Hierarchy Design with Socialism in Internal Capital Markets

Eric S. Chou

Department of Economics,

National Tsing Hua University,

Hsinchu, Taiwan, R.O.C.

August 14, 2007

Abstract

This paper compares the efficiency of flat and tall hierarchies from the perspective of 'socialism in internal capital markets' (SICM) – a recently documented problem of multi-segment firms in which high-profit segments tend to be underinvested and low-profit segments tend to be overinvested. SICM is characterized with the possibility of *divisionalization* – grouping elementary business segments into divisions and delegating decisions to division managers, which transforms flat hierarchies into tall hierarchies. We found that divisionalization aggravates the problem of SICM, suggesting that U-form (undivisionalized) organizations can outperform M-form (divisionalized) organizations – a result that is in contrast to Williamson's "M-form hypothesis." In addition, the SICM problem for a given divisionalized firm can be alleviated by redirecting more managerial ownership from high-tier managers to low-tier managers.

Keyword: M-form hypothesis, Hierarchy, Internal Capital Markets

JEL Classification No: G34. L22 (G3, G31)

1 Introduction

Understanding why and how organization forms matter has been a major research agenda in economics and finance. Although most organizations are hierarchical in general, they differ in many detailed respects. The primary focus of this paper is hierarchy height. Organizations can be categorized into 'flat' hierarchies – those with less layers of decision makers – and 'tall' hierarchies – those with more layers of decision makers. Note that flat hierarchies can be transformed into tall hierarchies through 'divisionalization' – grouping basic segments into divisions and delegating decision rights to division managers.

In the designing of hierarchies, efficiency in *capital allocation*, among many other activities, has long been a major concern.[1] In his seminal 'M-form hypothesis', Williamson [30] argues that multi-division (M-form) structures should be more efficient than unitary (U-form) structures in capital allocation. His main argument is that, with divisionalization, M-form structures can 'economize on bounded rationality,' and hence enjoy lower information, fine-tuning, and displacement costs. In other words, M-form structures can operate as 'miniature capital markets' and have

[1]The existing literature on hierarchies addresses many other important issues. For example, the height of a hierarchy can affect the efficiency of information processing (e.g. Bolton and Dewatripont[4]), intra-firm bargaining (Stole and Zwiebel [26]), capital budgeting (Harris and Raviv [12]), utilizing skilled workers (Beggs[1]), coordination and specialization (Hart and Moore [14]), protection of the sources of organizational rents (Rajan and Zingales[22]), and generating information and allocating capital (Stein[27]).

a better ability to assign capital to high-yield uses. Recently, this view that M-form structures have an information advantage has been echoed by Maskin, Qian and Xu [17], who argue that one of the advantages of M-form structures is that they can facilitate relative performance evaluation. Among many others, the degree of *delegation* is an important dimension on which M-form and U-form structures differ because more decisions are usually delegated down in the hierarchies in M-form structures. The current paper will shed new light on the M-form hypothesis by focusing on this particular distinction.

While the aforementioned view of the firms as miniature capital markets has been very successful in explaining merger waves and the formation of conglomerates, it is hard to reconcile with the recently identified 'diversification discounts' problem(see, for example Lang and Stulz [16] and Berger and Ofek [2]), in which diversified firms (operators of miniature capital markets) tend to be valued at a discount compared to a portfolio of compatible specialized firms. In addition, many studies (including Rajan, Servaes and Zingales [20], Scharfstein [23], and Shin and Stulz [25]) find that diversification discounts might be caused by a peculiar pattern of capital allocation in multi-segment firms – *low*-profit segments tend to be *overinvested* and *high*-profit segments tend to be *underinvested*. This is what Bolton and Scharfstein [5] refer to as 'socialism in internal capital markets' (henceforth SICM). Furthermore, Scharfstein finds that SICM is more severe if the firm is more diversified or the CEO's ownership of the firm is lower.[2]

In addition to the above findings that challenge the firms' ability in capital allocation, it has also been observed that corporate hierarchies are getting flatter (Rajan and Wulf [21] and Buble [6]). In fact, in the wave of 'corporate reengineering' or 'downsizing' that was inspired by Hammer and Champy [11], firms eliminated middle-layer managers. Regarding M-form structures, Itoh [15] finds that in Japan, M-form structures are not suitable for diversified firms. This implies *a potential link between the inefficiency of M-form structures and SICM* because diversified firms are exactly where SICM arises.

In sum, there should be a special cost for delegating capital-allocation decisions down the hierarchy. Understanding this cost is important because it seems to be significant enough to dominate the well-known informational benefit that is suggested by the M-form hypothesis. This paper starts the investigation from the particular perspective of SICM by asking the following questions. *Does the cost of SICM increase with hierarchy height? Does the M-form hypothesis hold under SICM?*

[2]There is still much active discussion in this literature. For example Graham, Lemmon, and Wolf [10] argue that the values of diversified firms are discounted because they are discounted before the diversification. In addition, Whited [28] shows that the finding of inefficient resource allocation might be due to a measurement error of the proxy for profitability – Tobin's q. However, Burch and Nanda [7] finds diversification discounts using evidence from spinoffs that are hard to explain by selection bias or measurement error.

We consider a firm with three different business segments that can be organized as a 'flat' hierarchy, which consists of one CEO directly managing all three segments, or as one of three different 'tall' hierarchies, each consisting of one division manager who directly manages two segments and then a CEO who manages the division manager and the segment that is not included in the division. Both the CEO and the division manager are in charge of 1) allocating capital to their direct subordinates and 2) retaining their direct subordinates by paying them cash. In addition, assume that the CEO and the division manager can divert cash saved in retaining her direct subordinates. Also, each segment manager cares about the cash that he receives and 'empire building,' which is modeled as an additive part of each segment manager's utility that is increasing and *concave* in the capital his segment gets allocated. Furthermore, following the existing literature, we consider production functions under which the production-profit-maximizing allocation for each segment is increasing in the segment's productivity.[3] Given this allocation, the manager of a more profitable segment will have a lower marginal utility with respect to capital, due to the concavity of empire-building utility. Consequently, the total cash payment can be reduced without altering the utility level of either segment managers if some capital is redirected from the segment manager of a more profitable segment to the segment

[3]This is in accordance with basic investment theories and adopted previously by Chou [9] and Scharfstein and Stein [24].

manager of a less profitable segment. Hence, this deviation from the production-profit-maximizing allocation – a cross subsidy from a more profitable segment to a less profitable one – can bring a benefit (at the cost of lowered production profit).

