Available online at www.sciencedirect.com

ScienceDirect

Journal of Economic Theory 147 (2012) 623–656

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Organizing growth ☆

Luis Garicano [a], Esteban Rossi-Hansberg [b,*]

[a] *Departments of Management and Economics and CEP, LSE, London, United Kingdom*
[b] *Department of Economics and WWS, Princeton University, Princeton, NJ 08540, USA*

Received 3 September 2009; final version received 18 November 2009; accepted 25 November 2009

Available online 4 January 2010

Abstract

We propose a framework to study the impact of information and communication technology on growth through its impact on organization and innovation. Agents accumulate knowledge to use available technologies and invent new ones. The use of a technology requires the development of organizations to coordinate the work of experts, which takes time. We find that while advances in information technology always increase growth, improvements in communication technology may lead to lower growth and even to stagnation, since the payoff to exploiting available technologies through organizations increases relative to the payoff from developing new innovations.
© 2009 Elsevier Inc. All rights reserved.

JEL classification: D24; L23; O33; O40

Keywords: Organization; Knowledge; Productivity growth; Information and communication technology; Innovations

☆ 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 two anonymous referees, Daron Acemoglu, Philippe Aghion, Tim Besley, Xiao Chen, Boyan Jovanovic, Per Krusell, Bob Lucas, Torsten Persson, Andrea Prat, Nancy Stokey and numerous seminar participants for useful comments, and Lorenzo Caliendo for excellent research assistance. Garicano acknowledges the financial support of the Toulouse Network on Information Technology. Rossi-Hansberg acknowledges the generous support of the Sloan Foundation.

* Corresponding author.
E-mail addresses: luis.garicano@gmail.com (L. Garicano), erossi@princeton.edu (E. Rossi-Hansberg).

0022-0531/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jet.2009.11.007

1. Introduction

Advances in information and communication technology (ICT) have been perhaps the most talked about technological improvement in the last three decades. Economies have invested massive amounts of money in setting up the necessary infrastructure to access and use these new technologies. Naturally, the empirical literature has tried to find a consistent effect of improvements in ICT on productivity growth over time and across countries. Surprisingly, no consensus on the size and magnitude of the effect has emerged, particularly outside of the ICT-producing sectors.[1] In this paper we argue that the effect of ICT on growth is complex and not necessarily positive. To do so, we propose a theory of economic growth through organization that incorporates ICT explicitly.

ICT affects how individuals acquire and use their knowledge of existing technologies to produce goods. 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. The cost of such specialization is that it requires communication and coordination among different individuals. That is, it requires "organization." Information and communication technology affects 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 a new technology. We study how the cost of acquiring and communicating information affects growth through its effect on organization and innovation. We will show that the interaction between ICT and organizations determines the direction (positive or negative) of the impact of ICT on growth.

When a new technology is introduced, agents learn only the most common problems associated with it. As time passes, organizations, in the form of knowledge-based hierarchies [9], 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 [10]. The dynamics of our model result from the time to build these markets. 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 that 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

[1] Pilat et al. [26] conclude that "...the United States and Australia are almost the only OECD countries where there is evidence at the sectoral level that ICT use can strengthen labour productivity and MFP growth. For most other OECD countries, there is little evidence that ICT-using industries are experiencing an improvement in labour productivity growth, let alone any change in MFP growth." A 2004 update arrives at the same conclusion. Jorgenson [14] also concludes that the effect of ICT on productivity growth has been positive, while Basu et al. [3] argue that the same was not true for the UK. Timmer and van Ark [30] and Gust and Marquez [12] argue that the impact of ICT on productivity growth in Europe has been small or nonexistent.

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, since eventually most problems are well known and the knowledge acquired is less and less valuable.

Second, innovation is the result of agents' decisions about 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 economy's ability 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.[2]

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 on exploitation as they first acquire the more productive pieces of knowledge about this technology and then the rarer ones. 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, since agents must choose how much and when to invest in radical innovation. As long as the value of continuing with 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.[3] Hence, inherent in new knowledge is a process of creative destruction [28] whereby adopting a radical innovation makes the existing organization obsolete.[4] That is, we build on the insight of Arrow [2] that organizations are specific to a particular technology.

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. The radical innovation process never gets started. When the current technology is already well exploited, agents do not have any development of the alternative radical technology to build on and prefer to exert their efforts on small advances of the existing technology. The result is stagnation.

[2] Throughout, we assume that knowledge is appropriable. In the case of problem-solving and production knowledge we assume so because communicating it takes time, and individual's 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, since it gives agents incentives to acquire more production knowledge.

[3] Indeed, empirically, new technologies are usually associated with new organizations. For example, associated with the arrival of electricity at the end of the 19th 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.

[4] Previous models of creative destruction, following on the pioneering work of Aghion and Howitt [1] and Grossman and Helpman [11], do not take organizations into account — new products substitute for the old, but organizations play no role.

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.

Better communication technology reduces the time spent by agents communicating with others. This unambiguously increases welfare, but has an ambiguous effect on growth: for intermediate communication costs — and thus for intermediate cost of exploiting a particular technology — organizations may get stuck in an old technology forever. The intuition is quite clear. When communication technology is expensive, organizations do not grow very large and so the amount of 'organizational learning' for a given technology is small and so organization does not give old technologies a considerable advantage over new technologies. Similarly, when communication technology is inexpensive, technologies are extensively exploited via organization and so technological improvements are very valuable. Thus, switching to a new technology is attractive and frequent. For intermediate values of communication costs, the old technology with organization may always be preferred to the new technology without any organization, which prevents agents from switching and leads to stagnation. For this intermediate range of communication costs, large organizations are useful for exploiting technology, and the potential improvement in technology is not valuable enough to induce switching. Hence, making organizations more efficient, by reducing communication costs from high to intermediate levels, shifts the balance of economic activity from investing in new innovations to exploiting better existing innovations, and this may reduce long-term economic growth potentially to zero.

The mechanism that leads to stagnation for intermediate values of communication costs would, we believe, be operative for a broader class of models of organization than ours. Specifically, stagnation of the kind we describe would exist in any model of organization for which improvements in communication technology results in "more specialization," or equivalently in "more organization." While this positive relation between communication and specialization is quite general (e.g., Becker and Murphy [4], or Bolton and Dewatripont [5]), it is not a feature of all theories of organization. For example, in Dessein and Santos [8] improvements in coordination (and communication) technologies may lead to the adoption of more coordination-intensive technologies, which may result in less (rather than more) specialization.

The idea that large organizations are detrimental to radical (as opposed to incremental) innovation, which is the key causal mechanism that may lead to the negative impact of ICT on growth, is supported by two main stylized facts.[5] First, radical innovations often take place outside existing firms. Microsoft, Apple, E-Bay and Google are all recent examples of new and large organizations that started small, have grown, and have replaced the old large organizations. An expression of this phenomenon is the common story in which a firm's founders (in Hewlett Packard'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 in a study by Rebecca Henderson [13], in which she shows that in the photolithographic equipment industry (producing pieces of capital equipment to manufacture solid-state semiconductors), each stage of technolog-

[5] The two following stylized facts concern firms; our model is broader than that, since 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 and outside firms are correlated, the evidence is still illuminating.

Table 1
Evolution of market shares in photolithographic equipment industry by type of technology. S&R is step and repeat first and second generations.

	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	87			
Canon		67	21	9	
P-Elmer			78	10	<1
GCA				55	12
Nikon					70
Total	61	75	99+	81	82+

Source: Rebecca Henderson [13].

ical 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.

Second, firm growth is fast as the new innovation takes place and slows down as it is exploited. Among others, Luttmer [23] documents that firm growth declines as firms become old and large. Hence, large firms delay innovation by exploiting old technologies more efficiently and for longer periods of time.

Our work has several precedents on top of the seminal endogenous growth theories of Lucas [22], Romer [27], Grossman and Helpman [11], and Aghion and Howitt [1].[6] Jovanovic and Rob [17] first 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 [16] present a model in which communication technology also affects growth through the search process, not, again, through organizations. Chari and Hopenhayn [6] propose a vintage human capital model in which older individuals operate in old vintages of technologies; if skilled and unskilled labor are complementary, switching to a new technology is costly and, in contrast to our model, several vintages operate simultaneously.[7] Parente [24] proposes a growth theory in which firms develop expertise through learning-by-doing and endogenously adopt new technologies. The evolution of technology in his model is similar to ours, but his model does not model organizations explicitly and, as a result, is not able to study either the characteristics of organizations through the development process nor the effect of ICT on growth. Finally, in both Krusell and Rios-Rull [19] and Jovanovic and Nyarko [15] economies can stagnate. 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. The latter is closest to our approach, but it is a single-agent model (there is no organization) and ICT plays no role. In contrast, we argue that stagnation can be the result of a

[6] Our theory also builds on the classic work of Penrose [25], 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.

