NBER WORKING PAPER SERIES

ORGANIZING GROWTH

Luis Garicano
Esteban Rossi-Hansberg

Working Paper 13705
http://www.nber.org/papers/w13705

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2007

A first draft of this paper was prepared for the Conference in honor of Robert E. Lucas Jr. at Clemson University, September 2007. We thank Lorenzo Caliendo for excellent research assistance and Philip Aghion, Tim Besley, Per Krussel, Bob Lucas, Torsten Persson, Andrea Prat, Nancy Stokey and seminar participants at Chicago GSB, Clemson, ECARES, Erasmus School of Economics, LSE, Northwestern and Princeton for useful comments. Garicano acknowledges the financial support of the Toulouse Network on Information Technology. Rossi-Hansberg acknowledges the generous support of the Sloan Foundation. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

Organizing Growth
Luis Garicano and Esteban Rossi-Hansberg
NBER Working Paper No. 13705
December 2007
JEL No. D23,E32,L23,O31,O40

ABSTRACT

We study the impact of information and communication technology on growth through its impact on organization and innovation. Agents accumulate knowledge through two activities: innovation (discovering new technologies) and exploitation (learning how to use the current technology). Exploitation requires the development of organizations to coordinate the work of experts, which takes time. The costs and benefits of such organizations depend on the cost of communicating and acquiring information. We find that while advances in information technology that lower information acquisition costs always increase growth, improvements in communication technology may lead to lower growth and even to stagnation, as the payoff to exploiting innovations through organizations increases relative to the payoff of new radical innovations.

Luis Garicano
School of Business
University of Chicago
1101 East 58th Street
Chicago, IL 60637
luis.garicano@chicagogsb.edu

Esteban Rossi-Hansberg
Princeton University
Department of Economics
Fisher Hall
Princeton, NJ 08544-1021
and NBER
erossi@princeton.edu

1. INTRODUCTION

The development of a new technology brings with it a new set of production challenges that must be dealt with for production to take place. Tackling these new problems requires individuals to acquire specialized expertise and to work with each other. Specialization allows individuals to acquire knowledge only about a narrow set of problems, which they then utilize intensively. The cost of such specialization is that it requires communication and coordination among different individuals. That is, it requires 'organization'. Information and communication technology affect the costs and benefits of organization and, through them, the extent to which a new technology can be exploited. Thus information technology is a 'meta-technology': a technology that affects the costs and benefits of investing in technology. In this paper we study how the cost of acquiring and communicating information affects growth through its effect on organization and innovation.

We propose a theory of economic growth through organization. When a new technology is introduced agents learn only the most common problems associated with it. As times passes, organizations, in the form of knowledge-based hierarchies (Garicano, 2000) are created. In them, some agents ('problem solvers' or experts) specialize in dealing with exceptional problems and other agents specialize in production and learn the routine problems. We model this process as the emergence of a collection of markets for expert services (referral markets) where agents sell the problems they cannot solve to other agents. These referral markets could be equivalently seen as consulting market arrangements or inside-the-firm hierarchies, as we have shown elsewhere (Garicano and Rossi-Hansberg, 2006). The dynamics of our model result from the time to build these markets. As we discuss later, in our view it is critical to growth dynamics that multiple complementary specialists do not emerge instantaneously, but that they take time to emerge. Specifically, we assume agents have to see the problems that remain unsolved at some moment in time before they decide to specialize in those problems. Thus, only one expert market can be created per period.

Our model differentiates two knowledge generating activities: exploiting existing technology and innovating to develop new technologies. First, exploitation takes place as organizations undertake production over time and add new layers (new markets) of experts. By allowing these new experts to leverage their knowledge about unusual problems, the new layers allow for more knowledge to be acquired and make production more efficient under the current technology. This process exhibits decreasing returns, as eventually most problems are well known and the knowledge acquired is less and less valuable.

Second, innovation is the result of agent's decisions of how much to invest to create radically new technologies. This investment process exhibits adjustment costs, so that the investment, if it happens, takes place smoothly over time. Of course, the ability of the economy to exploit the new technology through organization determines the profitability of innovation investments. The rate of innovation, the extent of exploitation, and the amount of organization in the economy are jointly determined in our theory, and depend on the cost of acquiring and communicating knowledge.

If it happens, progress in our model takes place in leaps and bounds. After a new technology is adopted, investment in innovation decreases and agents concentrate in exploitation as first the more productive pieces of knowledge about this technology and then the rarer ones are acquired. Radical innovation will not take place again until the current innovation has been exploited to a certain degree. Both the timing of the switch to a new technology, and the size of the jump in the technology are endogenous, as agents must choose how much and when to invest in radical innovation. As long as the value of continuing on an existing innovation is sufficiently high, the switch to the new technological generation does not take place. Adopting the new technology makes the knowledge acquired about the previous technology obsolete, and thus requires agents to start accumulating new knowledge and start building new organizations.[1] Thus inherent in new knowledge is a process of creative destruction (Schumpeter, 1942) whereby adopting a radical innovation makes the existing organization obsolete.[2] That is, we built on the insight of Arrow (1974) that organizations are specific to a particular technology.

[1]Indeed, empirically, new technologies are usually associated with new organizations. For example, associated with the arrival of the electricity at the end of the XIX century were notably Edison General Electric (now GE) and Westinghouse; associated with the development of the automobile a few years later were Ford and General Motors; with the development of film, Kodak; with the arrival of the computer and microprocessors first IBM and then Intel; with the development of the World Wide Web, Google, Yahoo, Amazon and E-Bay. We discuss these stylized facts in the next section.

[2]Previous models of creative destruction, following on the pioneering work of Aghion and Howitt (1992) and Grossman and Helpman (1991) do not take organizations into account– new products substitute the old, but organizations play no role.

Progress may also come to a halt if agents decide not to invest in radical innovation. Specifically, the payoff of exploiting existing technologies may be such that agents optimally create very large organizations, composed of a large set of referral markets and a large number of different specialized occupations. Such organizations take a long time to build, and thus agents choose to postpone forever the moment in which a new technology would be exploited. That is, the radical innovation process never gets started. When the current technology is fully exploited, agents do not have any development in the alternative radical technology to build on, and prefer to place their efforts on small advances of the existing technology. The result is stagnation.

Information and communication technology determines the depth to which an innovation is exploited, and thus the rate of growth. Consider first information technology. The main benefit of organization is that individuals can leverage their cost of acquiring knowledge over a larger set of problems, increasing the utilization rate of knowledge. Information technology advances that reduce the cost of acquiring and accessing knowledge (e.g. databases), decrease the need for organizational complexity, shorten the exploitation process, and unambiguously increase growth. Alternatively, increases in the cost of acquiring and accessing knowledge may dramatically slow long run growth, maybe to the extend of stopping it, as they require the creation of large, complex organizations with a long time to build.

Better communication technology reduces the time spent by agents communicating with others. This unambiguously increases welfare, but has an ambiguous effect on growth. Lower communication costs increase the value of both current technologies (which are more deeply exploited as organizations are more efficient) and of future technologies. If agents value the future sufficiently (because either they do not discount it much, or because their ability to sell their innovations is limited: 'appropiability') this increases the value of innovation. However, if individuals do not value the future much, particularly when communication is costly, better communication technology makes investment in exploiting current technology more appealing which means that organizations spend more effort on the process of deepening their production knowledge. But this postpones the moment when future technology will reap its rewards and may lead to lower investment in radical innovations and less future growth. In fact, it may lead to stagnation. In other words, making organizations more efficient may shift the balance of economic activity from investing in new innovations to exploiting better existing innovations, and that may reduce economic growth, potentially to zero. Thus improvements in communication technology could leave a less developed country, where communication costs and discount rates are high, with a lower rate of innovation, whereas they would tend to have a positive effect on the growth of more developed countries. Eventually, sufficiently large or continuous drops in communication costs are favorable to growth, as spans of control become very large and organizations very simple.

Throughout, we assume that knowledge is appropriable. In the case of problem solving and production knowledge because communicating it takes time, and individuals available time is limited. In the case of innovation knowledge, we assume individuals invest because they will appropriate the results of their investment by selling or using the future technologies. Alternatively, we could assume that there are externalities and that, for example, radical innovations are a by-product of the accumulation of production knowledge. This would also result in an endogenous growth theory, but one in which better communication technology trivially leads to faster growth as it incentivates agents to acquire more production knowledge.

Our theory has some other implications for growth and organization. Scale effects are absent. Empirically, as Jones (1995) first observed, adding more agents to the research side of the economy (in our case the allocation of more agents as problem solvers) does not necessarily increase the rate of growth. In our model, there are no scale effects on average. In fact, we actually obtain larger shares of agents in the 'knowledge' sector in the final periods of an existing technology, where the value of the extra innovations is the lowest.[3] Moreover, the rate of investment in innovation is lowest when the growth rate is highest (the existing technology is being exploited increasingly well) while it is highest then the growth rate is lowest. Second, productivity moves in cycles, involving large gains at the initial stages of the organization of a radical innovation, and decelerating for a long period of time as the decreasing returns involved in increasingly deep exploitation of existing technologies take place.

[3] The main previous existing explanations of this puzzle are Kortum (1997), Young (1998) and Howitt (1999). The first one's explanation is that the first ideas are low hanging fruit, and as these ideas are exploited, future innovations become increasingly costly to find. The other two papers focus on the increase in the amount of varieties as innovation increases, which leads to an increase in the innovation cost, as workers must improve a larger number of products.

Our work has several precedents on top of the seminal endogenous growth theories of Lucas (1988), Romer (1990), Grossman and Helpman (1991), and Aghion and Howitt (1992). Becker and Murphy (1992) first studied the connection between coordination costs and growth through economic organization. Unlike in their model, we specifically take into account the knowledge accumulation process and the occupational distribution and organization that results. We also depart in differentiating between exploitation and innovation. Second, Jovanovic and Rob (1990) develop a theory in which growth is generated by small innovations within a technology and large innovations across technologies. In their framework, alternative technologies are random and are not affected by the choices made within the current technology. In this sense, our theory endogenizes the quality of alternative technologies and adds organization as a source of growth. Jovanovic and Rob (1989) present a model in which communication technology also affects growth through the search process, not, again, through organizations. Our theory also builds on the work of Penrose (1959) who first (informally) identified the role of firms in inducing growth through their impact on knowledge accumulation, and emphasized how the constraints on the growth of managerial hierarchies constrain firm growth.

Our work is also related to the work on vintage human capital (Chari and Hopenhym, 1991) where older individuals operate in old vintages of technologies; if skilled and unskilled labor are complementary, it is hard to switch to a new technology. Also, Comin and Hobijn (2007) present a model where a technology must not just be discovered, but it must also be implemented in order to be beneficial and learning must take place after implementation. Organization does not play any role in this model, nor does information technology affect how far can organization exploit the new technology. Firms only choose where to start the implementation process; from then on it is exogenous. Organization does play a role in Legros, Newman and Proto (2007) where, like in our work, the division of labor is endogenously determined with the growth path. The approach in that paper is different, and complementary, to ours. While the key organizational issue in their paper is to monitor workers, the purpose of organization in our work is to increase the utilization of knowledge. The idea that information and communication technology is a 'meta-technology' (a technology to implement technology) so that better information technology affects, through organization, the efficiency of exploitation is, we believe, novel.

Our theory can explain why some economies stagnate. It shares this feature with Krusell and Rios-Rull (1996) and Jovanovic and Nyarko (1996). In both of those papers increased incentives to exploit the current technology may lead to stagnation. In the former paper due to the political economy problem of insider versus outsider rents and in the latter because of the cost implied by the loss of expertise when switching technologies. We argue that stagnation can also be the result of a lengthy proses of building organizations to exploit the current technology.

One way to understand our contribution is as a theory of endogenous growth in labour-augmenting technology. Take a standard production function of the form $AF(K, HL)$ where A denotes TFP, K capital, H labor-augmenting technology (human and organizational capital) and L labor. Then, we present a theory of the evolution of H. It is straightforward to embed it in a standard neoclassical model so we abstain from doing so. In Section 2 2. we present a theory of the evolution of H, where H increases as deeper organizations are formed and the technology is exploited more efficiently. The evolution of H exhibits decreasing returns and so permanent growth can never be the result of more complex and efficient organizations. In Section 2.3. we allow agents to invest in order to improve technology too. Even though H exhibits decreasing growth rates, agents will sporadically make radical innovations that will make the growth of H increase discontinuously and then follow a similar decreasing path. Section 3 analyzes the model and Section 4 presents some empirical evidence on technological cycles. Section 5 concludes.

2. THE MODEL

2.1. Preferences and Technology

The economy is populated by a mass of size 2 of ex-ante identical agents that live for two periods. Every period an identical set of agents is born. Agents work when they are young only and they have linear preferences so they maximize the discounted sum of income or consumption of the unique good produced in the economy. That is, agents preferences are given by $U(c_t, c_{t+1}) = c_t + \beta c_{t+1}$.

At the start of the period agents choose an occupation and a level of knowledge to perform their job. Agents can either work in organizations that use the current prevalent technology, or they can decide to implement a new technology. The quality of the new technology will depend on past investments in innovation knowledge which will be the only technology available to save for the retirement period.

A technology is a method to produce goods using labor and knowledge. One unit of labor generates a project or problem. To produce, agents need to have the knowledge to solve the problem. If they do, they solve the problem and output is produced. If the worker does not know the solution to the problem, she has the possibility to transfer or sell the problem or project to another agent that may have the knowledge to solve it. Organizations are hierarchical, they have one layer of workers and potentially many layers of problem solvers (as in Garicano and Rossi-Hansberg, 2006). Problem solvers have more advanced knowledge than workers and so are able to solve more advanced problems, but they need to 'buy' these problems from workers or lower layer problem solvers since they do not spend time producing but communicating existing problems.

