DISCUSSION PAPER SERIES

No. 6109

INFORMATION TECHNOLOGY, EFFICIENT RESTRUCTURING AND THE PRODUCTIVITY PUZZLE

Hans Peter Grüner

INDUSTRIAL ORGANIZATION and INTERNATIONAL MACROECONOMICS

Centre for Economic Policy Research

www.cepr.org

Available online at: www.cepr.org/pubs/dps/DP6109.asp

ISSN 0265-8003

INFORMATION TECHNOLOGY, EFFICIENT RESTRUCTURING AND THE PRODUCTIVITY PUZZLE

Hans Peter Grüner, Universität Mannheim and CEPR

Discussion Paper No. 6109
February 2007

Centre for Economic Policy Research
90–98 Goswell Rd, London EC1V 7RR, UK
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 6109

February 2007

ABSTRACT

Information Technology, Efficient Restructuring and the Productivity Puzzle*

Labour productivity in the US has recently grown more strongly than in most European countries. It is often argued that the American productivity increase is due to the widespread introduction of new information and communication technologies (ICT). But why have the same technologies not similarly increased Europe's labour productivity? This paper provides a theoretical explanation for this productivity puzzle based on an extension of Radner's (1992) model of hierarchical information aggregation. The introduction of new ICTs enables organizations to process any given amount of information with a shorter delay. This enables organizations to restructure and solve incentive problems without risking excessive delay. Even a marginal improvement in the ICT can yield significant increases in labour productivity if – and only if – the organization is drastically restructured. Restructuring yields hierarchies with fewer layers and fewer managers, all working under incentive pay and providing first best effort. However, managers need not participate in the gains associated with the restructuring of their business firms.

JEL Classification: D23, D70, D83, L22 and P51
Keywords: hierarchies, ICT, information processing, labour productivity and restructuring

Hans Peter Grüner
Lehrstuhl für Wirtschaftspolitik
Universität Mannheim
68131 Mannheim
GERMANY
Email: hgruener@rumms.uni-mannheim.de

* I thank Peter Funk, Eckhard Janeba, Mical Castanheira, Axel Ockenfels, Elisabeth Schulte and seminar participants at various institutions for their helpful comments and suggestions.

Submitted 15 January 2007

1 Introduction

Over the last decade labor productivity in the US has grown more strongly than in most European countries. Many observers argue that the American productivity increase is due to the widespread introduction of new information and communication technologies[1]. This includes the use of new communication tools such as the internet - and in particular e-mail, mobile phones, mobile computers, wireless internet access and new more comfortable software products such as powerpoint, excel and many others. However, if such innovations are the key explanation for the impressive American productivity increase, why haven't the same technologies increased European labour productivity to a comparable extend? The fact that Europe has not performed as well in terms of productivity as the US despite a similar access to these new technologies is what I would like to call the productivity puzzle.

This paper provides a theoretical explanation for this productivity puzzle, based on a model that analyzes the way in which innovations in information and communication technologies affect the optimal design of organizations. The explanation is based on two observations. First the introduction of new IC-technologies enables organizations to process any given amount of information with a shorter delay. This means that organizations can reduce the division of labor and solve incentive problems without risking to produce with excessive delay. Restructuring generates a higher surplus when the number of employees involved in a particular process is reduced. This is due the fact that individual effort in production can be viewed as a contribution to the provision of a public good. Under a proper incentive contract, the reduction of the number of employees increases equilibrium effort. Therefore, an improved IC-technology enables restructured organization to produce more efficiently at a given delay.

In this paper I show that even a marginal improvement in the IC-technology can yield significant increases in labour productivity if the organization is structured appropriately. Moreover, marginal changes in ICT productivity may result in a significant restructuring of organizations: several hierarchy levels may disappear and the optimal organization hires less employees. An improved IC technology without an appropriate restructuring instead

[1]Stephen D. Oliner and Daniel E. Sichel argue in "The resurgence of growth in the late 1990s: is information technology the story?" that "All in all, we estimate that the use of information technology and the production of computers accounted for about two-thirds of the 1 percentage point step-up in productivity growth between the first and second halves of the decade. Thus, to answer the question posed in the title of this paper, information technology largely is the story." Gordon (2004) argues that the US-EU productivity gap has been largest in the IT- intensive sectors.

only yields marginal productivity gains. Hence, there is a strong complementary between IC technology improvements and a lean organizational design. If the organization avoids restructuring then the productivity increase remains marginal. An explanation for the productivity puzzle is that European companies are more reluctant to restructure their production process in an appropriate manner. If complementarities are not used then the productivity effect of ICT improvements cannot fully be exploited.

In the model a principal hires a number of agents called managers who work on a given number of information items. Each agent has to provide effort in order to properly understand the information that he has supposed to handle. Agents have to process all items in order to have a chance of realizing a high output level. Unobservable additional effort on objects increases the likelihood of the good outcome. First-best effort cannot be implemented because all agents are wealth-constrained. The overall success probability of the organization is increasing in the aggregate amount of effort provided by all employees. An organizational form and the underlying incentive contract induce a game among the participating agents.

In the first part of the paper I impose an equal sharing rule for the management team. Under this rule most of our results can be derived quite easily. Generally, it may be optimal to provide only a subset of managers with incentives and to distribute the surplus unevenly. In section 5 I consider the case with general sharing rules. Without any restrictions it is optimal to provide a fraction of the employees with incentives for extra effort. The restructuring process that follows an ICT productivity increase may lead to a situation in which all remaining employees work under monetary incentives.

The paper also addresses the question why workers or managers may be willing to forego the efficiency gains that are associated with the restructuring of their firms. Can't they be compensated if the firm produces more efficiently? And won't new firms hire those workers who have been dismissed by others? We address these questions in a macroeconomic extension of the basic model in which capital - or the number of available projects - is fixed. In such a setup optimal restructuring can lead to an excess supply of workers which reduces labor rents. Therefore, countries in which political power is with employed labor risk to end up with inefficient production processes that hire too many workers and artificially generate labor rents.

1.1 Related literature

The model in this paper analyzes the way in which innovations in information and communication technologies affect the design of organizations. The paper is related to two recent strands of the literature on organizational design. The first one imports insights from computer science into economic theory. This literature introduces a delay of information processing into economic models, it started with Radner (1992, 1993) and Radner and Van Zandt (1992). The resulting optimal "reduced tree" structure is designed for one-shot problems in which there is only one set of data to be processed, or the processing of the data is finished before another calculation task occurs. I restrict attention to this case in most of the paper. Van Zandt (1997, 1998) and Meagher, Orbay and Van Zandt (2001) study the case when new data comes in before the processing of the old set is finished. Orbay (2002) adds the frequency with which new data arrives as a new dimension to the analysis of efficient hierarchies. A similar problem is addresses in the appendix of this paper.