[7] Kredler [18] generalizes Chari and Hopenhayn [6] to many inputs and lifetimes.

lengthy process of building organizations to exploit the current technology and relate stagnation to communication technology.

More recently, Comin and Hobijn [7] present a model in which 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. Again, organization does not play any role in this model, nor does information technology affect how far organization can exploit the new technology. Firms choose only where to start the implementation process; from then on, it is exogenous. Organization does play a role in Legros, Newman and Proto's [21] study where, as 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.

Our work is thus novel in two ways: first, it explicitly models organizational learning — how organizations get built to exploit a given technology; and second it considers information and communication technology as a "meta-technology" that affects economic growth through its impact on the gains from building organizations to exploit new technological innovations. Thus, the key result of the paper, concerning the ambiguous effect of communication on growth depending on the amount of organizational learning that takes place in equilibrium, is novel as well.

We use the technology in Garicano [9] to build organizations, but we go beyond that paper (which analyzes the static problem of an individual hierarchy) in that we study a general equilibrium problem in which earnings and specialization patterns are determined in equilibrium. We have also studied the labor market equilibrium with hierarchies in Garicano and Rossi-Hansberg [10]. Relative to that paper we present a model without heterogeneity but also with multiple-layer hierarchies that we characterize more fully and simplify sufficiently to embed in a dynamic framework. As far as we know, this is the first attempt to provide a dynamic theory of hierarchical organization.

The rest of the paper is organized as follows. In Section 2 we present the model, where we focus first on the creation of organizations given a technology and then proceed to introduce innovation and growth. Section 3 characterizes the equilibrium and analyzes the effect of ICT on growth. Section 4 concludes. Appendix A presents the technical proofs not included in the main text and Appendix B provides some extensions of the model of organization discussed in Section 2.2.

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 only when they are young and they have linear preferences so that 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 at which they 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. Selling and buying this technology will be the only mechanism 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 of transferring or selling 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 [10]). 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.

The key assumption in our model is that specialization requires organization and organizations cannot be constructed instantly. Organizations take time to build (see Kydland and Prescott [20] for a similar argument for the case of physical capital). In order for people to specialize in different types of problems, they must have mechanisms that allow them to know who knows what and how to ask the right questions 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 problems 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 radiologists who specialize in 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 if there are more layers in the organization. The first period a technology is used 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 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 [10], there are many equivalent ways of decentralizing these organizations. First, as here, there can be a market for problems where agents sell and buy problems from each other at a 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 exist 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.

We now turn to the description of the formation of an organization and the use of a technology. Later we study the decisions of agents to drop the current technology and make a radical innovation.

2.2. Organizing a technology

Suppose a new technology $A \geqslant 1$ 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.[8] In order to produce, the problem drawn must be within the workers' knowledge set; if it is not, no output is generated. Knowledge can be acquired at a constant cost $\tilde{c} > 0$, so that acquiring knowledge about problems in $[0, z]$ costs $\tilde{c}z$. Denote the wage of an agent working in layer $\ell \in \{0, 1, \ldots\}$ 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:

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

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 $\tilde{c}z$ is the cost of acquiring the knowledge in an interval of length 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 agents work only 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 them from workers as long as they can then solve some of them and obtain higher earnings. The assumption is that these agents need to first see that valuable problems are discarded in order to take advantage of them next period. 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, since they do not need to spend time generating the problems, only communicating with the seller. Organization makes learning unusual problems potentially optimal, since 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

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

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

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 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

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

Namely, they buy $1/h$ problems at price r_0 and solve a fraction $(F(z_1^0+z_1^1)-F(z_1^0))/(1-F(z_1^0))$ 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 $1-F(z_1^0+z_1^1)$ of problems. Hence, some agents enter as layer-two managers to buy these problems from layer-one problem solvers. 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

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

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

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

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 $(1-F(Z_L^{\ell-1}+z))/(1-F(Z_L^{\ell-1}))$) at price r_L^ℓ. Problem solvers in the highest layer L cannot sell their problems, since there are no buyers, so their earnings are just given by

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

In what follows we will use an exponential distribution of problems. This will allow us to simplify the problem 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

$$\begin{aligned} w_L^0 &= \max_z \bigl(A - e^{-\lambda z}\bigl(A - r_L^0\bigr)\bigr) - \tilde{c}z, \\ w_L^\ell &= \max_z \frac{1}{h}\bigl(\bigl(A - r_L^{\ell-1}\bigr) - e^{-\lambda z}\bigl(A - r_L^\ell\bigr)\bigr) - \tilde{c}z \quad \text{for } 0 < \ell < L, \\ w_L^L &= \max_z \frac{1}{h}\bigl(\bigl(A - r_L^{L-1}\bigr) - e^{-\lambda z}A\bigr) - \tilde{c}z. \end{aligned} \tag{2}$$

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 problem imply that

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

Note that the knowledge acquired is increasing in A and decreasing in $\tilde{c}$, h (for problem solvers) and the price obtained for selling problems. The intuition for the effect of A and $\tilde{c}$ 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 fewer 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 of 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 Eqs. (3) and

$$w_L^\ell = w_L^{\ell+1} \equiv \tilde{w}(A, L) \quad \text{for all } \ell = 0, \ldots, L-1. \tag{4}$$

This condition is equivalent to an equilibrium condition requiring that the supply and demand of problems at every layer equalize at the equilibrium prices $\{r_L^\ell\}_{\ell=0}^{L-1}$. 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

$$n_L^\ell = h\bigl(1 - F\bigl(Z_L^{\ell-1}\bigr)\bigr)n_L^0 = he^{-\lambda Z_L^{\ell-1}}n_L^0$$

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

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

So given t, A, and L an equilibrium for one generation of agents is a collection of L prices $\{r_L^\ell\}_{\ell=0}^{L-1}$ and $L+1$ knowledge levels $\{z_L^\ell\}_{\ell=0}^{L}$ that solve the $2L+1$ equations in (3) and (4). 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

$$z_\infty^0 = -\frac{1}{\lambda}\ln\frac{\tilde{c}}{\lambda(A-r_\infty)},$$
$$z_\infty^\ell = -\frac{1}{\lambda}\ln\frac{\tilde{c}h}{\lambda(A-r_\infty)} \quad \text{for all } \ell > 0.$$

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

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

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

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

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

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

and the knowledge acquired by agents is given by

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

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, agents 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 $L-2$, since they can sell them only for a low price to agents in layer L. This logic pertains to all layers. The more layers on top of an agent, the more valuable the problem, since 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, since 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 those of 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 workers can span their knowledge over only one problem instead of $1/h$ of them (since they use their time to produce). Of course, as $L \to \infty$ there is no final layer and so 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. $A \geqslant 1$, $h < 1$ and A, λ, $\tilde{c}$ and h satisfy

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

Proposition 2. *Under Condition* 1*, for any A and L finite, there exists a unique equilibrium determined by a set of prices $\{r_L^\ell\}_{\ell=0}^{L-1}$ and a set of knowledge levels $\{z_L^\ell\}_{\ell=0}^{L}$ such that $r_L^\ell > 0$ is strictly decreasing in ℓ and $z_L^\ell > 0$ is strictly increasing in ℓ.*

Proof. See Appendix A. □

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. The next proposition shows that as the number of layers increases, so do wages (or output per capita if knowledge costs 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 fewer problems, since 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 imply less knowledge acquisition, since 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 increases, w_t increases and $\lim_{L\to\infty} \tilde{w}(A, L) = w_\infty$. Furthermore, as the number of layers L increases, prices r_L^ℓ increase for all $\ell = 0, \dots, L-1$ and knowledge levels z_L^ℓ decrease for all $\ell = 0, \dots, L$. As $L \to \infty$, $r_L^\ell \to r_\infty$ for all $\ell = 0, \dots, L-1$ and $z_L^\ell \to z_\infty^0$ all $\ell = 0, \dots, L$.*

Proof. See Appendix A. □

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 the evolution over time of organizations with a given technology A in a growth model in which agents will have the choice of switching 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 agents' ability to organize.

Figs. 1 and 2 illustrate the results proven in the previous propositions. Fig. 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_∞. Fig. 2 shows the evolution of the size of a typical hierarchy. The top layer size is normalized to one agent. Clearly, as the economy adds more markets for expertise, or layers, the lower layers expand more than proportionally.