The key assumption in our model is that specialization requires organization and organization cannot be built instantly. Organization takes time to build (see Kydland and Prescott (1983) for a similar argument for the case of physical capital). For people to specialize in different types of problems, they must have mechanisms that allow them to know who knows what and to ask the right question from the right person—specialization requires organization. Two main factors prevent the simultaneous development of specialists in many different areas who can work together instantly. First, it is impossible to know what are the problems that will prove important in the next cycle of innovation, and the types of expertise that will be required. For example, experts in internet marketing or sophisticated wireless networks became available only after the internet was developed and there was a demand for their services. Second, agents have to be trained in the basic knowledge of the current technology before others can be trained in the more advanced knowledge; in fact, learning how to deal with the rare and advanced problems may not be useful if there are no agents specialized in simple tasks who can actually ask the right questions. The appearance of sophisticated radiologist who specialize in some specific kinds of tumors requires the previous appearance of normal radiologists, and of cancer specialists who can use the information obtained in these X-rays.

A technology is used more intensely the more layers in the organization. In the first period a technology is in use, agents learn basic knowledge to develop it and they work as production workers. Since higher layers of management have not been developed, the problems they cannot solve go to waste. In the next period, agents observe that in the last period some valuable problems were thrown away and some of them decide to work as first layer problem solvers. These problem solvers, in turn, learn to solve some problems and throw away those that they cannot solve. This induces the entry of second layer experts in the next period. This process goes on making the hierarchy taller as time proceeds and the use of the prevalent technology more efficient through a better allocation of workers. Of course, the knowledge acquired by agents in all layers will depend on the number of layers in the organization as well as the fees or prices for transferring problems. The price at which an agent with a particular level of knowledge can sell a problem is a measure of the efficiency of the organizational structure in exploiting a technology. As we will see, the more organizational layers, the higher the price and so the more efficient is the organization in allocating labor and knowledge.

As emphasized in Garicano and Rossi-Hansberg (2006) there are many equivalent ways of decentralizing these organizations. First, as here, there can be a market for problems and agents sell and buy problems for each other at a market price. Alternatively, there can also be firms that optimally organize these hierarchies and hire workers and managers for particular positions at a wage given their knowledge level. Finally, organizations can also be decentralized as consulting markets in which workers hire knowledgeable agents as consultants to solve problems for them for a fee. All of these interpretations are equivalent and can exists at the same time. In all of them agents obtain the same earnings and perform the same roles. In what follows we model these hierarchies as markets for experts services – there are markets for problems and problem solvers buy them, but we may as well talk about firms and managers.

We now turn to the description of the formation of organization and the use of a technology. We then study the decisions of agents to drop the technology currently in use and make a radical innovation instead of going deeper in the development of the current technology (add a new layer).

2.2. Organizing a Technology

Suppose a new technology $A \geq 1$ is put in place at time $t = 0$. The evolution of this technology will be our main concern in the next section. For now we just keep it fixed and focus on how it is exploited. Obtaining A units of output from this technology requires a unit of time and a random level of knowledge. An agent specialized in production uses his unit of time to generate one problem, which is a draw from the probability distribution $f(z)$. We assume that $f(z)$ is continuous and decreasing, $f'(z) < 0$, with cumulative distribution function $F(z)$. The assumption that $f'(z) < 0$ guarantees that agents will always start by learning how to solve the most basic and common problems.[4] In order to produce, the problem drawn must be within the workers' knowledge set, if it is not, then no output is generated. Knowledge can be acquired at a constant cost $\widetilde{c} > 0$, so that acquiring knowledge about problems in $[0, z]$ costs $\widetilde{c}z$. Denote the wage of an agent working in layer $\ell \in \{0, 1, ...\}$ of an organization with highest layer L (or in period L since the highest layer is, throughout this section, the time period) by w_L^ℓ. Then, the earnings of a production worker (layer 0) working on a new idea (so the highest layer in the organization is $L = 0$) at time 0 are:

$$w_0^0 = \max_z AF(z) - \widetilde{c}z,$$

where $AF(z)$ is total output by workers with ability z (they solve a fraction $F(z)$ of problems each of which produces A units of output) and $\widetilde{c}z$ is the cost of acquiring knowledge z. Denote by z_0^0 the level of knowledge that solves the problem above (where the notation is analogous to the one for wages). Note that an organization with only workers of layer zero will leave unsolved a fraction of problems $1 - F(z_0^0)$. These problems, if solved, would produce output $A(1 - F(z_0^0))$. But this simple organization, where workers only work by themselves, chooses optimally to discard them.

In order to take advantage of the discarded problems next period, $t = 1$, some agents will decide to buy the discarded problems from workers as long as they can

[4] That is, $f'(z) < 0$ will be chosen by agents if they can sequence the knowledge acquired optimally.

then solve some of them and obtain higher earnings. The assumption is that these agents need to first see that valuable problems are discarded to enter next period and take advantage of them. Agents can communicate the problems they did not solve in exchange for a fee or price. If communication is cheaper than drawing new problems, then some agents may find it in their interest to specialize in learning about unsolved problems; they pay a price for these problems, but in exchange they can solve many of them as they do not need to spend time generating the problems, only communicating with the seller. Organization makes it, potentially, optimal to learn unusual problems, as agents can amortize this knowledge over a larger set of problems.

Thus at time $t = 1$ agents have a choice between becoming production workers or specialized problem solvers. If they become production workers they earn

$$w_1^0 = \max_z AF(z) + (1 - F(z))r_1^0 - \widetilde{c}z \tag{1}$$

where $r_{1,1}^0$ is the equilibrium price at which workers in layer 0 sell their problems. As problem solvers they need to spend their time communicating with workers to find out about the problems they are buying. The number of problems a manager can buy is given by the communication technology. Let h be the time a problem solver needs to communicate with a worker about a problem. Then, a problem solver has time to find out, and therefore buy, $1/h$ problems. Clearly h is a key parameter of the model that determines the quality of communication technology. The manager knows that workers only sell problems that they cannot solve, so he knows that all problems sold by workers will require knowledge $z > z_{1,1}^0$ (where $z_{1,1}^0$ solves the problem above). Hence, the manager acquires knowledge about the more frequent problems above $z_{1,1}^0$. The wage of the layer one problem solver is then given by

$$w_1^1 = \max_z \frac{1}{h}\left(A\frac{F(z_1^0 + z) - F(z_1^0)}{1 - F(z_1^0)} - r_1^0\right) - \widetilde{c}z.$$

Namely, they buy $1/h$ problems at price r_0 and solve a fraction $\left(F(z_1^0 + z_1^1) - F(z_1^0)\right) / \left(1 - F(z_1^0)\right)$ of them, each of which produces A units of output. On top of this, they pay the cost of learning the problems in $[z_1^0, z_1^0 + z_1^1]$. As long as $r_1^0 > 0$, the value of the problems that were being thrown out was positive, and so $w_0^0 < w_1^0 = w_1^1$, where the last equality follows from all agents being identical ex-ante. Hence, if $r_1^0 > 0$ adding the first layer of problem solvers is optimal at time $t = 1$. We will show below that in equilibrium under some assumptions on F, r_1^0 is in fact positive. Note also that agents in layer 0 will choose to acquire less knowledge as we add a layer of problem solvers: It is not worth it to learn as much since unsolved problems can now be sold at a positive price.

Next period, $t = 2$, agents observe that some valuable problems were thrown away last period. Namely, a fraction $1 - F(z_1^1)$ of problems. Hence, some agents enter as managers of layer 2 to buy these problems from problem solvers of layer 1. This process continues, adding more layers each period, as long as some valuable problems are thrown away and agents can acquire enough knowledge to solve them and earn higher wages. Hence, each period this economy potentially adds another layer of problem solvers. More unusual and specialized problems are solved and society acquires a larger and larger range of knowledge.

To avoid repetition, we write the problem for period $t = L$ when the hierarchy has a maximum layer L. As described above, production workers earn

$$w_L^0 = \max_z AF(z) + (1 - F(z))r_L^0 - \widetilde{c}z.$$

Call Z_L^ℓ the cumulative knowledge of agents up to layer ℓ, in period L where the maximum number of layers is L : $Z_L^\ell = \sum_{i<\ell} z_L^i$. A problem solver of layer ℓ where $0 < \ell < L$ earns

$$w_L^\ell = \max_z \frac{1}{h}\left(\frac{A\left(F(Z_L^{\ell-1}+z) - F(Z_L^{\ell-1})\right) + \left(1 - F(Z_L^{\ell-1}+z)\right) r_L^\ell}{\left(1 - F(Z_L^{\ell-1})\right)} - r_L^{\ell-1}\right) - \widetilde{c}z,$$

where r_L^ℓ is the price of a problem sold by an agent in layer ℓ in an economy with organizations of $L+1$ layers at time t. Note that intermediate problem solvers both sell and buy problems. They buy $1/h$ problems at price $r_L^{\ell-1}$ and sell the problems they could not solve (a fraction $\left(1 - F(Z_L^{\ell-1}+z)\right) / \left(1 - F(Z_L^{\ell-1})\right)$ at price r_L^ℓ. Problem solvers in the highest layer L cannot sell their problems as there are no buyers, so their earnings are just given by

$$w_L^L = \max_z \frac{1}{h}\left(A\frac{F(Z_L^{L-1}+z) - F(Z_L^{L-1})}{1 - F(Z_L^{L-1})} - r_L^{L-1}\right) - \widetilde{c}z.$$

In what follows we will use an exponential distribution of problems. This will allow us to simplify the problem above substantially and will guarantee that the prices of problems at all layers are positive. Hence, absent a new technology, as time goes to infinity the number of layers also goes to infinity. In the next section we will introduce radical innovations that will prevent this from happening. For the moment, however, we continue with our technology A.

Let $F(z) = 1 - e^{-\lambda z}$. Then the earnings of agents in the different layers can be simplified to

$$\begin{aligned}
w_L^0 &= \max_z \left(A - e^{-\lambda z}\left(A - r_L^0\right)\right) - \widetilde{c}z, \\
w_L^\ell &= \max_z \frac{1}{h}\left(\left(A - r_L^{\ell-1}\right) - e^{-\lambda z}\left(A - r_L^\ell\right)\right) - \widetilde{c}z \text{ for } 0 < \ell < L, \\
w_L^L &= \max_z \frac{1}{h}\left(\left(A - r_L^{L-1}\right) - e^{-\lambda z}A\right) - \widetilde{c}z.
\end{aligned}$$

Thus, in a period where there are organizations with layer L as their highest layer (or organizations with $L+1$ layers), given prices, agents choose knowledge so as to maximize their earnings as stated above. The first order conditions from this problems imply that

$$\begin{aligned}
e^{-\lambda z_L^0} &= \frac{\widetilde{c}}{\lambda\left(A - r_L^0\right)}, \\
e^{-\lambda z_L^\ell} &= \frac{\widetilde{c}h}{\lambda\left(A - r_L^\ell\right)} \text{ for } 0 < \ell < L, \\
e^{-\lambda z_L^L} &= \frac{\widetilde{c}h}{\lambda A},
\end{aligned} \tag{2}$$

or

$$\begin{aligned}
z_L^0 &= -\frac{1}{\lambda}\ln\frac{\widetilde{c}}{\lambda\left(A - r_L^0\right)}, \\
z_L^\ell &= -\frac{1}{\lambda}\ln\frac{\widetilde{c}h}{\lambda\left(A - r_L^\ell\right)} \text{ for } 0 < \ell < L, \\
z_L^L &= -\frac{1}{\lambda}\ln\frac{\widetilde{c}h}{\lambda A}.
\end{aligned} \tag{3}$$

and so earnings in the economy are given by

$$w_L^0 = A - \frac{\widetilde{c}}{\lambda} - \widetilde{c} z_L^0 = A - \frac{\widetilde{c}}{\lambda}\left(1 - \ln \frac{\widetilde{c}}{\lambda\left(A - r_L^0\right)}\right), \tag{4}$$

$$w_L^\ell = \frac{A - r_L^{\ell-1}}{h} - \frac{\widetilde{c}}{\lambda} - \widetilde{c} z_L^\ell = \frac{A - r_L^{\ell-1}}{h} - \frac{\widetilde{c}}{\lambda}\left(1 - \ln \frac{\widetilde{c}h}{\lambda\left(A - r_L^\ell\right)}\right) \text{ for } 0 < \ell < L,$$

$$w_L^L = \frac{A - r_L^{L-1}}{h} - \frac{\widetilde{c}}{\lambda} - \widetilde{c} z_L^L = \frac{A - r_L^{L-1}}{h} - \frac{\widetilde{c}}{\lambda}\left(1 - \ln \frac{\widetilde{c}h}{\lambda A}\right).$$

Note that the knowledge acquired is increasing in A and decreasing in $\widetilde{c}$, h (for problem solvers) and the price obtained for selling problems. The intuition for the effect of A and $\widetilde{c}$ is immediate. For h, remember that a higher h implies a worse communication technology. So a higher h implies that problem solvers can buy fewer problems and so they can span their knowledge over less problems. Knowledge becomes less useful. As the price at which agents sell problems increases, agents have an incentive to sell their problems instead of learning more to squeeze all their value, which creates incentives to learn less.