The paper is also closely related to the recent seminal work by Garicano (2000) and Garicano and Rossi-Hansberg (2004,5) on the microeconomic and distributional consequences of IT innovations. These papers address related questions in a different framework. Garicano introduces a model of hierarchical information aggregation. In this model agents may either specialize in a problem solving or in providing advice to other problem solvers. Two variables measure the quality of information technology: the cost of acquiring knowledge and communication costs. While a reduction of the cost of knowledge unambiguously reduces the number of hierarchy layers, communication costs have an ambiguous effect because lower communication costs encourage firms to choose more division of labor. In the present paper a hierachy has the task of aggregating decentralized information. The paper focuses on the trade off between faster information aggregation and an incentive effect of the division of labor. Lower communication costs lower the benefits for the division of labor and organizations can provide better incentives. The size of a hierarchy always declines with information transmission costs.[2]

Garicano and Rossi Hansberg (2004,5) study a macroeconomic extension of Garicano's model that uses a matching equilibrium similar to the ones in Legros and Newman (2002) and in section 6. They identify a number of interesting effects of IT innovations on the equilibrium income distribution within skill groups. The present paper also studies a macroeconomic matching equilibrium but it concentrates on the distributional conflict between scarce capital and management or labor.

[2]Note that my notion of hierarchy layers is different from Garicano's. In his paper the number of hierarchy layers is the number of sets of agents with different knowledge. Not all items need to go through the entire hierarchy because information is not aggregated. In the present model the size of the hierarchy determines the maximum number of operations an initial item needs to go through until it reaches the top player.

Meagher, Orbay and Van Zandt (2003) study the impact of information technology on the organization of a business enterprise which collects data on changing market conditions. Collecting more data is useful per se but the delay of the analysis reduces the usefulness of the outcome when the environment changes over time.[3]

In Radner's model – and in most of the information processing literature which followed – individuals are thought of as machines, perfectly doing what they are programmed to do. The joint analysis of speed and quality of hierarchical decision processes has previously been studied in Grüner and Schulte (2004), Jehiel (1999) and in Schulte and Grüner (2004). Jehiel considers the case where some signals get lost in the hierarchy with an exogenous probability, depending on the size of the groups of which the hierarchy consists. Schulte and Grüner study the role of the hierarchy design when individuals make mistakes with an exogenously given probability. In the present paper, the quality of collected information is endogenously determined by the actions of self-interested agents. The result that the reduced tree provides a (weakly) better decision quality than other organizations is the same. Recently, Grüner and Schulte have provided the first game theoretic analysis of the interaction of agents with incentive problems in such an organization. The present paper adds to this literature and studies the way in which an information technology improvement affects the way in which an efficient hierarchy is organized.

The paper is also related to the extensive game theoretic literature that studies incentives in hierarchies, such as Aghion and Tirole (1997), Mookherjee and Reichelstein (1997), and Melhumad, Mookherjee and Reichelstein (1995). These papers consider problems in which certain tasks as well as authority have to be delegated (and sub-delegated) to (and by) agents whose interests diverge from that of the principal. Delegation involves a loss of control for the principal, but strengthens the incentives for the agent. In the present model, all agents only care about the amount of effort they provide, not about the decision per se.[4]

[3] Prat (1997) studies hierarchies in which some managers are able to work faster than others, and the wage a manager is paid is a function of his ability. It turns out that with these modifications – except for the one made by Prat (1997) – the reduced tree is still (close to) efficient. Bolton and Dewatripont (1994) allow for specialization, which reduces the time an agent needs to understand information he handles frequently. In their model, the trade-off between specialization and communication costs determines the extent and the form of decentralization.

[4] Another related paper is Winter (2004), who studies incentive provision in a hierarchy via a transfer scheme. In his paper, the n tasks are assigned to n agents right from the beginning. He does not allow for the possibility to assign tasks differently, nor that one agent performs all of them. Unlike Winter, I am interested in the effect of task assignment on effort provision.

A related moral-hazard-in-team problem is studied in Dewatripont and Tirole (2004). They consider a sender-receiver game, where the sender has payoff-relevant information for the receiver, and must invest unobservable effort for the receiver to understand the information. In my model, communication is costless for the sender, but he has to provide effort to acquire the information in the first place.[5]

Recently, Acemoglu, Aghion, Lelarge, Van Reenen, and Zilibotti (2006) analyzed the relationship between the diffusion of new technologies and the decentralization decisions of firms. Centralized control relies on the - publicly available - information of the principal. Decentralized control delegates authority to a manager with superior information and a private agenda. As the available public information about the specific technology increases, the trade-off shifts in favor of centralization. Firms closer to the technological frontier and firms in more heterogeneous environments are more likely to choose decentralization.

Finally, there are several recent papers that also address the productivity puzzle such as Gordon (2004), Prescott (2006) and Klump et al. (2006). Most closely related is the paper by Gordon (2004) who argues that poor institutions prevented the realization of IT induced productivity gains in Europe. In particular he lists policies that promote high-density metropolitain areas, a lack of competition, and a poor public system of research financing. Others focus on the role of labor supply versus labor productivity measured as output per hours worked. Interestingly, both the average number of hours worked and productivity per worker is higher in the US than in Europe. As will be argued below, both observations are compatible with the present theory. According to Prescott (2006) Americans work 50 percent more than do the Germans, French, and Italians. In this paper Prescott studies the role of taxes in accounting for the differences in labor supply across time and across countries, in particular, the effective marginal tax rate on labor income. The finding is that this marginal tax rate accounts for the predominance of the differences at points in time and the large change in relative labor supply over time. The present paper offers an additional explanation for the higher US labor supply. The relative importance of these explanations can only be studied on empirical grounds. Some existing empirical work and new empirical strategies will be discussed in the conclusion.

---

[5] Other papers derive decentralized (hierarchical) organizations from technology. Crémer (1980) considers a problem of resource allocation under constraints on managerial time and finds that hierarchical organizations increase the amount of information that can be applied to a particular decision. Rosen (1982) has a paper in which a hierarchical structure emerges due to the need to supervise production (and supervision).

2 The model

2.1 The information aggregation task

Consider the following information aggregation task. A principal hires a team of m agents who have to work on n items. Agents work in a programmed hierarchy. Each agent reports the result of his operations to a single superior. There is a top agent who makes the final decision. Each item is assigned to exactly one agent. The programme determines at what point of time an agent is supposed to report to his superior (see Radner 1993 for a detailed description of the concept of a programmed network). Each agent can privately choose to provide normal or extra effort on an item. The cost of normal effort is normalized to zero. Providing extra effort is costly, the per unit cost is measured in monetary units and denoted by $c > 0$. I call a the number of items on which agents have provided extra effort, and $b = n - a$ the number of items that have been read without extra effort. Whether or not an agent has provided normal effort is observable and verifiable. The decision on extra effort is not observable. For simplicity, I assume that any report about the aggregate of information items from a lower hierarchy level can fully be understood by the superior without extra effort.

Agents communicate effectively in the sense that all the available information on initial items or aggregates thereof is transmitted accurately to the top agent in the hierarchy. However, no agent is able to derive the amount of effort that has been provided by a subordinate from his message.