Fig. 3 illustrates the problem for prices in an economy with five layers and an economy with seven 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 problems is also decreasing as we move up the hierarchy since problems are more and more selected. Fig. 4 presents the knowledge acquisition of agents for the same two exercises. The picture is the reverse image of the prices in Fig. 3, since higher prices imply less knowledge as selling the problems becomes more attractive. So, given the layer, the smaller the maximum number of layers, the higher the acquisition of knowledge.

Another way of illustrating the implications of our model for knowledge acquisition given the level of technology A is presented in Fig. 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 Fig. 1.[9] Note also that the knowledge of the top problem solvers in hierarchies with two or more layers is constant. The figure also illustrates how, given the layer or occupation, the knowledge acquired by agents working in that layer decreases with time.

This model of organization maintains two important assumptions. First, knowledge is not storable. Each generation of agents has to acquire anew the knowledge of how to solve the problems related to the technology in use. Second, learning is not cumulative. An agent can learn how to solve rare problems without first learning the most frequent ones. In Appendix B we show that all our main results remain unchanged if we relax these two assumptions.

Before we end this section we discuss 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 managers at the highest layer are constant. In contrast, the distribution of

[9] It is easy to develop a model of endogenous growth in which the linear knowledge accumulation observed in Fig. 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).

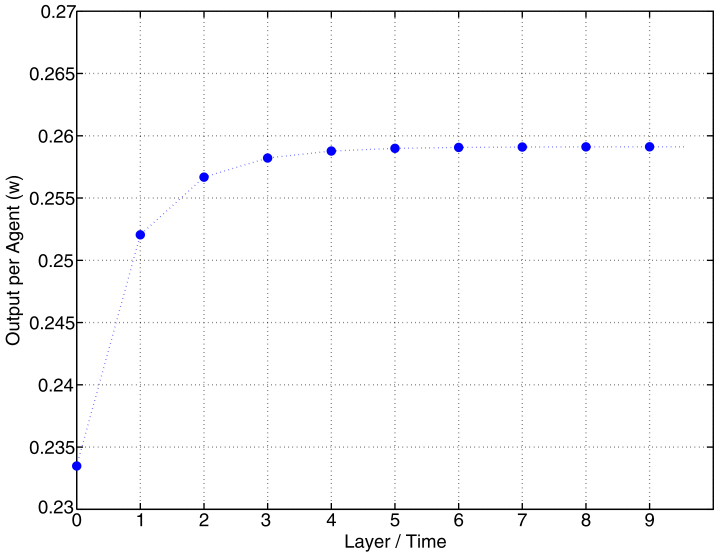

Fig. 1.

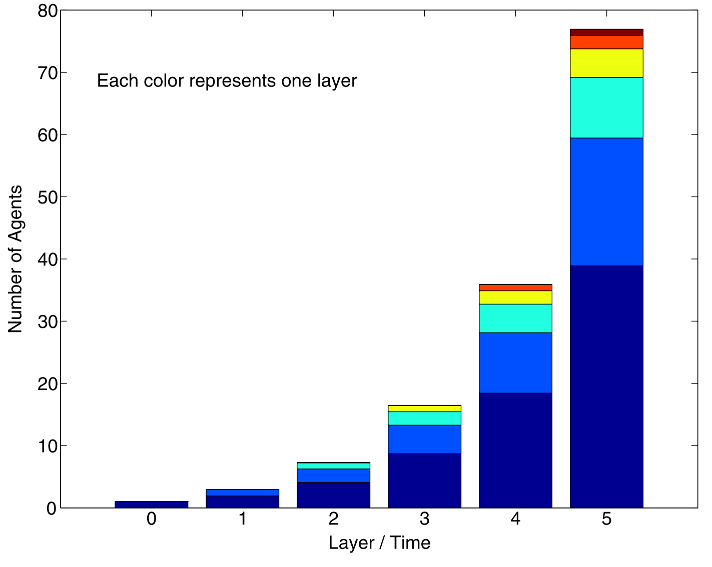

Fig. 2. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

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.

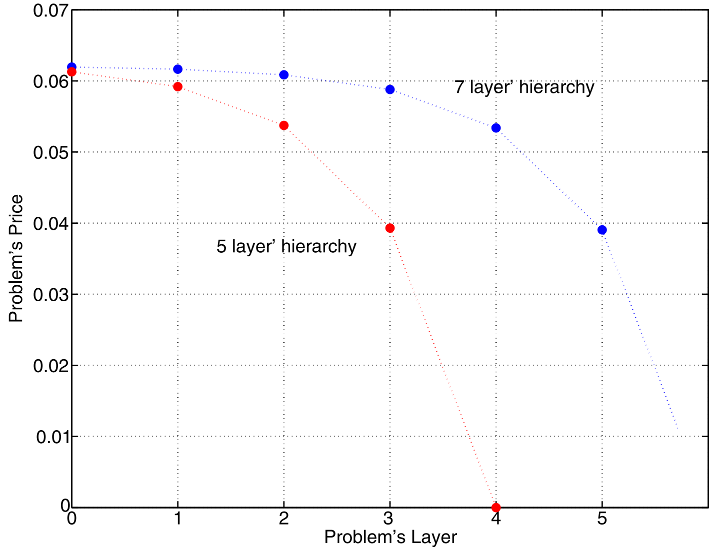

Fig. 3.

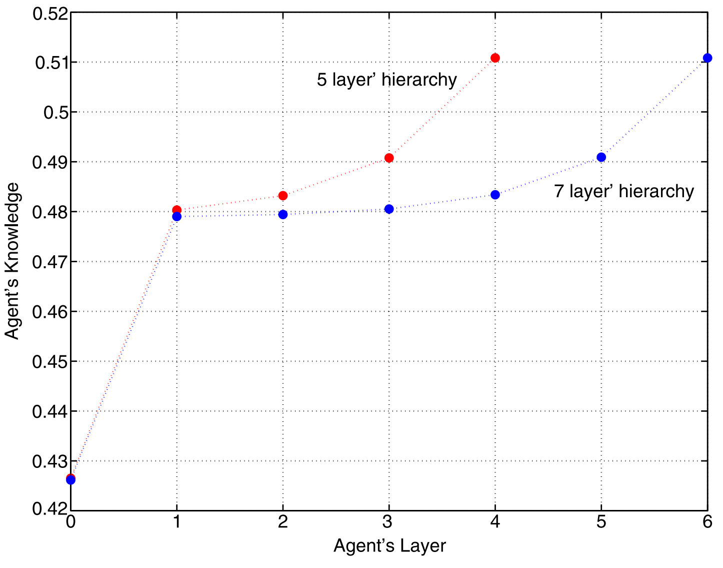

Fig. 4.

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, in the development of this problem-solving knowledge. We now turn to the evolution of technology A that we kept fixed in the previous

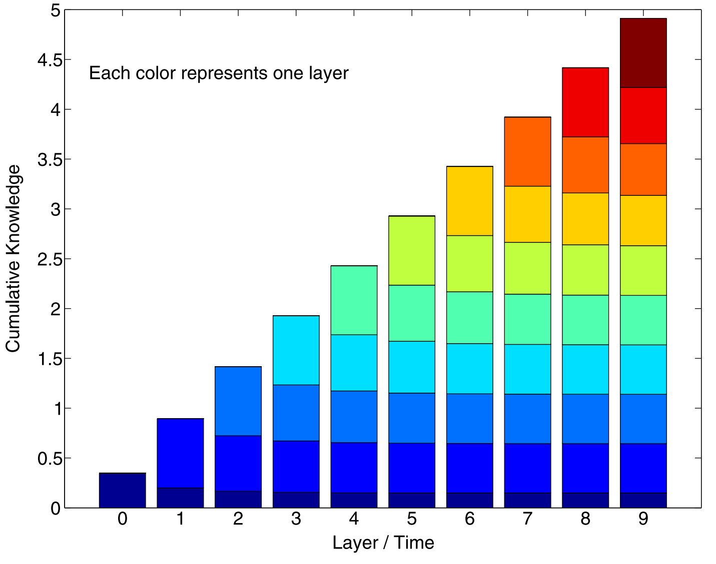

Fig. 5. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

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 agents' only savings mechanism). However, a technology has to be used in order to be productive and organizations have to be developed to exploit it. Naturally, only agents that use the same technology can work with each other.