At any point in time t an economy with technology A and organizations with $L+1$ layers is in equilibrium if the knowledge levels of agents solve Equations (3) and

$$w_L^\ell = w_L^{\ell+1} \equiv \tilde{w}\left(A, L\right) \text{ for all } \ell = 0, ..., L-1. \tag{5}$$

This condition is equivalent to an equilibrium condition requiring that the supply and demand of problems at every layer equalize at the equilibrium prizes $\left\{r_L^\ell\right\}_{\ell=0}^{L-1}$. The reason is that when wages are equalized, agents are indifferent as to their role in the organization, and thus they are willing to supply and demand positive amounts of the problems in all layers. Equilibrium in the markets for problems given L then implies that there are a number

$$n_L^\ell = h\left(1 - F(Z_L^{\ell-1})\right) n_L^0 = h e^{-\lambda Z_L^{\ell-1}} n_L^0$$

of agents working in layer ℓ. Since the economy is populated by a unit mass of agents, the number of workers is given by

$$n_L^0 = \frac{1}{1 + h\sum_{\ell=1}^{L}\left(1 - F(Z_L^{\ell-1})\right)}.$$

So given t, A, and L an equilibrium for one generation of agents is a collection of L prices $\left\{r_L^\ell\right\}_{\ell=0}^{L-1}$ and $L+1$ knowledge levels $\left\{z_L^\ell\right\}_{\ell=0}^{L}$ that solve the $2L+1$ equations in (3) and (5). Before we move on to characterize the solution to this system of equations consider the solutions of the system as $L \to \infty$. In this case, since there is no final layer, the system has a very simple solution. Guess that $r_\infty^\ell = r_\infty$ for all ℓ. Then, the first order conditions in (3) imply that

$$
\begin{aligned}
z_\infty^0 &= -\frac{1}{\lambda} \ln \frac{\widetilde{c}}{\lambda\left(A - r_\infty\right)}, \\
z_\infty^\ell &= -\frac{1}{\lambda} \ln \frac{\widetilde{c}h}{\lambda\left(A - r_\infty\right)} \text{ for all } \ell > 0.
\end{aligned}
$$

Note that, since $h < 1$, $z_\infty^0 < z_\infty^\ell$ for $\ell > 1$. That is, in the limit as the number of layers goes to infinity workers learn less than all other agents in the economy. Wages are then given by,

$$
\begin{aligned}
w_\infty^0 &= A - \frac{\widetilde{c}}{\lambda}\left(1 - \ln \frac{\widetilde{c}}{\lambda\left(A - r_\infty\right)}\right), \\
w_\infty^\ell &= \frac{A-r}{h} - \frac{\widetilde{c}}{\lambda}\left(1 - \ln \frac{\widetilde{c}h}{\lambda\left(A - r_\infty\right)}\right) \text{ for all } \ell > 0.
\end{aligned}
$$

Since r_∞ is not a function of ℓ, earnings of problem solvers are identical as is the amount of knowledge they learn. This verifies our guess if we can find an r such that $w_\infty \equiv w_\infty^0 = w_\infty^\ell$. It is easy to see that

$$
r_\infty = A\left(1-h\right) + \frac{\widetilde{c}h}{\lambda} \ln h
$$

solves this equation. Hence, earnings as $L \to \infty$ are given by

$$
w_\infty = A - \frac{\widetilde{c}}{\lambda}\left(1 + \ln\left(\frac{A\lambda h}{\widetilde{c}} - h \ln h\right)\right),
$$

and the knowledge acquired by agents is given by

$$
\begin{aligned}
z_\infty^0 &= \frac{1}{\lambda} \ln h \left(\frac{A\lambda}{\widetilde{c}} - \ln h\right), \\
z_\infty^\ell &= \frac{1}{\lambda} \ln \left(\frac{A\lambda}{\widetilde{c}} - \ln h\right) \text{ for all } \ell > 0.
\end{aligned}
$$

The case of $L \to \infty$ is helpful since it is evident that the economy will converge to it as the number of layers increases. Furthermore, when $L \to \infty$ no valuable problems are thrown away. Thus w_∞ bounds the level of earnings agents can achieve with technology A. We now turn to the characterization of an equilibrium given t, A and L finite. The next proposition shows that an equilibrium given A and L finite exists, is unique, r_L^ℓ is decreasing in ℓ, and z_L^ℓ is increasing in ℓ. The logic is straightforward. Start with layer L. These problem solvers cannot resell the problems to a higher layer. Hence, relative to agents one layer below, who can resell their problems, agent in L are willing to pay less for them than agents in layer $L-1$ are willing to pay for the problems they buy. Similarly, agents in layer $L-1$ are willing to pay less for the problems they buy than agents in layer $L-2$ as they can sell them for a low price to agent in layer L. This logic goes through for all layers. The more layers on top of an agent the more valuable the problem, as it can potentially be sold to all the layers above, up to L. Now consider the amount of knowledge acquired by agents. Agents in layer L cannot sell their problems and so they have an incentive to learn as much as possible to extract as much value as possible from each problem. In contrast, agents in layer $L-1$ are less willing to learn as they can sell their problems to agents in layer L. Agents in layer $L-2$ get a higher price for their unsolved problems so their incentives to learn are smaller than the agents above them. Again, this logic applies to all layers in the hierarchy, including layer 0 where the fall in knowledge is even larger since worker can span their knowledge over only one problem instead of $1/h$ of them (since they use their time to produce). Of course, as $L \to \infty$ this logic does not apply and all prices and knowledge levels of problem solvers are constant, since there is no final layer in which prices are equal to zero.

To prove the next proposition we will use the following parameter restriction which is necessary and sufficient for $z_L^\ell > 0$ for all ℓ and L.

Condition 1 *$A \geq 1$, $h < 1$ and $A, \lambda, \widetilde{c}$ and h satisfy*

$$\frac{A\lambda}{\widetilde{c}} > \frac{1}{h} + \ln h.$$

Proposition 2 *Under Condition 1, for any A, and L finite, there exists a unique equilibrium determined by a set of prices $\left\{r_L^\ell\right\}_{\ell=0}^{L-1}$ and a set of knowledge levels*

$\left\{z_L^\ell\right\}_{\ell=0}^{L}$ *such that* $r_L^\ell > 0$ *is strictly decreasing in* ℓ *and* $z_L^\ell > 0$ *is strictly increasing in* ℓ.

Proof. Use (3) to obtain the knowledge of each agent as a function of the price the agent receives for a problem passed. Letting $\alpha \equiv \frac{\widetilde{c}h}{\lambda}$ and $\beta \equiv A - \frac{\widetilde{c}h}{\lambda}$ and using (4) we obtain the following recursion for the set of prices:

$$\begin{aligned} r_L^{L-1} &= \beta - hw_L^L + \alpha \ln \frac{\alpha}{A} \\ r_L^{\ell-1} &= \beta - hw_L^\ell + \alpha \ln \frac{\alpha}{\left(A - r_L^\ell\right)} \text{ for } 0 < \ell < L. \end{aligned}$$

Imposing (5) for $\ell = 1, ..., L-1$ we obtain that

$$\begin{aligned} r_L^{L-1} &= \beta - h\tilde{w}\left(A, L\right) + \alpha \ln \frac{\alpha}{A} \\ r_L^{\ell-1} &= \beta - h\tilde{w}\left(A, L\right) + \alpha \ln \frac{\alpha}{\left(A - r_L^\ell\right)} \text{ for } 0 < \ell < L. \end{aligned} \tag{6}$$

For a given $\tilde{w}(A, L)$ there exists at most one $r_L^0 > 0$ such that the whole system holds. Specifically, note that given $\tilde{w}(A, L)$ we can determine r_L^{L-1}. So choose some $\tilde{w}(A, L) > 0$ such that the resulting price $r_L^{L-1} > 0$ (and abusing notation slightly denote by $r_L^\ell(\tilde{w})$ the solution of the system above given $\tilde{w}$). It is easy to see that, since $r_L^{L-1} > 0$, $r_L^{L-2}(\tilde{w}) > r_L^{L-1}(\tilde{w})$. Repeating this argument we can conclude that $\left\{r_L^\ell(\tilde{w})\right\}_{\ell=0}^{L-1}$ is decreasing in ℓ. It is also immediate from (3) that the higher the price the lower the corresponding knowledge level, so $\left\{z_L^\ell(\tilde{w})\right\}_{\ell=0}^{L}$ is increasing in ℓ (note that for z_L^0 there is an extra effect coming from the fact that workers cannot span their knowledge over many problems, a missing h in (3)). Condition 1 guarantees that the resulting values $\left\{z_L^\ell(\tilde{w})\right\}_{\ell=0}^{L}$ are positive, as $r_L^\ell(\tilde{w}) < r_\infty$ since when $L \to \infty$ all prices are positive (as opposed to zero in layer L) and, as can be readily observed in the system of equations above, prices in layer $\ell - 1$ are increasing in prices in layer ℓ. Note also that as the price at which agents in layer L can sell problems is equal to zero, the prices for all other layers are strictly positive.

Note that $r_L^0(\tilde{w})$ is decreasing in $\tilde{w}$ as

$$\frac{dr_L^0(\tilde{w})}{d\tilde{w}} = -h + \frac{\alpha}{A - r_L^1(\tilde{w})} \frac{dr_L^1(\tilde{w})}{d\tilde{w}}$$

and

$$\frac{dr_L^{L-1}(\tilde{w})}{d\tilde{w}} = -h,$$

so

$$\frac{dr_L^0(\tilde{w})}{d\tilde{w}} = -h\left(1 + \sum_{\ell=1}^{L-1}\prod_{k=1}^{\ell}\frac{\alpha}{A - r_L^k(\tilde{w})}\right) < 0,$$

and we can therefore invert it to obtain $w_L^s(r_L^0)$ which is also a continuous and strictly decreasing function.

Now consider the equation determining the wages of production workers and define

$$w_L^p(r_L^0) = A - \frac{\widetilde{c}}{\lambda}\left(1 - \ln\frac{\widetilde{c}}{\lambda(A - r_L^0)}\right) \tag{7}$$

which is a continuous and strictly increasing in r_L^0.

The last equilibrium condition is given by (5) for $\ell = 0$, and so $w_L^s(r_L^0) = w_L^p(r_L^0)$ for the equilibrium r_L^0. Since w_L^s is strictly increasing and w_L^p is strictly decreasing, if a crossing exists it is unique. But note that at $r_L^0 = A - \widetilde{c}/\lambda$, $w_L^p(A - \widetilde{c}/\lambda) = A - \widetilde{c}/\lambda$ and

$$\begin{aligned} w_L^s(A) &= \frac{\widetilde{c}}{\lambda h} - \frac{\widetilde{c}}{\lambda} + \frac{\widetilde{c}}{\lambda}\ln\frac{\widetilde{c}h}{\lambda(A - r_L^\ell)} \\ &< \frac{\widetilde{c}}{\lambda}\left(\frac{1}{h} + \ln h - 1\right) \\ &< A - \widetilde{c}/\lambda \end{aligned}$$

by Condition 1 and $r_L^0 > r_L^1$. Hence, $w_L^p(A - \widetilde{c}/\lambda) > w_L^s(A - \widetilde{c}/\lambda)$.

Now let $r_L^0 = 0$. Then

$$w_L^p(0) = A - \frac{\widetilde{c}}{\lambda} + \frac{\widetilde{c}}{\lambda}\ln\frac{\widetilde{c}}{\lambda A}$$

and note that

$$\begin{aligned} w_L^s(0) &= \frac{A}{h} - \frac{\widetilde{c}}{\lambda} + \frac{\widetilde{c}}{\lambda}\ln\frac{\widetilde{c}h}{\lambda(A - r_L^1)} \\ &> \frac{A}{h} - \frac{\widetilde{c}}{\lambda} + \frac{\widetilde{c}}{\lambda}\ln\frac{\widetilde{c}}{\lambda A} \end{aligned}$$

since $h < 1$ and $r_L^1 \geq 0$. Thus, $w_L^p(0) < w_L^s(0)$. The Intermediate Value Theorem then guarantees that there exists a unique value r_L^0 such that $w_L^s(r_L^0) = w_L^p(r_L^0)$ and so a unique equilibrium exists. ■

We now turn to the properties of this economy as we change the highest layer L. Note that for now, without radical innovations, changes in L happen as time evolves and so studying the properties of our economy as we change the number of layers is equivalent to studying the properties of our economy as time evolves. This equivalence will change in the next section once we introduce radical innovations as we will have organizations evolving for different technologies across time. The next proposition shows that as the number of layers increases so do wages (or output per capita if knowledge cost are considered forgone output). Furthermore since wages are bounded by w_∞, there are eventual decreasing returns in the number of organizational layers. This is just the result of higher layers dealing with less problems as they are more rare. So adding an extra layer contributes to output per capita (since more problems are solved) but it contributes less the higher the layer since there are fewer and fewer problems that require such specialized knowledge.

The proposition also shows that as time evolves and the number of layers increases, r_L^ℓ increases and z_L^ℓ decreases for all ℓ. The first result is a direct consequence of the logic used in the previous proposition. As time elapses and the number of layers increases the number of layers above a given ℓ increases, which implies that r_L^ℓ increases, since the problems can be resold further if not solved. In turn, higher prices in turn imply less knowledge acquisition as the opportunity to resell problems is a substitute for solving them.

Proposition 3 *Under Condition 1, for any technology A, as the number of layers L increase, w_t increases and $\lim_{L\to\infty} \tilde{w}(A, L) = w_\infty$. Furthermore, as the number of layers L increase, prices r_L^ℓ increase for all $\ell = 0, ..., L-1$ and knowledge levels z_L^ℓ decrease for all $\ell = 0, ..., L$. As $L \to \infty$, $r_L^\ell \to r_\infty$ for all $\ell = 0, ..., L-1$ and $z_L^\ell \to z_\infty^0$ all $\ell = 0, ..., L$.*

Proof. Consider the individual incentives of an agent in period t to form layer $L+1$ given that the economy's highest layer is L. Such an agent can use the problems thrown away by the agents in layer L. The wage such an agent in layer $L+1$ would command is given by

$$\frac{A}{h} - \frac{\widetilde{c}}{\lambda}\left(1 - \ln \frac{\widetilde{c}h}{\lambda A}\right)$$

which is always greater than the equilibrium wage in the economy given by

$$w_t = \frac{A - r_L^{L-1}}{h} - \frac{\widetilde{c}}{\lambda}\left(1 - \ln \frac{\widetilde{c}h}{\lambda A}\right),$$

since as shown in the previous proposition $r_L^{L-1} > 0$. Therefore, in the next period such an agent has incentives to enter and form layer $L+1$. Of course, once he enters, agents in layer L will demand a positive price for their problems and so some of the surplus will be distributed to other agents in the economy. However, the economy as a whole will produce more output as the higher price is only a redistribution of wealth between agents. Agents will also re-optimize and choose different levels of knowledge $\{z_L^\ell\}_{\ell=0}^{L}$ which will increase the surplus, as they have the option to choose the same level of knowledge they chose before. Hence, $\tilde{w}(A, L) > \tilde{w}(A, L+1)$ for all L.