The hierarchy's output x can take two values, 1 or 0. The probability of success π directly depends on the amount of effort that has been provided in the hierarchy. The chance of realizing a high output is linear in a and b. Moreover, it decreases in delay d.

$$\pi = \begin{cases} ap + bq - f(d) & \text{if } a + b = n, \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

with

$$f(0) = 0,\ \ f'(0) = 0,\ \ f \text{ non-decreasing,} \tag{2}$$

and

$$1/n \geq p > q > 0. \tag{3}$$

Effort is efficient, i.e. $p - q > c$. If one item is not read then the probability of success is zero. If all items are read but no agent provides extra effort then the probability of success is $nq > 0$. Thus, we may also write expected profits as:

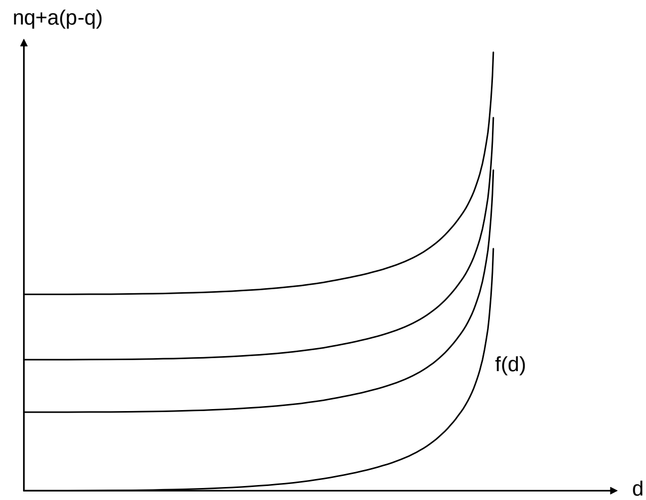

Figure 1: Iso-probability-of-success lines.

$$\pi = nq + a(p-q) - f(d). \tag{4}$$

Both a shorter delay and more effort provision increase expected profits. Figure 1 displays the corresponding iso-probability-of-sucess curves that emerge from a convex cost of delay function. The vertical axes measures the effort induced component of the success probability.

As one measure of firm productivity I use the expected output per unit of time

$$\rho := \frac{\pi}{d} = \frac{nq + a(p-q) - f(d)}{d}. \tag{5}$$

Note that one does not need to normalize for the total workload n because this workload is given exogenously. The surplus per unit of time is denoted by

$$\sigma := \frac{\pi - ac}{d}. \tag{6}$$

2.2 Preferences and wealth constraints

Agents working in the hierarchy do not care about output per se. They care about their monetary income and derive a disutility from providing effort on initial items. All individuals are wealth constraint: principals (the owners of projects) and all agents hold no cash initially.[6] Therefore, a contract can only specify how output will be shared in

[6]Relaxing the simplifying assumption that all agents hold zero wealth would permit to study the relationship between the wealth distribution and the organizational structure.

case that the project is a success. Consequently, contracts, which provide all agents with proper incentives in exchange for an appropriate entry fee, are not feasible. The outside payoff of agents is normalized to zero.

In most of the paper I concentrate on the simple case in which there is a maximum delay D beyond which the payoff is zero with certainty and below which losses are zero. I define:

$$f^D(d) := \begin{cases} 0 & \text{if } d \leq D \\ -\infty & \text{if } d > D \end{cases} . \tag{7}$$

Example 2 in Section 3 makes clear that the basic results of the paper also hold for the case with a convex and differentiable cost function.

2.3 The programmed hierarchy

The processing task is performed in a programmed hierarchy of agents (see Radner 1993, and, for a game theoretic extension Grüner and Schulte 2004). In such a network each information processing act (i.e. an agent reads an initial item or a message from some other agent) takes a given amount of time which is normalized to $\delta > 0$. The programme of the hierarchy specifies at which date an agent is supposed to handle an item. Moreover, the hierarchy describes a reporting structure.[7] Any report about the aggregate of information items from a lower hierarchy level can directly be understood by the superior. Reading the aggregate of several information items only consumes time δ and does not require effort.[8]

Figure 2 displays two organizations that aggregate 4 information items. The first one consists of one agent who reads all four items and produces with delay 4δ. The second one consists of two agents who read two items each and produces with delay 3δ.

The programmed hierarchy, together with the incentive scheme leads to an extensive form game. In this game the strategy fixes how many times an agent provides extra effort on his initial items. Note that only players who handle initial items have to make such an effort choice - all other agents always provide normal effort[9]. The solution concept

[7] Details can be found in Grüner and Schulte (2004).

[8] This assumption is made for simplicity. All results of this paper hold if one imposes a cost of effort for the processing of aggregates as well. Grüner and Schulte (2004) analyze this case in detail. The introduction of effort costs on higher levels leads of the hierarchy makes it somewhat more difficult to characterize equilibrium effort levels. The corresponding arguments are not of central importance in the present paper.

[9] Grüner and Schulte (2004) explicitely model the reporting space of agents. This leads to a richer

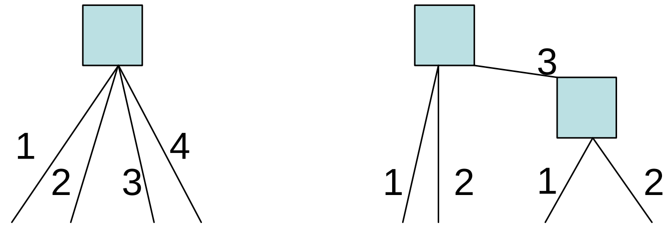

Figure 2: Two hierarchies with 1 (2) agents and 4 objects with delay 4(3).

is Nash equilibrium. Given the simple additive structure of the output function (1), an agent provides extra effort in equilibrium if and only if his variable share of output α_i is above $c/(p-q)$.

An improvement of ICT efficiency is an exogenous reduction of δ.

3 ICT innovations and productivity gains

3.1 Example 1

It is instructive to first consider an example in which the principal uses an equal sharing contract. This contract pays each manager the same fraction α/m of total output[10]:

$$y_i = \frac{\alpha}{m}x,\ \alpha \in [0,1]\,. \tag{8}$$

Consider a team that has to work on $n = 4$ items. Suppose that there is a maximum delay D beyond which the payoff is zero with certainty, i.e. let $f(d) = f^D(d)$. I fix $D = 4$ and $\delta = 1 + \varepsilon$. Figure 3 displays two corresponding iso-success-probability lines. The two horizontal iso-success-probability curves become vertical at $d = D$. The two straight lines represent the value added (in terms of success probability) that is generated by a single agent who does $(p \cdot d)$ or does not provide $(q \cdot d)$ effort. The principal maximizes expected profits measured by

$$\Pi = \pi - E(t) \tag{9}$$

where $E(t) = E\left(\frac{\alpha x}{m}\right)$ denotes expected transfers to the agent(s). He has to consider two options. One option is to divide the task evenly among two agents. Both agents then

possible history of the game and a more complex strategy space. The corresponding arguments are not central in the present analysis.

