Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Optimal Hierarchies in Firms: a Theoretical Model

Mikhail V. Goubko*, Sergey P. Mishin*

**Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russian Federation (e-mail: mgoubko@mail.ru).*

Abstract: A normative economic model of management hierarchy design in firms is proposed. We seek for the management hierarchy that minimizes the running costs. Along with direct maintenance expenses these costs include wastes from the loss of control. The results include analytic expressions for the optimal hierarchy attributes: span of control, headcount, efforts distribution, wages differential, etc, as functions of exogenous parameters. They are used to analyze the impact of environment parameters on a firm's size, financial results, employees' wages and shape of hierarchy. The detailed analysis of this impact can help draw up policy recommendations on rational bureaucracy formation in firms.

The research is supported by the grant 07-07-0078-a of Russian Foundation for Basic Research.

1. INTRODUCTION

The notion of transaction costs (or "economic system exploiting costs") forms the basis of neoinstitutional economic theory and the modern theory of the firm. As O.E. Williamson (1975) notes, economizing on transaction costs is the main goal of any economic institution. The internal structure of modern firms usually takes the form of management hierarchy. Transaction costs are produced inside the hierarchy and greatly influenced by its shape and other attributes. At present the attributes of management hierarchy are universally recognized to exert key influence on the effectiveness of management (Mintzberg H., 1983). Thus, the analysis of management hierarchies (organization structures) gives clues to deeper understanding of the nature of the firm.

Interest in normative models of management hierarchies increases in the context of the continuing processes of business globalization (mergers, absorptions, vertical and horizontal integration). The crucial problem of huge modern corporations is the rational organization of their bureaucracy. Severe competition for global markets makes not only the financial results but the very existence of a corporation dependent on the efficiency of its management structure. The increasing pace of change in production and management technologies requires fast and adequate changes in the organization structure of a firm, and normative models of a hierarchy design must provide the aid in the solution of these sophisticated problems.

In this paper the transaction costs approach is combined with the original mathematical results in an optimal hierarchy design (Mishin S.P., 2004; Goubko M.V., 2006) to formulate and study the models of multi-layer management hierarchies. Along with "direct" maintenance expenses (salaries, bonuses, options, office rents, stationary, etc) due to the management staff, transaction costs in this model also include wastes from the so-called "loss of control" (Williamson O.E., 1967). The questions posed by the model are: how many managers the firm must hire, when headcount should be increased or decreased, how managers wages depend on their positions, whether the implementation of corporate information systems results in a flatter management hierarchy, when the growth of the firm is advantageous, etc.

2. BRIEF LITERATURE REVIEW

Since the beginning of the 20th century transaction costs have become the central point of a new approach to the theory of the firm. Market mechanisms were recognized to lead to costs, thus allowing the rationality of alternative forms of institutions. The main topics of interest were "When do markets fail? What alternative modes of organization are available? What are the limits of these alternative modes?" Answering these questions was focused on the main alternative organization form – the management hierarchy – and demanded the advanced modeling of the internal structure of a firm.

One of the early formal models of intra-firm hierarchy was introduced by M. Beckmann (1960). He limits administrative costs to managerial wages. Imposing restrictions on the minimum *span of control* (the number of immediate subordinates of a manager) and the maximum *wage differential* between subsequent layers of hierarchy he proves that administrative costs rise approximately linearly with the firm growth. So, he concludes, these costs cannot limit the maximum size of a firm.

Later in his famous article O.E. Williamson (1967) introduces the important notion of *loss of control*. He argues that in real world the efficiency of a manager's control is limited by natural bounds of human attention and communication. It is claimed that only some fraction $\alpha < 1$ of a manager's orders and directions can be successfully implemented by his subordinates. Williamson supposes the output of any productive worker to be directly governed by the cumulative loss of control through the *chain of command* – the chain of managers "above" this worker. Assuming constant span of control and wage differential O. Williamson shows that the loss of control makes the revenue of a firm to be concave in its size and results in a finite optimal firm size even with linear (in size) administrative costs.

