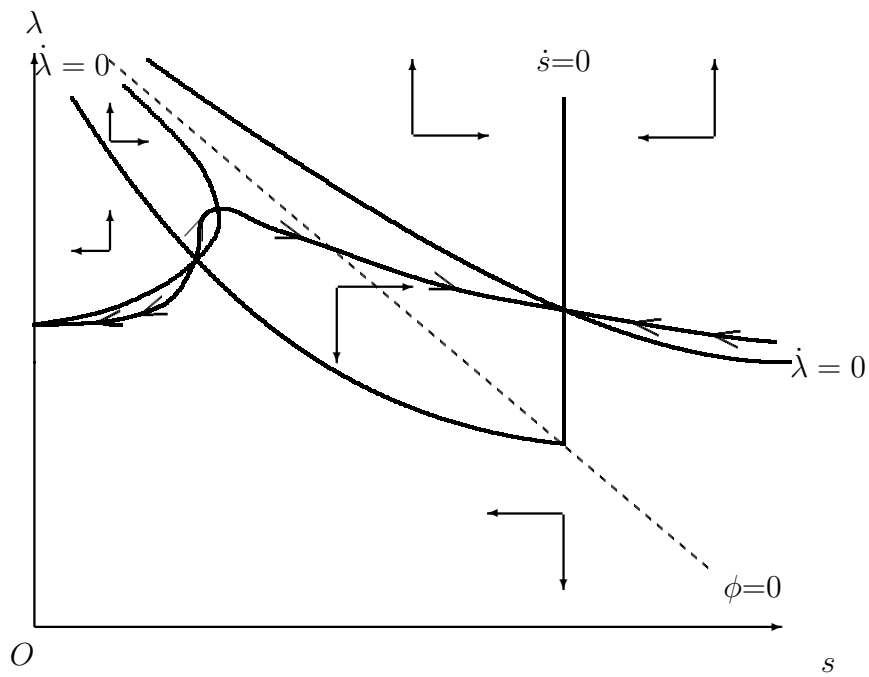


Figure 5