This result can be formally proven as follows. Consider $r_L^0(\tilde{w})$ defined in the proof of Proposition 2. As $r_L^L = 0$ (the last layer throws problems away) but $r_{L+1}^L > 0$ and since for a given w, by Equations (6), $r_{L+1}^{\ell-1}$ is increasing in r_{L+1}^ℓ, we obtain that $r_L^0(w) < r_{L+1}^0(w)$. Now define the function r^p using Equation (7), the price of problems sold by workers, as

$$r_p(w) \equiv A - \frac{\widetilde{c}}{\lambda} e^{\frac{\lambda}{\tilde{c}}(A-w)-1}.$$

In an equilibrium with L layers we know that $r_p(\tilde{w}(A, L)) = r_L^0(\tilde{w}(A, L))$ and in an equilibrium with $L+1$ layers $r_p(\tilde{w}(A, L+1)) = r_{L+1}^0(\tilde{w}(A, L))$. Since $r_L^0(w) < r_{L+1}^0(w)$ and $r_p'(w) > 0$ and $r_L^{0\prime}(w) < 0$, this implies that $\tilde{w}(A, L+1) > \tilde{w}(A, L)$ and that $r_{L+1}^0 > r_L^0$. By (6) this in turn implies that $r_{L+1}^\ell > r_L^\ell$ for all $\ell < L-1$. Note also that by (3) this implies that $z_{L+1}^\ell < z_L^\ell$ for all $\ell < L-1$.

Note that as we have shown in Proposition 2, $r_L^\ell < r_\infty$ for all ℓ and L finite. Hence, since $\{r_L^\ell\}_{L=0}^{\infty}$ is a strictly increasing and bounded sequence it has to converge for all ℓ. Since the equilibrium is unique as shown in Proposition 2 the limit is r_∞. Hence, as $L \to \infty$, $\{r_L^\ell\}_{L=0}^{\infty}$ approaches r_∞ from below. Equations (3) then imply that $\{z_L^\ell\}_{L=0}^{\infty}$ converges to z_∞^ℓ from above. ■

The previous proposition shows that our economy will grow. But it also shows that the level of wages is bounded. Hence, growth in wages (or per capita output) will converge to zero. That is, the economy does not exhibit permanent growth. We now turn to embed this evolution over time of organizations with a given technology A in a growth model in which agents will have a choice to switch to better technologies as they learn. This will yield a long-run growth model that will exhibit permanent growth and where this growth will be driven by the ability of agents to organize.

The following graphs illustrate the results proven in the previous propositions. Figure 1 shows an example of a wage path. The properties we have proven are easy to identify: wages increase at a decreasing rate and have an asymptote at w_∞. Figure 2 shows the evolution of the size of a typical hierarchy. The top layer size is normalize to one. Clearly, as the economy adds more markets for expertise, or layers, the lower layers expand more than proportionally.

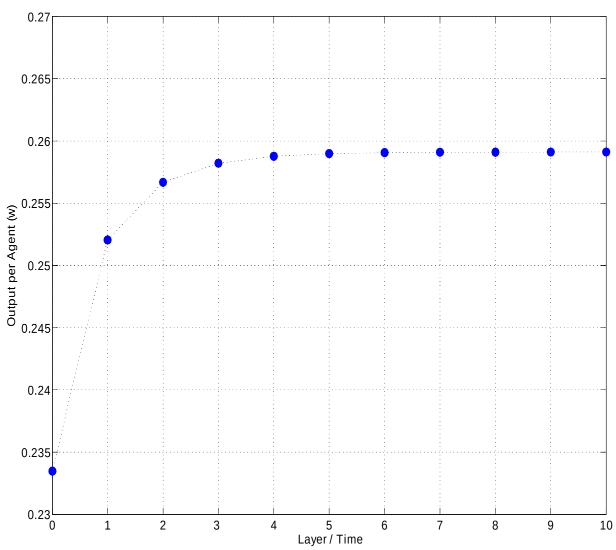

Figure 1

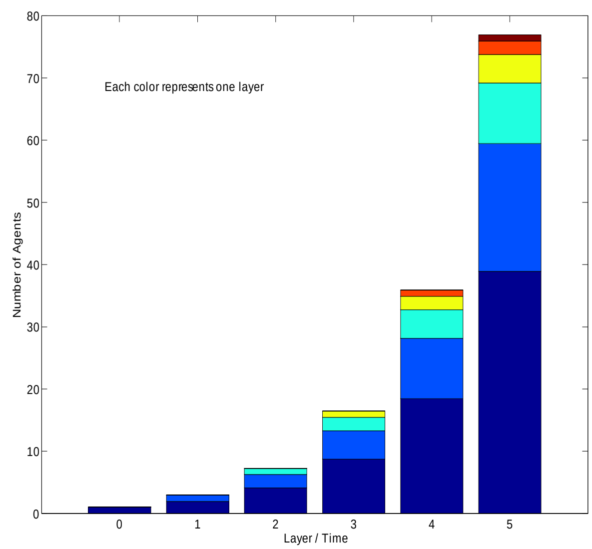

Figure 2

Figure 3 illustrates the problem for prices for an economy with 5 layers and an economy with 7 layers. As proven above, prices of problems are higher in an economy with a higher maximum number of layers: Problems can be exploited further and so conditional on someone having tried to solve them and failed, they preserve more value. Prices of problems at the highest layer are equal to zero by assumption, since there is no higher layer to sell them to. The price of problem is also decreasing as we move up the hierarchy since problems are more and more selected. Figure 4 present the knowledge acquisition of agents for the same two exercises. The picture is the reverse image of the prices in Figure 3, since higher prices imply less knowledge as selling the problems becomes more attractive. So, given the layer, knowledge acquisition is higher the smaller the maximum number of layers.

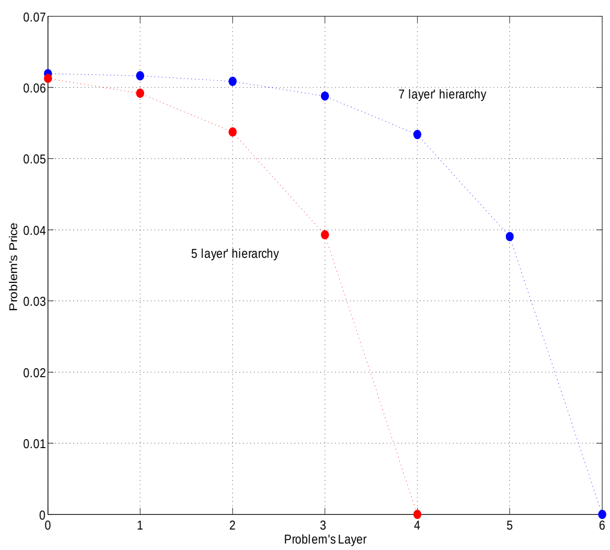

Figure 3

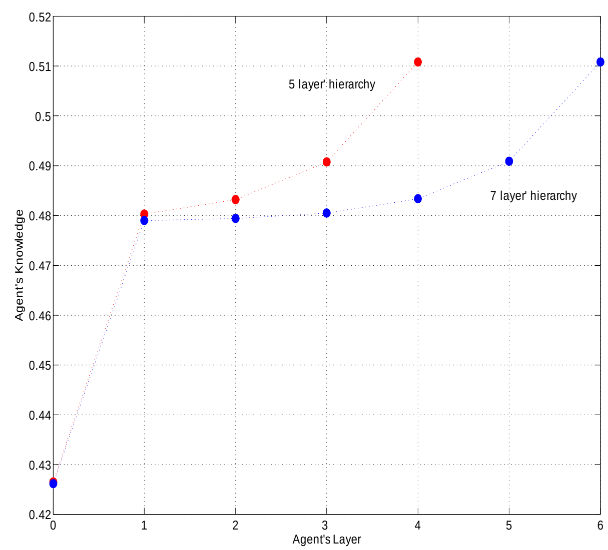

Figure 4

Another way of illustrating the implications of our model for knowledge acquisition given the level of technology A is presented in Figure 5. The figure presents cumulative knowledge in the economy. Clearly, as the number of layers increases the cumulative amount of knowledge acquired increases. Note that this is happening even though the growth in output per capita is converging to zero as illustrated in Figure 1.[5] Note also that the knowledge of problem solver in hierarchies with two or more layers is constant. The figure also illustrates how, given the layer or occupation, the knowledge acquired decreases with time (or the maximum number of layers).

[5] It is easy to develop a model of endogenous growth where the linear knowledge accumulation observed in Figure 1 creates, through an externality, improvements in A. These improvements would make agents switch to a new technology when the value of developing an extra layer with the current technology is smaller than the value of starting with the new technology but zero layers (no organization).

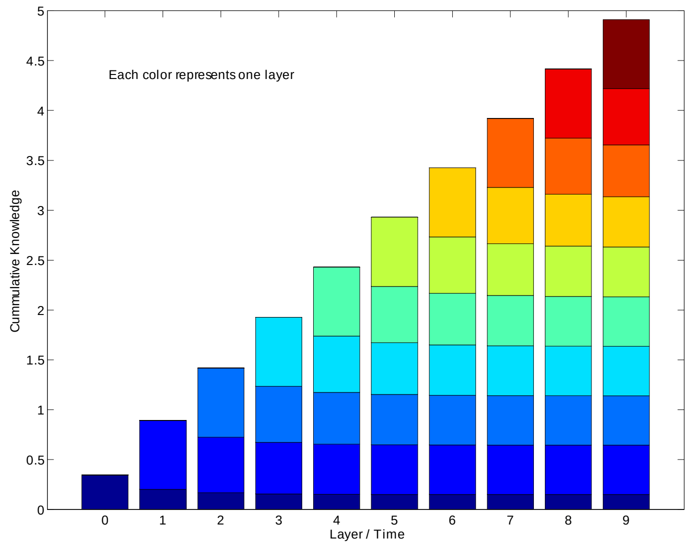

Figure 5

Before we end this section it is important to make one remark about the evolution of the distribution of gross wages (without subtracting learning costs). Overall wage inequality, as measured by the ratio of the gross wages of the highest level problem solvers to the gross wages of workers, increases over time as a technology is more efficiently organized. To see this note that everyone gains the same net of learning costs, and knowledge levels of workers decrease with the number of layers, while knowledge levels of entrepreneurs at the highest layer are constant. In contrast, the distribution of gross wages among problem solvers becomes less dispersed. The reason is that more layers are added and knowledge levels of intermediate problem solvers converge to z_∞. Thus, as organizations develop over time, inequality between workers and managers increases while inequality within problem solvers decreases.

2.3. Technological Innovation

In the previous section we studied how an economy organizes given a technological level A. In our economy, as organizations grow and become more complex, society learns how to solve a wider set of problems faced when using this technology. This knowledge is fully appropriable and society invests optimally, conditionally on A, on the development of this problem solving knowledge. We now turn to the evolution of the technology A that we kept fixed in the previous section. For this, we will assume that the technology is a fully private and rival good. Agents can invest in improving their technology and sell their technology to other agents (in fact, this will be the only savings technology of an agent). However, a technology has to be used in order to be productive and organizations have to be developed to exploit it. Of course, only agents that use the same technology can work with each other.

We studied above the problem of exploiting a given technology. This problem produced the equilibrium earnings that all agents in the economy working with a particular technology receive in equilibrium. These earnings changed every period as new layers of expertise develop in the economy to exploit this technology. As in the previous section, denote the equilibrium earnings of agents working with technology A in an economy with $L+1$ layers of expertise by $\tilde{w}(A,L)$. $\tilde{w}(A,L)$ contains all the information we need about the organization of the economy to study the decisions of agents to invest in innovation.

The cost of learning new technologies, measured in terms of foregone income, increases with the level of the new technologies —the more productive the technology the higher the cost of spending time learning to solve problems. Thus we specify the learning cost of the new technology as $\widetilde{c} = cA$. Then the cost of learning how to solve problems in an interval of size z is equal to Acz where A is the technology currently in use (not necessarily the best one). All the analysis in the previous section remains unchanged apart from the parameter $\widetilde{c}$ now becoming Ac.[6] Then, it is easy to see from Equations (4) and the equilibrium conditions in (5) that the prices of problems are proportional to the level of technology A so r_L^ℓ/A satisfies the problem in the previous section with $A=1$. Thus, earnings are proportional to the level of technology. Namely, $\tilde{w}(A,L) = Aw(L)$, where $w(L)$ is the equilibrium wage given that the economy has organized $L+1$ layers and $A=1$. So earnings are linear in the level of the technology in use. This is a property that we will exploit heavily below.

An agents with technology A faces a cost of $A\psi\zeta^2$ of improving the best technology he can use to $(1+\zeta)A$. So we are assuming that there are quadratic adjustment

[6] Note that if we do not scale c by the level of technology, as the economy grows the cost of acquiring knowledge relative to output will converge to zero. This would introduce an obvious scale effect in the model.

cost of improving the frontier technology. The frontier technology is not necessarily the technology in use, since changing technology implies wasting the organizations created to exploit the older technology. So agents will improve the frontier technology and, once it is good enough and the benefits from developing the old technology have died out, agents will switch to the frontier technology and will start creating new organizations to exploit this technology. The role of the quadratic adjustments costs is to have smooth investments in the frontier technology. Without them, agents would invest in the new technology only the period before a new technology is put in place.

As discussed before, agents order consumption using a linear utility function with discount factor β. Because of this preference specification, the marginal utility of consumption in every period and for every agent in the economy is equal to one. Hence, the solution to the problem of a sequence of two period lived agents that work when they are young, is identical to the solution to the problem of an infinitely lived agent, up to consumption transfers between generations. Thus, the solution to the infinitely lived agent problem will lead to innovation levels and a timing of radical innovation that are identical to the solution to the two period lived agents problem. That is, both problems will maximize the discounted sum of consumption. The actual consumption levels every period will, however, depend on the level of prices at which individuals can sell their technology. In what follows we discuss the infinite horizon problem to analyze the equilibrium allocation of this economy. We then discuss the price of technology and the consumption pattern across generations.