Relying on this approach G.A. Calvo and S. Wellisz (1978) proposed the model of hierarchical monitoring. They internalized α interpreting it as an employee effort (e.g. the time the manager engages in monitoring or worker spends in production). They were the first to set the formal problem of optimal hierarchy design: to determine the number of productive workers, the number of layers in a hierarchy, the span of control and the wage for every layer to maximize the profit of a firm. The authors do not solve this problem explicitly but prove that a firm's ability to grow crucially depends on the details of a monitoring mechanism in use.

Y. Qian (1994) employs the hierarchy design technique developed by M. Keren and D. Levhari (1983) to analyze the model where managers in a hierarchy engage both in monitoring and in production activities. Among the other results Y. Qian shows that the optimal employee's wage and effort level rise from the bottom to the top of a hierarchy, and the optimal span of control is always greater than $e \approx 2.71$. He also proves the profit of a firm to be a concave increasing function of a firm size. Thus, Y. Qian agrees with M. Beckmann in that the loss of control in a management hierarchy cannot limit the growth of a firm.

Another model of hierarchical authority and control was introduced by S. Rosen (1982). He incorporates the labor market (the market of managerial skills) into the model of hierarchy design. The goal is to describe an equilibrium distribution of firms by their size along with explaining the market mechanisms for manager wages formation. The distinctive feature of his model is that every potential employee has a unique vector of skills that influences his effectiveness as a productive worker, a first-layer manager, a second-layer manager, etc. For a special case of two-layer firms with constant returns technology S. Rosen finds the equilibrium prices for the worker and manager skills. He shows that in equilibrium more able managers govern the firms of a greater size.

Among the recent publications on optimal hierarchy design one can also mention K.J. Meagher (2003) and A. Patacconi (2005). The other approaches to the optimal hierarchy problem (see the survey in (Goubko M.V., 2006)) include "knowledge hierarchies" of L. Garicano and A.W. Beggs, "computer science approach" of R. Radner, T. Van Zandt, P. Bolton and M. Dewatripont, "teams theory approach" of J. Cremer, M. Aoki, J. Geanakoplos and P. Milgrom, "decision-making hierarchies" of R.K. Sah and J.E. Stiglitz, and contracts theory approach.

3. THE MODEL

Define a production technology of a firm. A manufacturing firm chooses what to produce from a set of final products (goods or services). A production technology for every product p requires a set $N(p)$ of productive workers. Assume every product requires a distinct set of workers, so we can use N as a synonym of a final product. Consider a single-product firm that can choose only one product at a moment.

The revenue function $R(N, z)$ depends on the product N and its output volume z. No matter what product and volume the firm chooses, it bears two types of costs. The first are product-specific costs that do not depend on the internal structure of the firm (these could be raw material costs, marketing expenses, etc). The second are structure-specific costs that (along with the product N) may depend on how the firm has organized its production (e.g. employees' wages). Since the point of this paper is the internal structure problem, suppose the product-specific costs are already accounted for in the revenue function $R(\cdot)$.

The simplest revenue function usually employed in the literature is a linear one: $R(N, z) = \pi(N)\cdot z$ (the firm buys raw materials and sells a final product at a constant price). In our model a bit more complicated revenue function is adopted: $R(N, z) = \pi(N)\cdot\ln(a(N)\cdot z)$ where $a(N)$ and $\pi(N)$ are some product-specific parameters. This function is concave in output and captures the narrowness of a market for any given product. In general, the shape of a revenue function may be more complicated but, as it is shown below, the logarithmic relation greatly simplifies the formal analysis.