The previous section studied the earnings that all agents in the economy, working with a particular technology A, receive in equilibrium, $\tilde{w}(A, L)$. These earnings changed every period as new layers of expertise developed in the economy to exploit this technology (L changed). $\tilde{w}(A, L)$ 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 forgone income, increases with the level of technology — 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 $\tilde{c} = cA$. 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 (which is not necessarily the best one). All of the analysis in the previous section remains unchanged apart from the parameter $\tilde{c}$ now becoming Ac.[10] Then, it is easy to see from Eq. (2) and the equilibrium conditions in (4) 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.

[10] 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.

An agent with technology A faces a cost of $A\psi\zeta^2$ of improving the best technology he can use to $(1+\zeta)A$. Namely, we are assuming that there are quadratic adjustment costs 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 it. The role of the quadratic adjustment 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.

Let $\Upsilon(A, A', L)$ denote the lifetime (he lives two periods) utility of an agent born in an economy using technology A, with a frontier technology A', and maximum layer L. Such an agent will buy at price $\tilde{V}(A, A', L)$ a frontier technology (which allows him to produce with any technology below that) when he is 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 set up the problem implies that the only way in which this agent can save is via selling and buying technology. 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). Then

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

That is, an agent can work using the current technology A with organizations of $L+1$ layers (remember that there is a layer zero) and receive wage $\tilde{w}(A, L)$, invest ζ at cost $A\psi\zeta^2$, pay the value of the current technology $\tilde{V}(A, A', L)$, and tomorrow he can sell the technology with a new frontier technology $A'' = A'(1+\zeta)$ and organizations with $L+2$ layers at price $\beta\tilde{V}(A, A'', L+1)$. 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, invests ζ at cost $A'\psi\zeta^2$, pays the cost of the current technology $\tilde{V}(A, A', L)$ and tomorrow sells the technology at price $\tilde{V}(A', A'', 1)$.

Subtracting $\Upsilon(A, A', L)$ and adding $\tilde{V}(A, A', L)$ from both sides of the functional equation above, we obtain a recursive equation that the value of technology has to satisfy in equilibrium, namely,

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

Note from this equation that the decisions of how much to innovate and when to switch technology are independent of $\Upsilon(A, A', L)$. Since agents order consumption using a linear utility function with discount factor β, 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 agents that live two periods and work when they are young is determined only up to consumption transfers between generations. The actual consumption levels every period will, however, depend on the level of prices at which individuals can sell their technology. If, for example, we let $\Upsilon(A, A', L) = 0$ then the prices of technology are as high as possible in equilibrium and all consumption is assigned to the first generation that sells a technology at $\tilde{V}(1, 1, 0)$. If, in contrast, $\Upsilon(A, A', L) = uA$ for some constant u, future generations obtain a level of utility proportional to the technology in use and so the level of prices of the current technology is zero. Given our assumption of linear utility, these distributional concerns across generations are irrelevant to determining choices in this economy, and therefore growth, the length of cycles, or the organization of production, and so we just let $\Upsilon(A, A', L) = 0$ in what follows. $\tilde{V}(A, A', L)$ can also be interpreted as the total present value of a technology and, as such, as the measure of welfare in this economy.[11]

Using $\Upsilon(A, A', L) = 0$, $\tilde{V}(A, A', L)$ solves

$$\tilde{V}(A, A', L) = \left[\max\left[\begin{array}{c}\max_\zeta \tilde{w}(A, L) + \beta\tilde{V}(A, A'', L+1) - A\psi\zeta^2, \\ \tilde{V}(A', A', 0)\end{array}\right]\right] \tag{5}$$

where

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

Note that we are assuming that everyone in this economy has the same A and A', and therefore, since agents are identical, they choose the same patterns of innovation. 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.[12] Of course, if we allow the economy to start with agents that are heterogeneous in their level of technology, there may be other nonsymmetric 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.

We can use Eq. (5) to understand the role that appropriability of innovation plays in this economy. By appropriability of innovation we refer to the fraction of the value of innovation that agents receive from future generations, namely, $\beta\tilde{V}(A, A'', L+1) - \tilde{V}(A, A', L)$. Hence, the level of appropriability is governed by the discount factor β. For example, suppose that appropriability is low, 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 the appropriability 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. So this parameter plays a double role: as a discount factor and as the level of appropriability in an economy. The higher β, the higher the appropriability and the more incentives agents have to invest in innovating the current technology.

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

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

where

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

Note that $G = A'/A$ now 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.

[11] In this case, $\tilde{V}(A, A', L)$, the price of technology, can also be interpreted as the utility of an infinitely lived agent starting in state (A, A', L).

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

Now remember that as $L \to \infty$, $w(L) \to w_\infty/A = 1 - \frac{c}{\lambda} - \frac{c}{\lambda}\ln(\frac{\lambda h}{c} - h\ln h)$. 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 we need to consider. First is the case where there exists a unique finite L^* such that next period a new technology is put in place, namely, $V(G, L^*+1) = GV(1,0)$. This is the case in which the value of innovation converges to zero more slowly than the value of an extra layer. Second is the 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 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 exists 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 price function as

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

and so

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

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

$$V^* \equiv V(1,0) = \max_{L,\{\zeta(\ell)\}_{\ell=0}^{L}} \frac{\sum_{\ell=0}^{L} \beta^{\ell}(w(\ell) - \psi\zeta(\ell)^2)}{1 - \beta^{L+1}\prod_{\ell=0}^{L}(1+\zeta(\ell))}, \tag{6}$$

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 restriction is that innovation costs ψ are such that $\psi > (\beta/(1-\beta))^2(w_\infty/A)$. It is immediate from Eq. (6) and the envelope theorem that a higher β (an economy in which technology is more appropriable) implies a higher V^*, namely, a higher value of the technology and welfare. As we will discuss in the next section, a higher β also implies a smaller set of parameter values for which the economy stagnates.

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

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

so

$$V^*\beta^{L^*+1}\prod_{\ell=0}^{L^*}\big(1+\zeta^*(\ell)\big) = \beta^l\psi 2\zeta^*(l)\big(1+\zeta^*(l)\big) \quad \text{all } \ell. \tag{7}$$

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

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

and β^l decreases with l since $\beta < 1$,

$$\zeta^*(l) \leqslant \zeta^*(l') \quad \text{all } l < l',$$

with equality when $\zeta^*(l) = 0$ for all l in the case where $L^* = \infty$ (stagnation). 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 invest only in the last period before the switch to a new technology. Note also that the above equation implies that $\zeta(0) > 0$ if L^* is finite. That is, the economy invests 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 $\zeta^*(l) = 0$. Note that even in this case the solution to the innovation 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

$$\frac{\sum_{\ell=0}^{L} \beta^\ell (w(\ell) - \psi \zeta_L^*(\ell)^2)}{1 - \beta^{L+1} \prod_{\ell=0}^{L} (1 + \zeta_L^*(\ell))} - \frac{\sum_{\ell=0}^{L-1} \beta^\ell (w(\ell) - \psi \zeta_{L-1}^*(\ell)^2)}{1 - \beta^{L} \prod_{\ell=0}^{L-1} (1 + \zeta_{L-1}^*(\ell))} > 0 \tag{8}$$

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 effects. First, as we increase the number of periods in which 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 $w(\ell)$ is monotone in ℓ we know that the difference in (8) is monotone too, which implies that there is a unique L^* as we conjectured above (see the upper-left panel of Fig. 6). Note, however, that this difference being monotone does not rule out the possibility that $L^* = \infty$. This is the case in which 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 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. These 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 exhibits 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.*

Fig. 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

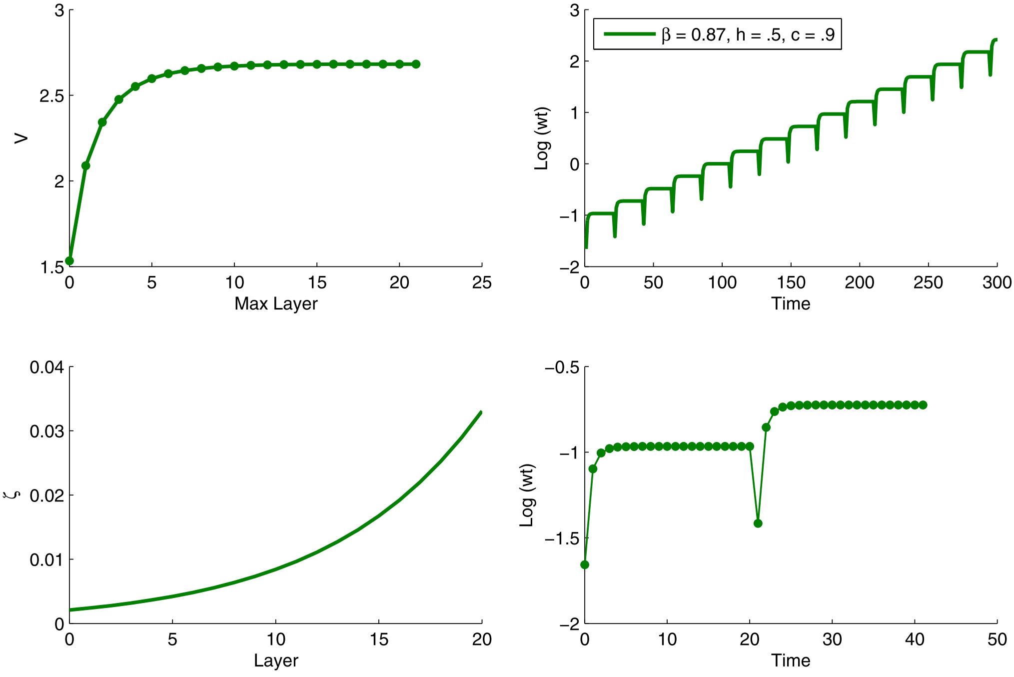

Fig. 6.