The problem of an infinitely lived agent (or a sequence of two period lived agents that sell each other their technology) is therefore given by

$$\tilde{V}\left(A, A', L\right) = \left[\max \left[\begin{array}{c} \max_{\zeta} \tilde{w}\left(A, L\right) + \beta \tilde{V}\left(A, A'', L+1\right) - A\psi\zeta^2, \\ \tilde{V}\left(A', A', 0\right) \end{array} \right] \right]$$

where

$$A'' = A'\left(1+\zeta\right).$$

$\tilde{V}\left(A, A', L\right)$ denote the value function of an agent using technology A, with a frontier technology A' in an economy with maximum layer L. Note that we are assuming that all individuals (or sequence of individuals) have the same technology so their technology, and the frontier technology, are the same for ever agent in the economy (more on this below). An agent can work using the current technology A with organizations of $L+1$ layers (remember that there is a layer zero) and receive $\tilde{w}(A,L)$, invest ζ at cost $A\psi\zeta^2$, and tomorrow repeat the problem with a new frontier technology $A'' = A'(1+\zeta)$ and organizations with $L+2$ layers. He can also decide to drop the available organizations and do a radical innovation, in which case he uses the frontier technology with zero layers and gets value $\tilde{V}(A', A', 0)$.

We are assuming that everyone in this economy has the same A and A' and therefore, as agents are identical they choose the same pattern of innovations. This requires that, at period zero, all agents start with the same frontier technology. If this is the case, then no agent would want to deviate by himself since this would force him to work on his own, therefore reducing his reliance on organization and therefore his income.[7] Of course, if we allow the economy to start with agents that are heterogenous in their level of technology, there may be other non-symmetric equilibria. The study of these equilibria may be interesting, but we leave it for future research and, in this paper, focus on symmetric equilibria only.

The fact that $\tilde{w}(A,L) = Aw(L)$ implies that $\tilde{V}(A,A',L)$ is also homogenous of degree one in A and so $\tilde{V}(A,A',L) = A\tilde{V}\left(1, \frac{A'}{A}, L\right) = AV\left(\frac{A'}{A}, L\right)$ where

$$V(G,L) = \max \begin{bmatrix} \max_{\zeta} w(L) + \beta V(G', L+1) - \psi\zeta^2, \\ GV(1,0) \end{bmatrix}$$

where and

$$G' = G(1+\zeta).$$

Note that $G = A'/A$ know has the interpretation of the technological gap. The ratio of the frontier technology to the technology in use. This is the only state of technology that is relevant for our problem.

Now remember that as $L \to \infty$, $w(L) \to w_\infty/A = 1 - \frac{c}{\lambda} - \frac{c}{\lambda}\ln\left(\frac{\lambda h}{c} - h\ln h\right)$. Hence, for L large the discounted benefit of developing an extra layer converges to zero. Similarly, the benefits of innovation also converge to zero as L becomes large since they are proportional to $\beta^L V(1,0) \to 0$ as $L \to \infty$. This implies that there are two cases that we need to consider. First, the case where there exists a unique finite L^* such

[7] This follows immediately from $w(A,0) \leq w(A',L)$ for all L and $A \leq A'$ and our assumption that production knowledge is specific to the technology in use.

that next period a new technology is put in place, namely, $V(G, L^*+1) = GV(1,0)$ for all G. This is the case in which the value of innovation converges to zero more slowly that the value of an extra layer. Second, a case where the value of innovation converges to zero faster than the value of an extra layer and so L^* is infinity. In this case a new technology is never put in place and so the economy stagnates and the growth rate converges to zero as we will discuss below.

Since, given a finite L^*, the problem of choosing an innovation level ζ is a well behaved concave problem by design, an optimal innovation level exits and is unique. Denote it by $\zeta^*(L)$. Of course, if $L^* = \infty$, $\zeta^*(L) = 0$ all L, since there are no incentives to invest in innovation capital.

In any case, we can write the value function as

$$V(G, L) = \sum_{\ell=L}^{L^*} \beta^{\ell-L} \left(w(\ell) - \psi\zeta^*(\ell)^2\right) + \beta^{L^*+1-L} V(1,0)\, G \prod_{\ell=L}^{L^*} (1+\zeta^*(\ell))$$

and so

$$V(1,0) = \sum_{\ell=0}^{L^*} \beta^{\ell} \left(\tilde{w}(\ell) - \psi\zeta^*(\ell)^2\right) + \beta^{L^*+1} V(1,0) \prod_{\ell=0}^{L^*} (1+\zeta^*(\ell))$$

or

$$V^* = V(1,0) = \frac{\sum_{\ell=0}^{L^*} \beta^{\ell} \left(w(\ell) - \psi\zeta^*(\ell)^2\right)}{1 - \beta^{L^*+1} \prod_{\ell=0}^{L^*} (1+\zeta^*(\ell))}.$$

So we can just restate the original problem in sequential form as finding L^* and $\zeta^*(\ell)$ for $\ell = 0, ..., L^*$ that solve

$$V^* = \max_{L, \{\zeta(\ell)\}_{\ell=0}^{L}} \frac{\sum_{\ell=0}^{L} \beta^{\ell} \left(w(\ell) - \psi\zeta(\ell)^2\right)}{1 - \beta^{L+1} \prod_{\ell=0}^{L} (1+\zeta(\ell))}, \tag{8}$$

where we are assuming that ψ is high enough to guarantee that $\beta^{L+1} \prod_{\ell=0}^{L} (1+\zeta(\ell)) < 1$. A sufficient condition to guarantee this condition is that innovation costs ψ are such that $\psi > (\beta/(1-\beta))^2 (w_\infty/A)$.

The investment in innovation are driven by the adjustment cost technology that we have assumed. The first order conditions with respect to $\zeta(l)$ are given by

$$\frac{\sum_{\ell=0}^{L} \beta^{\ell} \left(w(\ell) - \psi\zeta^*(\ell)^2\right)}{1 - \beta^{L+1} \prod_{\ell=0}^{L} (1+\zeta^*(\ell))} = \frac{\beta^{l} \psi 2 \zeta^*(l)}{\beta^{L+1} \prod_{\ell=0, \ell\neq l}^{L} (1+\zeta^*(\ell))} \text{ all } \ell$$

so

$$V^{*}\beta^{L^{*}+1}\prod_{\ell=0}^{L^{*}}\left(1+\zeta^{*}\left(\ell\right)\right)=\beta^{l}\psi 2\zeta^{*}\left(l\right)\left(1+\zeta^{*}\left(l\right)\right)\ \text{ all }\ell \tag{9}$$

Clearly, the left hand side does not depend on l and so since

$$\frac{\partial\zeta\left(l\right)+\zeta\left(l\right)^{2}}{\partial\zeta\left(l\right)}=1+2\zeta\left(l\right)>0$$

and β^{l} decreases with l as $\beta<1$,

$$\zeta^{*}\left(l\right)\leq\zeta^{*}\left(l'\right)\ \text{ all }l<l',$$

with equality when $\zeta^{*}\left(l\right)=0$ all l in the case of stagnation when $L^{*}=\infty$. Note that the fact that the investment is positive in all periods (when there is no stagnation) is only the result of discounting and the adjustment costs. Without adjustment costs we would only invest in the last period before the switch to a new technology. Note also that the above equation implies that $\zeta\left(0\right)>0$ if L^{*} is finite. That is, the economy will invest positive amounts in innovation capital, as long as it eventually switches to a new technology. Innovation is positive every period in the case of no stagnation. In the stagnation case $\beta^{L^{*}+1}=0$ and so the left hand side of the first order condition is equal to zero and so $\zeta^{*}\left(l\right)=0$. Note that even in this case the solution to the innovation knowledge problem satisfies the first order condition since the marginal cost of zero investment is zero.

It is optimal to add another layer of expertise as long as

$$\frac{\sum_{\ell=0}^{L}\beta^{\ell}\left(w\left(\ell\right)-\psi\zeta_{L}^{*}\left(\ell\right)^{2}\right)}{1-\beta^{L+1}\prod_{\ell=0}^{L}\left(1+\zeta_{L}^{*}\left(\ell\right)\right)}-\frac{\sum_{\ell=0}^{L-1}\beta^{\ell}\left(w\left(\ell\right)-\psi\zeta_{L-1}^{*}\left(\ell\right)^{2}\right)}{1-\beta^{L}\prod_{\ell=0}^{L-1}\left(1+\zeta_{L-1}^{*}\left(\ell\right)\right)}>0 \tag{10}$$

where ζ_{L}^{*} in the first term denotes the optimal innovation policy given that we switch technologies every $L+1$ periods. Note the two compensating effect. First, as we increase the number of periods we exploit a given technology we increase the total income we obtain from it. Second, as we delay the switch of technology we discount for a further period and have one extra period to invest in innovation. Since $w\left(\ell\right)$ is monotone in ℓ we know that the difference in (10) is monotone too which implies that there is a unique L^{*} as we conjectured above (see the upper-left panel of Figure 6 below). Note however, that this difference being monotone does not rule out the possibility that $L^* = \infty$. This is the case where the second effect (the discounting effect) dominates the first (each technology is more valuable because it is exploited for more periods) for all layers and so only one technology is put in place. As discussed above in this case there is no investment in innovation capital and so the long-run growth rate is zero.

We summarize our findings in the next proposition:

Proposition 4 *Given a common technology in period zero, A_0, there exists a unique competitive equilibrium of one of two types:*

1. *Permanent Growth: The equilibrium exhibits technological cycles of finite length. This cycles repeat themselves at output per capita that is $\prod_{\ell=0}^{L^*} (1 + \zeta^* (\ell))$ times higher each cycle. Investments in innovation increase with the number of layers in the organization, $\zeta^* (l) < \zeta^* (l')$ all $l < l'$. Furthermore $\zeta^* (0) > 0$, so the economy exhibits positive permanent growth. The average long-run growth rate is given by $\sqrt[L^*]{\prod_{\ell=0}^{L^*} (1 + \zeta^* (\ell))}$.*
2. *Stagnation: The equilibrium decreasing growth rates that converge to zero. Output per worker converges to w_∞. Investment in innovation is equal to zero each period and the long run growth rate is equal to zero as well.*

Figure 6 presents one example of an equilibrium allocation with positive permanent growth (first type of equilibrium). In the upper-left corner we show the total value of switching technologies every certain number of layers. The maximum of this curve is V^* and is obtained at L^*. The plot stops at layer 21, and reaches a maximum at $L^* = 20$. So the equilibrium allocation for these parameter values ($\beta = .87$, $h = .5$, $c = .9$, and $\psi = 50$ (we keep this value constant throughout the paper)) exhibits technological cycles every 21 periods. Note how the value increases the most when we lengthen the technology cycles from zero to one period, as developing the first organizational layer implies the largest gains. The lower-left panel shows the investment in innovation during the 21 periods this technology is exploited. At the beginning when only a few layers of organization have been formed, the switch to the boundary technology is far away in the future, so agents invest little in innovation. As the time of the switch approaches, agents invest progressively more. As Equations (9) reveal, this path is driven essentially by the discount factor β. The higher β the more even are investment over the cycle. Note that agents invest a positive amount each period so the frontier technology is constantly improving.

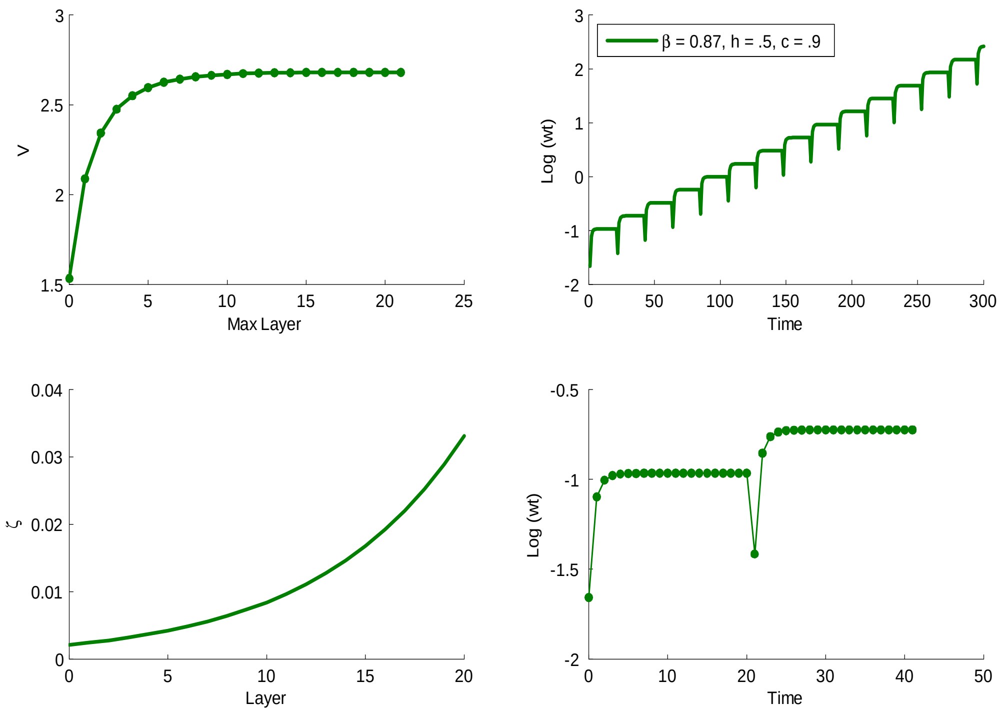

Figure 6

The right column of Figure 6 presents the natural logarithm of output per capita or wages ($\tilde{w}(A, L)$). The upper panel presents the long term view of this variable over 300 periods. It is clear that although output per capita grows in cycles the economy exhibits constant long term growth as we formalized in the previous proposition. The lower panel present a close up of equilibrium wages for two technological cycles. In the previous section we discussed the general shape of $\tilde{w}(A, L)$. This figure illustrates the time of the switch between technologies. As it is clear from the period, wages fall for at least a period before recovering and surpassing the old technology wages. The reason for the drop in wages is that the economy is losing the stock of organization (the old markets for expertise are not longer used). This implies a short term reduction in output per capita, although, of course, the economy wins in the long run, as the timing of the switch is optimal.

The second type of equilibrium, the one with stagnation, start similarly to the equilibrium in Figure 6, but instead of the value function reaching a maximum, the value function is strictly increasing but at a decreasing rate. Increases in the value function are always positive but converge to zero as we increase the maximum number of layers. Hence, the economy stays with the same technology forever, and $\ln \tilde{w}(A, L)$ converges, so growth rates go to zero.