Now describe a product N manufacturing technology. In the literature (Williamson O.E., 1967; Calvo G.A., Wellisz S., 1978; Rosen S., 1982; Qian Y., 1994) every worker $w \in N$ is usually assumed to produce a uniform output z_w, so the total output z is just a sum: $z = \sum_{w\in N} z_w$. This approach ignores the complementarity of employees' contributions. At the same time such complementarity is universally recognized (see Milgrom P.R., Roberts J., 1992) to be the main reason for the existence of firms *per se*. In contrast, we adopt an extreme case of very strong complementarity – the Leontief technology $z = \min_{w\in N} z_w$. It supposes every worker to provide a single unit of a local product for a single unit of a final product to be produced (the units of measure for the local outputs are assumed to be chosen accordingly). "Local" outputs are non-substitutable.

Planning in the firm is highly centralized, i.e. the principal (the owner of the firm) chooses the plan of production x to be executed by the firm. However, the worker $w \in N$ affects his output z_w by choosing the effort level $\xi_w \in [0, 1]$ (non-maximal effort $\xi_w < 1$ means some degree of shirking). Workers' effort levels are not directly observed by the principal, so, *monitoring* is needed to build effective incentive schemes for workers (see the discussion in Calvo G.A., Wellisz S., 1978; Qian Y., 1994). This monitoring task is due to the managerial hierarchy built over the set of workers. In the present paper this hierarchy is modeled by a directed tree, with productive workers being its leaves, managers being its intermediate nodes, and the top-manager being its root, while the edges showing subordination.

Let M denote a set of managers in a hierarchy. Every manager has a set of immediate subordinates (they could be workers or other managers). Suppose a manager $m \in M$ has k immediate subordinates. Then let $(\xi_j(m))_{j=1,\dots,k}$ be the vector of manager's m efforts ($\xi_j(m) \geq 0$, $j = 1, \dots, k$), j-th component being the effort referred to monitoring and control of *j*-th immediate subordinate. Along with a monitoring function the manager's effort plays an immediate role in production. It acts upon the output of all the workers who are directly or indirectly (through the chain of managers) controlled by the manager. So if the worker $w \in N$ chooses the effort level ξ_w, his immediate superior m_1 chooses the effort level ξ_1 to control the worker *w*, manager's m_1 superior m_2 chooses the effort level ξ_2 to control m_1, and so on up to the top-manager who chooses the effort level ξ_l, then the output of worker *w* is given by $z_w = x \cdot \xi_w \cdot \xi_1 \cdot \xi_2 \cdot \ldots \cdot \xi_l$.[1]

Now introduce the utility functions of employees. A productive worker $w \in N$ seeks to maximize the difference $u_w = \sigma_w - c(x, \xi_w)$ between his wage σ_w and the cost function $c(x, \xi_w)$ that depends both on a plan (what the worker is expected to do) and the worker's effort level. Such cost function arises naturally as an individual rationality constraint in the presence of labor market – both the worker and the principal know well how high certain responsibilities (plan *x*) and effort levels are valued by market. Similarly, every manager $m \in M$ maximizes the difference $u_m = \sigma_m - K(m, H)$, where $K(m, H)$ is the cost of maintaining a manager *m* in hierarchy *H*.

Costs $K(m, H)$ of a manager *m* may depend both on his position in hierarchy *H* and on the effort levels he exerts. Consider the manager *m* governing (directly or indirectly) a group of workers $s \subseteq N$. Suppose the manager *m* has *k* immediate subordinates that govern groups of workers $s_1, \ldots, s_k$ ($s = s_1 \cup \ldots s_k$) and the manager *m* has chosen the vector of efforts $(\xi_1, \ldots, \xi_k)$ to control them. The costs of the manager *m* may depend both on the set *s* (the larger is the group under control, the more complicated is the task of the manager) and the planned production volume *x* (the control of the execution of a more ambitious plan requires more efforts and costs). The costs must also depend on the span of control *k* (it can be very costly to directly control, for example, 1000 immediate subordinates). Also allow the costs of a manager to depend on how the group *s* is divided among his immediate subordinates. At the end, the costs must increase in manager's efforts. So one can write