[10]Some team production processes are actually working under similar incentives schemes. Equal sharing also arises in situations in which hard incentives are excluded and success is a publicly provided private good.

simultaneously read two objects. This yields a delay of 2δ. Next one agent sends a report to his colleague which yields a total delay of $3\delta < 4$ (point A in figure 3). There is an incentive contract such that both agents provide effort if

$$\frac{1}{2}(p-q) \geq c. \tag{10}$$

If instead c is above this threshold there is no way to decentralize the task efficiently. The second option is that one agent performs the task alone. The corresponding delay is $4\delta > D$ (point B). Hence, centralization is no good option. At $\delta = 1 + \varepsilon$ the optimal organization decentralizes the processing task, and the delay is $3\delta < 4$. At costs $c > \frac{1}{2}(p-q)$ both agents do not provide effort in the profit maximizing hierarchy.

Next consider an improvement in information technology that reduces delay of a single processing step to $\delta = 1 - \varepsilon$. Centralization now works sufficiently fast because $4\delta < D$ (point C in figure 3). Effort is provided if

$$p - q \geq c. \tag{11}$$

Assume now that

$$p - q \geq c > \frac{1}{2}(p-q). \tag{12}$$

For sufficiently small values of q, the profit maximizing hierarchy works with only one manager. He is in charge of the entire project. His incentive payment t ensures that $(p-q)\,t > c$. The agent is willing to participate since $4p\frac{c}{(p-q)} - 4c > 0$. Profits are $4p\left(1 - \frac{c}{p-q}\right) > 0$. This is more than $4q$ if $p\left(1 - \frac{c}{p-q}\right) > q$. Under the new contract expected output rises from $4q$ to $4p$ and delay increases from $3(1+\varepsilon)$ to $4(1-\varepsilon)$. The firm's output per period changes from $\frac{4q}{3(1+\varepsilon)}$ to $\frac{p}{1-\varepsilon}$. The surplus per unit of time goes from $\sigma = \frac{4q}{3(1+\varepsilon)}$ to $\frac{p-c}{1-\varepsilon}$.

One can now easily verify that marginal ICT improvements may result in significant productivity gains. Take as an example the case where c is slightly above $p/2$ and q is close to zero. The surplus per unit of time before restructuring is close to zero while, after restructuring it is approximately $4(p-c) \approx 2p$. The previous results can be summarized as follows:

Proposition 1 *There are parameters (n, c, p, q, D, δ) such that marginal improvements of ICT efficiency may result in (i) positive productivity gains ρ and σ, (ii) positive increases of profits, and (iii) a strongly reduced number of hierarchy layers and employees.*

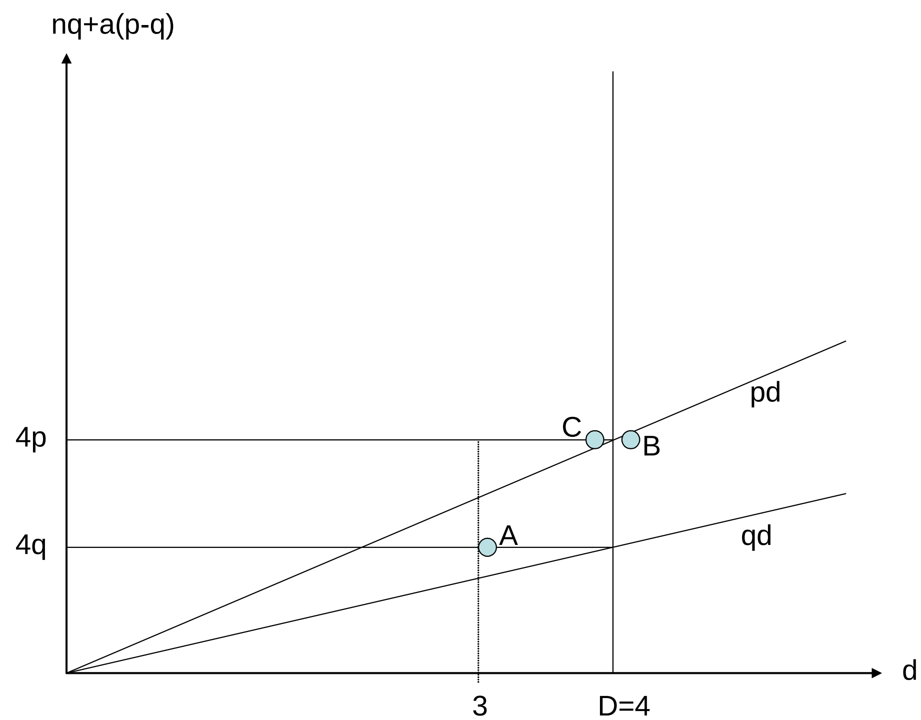

Figure 3: Example 1.

3.2 Example 2: convex cost

The purpose of the second example is to show that our previous results do not rest on the particular non-continuous cost structure that we considered in Example 1. They may also obtain when the cost of delay function is differentiable and convex. Consider the same environment (4 items) as in the previous example but with a convex cost function $f(d)$ with

$$f'(0) = 0,\ f''(d) > 0. \tag{13}$$

Again assume that

$$p - q \geq c > \frac{1}{2}(p - q). \tag{14}$$

Figure 4 explains why the outcomes from Example 1 may also realize with convex costs. The figure displays two iso-success-probability lines. Suppose that the initial per unit delay is $\delta = 1$. The two straight lines represent the value added (in terms of success probability) that is generated by a single agent who does $(p \cdot d)$ or does not provide $(q \cdot d)$ effort. Consider the case in which q is so low that the probability of success of a reduced tree with two agents is zero (point A in figure 4):

$$q = \frac{f(3)}{4}. \tag{15}$$

Moreover, assume that a single agent does not produce a positive probability of success: $4p - f(4) = \varepsilon < 0$ (point B). After an ICT improvement the optimal hierarchy is a single agent hierarchy (point C). The probability of success is

$$\pi = 4p - f(4\delta). \tag{16}$$

Expected profits are:

$$\Pi = 4p\left(1 - \frac{c}{p-q}\right) - f(4\delta). \tag{17}$$

Before restructuring the surplus per unit of time is given by:

$$\sigma = \frac{\pi - ac}{d} = \frac{0}{3} = 0. \tag{18}$$

After restructuring it is:

$$\sigma = \frac{4p - f(4\delta) - 4c}{4} = p - c - \frac{f(4\delta)}{4}. \tag{19}$$

Figure 4 characterizes a situation in which the new hierarchy generates a higher surplus. Point D in figure 3 corresponds to a situation in which the old hierarchy is kept while ICT improves. The convexity of $f(\cdot)$ guarantees that there is an extra gain in the probability of success that is due to the restructuring process. Points E and F identify the combinations of surplus and delay that are realized in the two different hierarchies. As one can easily verify, the surplus per unit of time is larger when the task is delegated to a single agent. Again, a marginal ICT improvements may result in significant productivity gains.