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 price increases the most when we lengthen the technology cycles from zero to one period, since developing the first organizational layer implies the largest gains. The lower-left panel shows the investment in innovation during the 21 periods in which 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 in the future, so agents invest little in innovation. As the time of the switch approaches, agents invest progressively more. As Eq. (7) reveals, this path is driven essentially by the discount factor β. The higher β the more even is investment over the cycle. Agents invest a positive amount each period so the frontier technology is constantly improving.

The right column of Fig. 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 presents 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 is clear from the figure, 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 no longer used). This implies a short-term reduction in output per capita, although, of course, the economy wins in the long run, since the timing of the switch is optimal.

Similarly, worker–manager wage inequality moves in cycles. As organizations develop within a technology, inequality between workers and managers increases, as we argued in Section 2.2, since more organization means workers specialize more in production and substitute the knowledge of experts for their own. However, when the technological switch takes place, inequality drops discretely and then starts increasing again.[13]

The second type of equilibrium, the one with stagnation, starts similarly to the equilibrium in Fig. 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 converge to zero.

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 Section 3.2 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 effects 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 (6), 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 (6), 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

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

since $dw(\ell)/dh \leqslant 0$ with equality if $L^* = 0$. The expressions are only approximate, since L is a discrete 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 for the case of h).

Even though the effect of ICT on welfare is to unambiguously increase it, the same is not necessarily true for growth. Reductions in growth can be welfare enhancing because the future is discounted by preferences and imperfect appropriability and agents invest less and consume more early on. However, this is never the case for reductions in the cost of acquiring information, c. As Fig. 7, where we fix $h = .8$, illustrates, a reduction in c leads to shorter cycles and less organization, but faster growth. A reduction in c makes organization less necessary, since

[13] This is in contrast to Thesmar and Thoenig [29], where creative destruction in general increases the skill premium.

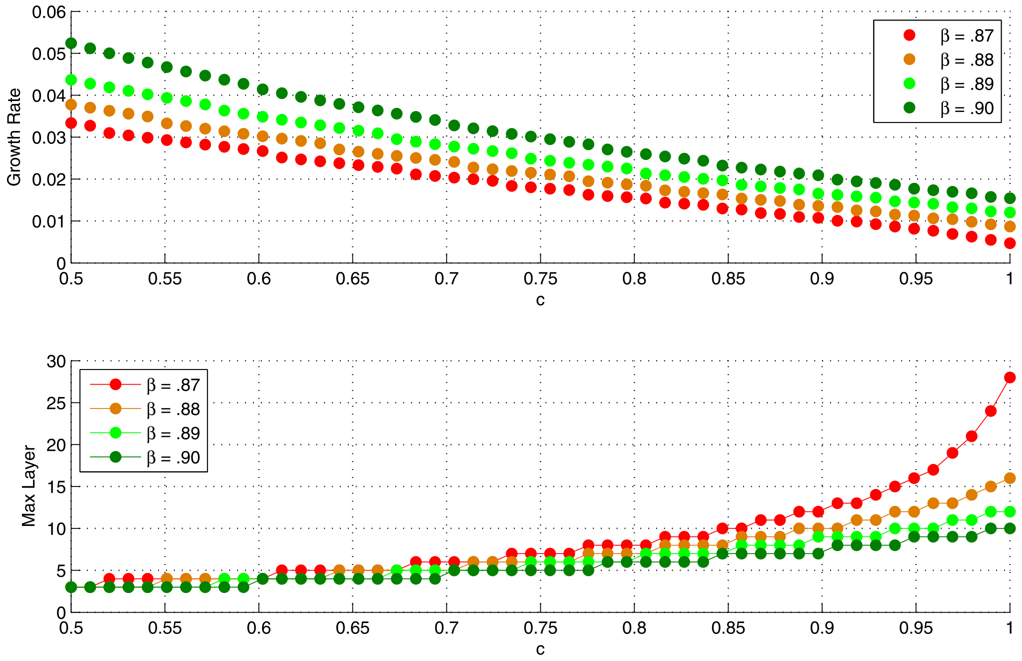

Fig. 7. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

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 gives agents incentives 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 appropriable) the growth rate increases.

Fig. 8 shows two examples of the simulations presented in Fig. 7. We fix $\beta = .87$ and change c from .9 to .7. Fig. 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 fewer periods, since 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.

The effect of changes in communication technology on growth is more complicated. The main reason is that a reduction in h changes output per capita more, the more organized is the technology. That is, the more intensive is the exploitation of technology on communication. For example, when agents are self-employed and there is no organization, changes in h do not affect output per capita.

Fig. 9 illustrates the effect of reductions in h on growth. Note that reductions in h increase the length of technology cycles and the use of organization when communication costs are high. This is intuitive, since 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. Since technology can be exploited more efficiently because building