$K(m, H) = K(x, s_1, \ldots, s_k, \xi_1, \ldots, \xi_k)$.[2]

Take for simplicity a special shape of a manager cost function, one of that allowing complete analytic calculation of optimal hierarchy attributes. For an arbitrary group of workers *s* define its *measure* by $\mu_s = x|s|$ (it increases both in plan *x* and in group's size $|s|$). We imply that the costs of the manager depend on the groups $s_1, \ldots, s_k$ measures rather than on the groups itself. Consider the constant elasticity of substitution cost function (see McFadden, 1963):

$$K(m,H) = K(\mu_1,\ldots,\mu_k,\xi_1,\ldots,\xi_k) = \left(\sum_{i=1}^{k} \mu_i^{\lambda} (-\ln \xi_i)^{-\delta}\right)^{\varepsilon},$$

where $\lambda \in [0, 1]$, $\varepsilon \in [0, 1]$, and $\delta \in [0, +\infty)$ are parameters (product-dependent in general).

This function satisfies the monotonicity conditions specified hereinabove. Also note the cost approaches infinity as any effort ξ_i tends towards the unity. This implies the impossibility of "total control". The parameter ε accounts for the cost function elasticity with respect to the workload $\sum_{i=1}^{k} \mu_i^{\lambda} (-\ln \xi_i)^{-\delta}$. One can think of $1/\varepsilon$ as of the manager effectiveness measure. The parameter λ describes the elasticity of workload with respect to the size of the group under control. In (Mishin S.P., 2004; Goubko M.V. 2006) λ is interpreted as a degree of standardization of management information in a firm – the less λ is, the more standardized the manager's work is, thus the manager's workload increases more slowly in the size of a unit (problems in big units become "typical"). Lastly, δ accounts for the workload elasticity with respect to the managerial effort.

The wages for all employees are set centrally by the principal on the basis of information elicited from monitoring, so the wage of an employee depends on his observed effort. In general, monitoring may be imperfect so the effectiveness of an incentive scheme σ for an employee may depend on the degree of monitoring inaccuracy. Herein the case of perfect monitoring is considered, i.e. managers elicit true efforts of their immediate subordinates and pass this information to the principal with no distortion. Although not benevolent, managers do not distort the information (in non-cooperative framework), as their compensation does not depend on their reports, but solely depends on their own efforts reported by their immediate superiors. The top-manager is monitored directly by the principal at zero cost.

Therefore, the principal faces a set of separate principal-agent incentive problems with perfect information. It is known (see Mas-Collel A. et al, 1995; Novikov D.A., Petrakov S.N., 1999) that in this setting an optimal incentive scheme gives a zero payment for all but one efforts vector where the compensation is equal to an employee's cost. Thus, the principal can gain any efforts from employees by just compensating for their costs. So the principal merely balances the output (revenue) and the total costs of the employees.

Now the optimal organization design problem can be stated formally: to choose the set of workers (product) *N*, the plan *x*, the hierarchy of managers *H*, and the effort levels for every manager and productive worker to maximize the profit

$$F = R(N, z) - \sum_{w \in N} c(x, \xi_w) - \sum_{m \in M} K(m, H).$$

[1] This is the formula of "cumulative loss-of-control" technology discussed in (Williamson O.E. 1967; Calvo G.A., Wellisz S., 1978; Rosen S., 1982; Qian Y., 1994). So the same argumentation may be used to justify it.

[2] As $s = s_1 \cup \ldots \cup s_k$, the cost also depends upon the whole group *s*. The function changes as the span of control *k* changes, so *k* is also accounted for in this notation.