We can now come back to the problem of a two period lived agent. Such an agent will buy a frontier technology (which allows him to produce with any technology below that) when his young and wants to produce. He will invest in innovation and improve his frontier technology and will sell it when he is old. In fact, the way in which we have setup the problem implies that the only way in which this agent can save is via technology (Of course, we could also add physical capital to the model in a standard way.) A technology includes all innovation knowledge acquired by society up to that point. In particular, it allows the agent to use the technology currently in place, but it also gives the agent the state of the art technology. The price of technology will be a function of the level of alternative technology being bought, the technology being used, and the number of layers in the current organization (the three state variables in our problem). Hence, the price of a technology is given by $\tilde{P}(A, A', L)$. Note that, as we argued for the value function above, this function needs to be homogenous of degree one in A and A'. Hence,

$$\tilde{P}(A, A', L) = A\tilde{P}\left(1, \frac{A'}{A}, L\right) \equiv AP\left(\frac{A'}{A}, L\right).$$

The problem of the agent is then

$$\max \begin{bmatrix} \max_{\zeta} w(L) + \beta P(G', L+1) - P(G, L) - \psi\zeta^2, \\ \max_{\zeta} G'(w(0) + \beta P(G', 1)) - P(G, L) - \psi\zeta^2, \end{bmatrix}$$

where

$$G' = G(1+\zeta).$$

Note that this problem is identical to the one we solved above if we let $P(G, L) = p + V(G, L)$ for any constant p. That is, prices will be identical to the value function up to lump sum transfers between generations. If $p = 0$ and we let prices be identical to the value function we are letting the first generation extract all the surplus that can be generated from the initial technology they own. If $p = -V(1,0)$ then this first generation does not get any of this surplus. For our purpose what is important is that this transfers do not change the equilibrium allocation in this economy apart from consumption. Hence, they do not affect growth, the length of cycles, or the organization of production.

Nevertheless, formulating the problem above is useful because it allows us to understand the role that technology appropiability plays in this economy. By appropiability we refer to the fraction of the payment for technology done by future generations that the owner of the technology obtains. That is, suppose that appropiability is very low, since we are modelling a country in which markets for technology are either taxed or very inefficient. Then there will be a gap between the price paid for technology by future generations and the price received by the old generation that owns the technology. The higher appropiability is in an economy, the more incentives agents have to innovate, since they will receive a higher price in return for their technology in the future. It is easy to see from the previous equation that the level of appropiability is technically equivalent to the discount factor β. So this parameter plays a double role, as a discount factor and as the level of appropiability in an economy. The higher β the higher appropiability and the more incentives do agents have to invest in innovating the current technology. It is immediate from Equation (8) and the envelope theorem that a higher β (an economy where technology is more appropriable) implies a higher welfare level. As we will discuss in the next section, a higher β also implies a smaller set of parameter values for which the economy stagnates.

3. THE EFFECT OF INFORMATION AND COMMUNICATION TECHNOLOGY ON GROWTH

In this section we study the effect of information and communication technology (ICT) on growth. We start by discussing the effect of ICT on growth in the case where the economy never stagnates. In the next subsection we study the circumstances in which economies never switch technologies and stagnate.

3.1. The Permanent Growth Case

We explore the model numerically, since as we will show some of the effect of information technology are quite complex. First note that the effect of the cost of acquiring information c and the cost of communication h change innovation and growth in the economy only through their effect on the wage schedule $w(L)$. This is evident from the problem in (8), as c and h do not enter directly in the problem. So ICT affects the dynamic technology innovation process only by changing the benefits of exploiting the current and future technologies. Essentially ICT is a technology that allows society to exploit other technologies. So better ICT, either through reductions in c or reductions in h, implies increases in $\tilde{w}(L)$ that, if we apply the envelope theorem to (8), result in increases in welfare. That is,

$$\frac{dV^*}{dc} \approx \frac{\sum_{\ell=0}^{L^*} \beta^{\ell} \frac{dw(\ell)}{dc}}{1 - \beta^{L+1} \prod_{\ell=0}^{L} (1 + \zeta^*(\ell))} < 0$$

since $d\tilde{w}(\ell)/dc < 0$ and

$$\frac{dV^*}{dh} \approx \frac{\sum_{\ell=0}^{L^*} \beta^{\ell} \frac{dw(\ell)}{dh}}{1 - \beta^{L+1} \prod_{\ell=0}^{L} (1 + \zeta^*(\ell))} \leq 0$$

since $dw(\ell)/dh \leq 0$ with equality if $L^* = 0$. The expressions are only approximate since L is a discreet variable so the envelope theorem does not apply exactly. These expressions then imply that welfare always increases with improvements of ICT independently of the source (unless $L^* = 0$ in which case welfare is unaltered).

Even though the effect of ICT on welfare is to unambiguously increase it, the same is not necessarily true for growth. Note that welfare could be increased by consuming more early and investing less in the future, as the future is discounted because of the standard reasons and, as we argue above, limits to the appropiability of technology. However, this is never the case for reductions in the cost of acquiring information c. As Figure 7, where we fix $h = .8$, illustrates, a reduction in c leads to shorter cycles and less organization, but faster growth. A reduction of c makes organization less necessary as acquiring information is cheaper and so agents acquire more. Hence, the number of layers of organization built to exploit a technology decreases. Even though there is less organization in the economy, growth increases as each technology is more valuable since solving problems is more affordable. This increases the value of present and future technologies which incentivates agents to innovate more since their innovations will be exploited more efficiently. Note also that if the value of future innovations is discounted less (or is more appropiable) the growth rate increases.

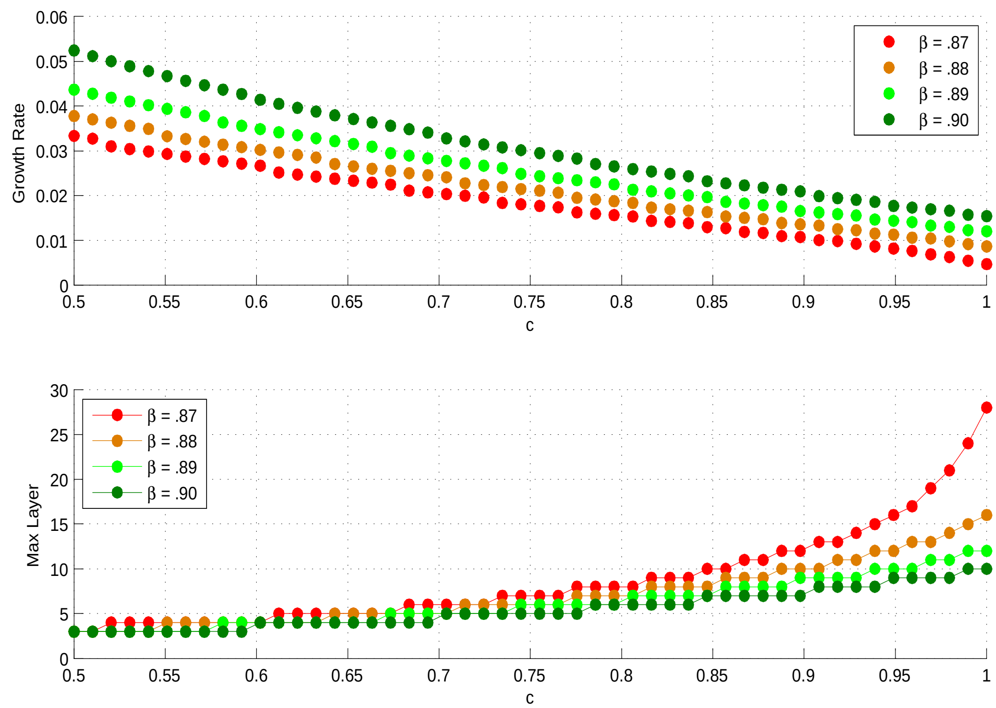

Figure 7.

Figure 8 shows two examples of the simulations presented in Figure 7. We fix $\beta = .87$ and change c from .9 to .7. Figure 7 shows how the value function increases almost proportionally as we decrease c, but the maximum number of layers decreases as organization is less useful since knowledge is cheaper. As c decreases investments in innovation increase in all periods although innovations are accumulated for less periods as the technology cycles are shorter. The net effect is an increase in the growth rate as can be seen in the upper-right panel. Note how reductions in c increase output per capita almost proportionally given a technology.

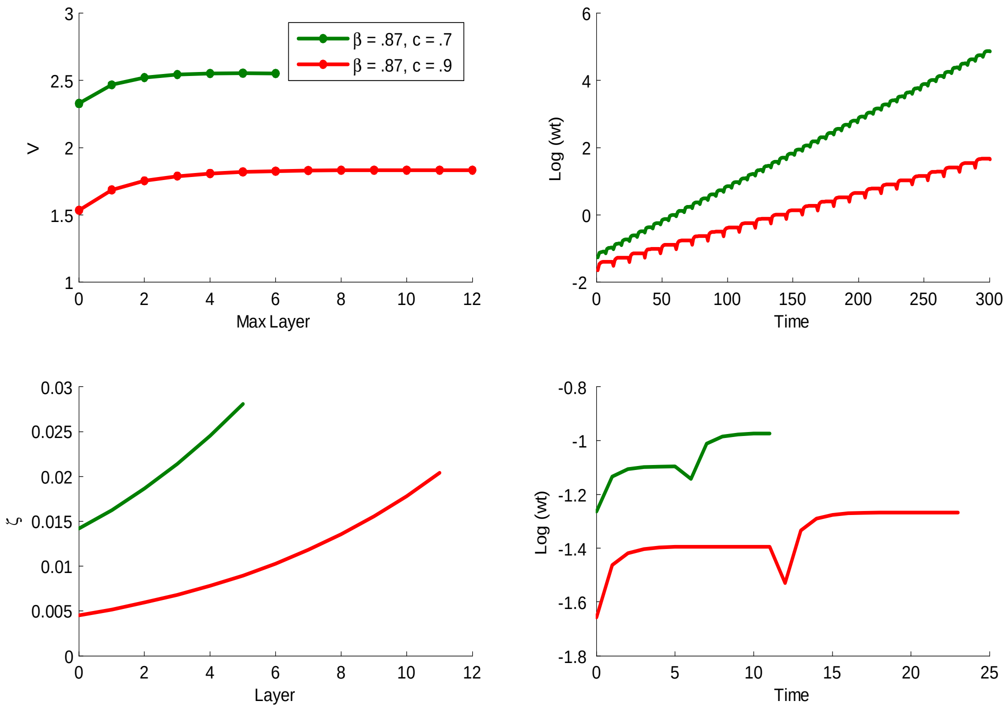

Figure 8

The effect of changes of communication technology on growth is more complicated. The main reason is that a reduction in h changes the output per capita extracted from a given technology more, the more organized is the technology. That is, the more heavily intensive is the exploitation of technology on communication. Hence, when agents are self-employed and there is no organization, changes in h do not affect output per capita. Figure 9 illustrates the effect of reductions of h on growth. Note that reductions in h increase the length of technology cycles and the use of organization when communication costs a high. This is intuitive, as smaller communication costs imply that building organizations is less costly. Agents can leverage their knowledge more, since they can deal with more problems as their span of control increases. However, as technology can be exploited more efficiently because building organizations is cheaper, the value of current and future technologies increases. On the one hand, increases in the value of present technologies reduce the incentives to innovate. On the other, increases in the value of future technologies increase the incentives to innovate. Note however, that since value is added by building larger organizations, the effects on future technologies are discounted more. So if β is small, the increase in the value of the present technology dominates, reduces investments in innovation and decreases growth. In contrast if β is large, the second effect dominates and the increase in the value of future technologies leads to more innovation and growth.[8] Figure 9 shows that reductions in h can reduce the length of the cycles if h is small (we let $c = 0.9$). This is because when h is small the value of organizing is concentrated more heavily in the first few layers.

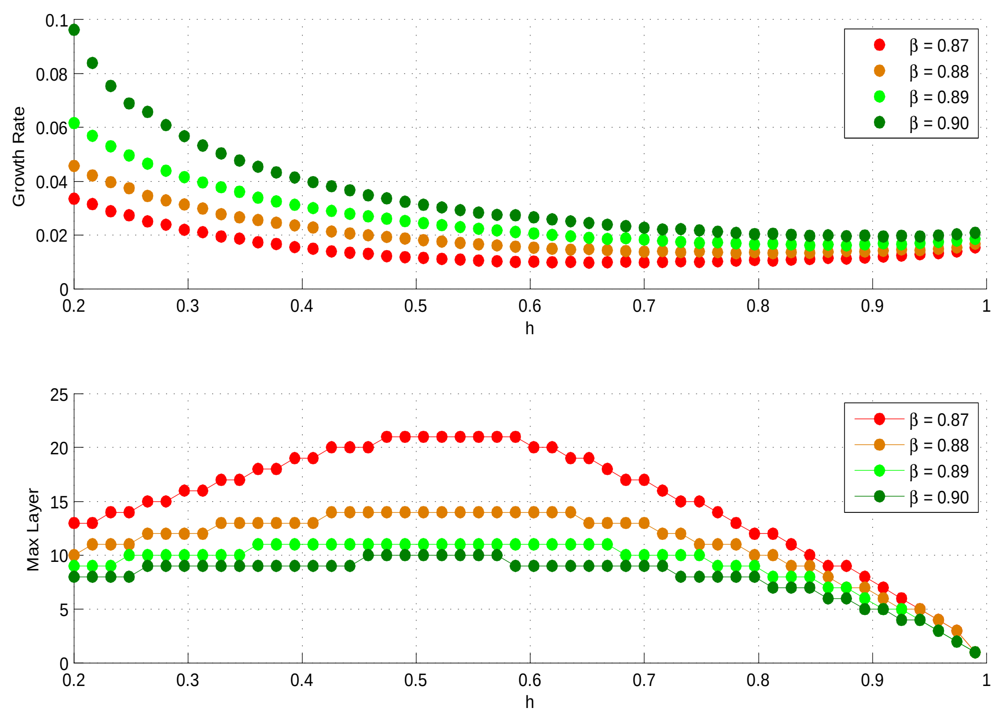

Figure 9

Figure 10 illustrates the point that improvements in communication technology can have negative effects on growth when appropiability is low (β is low). It shows the

[8]Note that changes in h do not change the speed at which one layer of organization can be built. If it did, decreases in h would lead to increases in growth for a wider range of parameter values.

effect of a reduction in communication costs when appropiability is high and when appropiability is low. It is clear from the picture that h does not affect wages in the first period, as we argued above. It is also clear that for both β's, investments in innovation are higher when h is high. However, agents also invest for fewer periods. The total accumulated effect is larger for lower h only when appropiability is high.