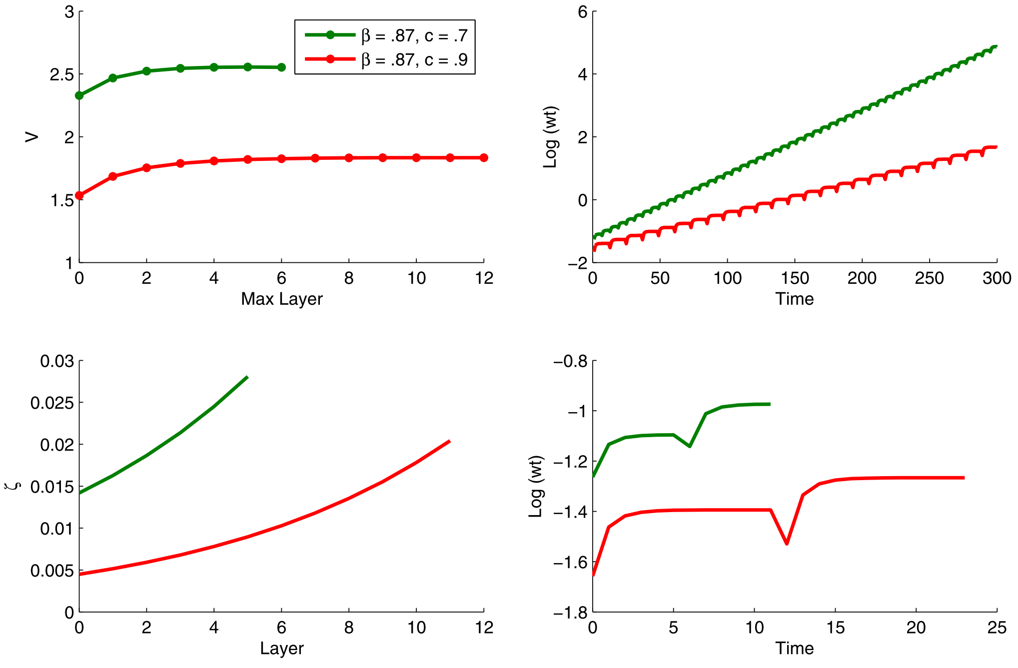

Fig. 8. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

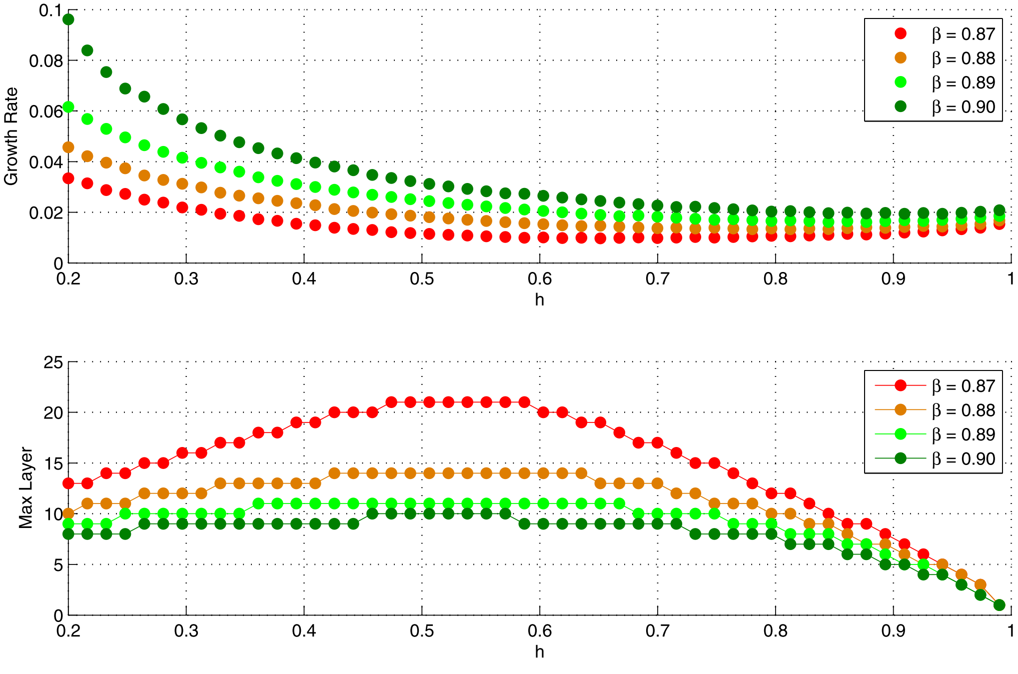

Fig. 9. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

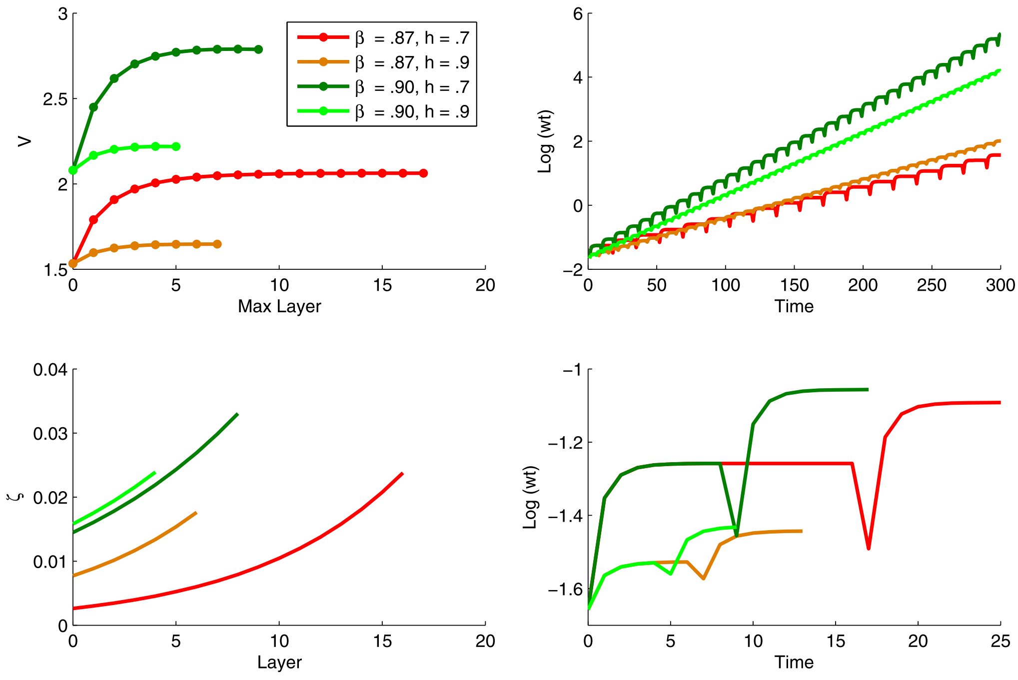

Fig. 10. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

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. However, 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.[14] Fig. 9 shows that reductions in h can reduce the length of the cycles if h is small (we let $c = .9$). This is because when h is small the value of organizing is concentrated more heavily in the first few layers.

Fig. 10 illustrates the point that improvements in communication technology can have negative effects on growth when appropriability is low (β is low). It shows the effect of a reduction in communication costs when appropriability is high and when appropriability 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 appropriability is high.

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, since it makes organization more efficient. This will increase wel-

[14] 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.

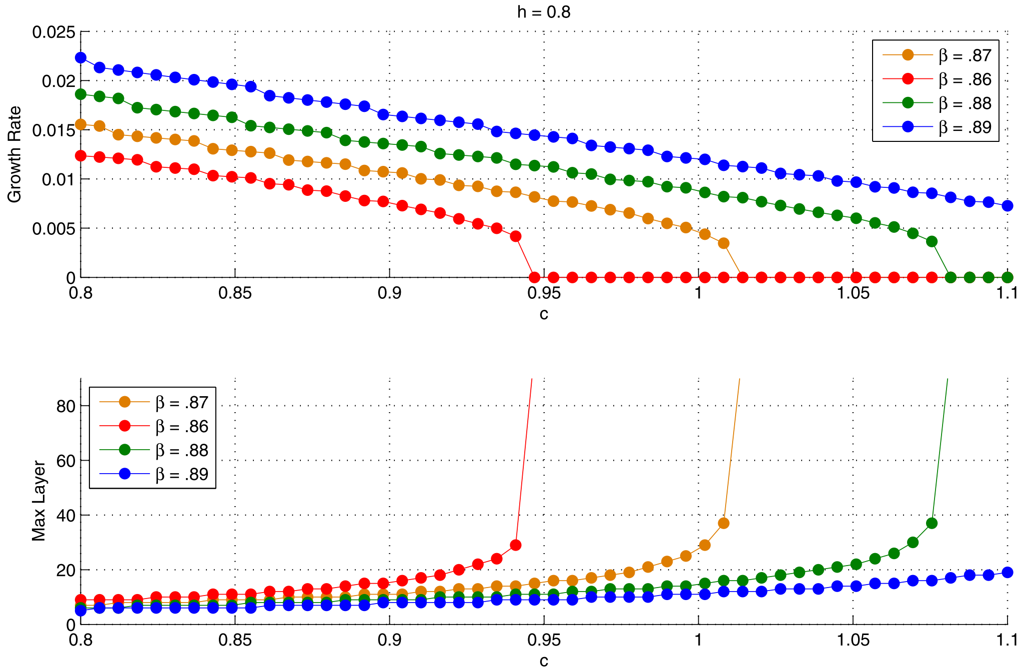

Fig. 11. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

fare but reduce growth if appropriability is low. Hence, what the model tells us is that countries in which technology markets are not well developed and appropriability is low will experience negative effects 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 increase.

3.2. Stagnation

As we discussed in the previous section, there are cases in which the economy 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,

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

by the envelope theorem but the value of never switching, given by $\sum_{\ell=0}^{\infty}\beta^{\ell}w(\ell)$, 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. Fig. 11 shows a graph similar to Fig. 7 but for a range of c's that includes larger values and for smaller values of β. As in Fig. 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 Fig. 11 shows how the number of layers explodes to infinity as we approximate the threshold.[15] 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 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 at 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.

The effect of communication technology on the set of parameters for which we obtain stagnation is again more complicated. The reason is, essentially, the nonmonotone 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, since 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 this first set of expertise markets. For intermediate values of h 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 Fig. 12, which parallels Fig. 8 but includes lower values of the discount factor β. The figure shows how, as we decrease β, the set of values of h for which we obtain stagnation increases. As we have argued, and as the lower panel shows, these are the parameter values for which $L^* = \infty$.

As we have discussed, decreases in h always increase welfare in this economy, independently of whether the economy stagnates. However, h will affect the growth rate dramatically. Communication technology is a technology to exploit today's and future 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 rather 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 the total value of technology from today's perspective. The savings from not investing in innovation today then overwhelm the negative implication that output per

[15] 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.