4. THE RESULTS

For the stylized setting defined above one can completely solve the optimal hierarchy problem. The Leontief technology along with the monotonicity of costs with respect to efforts implies the equality of the local outputs z_w ($w \in N$) in optimal hierarchy. The logarithmic revenue function then enables additive decomposition of the managers' contributions to the profit, so every manager's effort can be optimized separately[3]. Let us denote for short

$$\alpha := \frac{\lambda + \delta}{1+\delta}, \ \beta := \frac{\varepsilon(1+\delta)}{1+\varepsilon\delta}, \ \tau =: \frac{1}{1+\varepsilon\delta}, \text{ and } n = |N|.$$

The following result determines the optimal effort levels of a manager that the principal must elicit.

Lemma. If a manager m in some hierarchy controls groups of workers with measures $\mu_1, \ldots, \mu_k$, then his optimal efforts vector ($\xi_1, \ldots, \xi_k$) and the optimal contribution are given by

$$\xi_i = \exp\left[-\left(\frac{xn(1-\tau)}{\pi(N)\tau}\right)^{\tau} \mu_i^{\alpha-1}\left(\sum_{i=1}^{k}\mu_i^{\alpha}\right)^{\beta-1}\right], \ i = 1, \ldots, k,$$

$$f_m^{\max} = -\frac{1}{\tau}\left(\frac{\pi(N)\tau}{nx(1-\tau)}\right)^{1-\tau}\left(\sum_{i=1}^{k}\mu_i^{\alpha}\right)^{\beta}.$$

See the appendix A for the proof.

One can see the optimal contribution to obey constant elasticity with respect to the size of a unit under the manager's control. This allows us to use the analysis technique for the optimal hierarchy problem developed in (Goubko, 2006). It proves the optimal hierarchy to be uniform (i.e. every manager in a hierarchy has the same span of control) and symmetric (i.e. every manager seeks to divide the subordinate group of workers equally among his immediate subordinates)[4].

The optimal span of control r is then determined as[5]

$$r = \left[\beta(1-\alpha)/(\beta-1)\right]^{1/(1-\alpha\beta)},$$

while the estimate of optimal hierarchy costs (that include the wastes from the loss of control) is given by the following expression (for the most common case when $\alpha\beta \neq 1$):

$$\sum_{m \in M} K^*(m, H) = (nx)^{-(1-\tau)}[nx^{\alpha\beta} - (nx)^{\alpha\beta}] \cdot$$
$$\cdot \pi(N)^{1-\tau}\tau^{-\tau}(1-\tau)^{-(1-\tau)}\frac{r^{\beta}}{r - r^{\alpha\beta}}.$$

Thus, given the product N, a good estimate for the management headcount is $|M| = (n-1)/(r-1)$. Note that it does not depend on the value of plan x and is linear with respect to the number n of productive workers.

Having found the optimal span of control and the managers effort levels[6] one can analytically write down the expression for the profit $F(N, x)$ of a firm as a function of the product N and the plan x (the formula is omitted for short). Thus, planning of N and x becomes a standard optimization problem.

Also one can obtain some comparative static results on how the span of control, managers headcount, salary and efforts distribution depend on the model parameters (the degree of standardization $1/\lambda$ and the managers' ability $1/\varepsilon$).

Simple calculations give a surprising result: the optimal span of control increases with the decrease of standardization – the less standard problems managers solve, the fewer managers the firm must have. The explanation of such span of control behavior is that the manager cost function implies that the less standardization is (the more λ is), the greater is the manager's aspiration to pass the problems to his immediate superior. Thus the workload of top-management inevitably rises while the significance of middle-layer managers falls. So it becomes less costly for the firm to spare of some middle-tier managers even suffering from the top-management overload. The relation between the span of control and the managers' ability is more predictable – the span of control rises with the ability (i.e. with the decrease of ε).