Why is SICM inefficient? Because the CEO can divert cash saving, he will enjoy 100% of the benefit. However, he only bears a portion of the cost because he is not the owner. Hence, the CEO will abuse the cross subsidy. This is the rationale for the inefficient SICM proposed by Chou [9] and essentially what happens under the flat hierarchy in the current model.

Why does the cost of SICM increase with hierarchy height? Now suppose that the overall managerial ownership remains the same (say 20% of the firm) and consider any of the *tall* hierarchies in which two of the segments are grouped into a division. The benefit of SICM among these divisionalized segments will now accrue to the division manager and then the CEO when the retention constraints of the division and segment managers are binding. Suppose that each segment has the same allocation as it does in the flat hierarchy. If the division manager owns 100% of the *managerial ownership* (which equals to 20% of the firm), she will not alter the allocation because she faces the same trade off between the cost and benefit of SICM in the division as the CEO in the flat hierarchy. The CEO will not alter the allocation, either, because all of the benefit that accrues to the division manager also accrues to the

CEO, and hence the CEO in the tall hierarchy faces the same trade-off as the CEO in the flat hierarchy. Consequently, the capital allocation under the tall hierarchy is identical to that under the flat hierarchy in the extreme case when the division manager owns 100% of the managerial ownership. However, in generic cases when the division manager owns only a portion of the managerial ownership, SICM will be more serious under tall hierarchies. The reason is that even the CEO still faces the same trade-off, the division manager now cares less about the cost of SICM because she bears a smaller portion (less than 20%) of it.

The result that delegation aggravates the problem of SICM helps to explain the recent decreased layers of corporate hierarchies, and links that phenomena with the malfunction of internal capital markets. It also provides a view of organizations that is opposite but complementary to the M-form hypothesis. Note that this paper focuses on delegation as the key distinction between the U-form and M-form structures. We do not claim that M-form structures are less efficient than U-form structures in any other dimensions.

Quite a few studies try to explain why SICM arises. Most of them, such as those of Rajan, Servaes and Zingales [20] and Scharfstein and Stein [24], follow what we refer to as 'the rent-seeking approach,' which is broadly defined as models with the assumption that there is some socially wasteful activity that segment managers can engage in to seek personal rent. Relatively little is known about how SICM affects hierarchy design, which is the main goal of the current paper. To generate SICM, we follow a non-rent-seeking model of SICM by Chou [9] because attaching an auxiliary rent-seeking game to the main capital allocation game can render the results sensitive to the detailed design of the rent-seeking game (as noted by Scharfstein and Stein [24], page 2546) and is unnecessary (as shown by Chou [9]).

The rest of the paper is organized as follows. Section 2 sets up the model, and Section 3 analyzes it and presents the results. Section 4 concludes the paper.

2 Model Setup

2.1 Assumptions and Definitions:

The model is setup with general functional forms in this section and solved with specialized quadratic forms in the next. Consider a firm that consists of three business segments, indexed by i, $i \in \{1, 2, 3\}$. Denote the revenue function of segment i as $V_i(k_i)$, where k_i is the input of capital in segment i. Each V_i satisfies the following assumption.

Assumption 1: $V_i : \mathbf{R} \to \mathbf{R}$, $i \in \{1, 2, 3\}$, *is a single-picked, continuously differentiable, and strictly concave function that satisfies* $V_i(0) = 0$, $-\infty < V_i''(k) <$

0, and

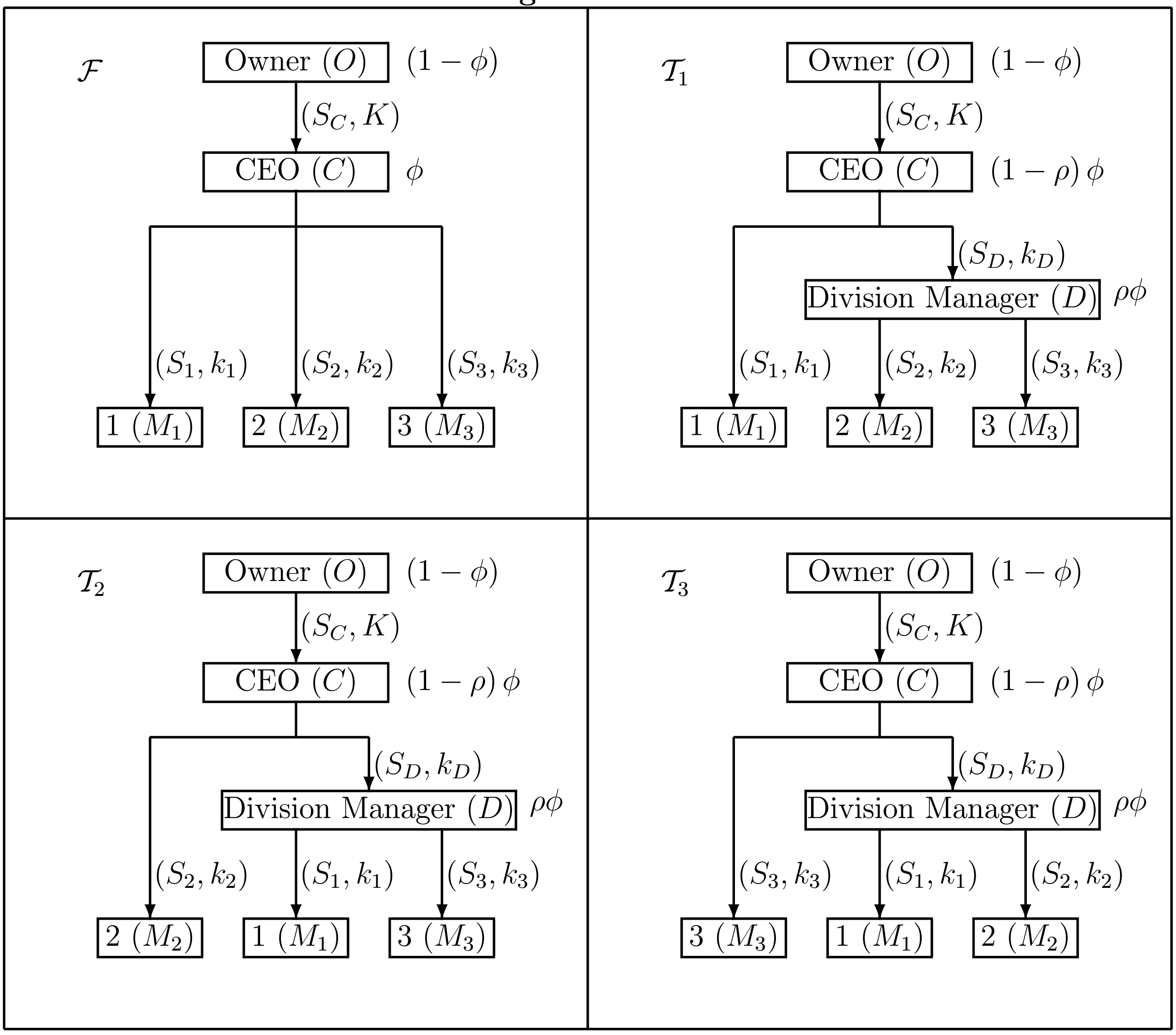
$$V_3'(k) < V_2'(k) < V_1'(k). \tag{1}$$