4 Optimal hierarchy, equal sharing rule

4.1 The reduced tree

In this section, I derive two more general results about the optimal hierarchy design for the case of an equal sharing rule. Under an equal sharing rule the optimal hierarchy can easily be determined. It has the form of a reduced tree as described in Radner (1993). The reduced tree is the hierarchy which minimizes the delay for a given number of objects and agents. It can be constructed as follows. All managers are numbered consecutively from 1 to m. Each manager is assigned the same number of initial items. If there are

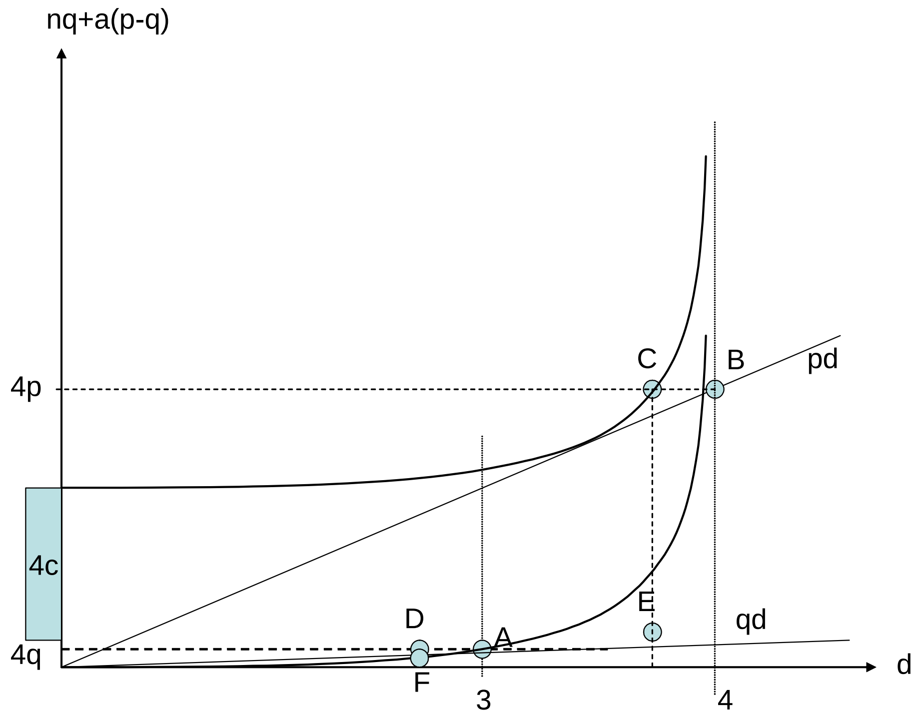

Figure 4: Example 2

single leftover objects, then the remaining items are distributed in such a way that no manager handles more than one additional item. After reading initial objects, managers with an even ordering number x report to manager $x-1$. After that the procedure is repeated with the remaining managers until the final result is obtained by manager 1.

Radner studied efficiency along the dimensions delay (d), number of managers (m) and number of items (n). He found that any efficient network is a reduced tree. Figure 5 provides an example of a reorganization that leads to less delay for a given number of processors. Figure 6 provides an example for the construction of a reduced tree with 8 processors and 24 items.

4.2 Optimal hierarchy

I consider the case where $f(d) = f^D(d)$. Two definitions are useful. There is less incentive to provide effort if the number of managers increases. Managers can only be induced to provide effort if $p - q/m > c$. Therefore one may define:

Definition 1 *Call* $m^e = \lfloor (p-q)/c \rfloor$ *the largest number of managers for which there is a full effort equilibrium under an equal distribution of revenues.*

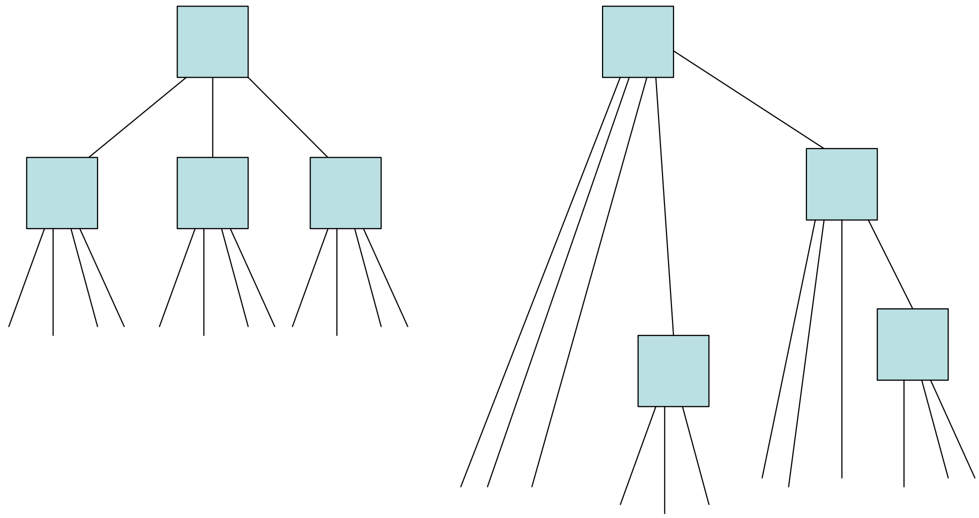

Figure 5: Two programmed hierarchies with 4 processors and 12 items.

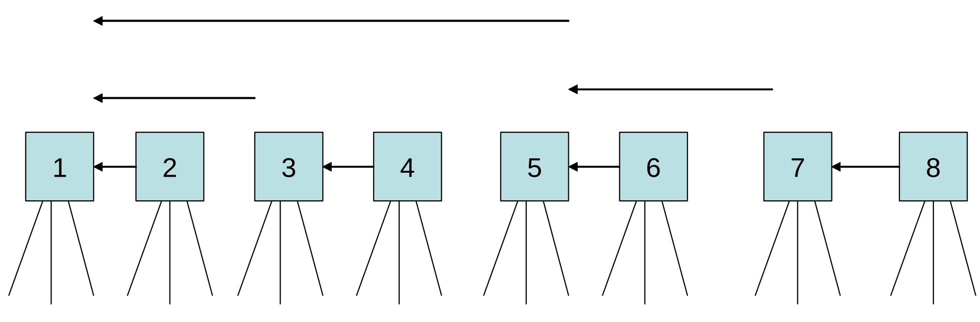

Figure 6: Construction of a reduced tree with 8 processors.

A delay of less than D requires a minimum number of managers instead. I define:

Definition 2 *Let* $f(d) = f^D(d)$*. Call* $m^d(n, \delta, D)$ *the smallest number of managers necessary to process* n *objects with a delay of less than* D*.*

The optimal hierarchy design depends on the relation of m^d and m^e.