It can also be shown that the monitoring effort increases in the measure μ of the unit under control if $\alpha\beta < 1$ and decreases otherwise. Therefore, the loss of control rises to the top of the hierarchy if $\alpha\beta = (\lambda+\delta)\varepsilon/(1+\varepsilon\delta) > 1$. This "pathological" behavior badly hurts the profit of the firm and, as it is shown below, the inequality $\alpha\beta < 1$ is the condition of the ability of unrestricted growth of the firm.

More precisely, given the linear price law $\pi(N) = \pi \cdot n$, if $\alpha\beta > 1$, then the profit is unimodal in n, so there exists the limit of the firm's growth, otherwise there may be no limit. Both cases are possible with reasonable values of parameters,

[3] Similar decomposition approach is used by J. Geanakolols and P. Milgrom (1991) in their analysis of hierarchical planning with bilinear production costs.

[4] The symmetry of the optimal hierarchy may seem obvious but surprisingly it holds only for a certain range of model parameters (fortunately, the most interesting one). Note that the purely uniform hierarchy may not exist due to the finiteness of the set N. But it is proven in (Goubko M.V., 2006) that, in any case, the optimal hierarchy is "roughly uniform" and the analytic formula for the uniform hierarchy costs is a good estimate for the costs of the optimal hierarchy in a big organization.

[5] This formula presents the estimate span of control. Real optimal span of control is one of the two nearest integers.

[6] The calculation of optimal workers efforts is obvious.

therefore deeper parameters identification is needed to specify the real situation.

At the end investigate the manager's wage dependence on the size of the unit he controls. As the manager's wage compensates his costs in equilibrium, so, it obeys the power law in the size of the unit with the exponent $\alpha\beta$. So, if $\alpha\beta < 1$, the wage is concave in the size of the unit under control (given the plan x), otherwise the wage is convex. From the empiric literature on managerial wages the exponent of managerial wage is known to be in the range [0.2, 0.4]. In most real-world organizations the span of control varies from 4 to 10. These observations along with the formula for the optimal span of control help us to identify the range of potentially relevant parameters λ and ε. The area of interest is defined by the intervals $\lambda \in [0.05, 0.25]$, $\varepsilon \in [1.15, 1.60]$ (given $\delta = 0.1$).

5. PERSPECTIVES

The prospective studies are devoted to the subsequent elimination of the model restrictions.

First, the assumption of a common plan x can be relaxed. We believe that allowing for individual plans x_w for every worker will not change the conclusions (although the formal proof may tangle). Then, every productive worker may be endowed with the individual technology-dependent cost function. The topical question is whether this complication results in the asymmetry of the optimal hierarchy. Also, different types of manager cost functions can be investigated along with giving them a fully-fledged economic explanation.

Imperfect and asymmetric information is known to be one of the main roots of the market failures the hierarchical control must resolve. So the most challenging line of the research is generalization of the model towards accounting for imperfect information. An obvious way is the introduction of imperfect monitoring. In this case the wage of an employee may depend not only on the efforts exerted but also on the accuracy of monitoring that depends on the workload and the efforts of his immediate superior (as shown in simple models of imperfect monitoring by G. Calvo and S. Wellisz, 1978; Y. Qian, 1994). In general, this complication may require the development of advanced techniques for the optimal hierarchy search.

Every manager may be endowed with personal characteristics (*a type*). The principal then faces an adverse selection problem (see Mas-Collel A. et al, 1995) when assigning compensations. A standard incentive compatible scheme then results in information rents. These rents influence a manager's "effective cost" for the principal. The point is how the degree of information asymmetry influences the shape of the optimal hierarchy.

Yet another topical line of the inquiry is the study of incentives decentralization (the situation when the principal gives managers rights and resources to implement incentive schemes for their subordinates) and its impact on management effectiveness. While in the world of perfect information costless mechanisms of such decentralization are possible (see Mishin S.P., 2004; Goubko M.V., 2006), this may not hold in the presence of asymmetric information. The models of incentive contracts decentralization in adverse selection and moral hazard environment do exist in the literature but they restrict attention to the study of the simplest hierarchies (two agents with one principal), and the generalization of these results to the case of complex managerial hierarchies is still a question at issue.