Inequality (1) reflects the diversity among the three segments, which is modeled by the differences in their marginal revenues. By this construction, we can say that segment 1 is the most profitable, segment 3 is the least profitable, and segment 2 is in between. The total *production* profit of the firm is $V_1(k_1)+V_2(k_2)+V_3(k_3)-k_1-k_2-k_3$ and denoted as $\Phi(k_1, k_2, k_3)$.

Players can be categorized by the four different roles: an 'owner' (henceforth O), a 'CEO' (henceforth C), a 'division manager' (henceforth D), and three 'segment managers' (henceforth M_i, $i \in \{1, 2, 3\}$). While O, C, and M_i's are assumed to be indispensable to the firm, the existence of D depends on whether there is divisionalization. For ease of reference, denote the segment that is directly managed by C (the 'independent' segment) as segment l and the segments that are grouped into a division and managed by D (the divisionalized segments) as segments m and n, with $n > m$ so that segment n (m) represents the less (more) profitable one of the divisionalized segments.

With three segments, there are four possible hierarchies, $\mathcal{F}$, $\mathcal{T}_1$, $\mathcal{T}_2$, and $\mathcal{T}_3$, as shown in Figure 1. Among them, $\mathcal{F}$ is the only *flat* hierarchy, representing U-form structures (organizations without divisionalization), whereas $\mathcal{T}_1$, $\mathcal{T}_2$, and $\mathcal{T}_3$ are *tall* hierarchies, representing M-form structures (organizations with divisionalization). When there is no divisionalization and hence D is not employed, the firm consists of a two-tier hierarchy (as shown by $\mathcal{F}$), with O at the top as C's principal, C at the middle as M_i's principal, and M_i's at the bottom as ultimate agents. When there is divisionalization and hence D is hired, the firm becomes a three-tier hierarchy (as shown by $\mathcal{T}_1$, $\mathcal{T}_2$, or $\mathcal{T}_3$), with O at the top as C's principal, C at the second level as the principal of D and M_l, D at third level as the principal of M_m and M_n, and all M_i's at the bottom as simply agents.

Figure 1.



All players except M_i's are potential shareholders of the firm.[4] Denote the total managerial ownership, the sum of shares owned by C and D, as $\phi < 1$. Consequently,

[4]For simplicity, we assume that M_i's have no ownership of the firm. Relaxing this assumption will not change the result because, as it turns out, it is the total managerial ownership and its distribution between the *decision makers*, C and D, that matter.

the fraction of shares that is owned by O is $1-\phi$. Denote the fraction of managerial ownership of D as ρ, so the ownerships of C and D are $(1-\rho)\,\phi$ and $\rho\phi$, respectively. Both ρ and ϕ are exogenous parameters. Arguably, they can be the decision variables of O. However, in practice, there could be many exogenous reasons why the owner of a firm is refrained from freely rewarding the management with shares. One example is that the owner might want to maintain her status as a majority shareholder. In addition, endogenizing ρ and ϕ renders the model intractable and should not bring any qualitative changes to the results, as suggested by Chou[9] in a simpler model. Note that 'ownership' here is defined as the shares people own that give them the 'residual income rights' but not 'residual control rights', as the separation between the two is one of the key driving forces of SICM (Scharfstein [23]).[5]

While O, C, D, and M_i's are risk neutral and seek to maximize their individual wealth (the sum of the values of cash and shares), M_i also enjoys a *private benefit*, $b(k_i)$, from managing a segment with capital k_i.[6] This private benefit reflects the *empire building tendency* of the segment managers and satisfies the following assumption.

[5]For more discussion on these rights, see Bolton and Scharfstein [5] and Hart [13].

[6]If the private benefit were to be dependent directly on production, i.e. $V_i\,(k_i)$, then, as shown in Scharfstein and Stein [24], an auxiliary rent-seeking game is needed to generate SICM, and hence, as discussed above,.rendering a less robust SICM foundation for the current analysis. The possibility that O, C, or D might also enjoy private benefits from capital will not change the current results qualitatively, and is ruled out for simplicity.

Assumption 2: $b : \mathbf{R} \to \mathbf{R}$ is strictly concave and increasing within the range of possible equilibrium allocation.