Proposition 2 *Consider a hierarchy working under an equal sharing rule and let* $f(d) = f^D(d)$.

(i) Let $m^d(n, \delta, D) = m^e$*. The reduced tree with* $m^d = m^e$ *managers maximizes the probability of success. All managers in this hierarchy provide effort.*

(ii) Let $m^d > m^e$*. Every hierarchy that produces with a delay of less than* D *maximizes the probability of success. Any hierarchy which maximizes the probability of success does not induce effort.*

(iii) Let $m^d < m^e$*. There is a full effort reduced tree which maximizes the probability of success for any number of managers* m *with* $m^d \leq m \leq m^e$*. All these hierarchies yield full effort and an identical probability of success.*

Proof (i) According to Radner 1993, Theorem 1, any faster hierarchy uses more managers for the given number of items. These managers have no incentive to provide effort. Any slower hierarchy takes too long for its processing task and the corresponding surplus would be zero. Any hierarchy with the same number of managers creates at most the same profit but cannot work faster. Parts (ii) and (iii) are obvious. *Q.E.D.*

Note that in cases (ii) and (iii) the hierarchy which maximizes the surplus is uniquely defined when normal effort is costly.

For small values of q Proposition 1 also describes the profit maximizing hierarchy. For any given set of parameters (p, c) the hierarchy which maximizes the probability of success also maximizes profits when q is sufficiently close to zero. This is so because the hierarchy that maximizes the probability of success always generates non-negative profits whereas the no-effort hierarchy generates the high return with a probability that can be made arbitrarily close to zero.

A consequence of Proposition 1 is that ICT improvements are likely to lead to a reduction of the number of hierarchy layers (measured by the highest number of superiors of all agents in the hierarchy). The reason is that without any institutional change delay is reduced. Reducing the number of hierarchy layers has the potential of increasing incentives without sacrificing too much delay. I call *the minimum number of hierarchy layers* the smallest number of hierarchy layers of any hierarchy maximizing the probability of success.

Proposition 3 *Let* $f(d) = f^D(d)$*. An increase in ICT productivity never increases the minimum number of hierarchy layers. When* $m^d < m^e$ *a sufficiently strong increase in ICT productivity decreases the minimum number of hierarchy layers.*

Proof An increase in ICT efficiency weakly reduces m^d. The rest is a corollary of the previous proposition and Radner (1993). *Q.E.D.*

Note that this result differs from Garicano (2000). In his model improvements in the processing technology tend to increase the number of hierarchy layers because the cost of the decentralization of knowledge declines. A consequence of the present result is that the optimal change of the organizational structure may be drastic in the sense that hierarchy levels may disappear and that a huge fraction of the employees may have to be dismissed. This holds for arbitrarily large hierarchies.

Observation 1 *For all* $k \in N$ and $n = 2^{k+j}$ *with* $j \in N \setminus \{1\}$ *there are values* c*,* D, and δ *such that (i) the profit maximizing hierarchy works with* $m = 2^k$ managers and (ii) *profit maximizing restructuring after a marginal increase in IT productivity includes that one hierarchy layer is skipped and half of the old employees are no longer employed.*

Proof First fix D, and δ so that $D/\delta = k + j + \varepsilon$ where $\varepsilon > 0$. In this case $m^d(n, \delta, D) = 2^k$. Next fix c so that $c = (p - q)/2^{k-1} > (p - q)/2^k$. The optimal hierarchy works with 2^k managers who do not provide effort. At $\varepsilon = 0$ a hierarchy with 2^{k-1} managers provides effort, is sufficiently fast and maximizes profits. *Q.E.D.*

Note that the last results does not critically rely on the particular formulation of the cost of delay function $f(d)$. One can easily verify that a convex and continuous cost function may also permit observation 1.

5 Unrestricted sharing rules

So far I have assumed that the principal has to share the output of $x = 1$ equally among all agents. Generally, this may be restrictive because the principal can remunerate different agents differently when a full effort equilibrium is not available under an equal sharing rule. I now consider the more general case with unequal sharing rules. An incentive contract fixes the output share $\alpha_i \in [0, 1]$ of each agent and an up front payment $\bar{y}_i$. An agent either provides effort on all his objects or on none of them. He provides effort if his share α_i satisfies

$$\alpha_i (p - q) \geq c. \tag{20}$$

It may now be optimal to provide only some agents with incentives. This ensures that those agents work hard on the initial items that have been assigned to them. The remaining agents only provide normal effort on their workload.

Observation 2 *With an unrestricted sharing rule it may be the case that*

(i) before the improvement of ICT efficiency some employees get an incentive payment and work hard while others do not.

(ii) after restructuring the hierarchy has less layers and less employees, incentive payments are made to all employees, and all agents provide first best effort.

Proof It suffices to provide an example. Consider again the first example from section 3. For large enough values of c, there is no full effort equilibrium contract in the decentralized organization with two agents. One option is not to provide any incentives to both agents and to realize profits of $4q$. However, at low values of q it pays to fix $\alpha_1 = c/(p-q)$ and $\alpha_1 = 0$ instead. In this case agent 1 provides effort while agent 2 does not. The delay is below D and the surplus is $2p + 2q$. Profits amount to $(2p + 2q)(1 - c/(p-q))$ which exceeds $4q$ when q is small enough. Reducing δ slightly may make it possible to delegate the entire task to one agent who then provides full effort and works with a sufficiently small delay. *Q.E.D.*

This last observation is related to a feature of many real world organizations: hard incentives are only provided to some employees - mostly executives - while others (PA's, secretaries) who also provide relevant input are working on a fixed salary. In such cases ICT innovations enable the principal to reduce the division of labor in the organization and to provide all remaining employees with proper incentives.

6 Macroeconomic consequences

We have so far established that a proper reorganization of business enterprises may significantly enhance the productivity effect of ICT innovations. In principle the resulting welfare gain could be used to make everybody better off. Can it be that some economies systematically forego some of the welfare gains which are associated with a restructuring of their business firms? In this section I study a simple macroeconomic extension of the previous model in order to address this question.

The model focuses on the distribution of rents in an economy in which several investment projects are carried out. I assume that there are two types of agents: a number of E entrepreneurs (principals) and M managers. Each entrepreneur has access to the same number $P \geq 1$ of potential projects. The total numbers of entrepreneurs, the number of managers and the number of available projects $E \cdot P$ is fixed. Again, for simplicity I assume that principals have to stick to an equal sharing rule when they design incentive contracts for managers.