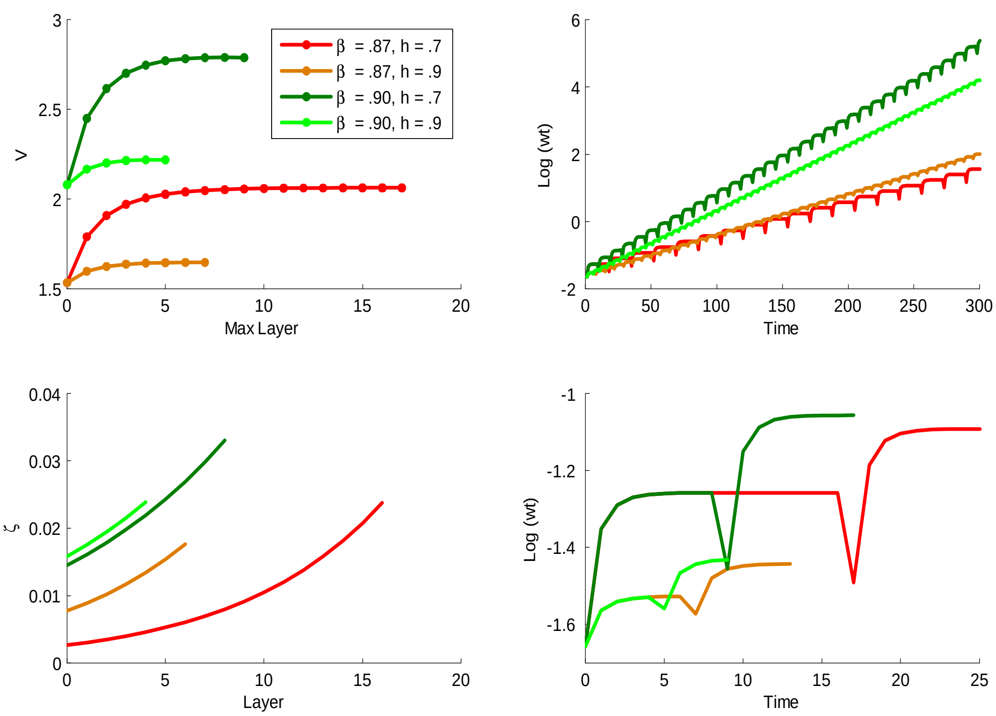

Figure 10

This exercise illustrates what we believe is an important point about the effect of communication technology on growth. Communication technology can improve current technology and perpetuate its use for longer as it makes organization more efficient. This will increase welfare but reduce growth if appropiability is low. Hence, what the model tells us is that countries where technology markets are not well developed and appropiability is low or deficient will experience negative effect of improvements in communication technology on their growth rates. However, as they improve communication technology, their technology cycles will eventually become shorter and growth will start to increase.

3.2. Stagnation

As we discussed in the previous section, there are cases in which economies stagnates as agents decide optimally not to invest in innovation. The reason is that agents prefer to extract more from the current technology in the near future through the development of organizations than to invest and innovate in the long term. Building organizations takes time and so there is a trade-off between exploiting technologies better and long term innovation.

It is easy to understand the effect of ψ on the possibility of stagnation. Clearly, the higher ψ the larger the set of parameters for which the economy will stagnate. This can be seen from the fact that the value of an agent in period zero is decreasing in ψ for all finite L, namely,

$$\frac{d\left(\frac{\sum_{\ell=0}^{L}\beta^{\ell}\left(w(\ell)-\psi\zeta(\ell)^{2}\right)}{1-\beta^{L+1}\prod_{\ell=0}^{L}(1+\zeta(\ell))}\right)}{d\psi}<0$$

by the envelope theorem but the value of never switching, given by $\sum_{\ell=0}^{\infty}\beta^{\ell}w\left(\ell\right)$ is independent of c.

The effect of the other three parameters is more complicated and cannot be signed analytically so we proceed with numerical simulations. Figure 11 shows a graph similar to Figure 7 but for a range of c's that includes larger values and for smaller values of β. As in Figure 7 we can see that a larger c implies a lower growth rate: The cost of acquiring knowledge has a negative effect on growth. However, now the Figure also illustrates how there is a threshold of c over which economic growth drops to zero. Further increases in the cost of knowledge have no effect on growth that stays at zero. That is, numerically we find a threshold for c over which there is stagnation. The lower panel of Figure 11 shows how the number of layers explodes to infinity as we approximate the threshold. [9] The logic for this result is straightforward. As c increases knowledge acquisition becomes more expensive which implies that organization becomes more useful to exploit technologies: Agents want to leverage knowledge more since knowledge is more costly. However, as agents increase the

[9] We calculate the equilibrium allowing for a maximum of 100 layers. If the value function is strictly increasing for all 100 layers we set $L^{*}=\infty$ and the growth rate to zero.

number of layers to leverage their knowledge, they also push back the date at which innovations would happen and so investing in innovation becomes less attractive. At some point the gains from innovation are so low (given that it will happen only in the very long run and agents discount the future) that it is better to keep exploiting the current technology. Of course, the more agents discount the future (or the less they can appropriate the value of innovation) the less valuable are future innovations and so the lower the threshold of c for which the economy stagnates.

This exercise as well as the one in the previous subsection indicates how in our model the growth rate of an economy and the cost of acquiring knowledge are related. However, the model also indicates why, in a cross-section of countries with different c's, output, knowledge acquisition and investments in innovation are not perfectly correlated.

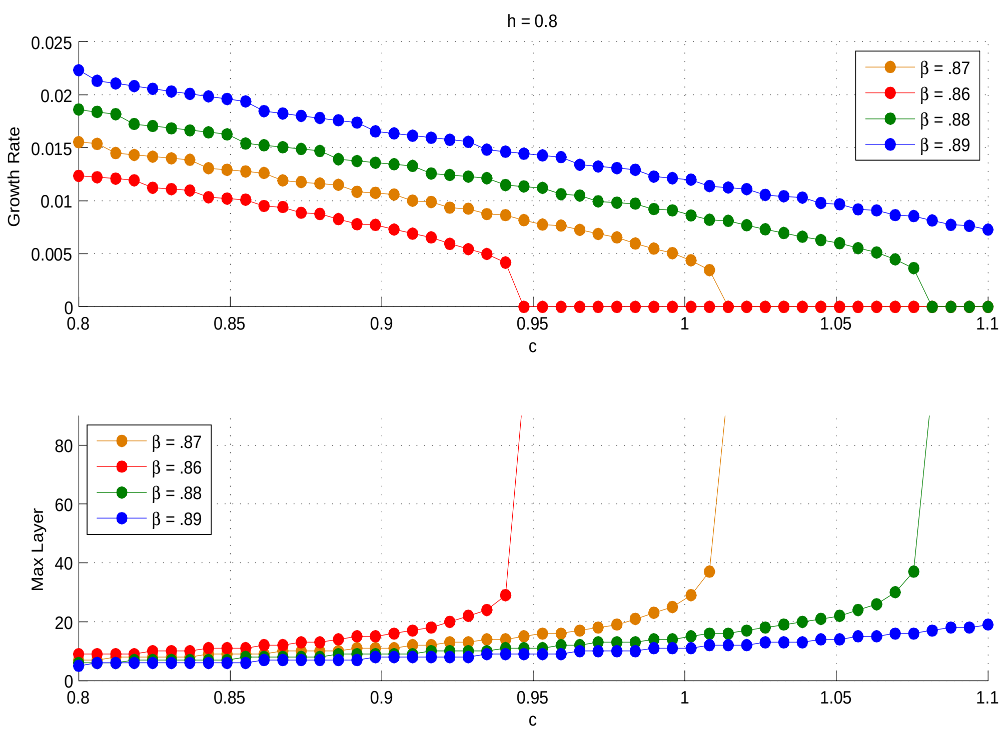

Figure 11

The effect of communication technology on the set of parameters for which we obtain stagnation is more complicated. The reason is, essentially, the non-monotone effect of h on the length of the technology cycle. For large values of h the maximum number of layers organized for a given technology is relatively low as the span of control of agents is small (and therefore their ability to leverage their knowledge through organizations). For small values of h the value of creating the first layers of organization is so large, that it is more valuable to keep innovating and organizing only these first set of expertise markets. For intermediate values of h the effect the number of layer is large as organization is useful because spans of controls are relatively large but the difference between the value of the first and later layers is not large enough to want to only organize the first layers and innovate quickly. Hence, it is for the intermediate values of h for which the number of layers might explode and we can obtain a stagnating economy. This is illustrated in Figure 12 which parallels Figure 8 but includes lower values of the discount factor or appropiability level β. The figure shows how as we decrease β the set of values of h for which we obtain appropiability increases. As we have argued, and as the lower panel shows, these are the parameter values for which $L^* = \infty$.

The welfare effects still hold. Decreases in h always increase welfare in this economy, independently of whether the economy stagnates or not. However, h will affect the growth rate dramatically. Communication technology is a technology to exploit today's and future's technologies, and the cost of using it is that it takes time since society has to organize these markets for expertise. If this technology is good, but not great, the value of organizing many of these markets is very large and so society prefers to do that than to invest in innovations that will benefit output only in the very far future. But how can an economy with zero growth maximize welfare? The key is that we are maximizing welfare from today's perspective of an infinitely lived agent (or a dynasty of two period lived agents). These agents value a lot the savings from not investing in innovation today and value little the fact that output per capita does not go up in the future. So, in this economy, stagnation is a choice of the country's ancestors. Current agents do not change the choice because the level of technology is too low relative to output (given the large organizations) to start investing in innovation; it would take too long (or it would be too expensive) to improve the alternative technology enough to make a technological switch valuable.

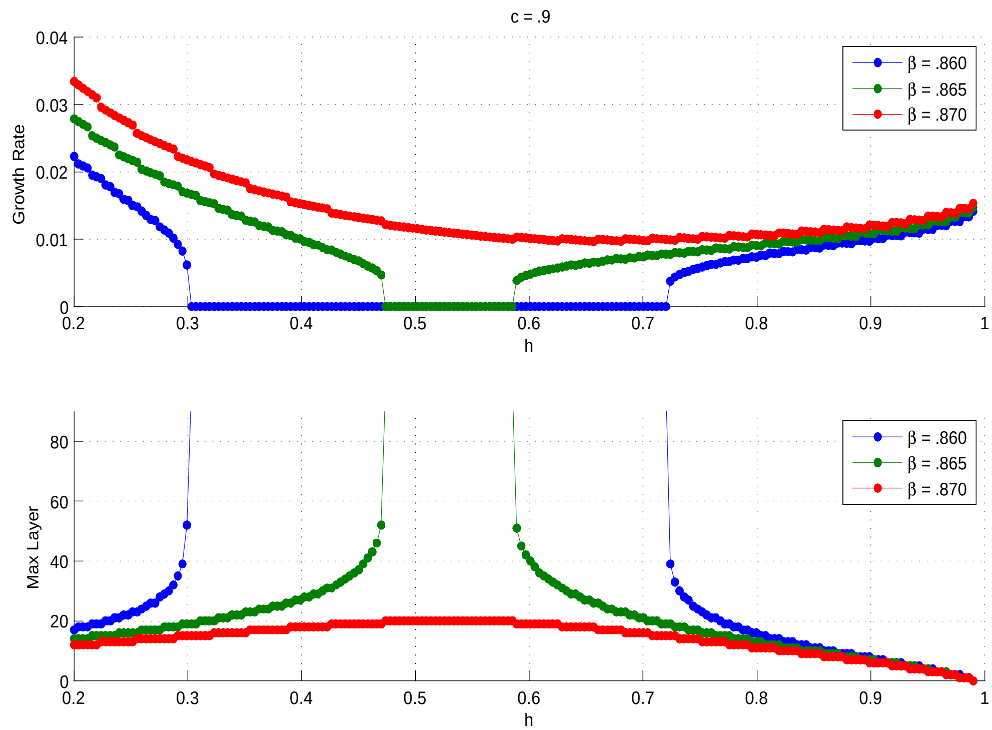

Figure 12

4. TECHNOLOGICAL CYCLES IN THE DATA

Our view is supported by two main stylized facts.[10] First, radical innovations often take place outside existing firms. Second, successful firms grow fast first but their growth slows down as their innovation is exploited. Consider the first one. If organizations are about acquiring knowledge, and knowledge is technology specific, we expect radical innovation to take place outside of existing organizations and to replace them. Casual observation is consistent with this, as the examples in the introduction suggest. Microsoft, Apple, E-Bay and Google, are all recent examples of

[10]The two following stylized facts concerns firms; our model is broader than that, as it applies to organizations mediated through markets, as well as firms. Firm birth and growth is easier to measure than the birth and growth of specialized consultants and experts in related technologies, and that is why this is the evidence that exists. To the extent that changes inside firms and outside are correlated, the evidence is still illuminating.

new and large organizations that started small, have grown, and have replaced the old large organizations. An expression of this phenomenon are the famous stories (such as Hewlett-Packard) where founders (in HP's case, Bill Hewlett and Dave Packard) developed the idea that was the germ of a large firm in their garage. Some systematic evidence in this respect is provided by a study by Rebecca Henderson (1993), where she shows that, in the photolitographic equipment industry (producing pieces of capital equipment to manufacture solid-state semiconductors), each stage of technological change was brought about by a new set of firms. Table 1, taken from her work, shows the evolution of market shares of firms in this market as the technology changed.

Cumulate share of sales of photolithographic alignment equipment, 1962-1986, by generation

	Contact	Proximity	Scanner	S&R (1)	S&R (2)
Cobilt	44		<1		
Kasper	17	8		7	
Canon		67	21	9	
P-Elmer			78	10	<1
GCA				55	12
Nikon					70
Total	61	75	99+	81	82+

Table 1: Evolution of Market Shares in Photolitographic Equipment Industry by type of technology. S&R is Step and Repeat first and second generations. Source: Rebecca Henderson (1993).