O alone owns an unlimited endowment of cash and capital[7], whereas C, D, and M_i's have no endowment of their own. In each tier of principal-agent relationship except the top one, the principal decides how much cash and capital to give his/her *direct* agents out of the cash and capital that is provided by the principal's principal. In the top tier, the principal (O) provides cash and capital out of her own endowment. To generate the scope of capital misallocation in general, we follow Scharfstein and Stein [24] in the following three assumptions.

Assumption 3: *The amounts of cash and capital that a principal provides to each of his/her agents are not contractible to the principal's principal.*

Assumption 4: *A principal can divert any remaining cash that is not paid out to his/her direct agents.*

Assumption 5: *The decision rights to allocate capital to a particular segment and to retain the manager of that segment cannot be separated.*

Assumptions 3, 4, and 5 are essential and commonly used ingredients in modeling SICM (see, for example, Chou [9] and Scharfstein and Stein [24].) Admittedly, they are all very strong assumptions in the sense that SICM will disappear as soon

[7] One can imagine O as a group of dispersed shareholders who act cooperatively.

as any one of them fails to hold.[8] However, SICM does not follow directly from these assumptions. Moreover, as mentioned earlier, the goal of this paper is to gain new insights from the particular but empirically relevant perspective of SICM. More precisely, the aim is to evaluate the performance of different hierarchies *under environments in which SICM is present*, which is hard to achieve without invoking these three assumptions.

Next, the benchmark and potential equilibrium situations will be defined. Denote the cash that C, D, and M_i are paid as, S_C, S_D, and S_i, respectively. According to the setup above, the utility functions for O, C, D, and M_i's are, respectively

$$\begin{aligned} U^O &\equiv (1-\phi)\left[\Phi(k_l,k_m,k_n)-S_C\right], \\ U^C &\equiv S_C - S_D - S_l + (1-\rho)\,\phi\left[\Phi(k_l,k_m,k_n)-S_C\right], \\ U^D &\equiv S_D - S_m - S_n + \rho\phi\left[\Phi(k_l,k_m,k_n)-S_C\right], \text{ and} \\ U^{M_i} &\equiv S_i + b(k_i). \end{aligned}$$

The first-best capital allocation is defined as follows.

[8]For example, allowing incentives contracts in the lines of Bernardo *et al* [3] will result in underinvestment in both segments.

Definition 1 *The first-best capital allocation is* (k_1^*, k_2^*, k_3^*)*, such that*

$$k_i^* = \arg\max_{k_i}[V_i(k_i) + b(k_i) - k_i], \text{ for } i = 1, 2, 3.$$

To rule out non-instructive corner solutions where C, D, or M_i's receive zero net cash, we also make the following assumption.

Assumption 6: *The reservation values of the CEO, the division manager, and the segment managers are high enough (in particular, strictly greater then zero) so that their participation constraints cannot be satisfied without being paid a positive net amount of cash.*

In addition to capital allocation, the comparison between $\mathcal{F}$ and $\mathcal{T}_l$ also depends on the reservation values of C and D, given Assumption 6. Because the firm needs to hire a division manager under $\mathcal{T}_l$ but does not need to do so under $\mathcal{F}$, the participation constraint of D imposes a fixed cost on $\mathcal{T}_l$, but not on $\mathcal{F}$. Because this fixed cost has nothing to do with the important issue we are interested in in this paper – capital allocation, we made the following assumption to avoid its impact on the presentation of the results.

Assumption 7:$\overline{u}_C^{\mathcal{T}_l} + \overline{u}_D = \overline{u}_C^{\mathcal{F}}$.

This assumption is not meant to be practical, even remotely. Rather, it is made only for clear presentation of the result (Proposition 4). Removing it will not change any result except Proposition 4 because we focus on capital allocation and the assumption only affects the wages. Without Assumption 7, the result that divisionalization destroys values in Proposition 4 continues to hold for a certain range of values of $\overline{u}_C^{\mathcal{I}_l}$, $\overline{u}_D$, and $\overline{u}_C^{\mathcal{F}}$. However, uninteresting cases under which divisionalization can create value when it happens that hiring a CEO and a division manager is substantially cheaper than hiring only a CEO (that is when $\overline{u}_C^{\mathcal{I}_l} + \overline{u}_D$ is very small compared with $\overline{u}_C^{\mathcal{F}}$) may arise.

2.2 The Flat Hierarchy ($\mathcal{F}$)

Under $\mathcal{F}$, C directly manages three segment managers without the help of D. Hence, in this case, ϕ is simply the ownership of C. Denote $\mathbf{S} \equiv (S_1, S_2, S_3)$, $\mathbf{k} \equiv (k_1, k_2, k_3)$. Denote also K and S_C as the overall capital investment and the overall cash investment, respectively. The equilibrium capital allocation is the solution to the following two-tier principal-agent problem:

$$\max_{(S_C, K)} (1-\phi)\left[\Phi(k_l, k_m, k_n) - S_C\right] \tag{2}$$

subject to

$$S_C - \sum_{i=1}^{3} S_i + \phi\left[\Phi(k_l, k_m, k_n) - S_C\right] \geq \overline{u}_C^{\mathcal{F}} \tag{3}$$

$$(\mathbf{S}, \mathbf{k}) = \arg\max_{(\widetilde{\mathbf{S}}, \widetilde{\mathbf{k}})} \left[S_C - \sum_{i=1}^{3} \widetilde{S}_i + \phi \left[\Phi(k_l, k_m, k_n) - S_C \right] \right] \tag{4}$$

subject to

$$\sum_{i=1}^{3} \widetilde{S}_i \leq S_C \tag{5}$$

$$\sum_{i=1}^{3} \widetilde{k}_i \leq K \tag{6}$$

$$\widetilde{S}_i + b(\widetilde{k}_i) \geq \overline{u}_i \text{ for } i \in \{1, 2, 3\}, \tag{7}$$

with all decision variables non-negative and $\overline{u}_C^{\mathcal{F}}$ and $\overline{u}_i$ representing the reservation values of C and the manager of segment i, respectively.