Entrepreneurs use incentive contracts in order to motivate their managers to provide effort in their productive relationship. Every such incentive contract together with equilibrium effort determines the equilibrium surplus which is generated in each firm. The surplus is divided into the entrepreneur's and managers' rents. We define w as the expected equilibrium labor rent per employed manager and r as the expected equilibrium rent of a principal. Similar to Legros and Newman (2000) an equilibrium is defined as follows:

Definition 3 *An equilibrium consists of (i) a hierarchy H of size m, (ii) an incentive contract C, (iii) wage w and rent r, and (iv) an allocation of managers to entrepreneurs' projects such that (a) at w, r, no entrepreneur can offer to hire some managers on one of his projects such that at least one member of this new coalition (capitalist plus managers) is better off without reducing the payoff of another member of the coalition and (b) labor rent w and principal's rent r obtain in a hierarchy of size m, working under incentive contract C, and (c) hierarchy H with contract C maximizes the entrepreneurs profits subject to agents' participation and incentive constraints.*

Before we analyze the distributional consequences of restructuring, we need to verify that an equilibrium generally exists. Consider the hierarchy that maximizes the expected surplus of a single project. Suppose that this hierarchy employs m managers and let $E \cdot P \cdot m < M$. There is an equilibrium in which all entrepreneurs choose the surplus maximizing hierarchy and make profits with all their projects. When managers in the surplus maximizing hierarchy do not provide effort they get zero rents in equilibrium. Otherwise they get exactly what is needed in order to guarantee that effort is provided because managers are abundant and some are unemployed. If instead $E \cdot P \cdot m > M$ workers are scarce in equilibrium. All workers are employed in surplus-maximizing firms and entrepreneurs' rents are zero. One can easily verify that both equilibria are unique when they exist.

Consider now a restructuring process which reduces the number of agents that work on a particular project. Such a restructuring process increases the number of projects that can be carried out successfully in the economy. It may therefore be the case that efficient restructuring leads to a situation in which projects become scarce while managers become abundant. Consequently, the distribution of rents in the economy may be strongly affected by the restructuring process. Restructuring eliminates all managers' rents. This leads to the following result:

Proposition 4 *Consider an economy in an equilibrium in which in all existing firms employ* $m > 1$ *agents and in which the economy's surplus is positive. Let* $2EP > M > EP$.

(i) In equilibrium $w > 0 = r$.

(ii) There is an improvement in IC technology such that in the new equilibrium less managers work in every hierarchy, providing effort. In the new equilibrium capitalists' rents are positive while manager's aggregate rents may decline.

Proof (i) In equilibrium there are not enough managers in order to endow all EP projects with more than one worker because $2EP > M$. Therefore equilibrium rents of principals must be zero. All rents go to managers. (ii) There always is a reduction of δ so that the surplus maximizing hierarchy turns out to be one in which full effort is provided. In deed, as δ goes to zero, all projects will optimally be carried out by single agents who provide effort and maximize the probability of success. An equilibrium in which all projects are carried out by single managers exists. Some managers are unemployed. In this equilibrium only the incentive constraint and the participation constraint of managers need to hold.

In the aggregate managers on a single project get lower rents than before. An employed manager's payoff is

$$np\frac{c}{p-q} - nc = \frac{q}{p-q}nc. \tag{21}$$

Before, the surplus per firm was nq and went entirely to the firm's managers. The rent of all managers in a single firm always declines since:

$$nq > \frac{q}{p-q}nc \Leftrightarrow p-q > c. \tag{22}$$

One can easily find cases in which the aggregate (and average) rents of managers in the economy are also reduced. This is the case when the total number of firms does not increase too strongly. Before restructuring there are M/m firms, after restructuring $EP < M$ firms. Hence, aggregate rents of managers decline when M is close enough to EPm. *Q.E.D.*

This simple macroeconomic extension of our basic model points out that economies in which managers (or, where appropriate, employed workers) are more powerful are less likely to implement the necessary efficiency-enhancing structural reforms in business enterprises. When firms are not restructured managers are likely to capture the increase of surplus because they are scarce. After restructuring managers are abundant and their rents will decline instead. Only a compensation of all managers for the restructuring process may yield a solution to this problem.[11]

7 Conclusion

The paper provides an explanation for the productivity puzzle. ICT improvements and organizational restructuring are highly complementary. An ICT improvement enables employees to perform more complex tasks alone or in small teams. The principal can then solve elementary incentive problems more easily with an appropriate incentive scheme. The number of employees may have to be reduced drastically in such a process. Europe is well known for rigid labor markets and in particular for high firing costs. This raises restructuring costs and may explain part of the puzzle. The restructuring process also entails a reduced division of labor. Hence, status concerns of employees who are supposed to take over tasks of their subordinates may yields additional management opposition.[12] These results indicate that IT-induced restructuring may not be in managers' or workers' interest unless appropriate compensatory payments are made. Otherwise management or labor will favor a restrictive regulation that raises the costs of restructuring processes.

The present theory is consistent with a number of recent macroeconomic and microeconomic observations. First, productivity per hour worked is higher in the US than in Europe. This is a feature of the current model when workers time input is measured by the number of operations which are performed in the hierarchy. Second, the individual workload is larger in the optimal hierarchy which corresponds well to the fact that in the US worktime is larger (see figure 5).

The paper is also in line with recent micro-data on the development of hierarchies in US business companies. In an analysis of a panal of about 300 publicly traded firms Rajan

[11] Note that Fernandez and Rodrik (1991) have pointed out some of the difficulties that are associated with such compensation packages. A particular problem is that the credibility of any promised transfer is low when political power is on the side of the agens that capture the rents.

[12] There is anectdotal evidence that some executives refuse to use software such as powerpoint and prefer to give handwritings to secretaries or PAs. Others do not have an own email acount but prefer to send or receive emails through their secretary.

and Wulf (2003) find that in those companies about 25 % of the layers of intermediate management have on average been removed between 1986 and 1999. In the United States, the decline in the average size of firms, as measured by employment, has been well-documented. Brynjolfsson et al (1994) find broad evidence that investment in information technology is significantly associated with subsequent decreases in the average size of firms. Most importantly Bresnahan et. al. (2002) find evidence of complementarities among information technology and workplace reorganization in factor demand and productivity regressions.

An interesting option for a deeper empirical analysis of the present theory is to study the differences in the structural adjustment to ICT improvements in firms in Europe and the US. Such an analysis could rely on occupational statistics that are widely available from sources such as the ILO or the ISSP survey programme. The testable hypothesis is that countries in which occupational shares react less flexible exhibit lower productivity levels and growth rates. An alternative option is to collect micro-data on the composition and organizational structure of firms and on their performance.

Some extensions of the present theoretical framework are also worth to be explored. The paper has considered an economy in which a given set of one shot tasks can be performed by a variable number of hierarchies. An alternative is to consider the case in which tasks arrive periodically at some rate. As has been shown by van Zandt (1997,8) this may yield a different shape of efficient institutions. Balanced hierarchies become more advantageous because different hierarchy layers can simultaneously work on different waves of items. Nevertheseless, one can construct examples in which the message of the present paper remains intact when items arrive periodically (see the appendix for such an example). A deeper analysis of this issue along the lines of van Zandt (1997,1998) and Orbay (2002) may provide some additional interesting insights.