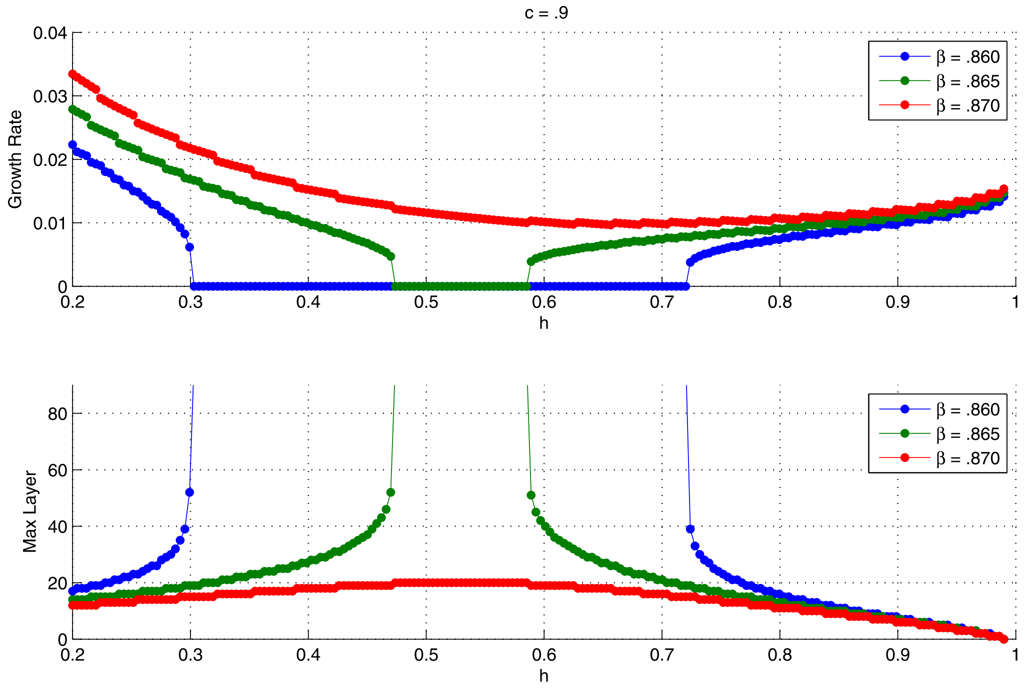

Fig. 12. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

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 this 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. Of course, current agents may be happy with this outcome if the price they paid for technology was low enough.

4. 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 been developed to allocate it and exploit the existing technology. Our paper provides a new framework to think about this process.

Our theory is useful to understand the effect of ICT on productivity growth. 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 do not value the future enough, decrease the rate of growth, even to the extent of stopping the growth process altogether. In contrast, reductions in the cost of acquiring information always increase growth, since organizations become less valuable as knowledge is less expensive. In sum, the effect of ICT on productivity growth depends crucially on whether ICT increases or decreases the value of organization. If it decreases it, we will see organizations with fewer layers, faster innovation, and growth. Otherwise, we will see larger organizations with more layers and slower growth or stagnation.

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 about 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.

This paper studies a one-good economy, and thus when a radical innovation is introduced, all of the existing knowledge, and the existing organization, is made obsolete. Clearly, this is an extreme conclusion. While it is quite reasonable to think that the development of the automobile 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 be to generalize the model to a world with differentiated commodities, which would yield a smoother prediction.

To conclude, 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. We are convinced that, in a world in which the sources of growth are the creativity and the ideas of individuals, understanding the way individuals organize to produce is fundamental to our understanding of the observed income differences across countries.

Appendix A

A.1. Proof of Proposition 2

Use (3) to obtain the knowledge of each agent as a function of the price the agent receives for a problem passed. Letting $\alpha \equiv \frac{\tilde{c}h}{\lambda}$ and $\beta \equiv A - \frac{\tilde{c}h}{\lambda}$ and using (2) we obtain the following recursion for the set of prices:

$$r_L^{L-1} = \beta - h w_L^L + \alpha \ln \frac{\alpha}{A},$$
$$r_L^{\ell-1} = \beta - h w_L^\ell + \alpha \ln \frac{\alpha}{(A - r_L^\ell)} \quad \text{for } 0 < \ell < L.$$

Imposing (4) for $\ell = 1, \dots, L-1$ we obtain that

$$r_L^{L-1} = \beta - h\tilde{w}(A, L) + \alpha \ln \frac{\alpha}{A},$$
$$r_L^{\ell-1} = \beta - h\tilde{w}(A, L) + \alpha \ln \frac{\alpha}{(A - r_L^\ell)} \quad \text{for } 0 < \ell < L. \tag{9}$$

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 $\{r_L^\ell(\tilde{w})\}_{\ell=0}^{L-1}$ is decreasing in ℓ. It is also immediate from (3) that the higher the price, the lower the corresponding knowledge level, so $\{z_L^\ell(\tilde{w})\}_{\ell=0}^L$ 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 $\{z_L^\ell(\tilde{w})\}_{\ell=0}^L$ 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

$$w_L^p\left(r_L^0\right) = A - \frac{\tilde{c}}{\lambda}\left(1 - \ln \frac{\tilde{c}}{\lambda(A - r_L^0)}\right) \tag{10}$$

which is continuous and strictly increasing in r_L^0.

The last equilibrium condition is given by (4) 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 $r_L^0 = A - \tilde{c}/\lambda$, $w_L^p(A - \tilde{c}/\lambda) = A - \tilde{c}/\lambda$ and

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

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

Now let $r_L^0 = 0$. Then

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

and note that

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

since $h < 1$ and $r_L^1 \geqslant 0$. 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. □

A.2. Proof of Proposition 3

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

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

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

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

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, since 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, since 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. Since $r_L^L = 0$ (the last layer throws problems away) but $r_{L+1}^L > 0$ and since for a given w, by Eq. (9), $r_{L+1}^{\ell-1}$ 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 Eq. (10), the price of problems sold by workers, as

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

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 (9) 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. Eq. (3) then implies that $\{z_L^\ell\}_{L=0}^\infty$ converges to z_∞^ℓ from above. □

Appendix B

In this appendix we show how the key characteristics of the organizational problem we developed in the main text are unaltered if we change some of the maintained assumptions on knowledge acquisition and accumulation. We discuss first the case in which knowledge about a technology is storable. We then proceed to discuss the case in which the cost of acquiring knowledge is cumulative.

B.1. Storable knowledge

Consider a modified version of the model presented in Section 2.2 in which we assume that the knowledge learned by the previous generation is now stored and freely available to everyone in the economy. Denote by $\bar{Z}_t$ all the knowledge about technology A stored in the economy up to period t. That is, $\bar{Z}_t = \sum_{L=0}^{t-1} Z_L^L$.

Then, for $t = L$, the earnings of workers will be given by

$$w_L^0 = \max_{z \geqslant 0} AF(\bar{Z}_L + z) + \big(1 - F(\bar{Z}_L + z)\big) r_L^0 - \tilde{c}z,$$

and those of problem solvers by

$$w_L^\ell = \max_{z \geqslant 0} \frac{1}{h} \left(\frac{A(F(\bar{Z}_L + Z_L^{\ell-1} + z) - F(\bar{Z}_L + Z_L^{\ell-1}))}{(1 - F(\bar{Z}_L + Z_L^{\ell-1}))} - \frac{(1 - F(\bar{Z}_L + Z_L^{\ell-1} + z)) r_L^\ell}{(1 - F(\bar{Z}_L + Z_L^{\ell-1}))} r_L^{\ell-1} - \tilde{c}z \right),$$

and

$$w_L^L = \max_{z \geqslant 0} \frac{1}{h} \left(A \frac{F(\bar{Z}_L + Z_L^{L-1} + z) - F(\bar{Z}_L + Z_L^{L-1})}{1 - F(\bar{Z}_L + Z_L^{L-1})} - r_L^{L-1} \right) - \tilde{c}z.$$

Using an exponential F, we obtain

$$w_L^0 = \max_{z \geqslant 0} \big(A - e^{-\lambda(z + \bar{Z}_L)} \big(A - r_L^0\big)\big) - \tilde{c}z,$$

$$w_L^\ell = \max_{z \geqslant 0} \frac{1}{h} \big(\big(A - r_L^{\ell-1}\big) - e^{-\lambda z}\big(A - r_L^\ell\big)\big) - \tilde{c}z \quad \text{for } 0 < \ell < L,$$

$$w_L^L = \max_{z \geqslant 0} \frac{1}{h} \big(\big(A - r_L^{L-1}\big) - e^{-\lambda z} A\big) - \tilde{c}z.$$

Comparing this system of equations with the one in Section 2.2, we see that the only change is in the expression for w_L^0. This implies that the first-order conditions that determine the knowledge of problem solvers, z_L^ℓ for $0 < \ell \leqslant L$, remain unchanged. In contrast, the knowledge of workers is now given by

$$z_L^0 = \max\left\{0, -\frac{1}{\lambda} \ln \frac{\tilde{c}}{\lambda(A - r_L^0)} - \bar{Z}_L \right\}.$$

Note first that the nonnegativity constraint on z_L^0 is now relevant for sufficiently high L, since $\lim_{L \to \infty} Z_L = \infty$. So workers will learn in the first periods after a technology is put in place, and then will stop learning actively, and solve only the problems they know how to solve because of the stored technology. This implies that $\lim_{L \to \infty} w_L^0 = A$ and so the equilibrium wage $\tilde{w}(A, L)$ converges to A from below as L increases. That is, eventually all problems have been solved and stored and everyone in the population uses their time to generate problems and produce A per problem. Hence, the number of problem solvers in upper layers converges to zero and $r_\infty = w_\infty = A$.