Because (6) will bind and (5) will be slack under Assumption 6, C's problem for a given K ((4)-(7)) can be simplified into finding k_1 and k_2 that satisfy

$$\max_{k_1, k_2} b(k_1) + b(k_2) + b(K - k_1 - k_2) + \phi \left[V_1(k_1) + V_2(k_2) + V_3(K - k_1 - k_2) \right]. \tag{8}$$

Before proceeding further, we turn to the case of $\mathcal{T}_l$ in section 2.3 and examine the main difference divisionalization makes in section 2.4.

2.3 The Tall Hierarchy ($\mathcal{T}_l$)

For $\mathcal{T}_l$, the equilibrium allocation is the solution of the following three-tier principal-agent problem,

$$\max_{(S_C,K)} (1-\phi)\left[\Phi(k_l, k_m, K-k_l-k_m) - S_C\right] \tag{9}$$

subject to

$$S_C - S_D - S_l + (1-\rho)\,\phi\left[\Phi(k_l, k_m, K-k_l-k_m) - S_C\right] \geq \overline{u}_C^{\mathcal{T}_l} \tag{10}$$

$$(S_D, S_l, k_l) = \arg\max_{\left(\widetilde{S_D},\widetilde{S}_l,\widetilde{k}_l\right)} S_C - \widetilde{S_D} - \widetilde{S}_l + (1-\rho)\,\phi\left[\Phi(k_l, k_m, K-k_l-k_m) - S_C\right] \tag{11}$$

subject to

$$\widetilde{S_D} + \widetilde{S}_l \leq S_C \tag{12}$$

$$\widetilde{S_D} - S_m - S_n + \rho\phi\left[\Phi(k_l, k_m, K-k_l-k_m) - S_C\right] \geq \overline{u}_D \tag{13}$$

$$\widetilde{S}_l + b(\widetilde{k}_l) \geq \overline{u}_l$$

$$(S_m, S_n, k_m) = \arg\max_{\left(\widetilde{S_m},\widetilde{S_n},\widetilde{k_m}\right)} \widetilde{S_D} - \widetilde{S_m} - \widetilde{S_n} + \rho\phi\left[\Phi(k_l, \widetilde{k_m}, K-\widetilde{k}_l-\widetilde{k_m}) - S_C\right] \tag{14}$$

subject to

$$\widetilde{S_m} + \widetilde{S_n} \leq \widetilde{S_D} \tag{15}$$

$$\widetilde{S_m} + b(\widetilde{k_m}) \geq \overline{u}_m \text{ and } \widetilde{S_n} + b(K - \widetilde{k}_l - \widetilde{k_m}) \geq \overline{u}_n, \tag{16}$$

with all decision variables non-negative and $\overline{u}_C^{\mathcal{T}_l}$, $\overline{u}_D$, and $\overline{u}_i$ representing the reservation values of C, D and the manager of segment i, respectively.

To illustrate the decision rule of C under a tall hierarchy – the central driving force of this paper, we solve the program above for $\mathcal{T}_1$ as an example. Under $\mathcal{T}_1$, C's problem for any given K ((11)-(16)) can be simplified as finding

$$\max_{k_1,k_2} b(k_1) + b(k_2) + b(K - k_1 - k_2) + \phi \left[V_1(k_1) + V_2(k_2) + V_3(K - k_1 - k_2) \right]. \tag{17}$$

subject to

$$\rho\phi \left[V_2'(k_2) - V_3'(K - k_1 - k_2) \right] + b'(k_2) - b'(K - k_1 - k_2) = 0, \tag{18}$$

where (18) is the first order condition for D's problem ((14)-(16)).

2.4 The Novel Effect of Divisionalization

Because the setup of the current model is similar to that of Chou [9], our main result that *divisionalization* destroys values may seem like a direct corollary of the main result of Chou [9] that *integration* destroys values. However, comparing C's problem under $\mathcal{F}$, that is problem (8), and C's problem under $\mathcal{T}_1$, that is problem (17)-(18), reveals the relative novelty of this paper compared with Chou [9]. While the former is a simple extension of Chou [9] into a three-segments environment, the latter is a new problem C faces under divisionalization. The two differs in the addition constraint (18) of the latter.[9] Clearly, given any K, C will be worse off under $\mathcal{T}_1$. However, this by no means implies that the *owner* necessarily will be worse off under $\mathcal{T}_1$, too. In fact, as pointed out in Chou [9] and perhaps can be seen from the analysis so far, there are two main factors behind SICM: (a) diversity in the segments' profitability and (b) separation between ownership and control. Given that the profitability of the three segments (pertaining to (a)) and managerial ownership (pertaining to (b)) are the same across all hierarchies, it is not clear how divisionalization affects either (a) or (b). For example, is (b) more serious in a flat hierarchy under which C owns 60% of the firm or in a tall hierarchy under which C owns 40% and D owns 20%?

From a technical point of view, switching from the flat hierarchy to any of the tall hierarchy entails switching from a two-tier principal-agent problem to a three-tier one. It is like splitting up the tasks that originally belong to one agent (C under the flat hierarchy) among two agents (C and D under the tall hierarchy). Without formal analysis, one simply can not rule out the possibility that the extra tier of

[9] As discussed after Proposition 3, this constraint is binding.

principal-agent problem can refrain the existing agency problem somehow so that it benefits the ultimate principal (the owner).

3 The Quadratic Specialization

While most of the results can be obtained with the general functional forms specified in Assumptions 1 and 2, we focus on the following quadratic specialization so that the results can be presented clearly.

$$V_1(k) \equiv -ak^2 + bk + \alpha k,\ V_2(k) \equiv -ak^2 + bk,\ V_3(k) \equiv -ak^2 + bk - \alpha k \quad (19)$$

$$b(k) \equiv -ck^2 + dk, \text{ where } a, b, c, d, \alpha > 0 \text{ and} \quad (20)$$

$$\frac{d}{c} > \frac{b+\alpha}{a}. \quad (21)$$

Clearly, (19) and (20) satisfy the strict concavity required in Assumptions 1 and 2. In addition, the quadratic form of $b(\cdot)$ in (20) will not violate the monotonicity required in Assumption 2 because (21) guarantees that $b(k)$ never peaks before $V_1(k)$, $V_2(k)$, or $V_3(k)$ does, and hence $b(k)$ is monotonic increasing within the possible range of equilibrium allocation.