6. CONCLUSIONS

The normative model of optimal hierarchy design in a firm is developed. The model accounts for revenue effects of the size of a firm, employees' costs and efforts, monitoring costs, etc. These features have not been combined before in the models of multi-layer hierarchies.

The results of the analysis include the optimal monitoring efforts subject to a manager's position in a hierarchy, the optimal managerial headcount and the span of control, efficient employees' wages and the optimal profit of a firm. These results allow us to analyze the impact of environment parameters on a firm's size, its financial results, employees' wages and the shape of the optimal hierarchy. The detailed analysis of this impact will help draw up policy recommendations on rational bureaucracy formation in firms, big corporations and holdings. For the specific enterprise the model can answer the following important questions of organizational design:

1. How many managers should an organization employ and how many subordinate workers should these managers have?
2. How much does the maintenance of control system cost?
3. How will the growth of an organization increase the management expenses? Does this growth require radical restructuring the control system?
4. How should an organizational structure change in response to the new management technologies, production modernization and standardization, environment changes, etc.?

REFERENCES

Beckmann M.J. (1960) Some aspects of returns to scale in business administration. *Quarterly J. of Econ.*, **V. 74**, **No 3.**, pp. 464-471.

Calvo G.A., Wellisz S. (1978) Supervision, loss of control and the optimal Size of the Firm. *J. of Political Econ.*, **V. 86, No 5.** pp 943–952.

Geanakoplos J., Milgrom P. (1991) A theory of hierarchies based on limited managerial attention. *J. of Japanese and International Economies*, **Vol. 5(3),** pp 205-225.

Goubko M.V. (2006) *Matematicheskiye modeli optimizatsii ierarkhichekikh struktur*. Moscow, LENAND.

Keren M., Levhari D. (1983) The internal organization of the firm and the shape of average costs. *Bell J. of Econ.*, **Vol. 14. No. 2.** pp 474-486.

Mas-Collel A., Whinston M.D., Green J.R. (1995) *Microeconomic theory*. N.Y.: Oxford Univ. Press.

McFadden D. (1963) Constant elasticity of substitution production functions. *Review of Econ. Studies*, **Vol. 30, No. 2.** pp 73-83.

Meagher K.J (2003) Generalizing incentives and loss of control in an optimal hierarchy: the role of information technology. *Economic Letters*, **78**, pp 273-280.

Milgrom P.R., Roberts J. (1992) *Economics, organization, and management*. Englewood Cliffs.

Mintzberg H. (1983) *Structure in fives: designing effective organizations*. N.J.: Prentice-Hall, Co.

Mishin S.P. (2004) *Optimal hierarchies in firms*. Moscow, PMSOFT.

Novikov D.A., Petrakov S.N. (1999) *Kurs teorii activnykh system*. Moscow, SINTEG.

Patacconi A. (2005) Optimal coordination in hierarchies. *mimeo*. University of Oxford, June.

Qian Y. (1994) Incentives and loss of control in an optimal hierarchy. *Review of Econ. Studies*, **Vol. 61, No 3**, pp 527-544.

Rosen S. (1982) Authority, control, and the distribution of earnings. *Bell J. of Econ.* **Vol. 13. No. 2.** pp. 311-323.

Williamson O.E. (1967) Hierarchical control and optimum firm size. *J. of Political Econ.*, **Vol. 75, No. 2**, pp 123-138.

Williamson O.E. (1975) *Markets and hierarchies*. New York: Free Press.

Appendix A. THE PROOF OF THE LEMMA

Given the Leontief technology $z = \min_{w\in N} z_w$ the economy on nonproductive costs requires the optimal efforts to equalize the outputs of all workers, so every z_w must be equal to z.