Second, firm growth is fast as the new innovation takes place and slows down as it is exploited. Figure 14 presents evidence on firm sizes over time for a collection of Fortune 500 high-tech firms in the US (see also Luttmer, 1997). The data on the natural logarithm of the total number of employees in these firms comes from Compustat and includes all employees (domestic and foreign). The period depicted is governed by data availability and the initial period in which these firms became public. It is clear from the graph that the growth rate of these firms is decreasing over time.

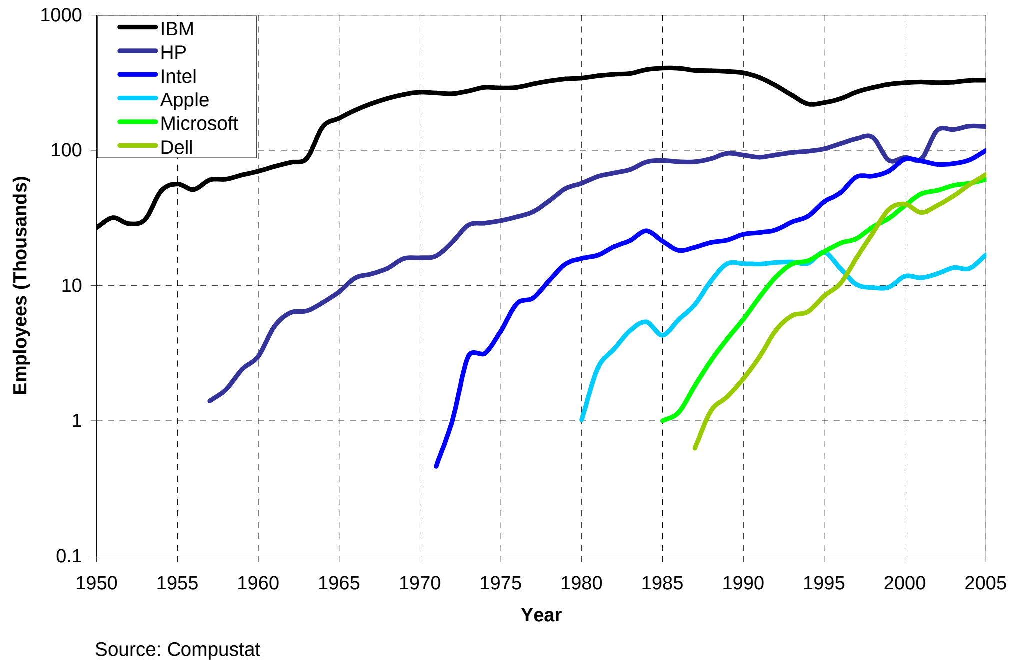

Figure 13: Evolution of the Numer of Employees of 6 Large High-Tech Firms

A third fact consistent with our theory is the existence of long run technological cycles. The existence of such cycles (longer than business cycles) in the data has been long documented (e.g. Schumpeter (1939)). In what Granger (1966) called the 'typical spectral shape', the data shows 'technological' (Kondratieff (40-60 years) and Kuznets (20 to 30 years)) cycles. This same patterns extend to our times.[11] The model also predicts that the technological cycle of output per hour should coincide with the cycle of organizational size. This seems to be the case for the manufacturing sector if we use average firm size data to proxy for the size of organizations as Figure 14, which uses data from the Historical Statistics of the United States, illustrates.[12]

[11] In a previous version of the paper, we used GDP per hour in the non-farm private sector from 1889 to 2000, from the Historical Statistics of the United States, Millennial Edition, and calculated the linear trend of log wages, as well as the Hodrick-Prescott (HP) filter with parameter 100 to remove the very low frequency component. The difference between both served to compute the long run cycles in the data. We identified three long term technological cycles. A large one at the turn of the century. A relatively small one in the middle of the century and a very large one in the second half. The one in the second half is related to what has been called the productivity slowdown. In the last part of the 90's we seem to be moving to a new technological cycle.

[12] We calculate a 'technology' cycle as the difference between the linear and HP (100) trends for a output per hour and as the HP (100) trend for firm size (since the model implies no long term growth in the size of organizations).

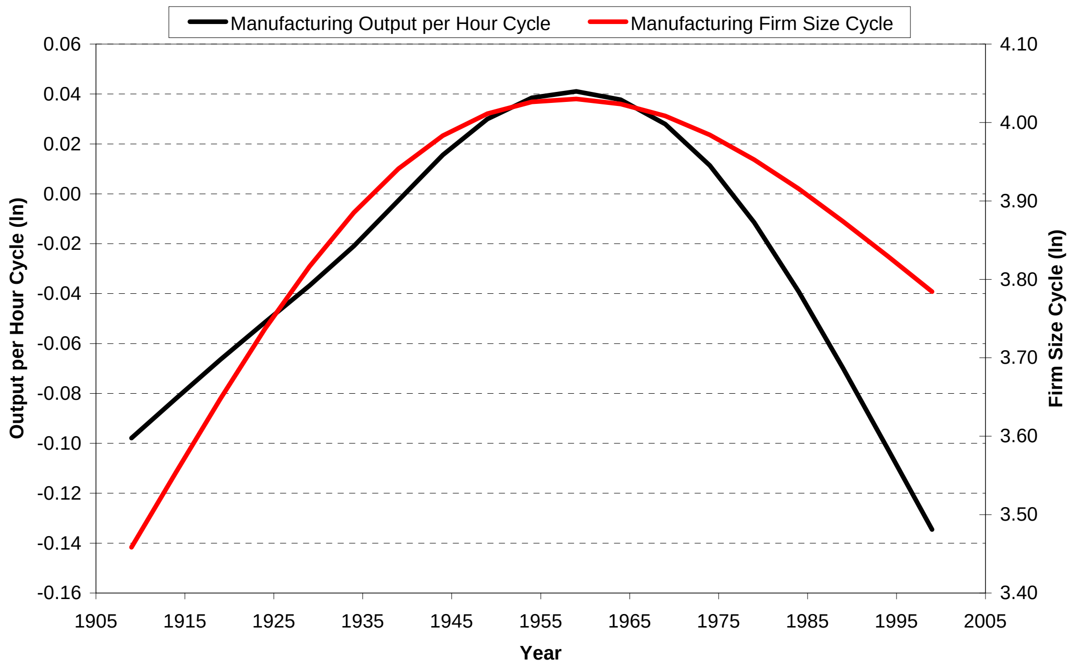

Figure 14: Manufacturing Firm Size and Output per Hour Cycles

5. CONCLUSIONS

Change is a fundamental aspect of growth. Growth is not a smooth continuous process of accumulation. As the 'new growth literature' has recognized, it involves the creation and destruction of products and, as we underscore, of organizations. Each new idea requires a particular type of knowledge, and new organizations —with experts in the relevant sets of knowledge— gradually emerge to exploit this knowledge. When technology is revolutionized, knowledge becomes obsolete and so do the organizations that have developed to allocate it and exploit the existing technology. Our paper provides a new framework to think about this process. We distinguish between the accumulation of knowledge about how to use the existing technology more productively, which we term, following Schumpeter 'exploitation' and the development of new technologies through innovation. Exploitation of existing technologies takes place within existing organizations. As organizations grow by adding more hierarchical layers, they steadily improve the efficiency of the allocation of agents in production. Radical innovation makes existing knowledge, and organizations, obsolete. Decreases in communication costs that allow for the establishment of deeper hierarchies permit more exploitation of existing knowledge and may increase the rate of innovation. However, by making exploitation more attractive, they may also, if agents don't value the future enough, decrease the rate of growth, even to the extent of stopping the growth process altogether.

We view our analysis as the start of an effort to understand, at a deeper microeconomic level, the use of the labor input usually introduced in aggregate production functions. What matters for development is not how many units of labor are used, but how these units are organized, and how this changes over time. The dynamics in our theory are due to the difficulty of building up organizations and of acquiring the relevant pieces of complementary knowledge. Or, in other words, the dynamics are the result of the difficulty of forming markets so that agents can sell their specialize knowledge and buy the knowledge of others. We believe that, in a world where the sources of growth are the creativity and the ideas of individuals, rather than raw materials and capital, understanding the way individuals organize to produce is fundamental to our understanding of the observed income differences across countries.

An interesting avenue for future research concerns the dynamics of switching to entirely new technologies. In our analysis, given that agents are homogeneous, all agents switch at the same time. An analysis with heterogeneous agents would bring about deeper, game theoretical, considerations to the switching decision that may be of interest. Agents need to form expectations on when and who will switch to new technologies; such expectations will affect their own investment paths and the moment of their own switching. We believe our framework is simple and flexible enough to permit an analysis of such issues.

Our analysis studies a one good economy, and thus when a radical innovation is introduced all the existing knowledge, and the existing organization, is made obsolete. Clearly, this is an extreme conclusion. While it is quite reasonable that the development of the automobiles wiped out the stagecoach industry, it is clearly not the case, in a multi-good economy, that all existing firms disappear. Thus another avenue for future research may generalize the model to a world with differentiated commodities, which would yield a smoother prediction.

Finally, our theory has a range of empirical implications that may be tested. First, our model predicts that long term cycles in output should be related to long term cycles in organization, and specifically (to the extent organizations and firms coincide) in average firm size. When a new technology appears, we should see fast productivity growth and a drop in average firm size. As the technology is exploited, firm size should grow and productivity should slow down. Second, our theory has implications for the impact of ICT on the growth rate and shape of organizations. The extent to which different economies exploit available technologies by organizing in complex organizations is mediated by communication technology, the cost of acquiring knowledge, and the distribution of problems faced in production. Any change in these parameters will change the number of periods τ that a technology is used, the average (over time) size of these organizations, as well as the output level and growth rate in the economy. We leave for future research the investigation of the relationship between information technology, organization and growth.

REFERENCES

[1] Acemoglu, Daron and Robinson, James, 2005a. "Institutions as the Fundamental Cause of Long Run Growth,", *Handbook of Economic Growth.* North Holland.

[2] Acemoglu, Daron and Robinson, James, 2005b. "Unbundling Institutions," *Journal of Political Economy*, 113: 5, 949-995.

[3] Aghion, Philippe and Howitt, Peter, 1992. "A Model of Growth Through Creative Destruction," *Econometrica*, 60:2, 23-351.

[4] Arrow, Kenneth J., 1974. *The Limits of Organization.* Norton, New York.

[5] Chari, V V and Hopenhayn, Hugo, 1991. "Vintage Human Capital, Growth, and the Diffusion of New Technology," *Journal of Political Economy*, 99:6, 1142-65.

[6] Comin, Diego and Bart Hobijn, 2007. "Implementing Technology," NBER Working Papers 12886.

[7] Garicano, Luis, 2000. "Hierarchies and the Organization of Knowledge in Production," *Journal of Political Economy*, 108:5, 874-904.

[8] Garicano, Luis and Rossi-Hansberg, Esteban, 2006. "Organization and Inequality in a Knowledge Economy," *Quarterly Journal of Economics*, 121:4, 1383-1435.

[9] Granger, Clive W.J. "The Typical Spectral Shape of an Economic Variable," *Econometrica,* 1966: ,150-161,

[10] Grossman, Gene M. and Helpman, Elhanan, 1991. "Quality Ladders in the Theory of Growth," *Review of Economic Studies*, 59:1, 43-61.

[11] Howitt, Peter, 1999. "Steady Endogenous Growth with Population and R & D Inputs Growing," *Journal of Political Economy*, 107:4, 715-730.

[12] Jones, Charles I., 1995. "R & D-Based Models of Economic Growth," *Journal of Political Economy*, 103:4, 759-84.

[13] Jovanovic, Boyan and Nyarko, Yaw, 1996. "Learning by Doing and the Choice of Technology," *Econometrica*, 64:6, 1299-1310.

[14] Jovanovic, Boyan and Rob, Rafael, 1990. "Long Waves and Short Waves: Growth through Intensive and Extensive Search," *Econometrica*, 58:6, 1391-1409.

[15] Jovanovic, Boyan and Rob, Rafael, 1989. "The Growth and Diffusion of Knowledge," *Review of Economic Studies*, 56:4, 569-82.

[16] Kortum, Samuel S, 1997. "Research, Patenting, and Technological Change," *Econometrica*, 65:6, 1389-1419.

[17] Krusell, Per and Ríos-Rull, José-Víctor, 1996. "Vested Interests in a Positive Theory of Stagnation and Growth," *Review of Economic Studies*, 63:2, 301-329.

[18] Kuznets, Simon., 1940. "Schumpeter's Business Cycles," *American Economic Review*, 30:2, 257-271.

[19] Kydland, Finn E. and Prescott, Edward C. 1982. "Time to Build and Aggregate Fluctuations," *Econometrica*, 50:6, 1345-70.

[20] Legros, Patrick Legros, Andrew F. Newman and Eugenio Proto, 1997. "Smithian Growth through Creative Organization," mimeo, ECARES.

[21] Lucas, Robert Jr., 1988. "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 22:1, 3-42.

[22] Luttmer, Erzo, G. J., 2007. "New Goods and the Size Distribution of Firms," mimeo, University of Minnesota.

[23] North, Douglass C., 1981. *Structure and Change in Economic History*, New York; W.W. Norton & Co.

[24] Penrose, Edith T., 1959. *The Theory of the Growth of the Firm,* Oxford, Oxford University Press.

[25] Romer, Paul M, 1990. "Endogenous Technological Change," *Journal of Political Economy*, 98:5, S71-102.

[26] Schumpeter, Joseph A., 1939. *Business Cycles: A Theoretical, Historical, and Statistical Analysis of the Capitalist Process*, Vols. I and 11. New York: McGraw-Hill.

[27] Schumpeter, Joseph A., 1942. *Capitalism, Socialism and Democracy*. New York: Harper and Brothers.

[28] Young, Allyn A., 1928. "Increasing Returns and Economic Progress," *The Economic Journal*, 38, 527-542.

[29] Young, Alwyn, 1998. "Growth Without Scale Effects," *Journal of Political Economy,* 106:1, 41-63.