Under $\mathcal{F}$, the solution needs to satisfy the following first order conditions of (8)

$$\phi\left[V_1'(k_1) - V_3'(K - k_1 - k_2)\right] + b'(k_1) - b'(K - k_1 - k_2) = 0 \quad (22)$$

$$\phi\left[V_2'(k_2) - V_3'(K - k_1 - k_2)\right] + b'(k_2) - b'(K - k_1 - k_2) = 0. \quad (23)$$

In addition, if we let $f(k_1, k_2)$ denote the objective function of (8), then the elements of the Hessian $D^2 f(k_1, k_2) = \begin{pmatrix} f_{k_1k_1} & f_{k_2k_1} \\ f_{k_1k_2} & f_{k_2k_2} \end{pmatrix}$ can be obtained as follows

$$\begin{aligned} f_{k_1k_1} &= \phi\left[V_1''(k_1) + V_3''(K - k_1 - k_2)\right] + b''(k_1) + b''(K - k_1 - k_2) \\ f_{k_1k_2} &= f_{k_2k_1} = \phi\left[V_3''(K - k_1 - k_2)\right] + b''(K - k_1 - k_2) \\ f_{k_2k_2} &= \phi\left[V_2''(k_2) + V_3''(K - k_1 - k_2)\right] + b''(k_2) + b''(K - k_1 - k_2) \end{aligned}$$

Because $f_{k_1k_1} < 0$ and $f_{k_1k_1}f_{k_2k_2} - f_{k_1k_2}f_{k_2k_1} > 0$ given that $V_i(\cdot)$'s and $b(\cdot)$ are strictly concave (which implies that $f_{k_1k_1} < f_{k_1k_2} < 0$ and $f_{k_2k_2} < f_{k_2k_1} < 0$), $D^2 f(k_1, k_2)$ is a negative definite symmetric matrix, and hence the second order conditions for (8) are satisfied. Therefore, under (19) and (20), (22) and (23) imply that the best response of C given any K is

$$k_1 = \frac{1}{3}K + \frac{\alpha\phi}{2c + 2a\phi}, \; k_2 = \frac{1}{3}K \text{ and } k_3 = \frac{1}{3}K - \frac{\alpha\phi}{2c + 2a\phi}. \quad (24)$$

Solving (2) by plugging (24) into it yields that $K = \frac{3(b+d-1)}{2(c+a)} = K^*$.

Under $\mathcal{T}_1$, (18) implies that the best response of D given any k_D is

$$k_2 = \frac{1}{2}k_D + \frac{\alpha\phi\rho}{4(c+a\phi\rho)} \text{ and } k_3 = \frac{1}{2}k_D - \frac{\alpha\phi\rho}{4(c+a\phi\rho)}. \tag{25}$$

This can be plugged into C's problem and transforms it into finding k_D which solves

$$\begin{aligned} & \max_{k_D} b(k_1) + b(k_2) + b(k_3) + \phi\left[V_1(k_1) + V_2(k_2) + V_3(k_3)\right] \\ = \quad & \max_{k_D} b(K - k_D) + b\left(\frac{1}{2}k_D + \frac{\alpha\phi\rho}{4(c+a\phi\rho)}\right) + b\left(\frac{1}{2}k_D - \frac{\alpha\phi\rho}{4(c+a\phi\rho)}\right) \\ & +\phi\left[V_1(K - k_D) + V_2\left(\frac{1}{2}k_D + \frac{\alpha\phi\rho}{4(c+a\phi\rho)}\right) + V_3\left(\frac{1}{2}k_D - \frac{\alpha\phi\rho}{4(c+a\phi\rho)}\right)\right]. \end{aligned}$$

The corresponding first order condition is

$$\begin{aligned} & \phi\left[\frac{1}{2}V_2'\left(\frac{1}{2}k_D + \frac{\alpha\phi\rho}{4(c+a\phi\rho)}\right) + \frac{1}{2}V_3'\left(\frac{1}{2}k_D - \frac{\alpha\phi\rho}{4(c+a\phi\rho)}\right) - V_1'(K - k_D)\right] \\ & +\frac{1}{2}b'\left(\frac{1}{2}k_D + \frac{\alpha\phi\rho}{4(c+a\phi\rho)}\right) + \frac{1}{2}b'\left(\frac{1}{2}k_D - \frac{\alpha\phi\rho}{4(c+a\phi\rho)}\right) - b'(K - k_D) = 0, \end{aligned} \tag{26}$$

which implies the following solution for k_D.[10]

$$k_D = \frac{2}{3}K - \frac{\alpha\phi}{2(c + a\phi)}.$$

This, together with (25), implies that

$$\begin{aligned} k_1 &= \frac{1}{3}K + \frac{\alpha\phi}{2(c+a\phi)}, \; k_2 = \frac{1}{3}K - \frac{\alpha\phi}{4(c+a\phi)} + \frac{\alpha\phi\rho}{4(c+a\phi\rho)}, \\ k_3 &= \frac{1}{3}K - \frac{\alpha\phi}{4(c+a\phi)} - \frac{\alpha\phi\rho}{4(c+a\phi\rho)} \end{aligned} \tag{28}$$

Solving (9) by plugging (28) into it yields that

$$K = \frac{3(b+d-1)}{2(c+a)} = K^*. \tag{29}$$

This concludes the solution under $\mathcal{T}_1$. The analysis for $\mathcal{T}_2$ and $\mathcal{T}_3$ is symmetric to the one above and hence relegated to the appendix.