8 Appendix: Sequential arrival of information

Consider a situation in which aggregation tasks are performed sequentially. As has been shown by van Zandt (1997, 1998) in this case balanced hierarchies in which superiors do not deal with initial objects may be efficient. These hierarchies permit that the hierarchy simultaneously works on two or more information aggregation task. In this appendix I provide an example in which - under sequential arrival - the basic insights from section 3 remain intact.

Consider an information aggregation task with four items. Let $D = 3.9$ and $\delta = 1$.

Suppose that there is an abundant quantity of tasks that may be processed. What is the efficient hierarchy design? A single decisionmaker is not fast enough. He produces an excessive delay of 4 and can not generate positive profits. The same holds for a balanced hierarchy of three agents. The only remaining option is the reduced tree with two agents. This hierarchy processes a new aggreagation task every 3 units of time.

Suppose now that sharing the surplus among two agents yields no incentives. A reduction of δ to 0.9 makes the single decision maker and the balanced hierarchy work with delay $3.6 < 3.9$. The balanced hierarchy processes twice as many tasks per unit of time as the single decision maker. However, only the single decision maker provides effort. The per unit of time surplus he generates is $4\left(p-c\right)/3.6$, and for the balanced hierarchy $\left(4q-4c\right)/1.8$. The single decision maker performs better if $p > 2q-c$. Note that there are values p, q, and c such that this condition is compatible with the two incentive constrainst of a one- and a two agent hierachy, $2c > p-q > c$.

References

[1] Acemoglu, D., Aghion, P., Lelarge, Van Reenen, and F. Zilibotti (2006): Technology, Information and the Decentralization of the Firm, CEPR Discussion Paper 5678.

[2] Aghion, Philippe and Jean Tirole (1997): Formal and Real Authority in Organizations, *Journal of Political Economy* 105(1):1-29.

[3] Bolton, Patrick and Mathias Dewatripont (1994): The Firm as a Communicating Network, *Quarterly Journal of Economics* 109:809-839.

[4] Bresnahan, Timothy F., Erik Brynjolfsson, and Lorin M. Hitt (2002) "Information Technology, Workplace Organization, and the Demand for Skilled Labor: Firm-Level Evidence". *Quarterly Journal of Economics*, 117, 339-376.

[5] Brynjolfsson, Erik, Thomas Malone, Vijay Gurbaxani, and Ajit Kambil (1994) Does Information Technology Lead to Smaller Firms?, *Management Science*, 40, 1645-1662.

[6] Crémer, Jacques (1980): A Partial Theory of the Optimal Organization of a Bureaucracy, *Bell Journal of Economics* 11:683-93.

[7] Dewatripont, Mathias and Jean Tirole (2004): Modes of Communication, IDEI Working Paper 323.

[8] Fernadez, Raquel und Dani Rodrik (1991) ”Resistance to Reform: Status Quo Bias in the Presence of Individual Specific Uncertainty,” *American Economic Review*, 81, 1146-55.

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

[10] Garicano, Luis and Esteban Rossi-Hansberg (2004) "Inequality and the Organization of Knowledge", *American Economic Review*, 94, 197-202.

[11] Garicano, Luis and Esteban Rossi-Hansberg (2005) "Organization and Inequality in a Knowledge Economy", *Quartely Journal of Economics*, forthcoming.

[12] Gordon, R.J. (2004) "Why was Europe Left at the Station When America's Productivity Locomotive Departed?", NBER Working Paper No. W10661.

[13] Grüner, Hans Peter and Elisabeth Schulte (2004): ”Speed and Quality of Collective Decision Making, II: Incentives for Information Provision”, CEPR Discussion Paper 4397.

[14] Jehiel, Philippe (1999): ”Information Aggregation and Communication in Organizations”, *Management Science*, 45:659-669.

[15] Klump, R., P. McAdam and A. Willman, (2006) "Factor substitution and biased technical change: New results for the euro area" mimeo.

[16] Legros, Patrick and Andrew Newman (2002): "Monotone Matching in Perfect and Imperfect Worlds," *Review of Economic Studies*, 69, 925-42.

[17] Luce, R. Duncan (1959): *Individual Choice Behavior*. New York: Wiley.

[18] McGrattan Ellen R. and Edward C. Prescott, 2004. "Productivity and the Post-1990 U.S. Economy", UCLA Department of Economics.

[19] Meagher, Kieron, Hakan Orbay and Timothy Van Zandt (2003): Hierarchy Size and Environmental Uncertainty, in: Advances in Economic Design, M.R. Sertel and S. Koray (Eds.), Springer-Verlag.

[20] Meagher, Kieron and Timothy Van Zandt (1998): Managerial Costs for One-shot Decentralized Information Processing, *Review of Economic Design* 3: 329-345.

[21] Melumad, Nahum D., Dilip Mookherjee and Stefan Reichelstein (1995): Hierarchical Decentralization of Incentive Contracts, *RAND Journal of Economics*, 26:654-672.

[22] Mookherjee, Dilip and Stefan Reichelstein (1997): Budgeting and Hierarchical Control, *Journal of Accounting Research* 35(2):129-155.

[23] Orbay, Hakan (2002): Information Processing Hierarchies, *Journal of Economic Theory* 105: 370-407.

[24] Prat, Andrea (1997): Hierarchies of Processors with Endogenous Capacity, *Journal of Economic Theory* 77:214-222.

[25] Prescott, Edward C. (2003) "Why do Americans work so much more than Europeans?", Staff Report 321, Federal Reserve Bank of Minneapolis.

[26] Radner, Roy (1992): Hierarchy: The Economics of Managing, *Journal of Economic Literature* 30:1382-1415.

[27] Radner, Roy (1993): The Organization of Decentralized Information Processing, *Econometrica* 61(5): 1109-1146.

[28] Radner, Roy and Timothy Van Zandt (1992): Information Processing in Firms and Returns to Scale, *Annales d'economie et de statistique* 25/26:265-298.

[29] Rajan R.G and J. Wulf (2003): "The Flattening Firm, Evidence from Panel Data on the Changing Nature of Corporate Hierarchies", NBER working paper, w9633.

[30] Rosen, Sherwin (1982): Autority, control, and the distribution of earnings, *Bell Journal of Economics* 13: 311-323.

[31] Schulte, Elisabeth and Hans Peter Grüner (2004): Speed and Quality of Collective Decision Making: Imperfect Information Processing, forthcoming in *Journal of Economic Theory*.

[32] Smorodinsky, Rann and Mose Tenneholtz (2003): "Overcoming Free Riding in Multi-Party Computations – The Anonymous Case", *mimeo*, Technion.

[33] Van Zandt, Timothy (1997): The Scheduling and Organization of Periodic Associative Computation: Essential networks, *Review of Economic Design* 3:15-27.

[34] Van Zandt, Timothy (1998): The Scheduling and Organization of Periodic Associative Computation: Efficient networks, *Review of Economic Design* 3:93-127.

[35] Winter, Eyal (2004): Incentives and Discrimination, *American Economic Review* 94:764-773.