Hence, as in the case developed in the main text, the economy converges in levels and so there is no permanent growth absent radical innovations. Given a technological level A, storing knowledge implies that the economy can exploit the technology faster and to a larger extent, but eventually everything about the technology is known, and the economy stagnates. All other comparative statics, including the ones associated with $\tilde{c}$ and h, remain qualitatively unchanged.

B.2. Cumulative costs of learning

So far we have assumed that agents can learn any interval of knowledge of length z at a cost $\tilde{c}z$. Consider now the case when learning how to solve a particular problem z involves learning how to solve all the problems up to z, namely, the interval $[0, z]$. Thus in this extension knowledge is cumulative: agents can learn how to solve the less frequent problems only if they learn how to solve the more common ones. Assume again that knowledge is not storable as in the main text, then we can write the system of wages, using an exponential F, as

$$w_L^0 = \max_{z\geqslant 0}\bigl(A - e^{-\lambda z}\bigl(A - r_L^0\bigr)\bigr) - \tilde{c}z,$$

$$w_L^\ell = \max_{z\geqslant 0}\frac{1}{h}\bigl(\bigl(A - r_L^{\ell-1}\bigr) - e^{-\lambda z}\bigl(A - r_L^\ell\bigr)\bigr) - \tilde{c}\bigl(Z_L^{\ell-1} + z\bigr) \quad \text{for } 0 < \ell < L,$$

$$w_L^L = \max_{z\geqslant 0}\frac{1}{h}\bigl(\bigl(A - r_L^{L-1}\bigr) - e^{-\lambda z}A\bigr) - \tilde{c}\bigl(Z_L^{L-1} + z\bigr).$$

Clearly, the first-order conditions with respect to the increment in knowledge across layers are identical to the ones in the main text. Two important changes are worth mentioning. First, since the costs of learning are now higher for everyone, wages and the price of problems at all layers will be lower for any given L than for the case developed in the text. Second, since $\lim_{L\to\infty} Z_L^L = \infty$, the rents of adding another layer eventually become negative. So there exists an $\bar{L}$ such that $\tilde{w}(A, \bar{L}) > \tilde{w}(A, L+1)$. Hence, no extra layers beyond $\bar{L}$ form and the economy stagnates absent any radical innovations.

If we combine storable knowledge with cumulative learning cost, given A, the economy converges to a layer $\bar{L}$, as above, but the number of problem solvers decreases as more knowledge is stored and eventually converges to zero. Therefore, in the limit all agents are workers and the economy converges to a level of output per capita equal to A. For our purposes, the key property of all these extensions is that they leave the main characteristics of the evolution of wages over time intact. Since wages converge in levels, the economy stagnates and so, under circumstances similar to the ones outlined in Section 2.3, agents invest in alternative technologies and switch to them sporadically. In all these extensions h and c play the same role as in the main text, so all the qualitative results continue to apply.

References

[1] Philippe Aghion, Peter Howitt, A model of growth through creative destruction, Econometrica 60 (2) (1992) 23–351.
[2] Kenneth J. Arrow, The Limits of Organization, Norton, New York, 1974.
[3] Susanto Basu, John G. Fernald, Nicholas Oulton, Sylaja Srinivasan, The case of the missing productivity growth: Or, does information technology explain why productivity accelerated in the US but not the UK?, NBER Working Paper 10010, 2003.
[4] Gary S. Becker, Kevin M. Murphy, The division of labor, coordination costs, and knowledge, Quart. J. Econ. 107 (4) (1992) 1137–1160.

[5] Patrick Bolton, Matthias Dewatripont, The firm as a communication network, Quart. J. Econ. 109 (4) (1994) 809–839.
[6] V.V. Chari, Hugo Hopenhayn, Vintage human capital, growth, and the diffusion of new technology, J. Polit. Economy 99 (6) (1991) 1142–1165.
[7] Diego Comin, Hobijn Bart, Implementing technology, NBER Working Paper 12886, 2007.
[8] Wouter Dessein, Tano Santos, Adaptive organizations, J. Polit. Economy 114 (5) (2006) 956–995.
[9] Luis Garicano, Hierarchies and the organization of knowledge in production, J. Polit. Economy 108 (5) (2000) 874–904.
[10] Luis Garicano, Esteban Rossi-Hansberg, Organization and inequality in a knowledge economy, Quart. J. Econ. 121 (4) (2006) 1383–1435.
[11] Gene M. Grossman, Elhanan Helpman, Quality ladders in the theory of growth, Rev. Econ. Stud. 59 (1) (1991) 43–61.
[12] Christopher Gust, Jaime Marquez, International comparisons of productivity growth: The role of information technology and regulatory practices, Lab. Econ. 11 (1) (2004).
[13] Rebecca Henderson, Underinvestment and incompetence as responses to radical innovation: Evidence from the photolithographic alignment equipment industry, RAND J. Econ. 24 (2) (1993) 248–270.
[14] Dale Jorgenson, Accounting for growth in the information age, in: Philippe Aghion, Steven Durlauf (Eds.), Handbook of Economic Growth, vol. 1A, North-Holland, Amsterdam, 2005, pp. 743–815.
[15] Boyan Jovanovic, Yaw Nyarko, Learning by doing and the choice of technology, Econometrica 64 (6) (1996) 1299–1310.
[16] Boyan Jovanovic, Rafael Rob, The growth and diffusion of knowledge, Rev. Econ. Stud. 56 (4) (1989) 569–582.
[17] Boyan Jovanovic, Rafael Rob, Long waves and short waves: Growth through intensive and extensive search, Econometrica 58 (6) (1990) 1391–1409.
[18] Matthias Kredler, Experience vs. obsolescence: A vintage-human-capital model, MPRA Paper 10200, University Library of Munich, 2008.
[19] Per Krusell, José-Víctor Ríos-Rull, Vested interests in a positive theory of stagnation and growth, Rev. Econ. Stud. 63 (2) (1996) 301–329.
[20] Finn E. Kydland, Edward C. Prescott, Time to build and aggregate fluctuations, Econometrica 50 (6) (1982) 1345–1370.
[21] Patrick Legros, Andrew F. Newman, Eugenio Proto, Smithian growth through creative organization, mimeo, ECARES, 2007.
[22] Robert Lucas Jr., On the mechanics of economic development, J. Monet. Econ. 22 (1) (1988) 3–42.
[23] Erzo G.J. Luttmer, On the mechanics of firm growth, mimeo, University of Minnesota, 2007.
[24] Stephen L. Parente, Technology adoption, learning-by-doing, and economic growth, J. Econ. Theory 63 (2) (1994) 346–369.
[25] Edith T. Penrose, The Theory of the Growth of the Firm, Oxford University Press, Oxford, 1959.
[26] Dirk Pilat, Frank Lee, Bart Van Ark, Production and the use of ICT: A sectoral perspective on productivity growth in the OECD area, Econ. Stud. 35 (2) (2002) 47–78.
[27] Paul M. Romer, Endogenous technological change, J. Polit. Economy 98 (5) (1990) S71–102.
[28] Joseph A. Schumpeter, Capitalism, Socialism and Democracy, Harper and Brothers, New York, 1942.
[29] David Thesmar, Thoenig Mathias, Creative destruction and firm organization choice, Quart. J. Econ. 115 (4) (2000) 1201–1237.
[30] Marcel P. Timmer, Bart van Ark, Does information and communication technology drive EU–US productivity growth differentials?, Oxford Econ. Pap. 57 (2005) 693–716.