Lemma 1 *The capital allocation under the first best and all the different hierarchies*

[10] The second order condition is satisfied because

$$\begin{aligned} &\phi\left[\frac{1}{4}V_2''\left(\frac{1}{2}k_D + \frac{\alpha\phi\rho}{4(c+a\phi\rho)}\right) + \frac{1}{4}V_3''\left(\frac{1}{2}k_D - \frac{\alpha\phi\rho}{4(c+a\phi\rho)}\right) + V_1''(K-k_D)\right] \\ &+\frac{1}{4}b''\left(\frac{1}{2}k_D + \frac{\alpha\phi\rho}{4(c+a\phi\rho)}\right) + \frac{1}{4}b''\left(\frac{1}{2}k_D - \frac{\alpha\phi\rho}{4(c+a\phi\rho)}\right) + b''(K-k_D) < 0. \end{aligned} \tag{27}$$

are presented in the following table.

	k_1	k_2	k_3
First best	$\frac{1}{3}K^* + \frac{\alpha}{2(c+a)}$	$\frac{1}{3}K^*$	$\frac{1}{3}K^* - \frac{\alpha}{2(c+a)}$
$\mathcal{F}$	$\frac{1}{3}K^* + \frac{\alpha\phi}{2(c+a\phi)}$	$\frac{1}{3}K^*$	$\frac{1}{3}K^* - \frac{\alpha\phi}{2(c+a\phi)}$
$\mathcal{T}_1$	$\frac{1}{3}K^* + \frac{\alpha\phi}{2(c+a\phi)}$	$\frac{1}{3}K^* - \frac{\alpha\phi}{4(c+a\phi)} + \frac{\alpha\phi\rho}{4(c+a\phi\rho)}$	$\frac{1}{3}K^* - \frac{\alpha\phi}{4(c+a\phi)} - \frac{\alpha\phi\rho}{4(c+a\phi\rho)}$
$\mathcal{T}_2$	$\frac{1}{3}K^* + \frac{\alpha\phi\rho}{2(c+a\phi\rho)}$	$\frac{1}{3}K^*$	$\frac{1}{3}K^* - \frac{\alpha\phi\rho}{2(c+a\phi\rho)}$
$\mathcal{T}_3$	$\frac{1}{3}K^* + \frac{\alpha\phi}{4(c+a\phi)} + \frac{\alpha\phi\rho}{4(c+a\phi\rho)}$	$\frac{1}{3}K^* + \frac{\alpha\phi}{4(c+a\phi)} - \frac{\alpha\phi\rho}{4(c+a\phi\rho)}$	$\frac{1}{3}K^* - \frac{\alpha\phi}{2(c+a\phi)}$

The overall capital investment is first-best, i.e. $K = K^*$*, under all hierarchies.*

The nice-bahaving result clearly hinges on the simple and specialized setup in (19) and (20). The patterns of SICM revealed in the table can be summarized as follows.

Proposition 1 *There is SICM at the firm's level* ($k_1^* > k_1$ *and* $k_3^* < k_3$) *under all hierarchies. In addition, there is SICM at the division's level* ($k_m^* > k_m$ *and* $k_n^* < k_n$) *under all tall hierarchies.*

The following results can be directly obtained from Lemma 1. It shows that the separation of ownership and control is one of the driving forces of SICM in any hierarchy, confirming the important empirical findings of Scharfstein [23].

Proposition 2 *Under all hierarchies, SICM at the firm's level will be alleviated as managerial ownership (ϕ) increases, i.e.*

$$\frac{\partial k_1}{\partial \phi} > 0 \text{ and } \frac{\partial k_3}{\partial \phi} < 0.$$

Under all tall hierarchies, SICM at the division's level will be alleviated as D's portion of managerial ownership (ρ) increases, i.e.

$$\frac{\partial k_m}{\partial \rho} > 0 \text{ and } \frac{\partial k_n}{\partial \rho} < 0.$$

According to Shin and Stulz [25], one of the potential problems in multi-segment firms is that the capital allocation to a segment not only depends on its own profit-maximizing concerns, but also depends on the capital, such as cash flow, that is available to other segments. This is verified by (24) in a very stark sense – it is the best response of C to share any change in K *equally* among the three segments. Such kind of socialism certainly is counter-productive when some segment receives too much investment and some other receives too little investment. While such a result under the flat hierarchy can be viewed as direct application of Chou [9], we learned in the current model that such a specific type of socialism continues to hold with the introduction of divisionalization. In fact, it holds both at the division's level, as can be seen in (25), and the firm's level, as can be seen in (28). We summarize these results as follows.

Proposition 3 (Socialism in the Internal Capital Market) *Under all tall hierarchies, the division manager's best response to any change in* k_D *is to to spread it equally among the two divisionalized segments, i.e.*

$$\frac{\partial k_m}{\partial k_D} = \frac{\partial k_n}{\partial k_D} = \frac{1}{2}.$$

Under all hierarchies, the CEO's best response to any change in K *is to to spread it equally among all three segments, i.e.*

$$\frac{\partial k_1}{\partial K} = \frac{\partial k_2}{\partial K} = \frac{\partial k_3}{\partial K} = \frac{1}{3}.$$

It is somewhat surprising that the *slopes* of the best response functions for C and division manager can be independent of ownership (ϕ or ρ) and organization structure. However, this does not mean that ownership and organization structure do not matter. They are, in fact, the key factors that determine the *differences* between k_1, k_2, and k_3. From (24), one can see that if C is directly allocating capital to segment 2 and 3, he would prefer to maintain a difference of $\frac{\alpha\phi}{2(c+a\phi)}$ between the two. However, if C delegates the capital allocation of segment 2 and 3 to D, (25) suggests that the difference becomes $\frac{\alpha\phi\rho}{2(c+a\phi\rho)}$ which is smaller than what C desires. Hence, divisionalization necessarily creates a binding constraint for C's allocation problem that is absent in Chou [9], and this constraint does make the SICM problem worse. This leads to the following central result of the paper.

Proposition 4 *Divisionalization exacerbates SICM and hence destroys values as long as* $\rho < 1$.