Identical transformation then yields:

$$F = \frac{\pi(N)}{n}\sum_{w\in N}[\ln a(N)x + \ln\xi_w + \sum_{j=1}^{l(w)}\ln(\xi_j(w))] -$$
$$\sum_{w\in N} c(x,\xi_w) - \sum_{m\in M}\left(\sum_{i=1}^{k(m)}\mu_i(m)^{\lambda}(-\ln\xi_i(m))^{-\delta}\right)^{\varepsilon}.$$

Note that if the immediate subordinates of the manager m in the hierarchy H control the groups of workers $s_1(m), \ldots, s_k(m) \subseteq N$, then the effort level $\xi_i(m)$ of manager m is accounted $|s_i(m)|$ times in the first sum of the above formula. So one can regroup the elements of managerial efforts and write

$$F = \pi(N)\ln(a(N)x) + \sum_{w\in N}\left[\frac{\pi(N)}{n}\ln\xi_w - c(x,\xi_w)\right] +$$
$$+\sum_{m\in M}[\frac{\pi(N)}{n}\sum_{i=1}^{k(m)}|s_i(m)|\ln\xi_i(m) -$$
$$-\left(\sum_{i=1}^{k(m)}\mu_i(m)^{\lambda}(-\ln\xi_i(m))^{-\delta}\right)^{\varepsilon}].$$

Remember that by definition $\mu_i(m) := x|s_i(m)|$ so, finally we obtain the following expression for the profit:

$$F = \pi(N)\ln(a(N)x) + \sum_{w\in N}\left[\frac{\pi(N)}{n}\ln\xi_w - c(x,\xi_w)\right] +$$
$$(1)\quad \sum_{m\in M}[\frac{\pi(N)}{nx}\sum_{i=1}^{k(m)}\mu_i(m)\ln\xi_i(m) -$$
$$-\left(\sum_{i=1}^{k(m)}\mu_i(m)^{\lambda}(-\ln\xi_i(m))^{-\delta}\right)^{\varepsilon}].$$

The problem is to maximize (1) by choosing the efforts levels given the outputs $z_w = x\cdot\xi_w\cdot\xi_1(w)\cdot \ldots\cdot\xi_{l(w)}(w)$ for all workers being equal. Below these constraints are omitted in the optimization. The unconstrained efforts levels for all managers are determined. Then the optimal hierarchy supported by these effort levels is found. Finally, this optimal hierarchy can be shown to obey symmetry, so the constraints of the local outputs equality are automatically satisfied in it.

The profit (1) is additive in the contributions of any worker or manager, so one can find the optimal efforts separately for any employee. Let k immediate subordinates of some manager m in the hierarchy H control the groups of workers with measures $\mu_1, \ldots, \mu_k$. Then to find the optimal effort levels $\xi_1, \ldots, \xi_k$ of manager m one must maximize the contribution f_m of this manager to the profit of the firm. From (1) the formula for the contribution is obtained:

$$(2)\quad f_m = \frac{\pi(N)}{nx}\sum_{i=1}^{k}\mu_i\ln\xi_i - \left(\sum_{i=1}^{k}\mu_i^{\lambda}(-\ln\xi_i)^{-\delta}\right)^{\varepsilon}.$$

From the first-order conditions the optimal efforts and the optimal contribution of a manager are calculated:

$$\xi_i = \exp\left[-\left(\frac{xn(1-\tau)}{\pi(N)\tau}\right)^{\tau}\mu_i^{\alpha-1}\left(\sum_{i=1}^{k}\mu_i^{\alpha}\right)^{\beta-1}\right],\ i = 1,\ldots,k,$$
$$f_m^{\max} = -\frac{1}{\tau}\left(\frac{\pi(N)\tau}{nx(1-\tau)}\right)^{1-\tau}\left(\sum_{i=1}^{k}\mu_i^{\alpha}\right)^{\beta}.[7]$$

[7] Remember the definitions of α, δ, and τ introduced above.