While Proposition 4 points out a potential cost of divisionalization (M-form), in reality there could be other benefits of divisionalization (such as the M-form hypothesis) at work and hence divisionalization should be adopted anyway when there is a net benefit in doing so. Under such a situation, it is useful to know how can the cost of divisionalization due to SICM can be reduced. For such a practical question, Lemma 1 suggested that a 'bottom-heavy' ownership structure (in which most of managerial ownership resides with low-level managers) is better than a 'top-heavy' one (in which most of managerial ownership resides with high-level managers), because the differences between the capital allocation of the two divisionalized segments ($\frac{\alpha\phi\rho}{2(c+a\phi\rho)}$ under $\mathcal{T}_1$ and $\mathcal{T}_3$, and $\frac{\alpha\phi\rho}{c+a\phi\rho}$ under $\mathcal{T}_2$) will increase (and hence get closer to the allocation under $\mathcal{F}$) as ρ increases.[11] We list this result as follows.

[11] This result also can be obtained by plugging back the solutions back into the owner's objective function and differentiating it with respect to ρ.

Proposition 5 (the Efficiency of Buttom-heavy Ownership) *For any given* ϕ*, increasing* ρ *creates values for all tall hierarchy. In the extreme case when* $\rho = 1$*, all four hierarchies are equally efficient.*

Proposition 5 suggests that for a given amount of managerial ownership, the degree of separation between ownership and control is smallest when D owns all the managerial ownership and C owns nothing – a rather counterintuitive result. After all, C does control a big part of capital allocation. Why does he never deserve to own some share of the firm to match with his control rights? The answer lies in one of the important condition for SICM to arise – Assumption 6, suggesting that the result is specific to the SICM problem. Under this assumption, the production-profit-maximizing concern of D given by his ownership ρ will also accrue to C and hence only the total managerial ownership matters in C's decision problem. In fact, for any given K, C will allocate $\frac{1}{3}K + \frac{\alpha\phi}{2(c+a\phi)}$ to segment 1 and $\frac{2}{3}K - \frac{\alpha\phi}{2(c+a\phi)}$ to the division regardless of ρ, even when he does not own any shares of the firm.

To our knowledge, Proposition 5 is probably among the first few results that can explain how does different ownership structure within a firm can matter, and hence warrants empirical testing. It also suggests that empirical research on SICM should use overall managerial ownership, instead of CEO ownership alone, to measure the separation between ownership and control.

4 Conclusion

The goal of this paper has been to understand how delegation or increasing the layers of a hierarchy affects the cost of SICM. Flat hierarchies are more efficient in capital allocation than tall hierarchies because increasing the layer of hierarchies will increase the separation of ownership and control, and hence will aggravate the problem of SICM. This result suggests that U-form structures can out-perform M-form structures in a diversified firm where SICM tends to arise.

One might argues that in the original text of Williamson, capital allocation is considered an example of the *strategic decisions* that need to be preserved for the CEO in both the U-form and M-form structures. In this sense, it seems that whether capital allocation is delegated cannot serve as a distinction between the U-form and M-form structures. This view, however, takes the text too literally. After all, it is hard to imagine the CEO of any M-form firm, especially a big conglomerate that consists of hundreds of segments, is personally in charge of fine tuning the capital allocation to each and every segment. In our opinion, strategic decisions should mean 'big' decisions such as 'allocation of capital among the competing operating *divisions*' (Williamson [30], page 137.) instead of among *segments* inside each division. The present results are relevant as long M-form structures facilitate a greater degree of delegation in capital allocation than U-form structures do.

5 Appendix

In this appendix, the solution for $\mathcal{T}_2$ and $\mathcal{T}_3$ is presented. Under $\mathcal{T}_2$, (18) implies that

$$k_1 = \frac{1}{2}k_D + \frac{\alpha\phi\rho}{2\,(c + a\phi\rho)} \text{ and } k_3 = \frac{1}{2}k_D - \frac{\alpha\phi\rho}{2\,(c + a\phi\rho)}. \tag{30}$$

Plugging this into C's problem and solving for k_D yield

$$\begin{aligned} k_D &= \frac{2}{3}K \text{ and hence} \\ k_1 &= \frac{1}{3}K + \frac{\alpha\phi\rho}{2\,(c + a\phi\rho)},\ k_2 = \frac{1}{3}K,\ k_3 = \frac{1}{3}K - \frac{\alpha\phi\rho}{2\,(c + a\phi\rho)}. \end{aligned}$$

Plugging this into the owner's problem and solving for K yield

$$K = \frac{3\,(b + d - 1)}{2\,(c + a)} = K^*. \tag{31}$$

Under $\mathcal{T}_3$, (18) implies that

$$k_1 = \frac{1}{2}k_D + \frac{\alpha\phi\rho}{4\,(c + a\phi\rho)},\ k_2 = \frac{1}{2}k_D - \frac{\alpha\phi\rho}{4\,(c + a\phi\rho)}. \tag{32}$$

Plugging this into C's problem and solving for k_D yield

$$\begin{aligned}
k_D &= \frac{2}{3}K + \frac{\alpha\phi}{2\left(c+a\phi\right)} \text{ and hence} \\
k_1 &= \frac{1}{3}K + \frac{\alpha\phi}{4\left(c+a\phi\right)} + \frac{\alpha\phi\rho}{4\left(c+a\phi\rho\right)},\ k_2 = \frac{1}{3}K + \frac{\alpha\phi}{4\left(c+a\phi\right)} - \frac{\alpha\phi\rho}{4\left(c+a\phi\rho\right)}, \\
k_3 &= \frac{1}{3}K - \frac{\alpha\phi}{2\left(c+a\phi\right)}
\end{aligned}$$

Plugging this into the owner's problem and solving for K yield

$$K = \frac{3\left(b+d-1\right)}{2\left(c+a\right)} = K^*.$$

References

[1] A. W. Beggs, Queues and Hierarchies, *Review of Economic Studies*, **68** 2001, 297-322.

[2] P. Berger and E. Ofek, Diversification's effect on firm value, *Journal of Financial Economics,* **37** 1995, 39-65.

[3] A. Bernardo, H. Cai, J. Luo, Capital Budgeting and Compensation with Asymmetric Information and Moral Hazard, *Journal of Financial Economics,* **61** 2001, 311-344.

[4] P. Bolton and M. Dewatripont, The Firm as a Communication Network, *Quarterly Journal of Economics*, **109** (4) 1994, 809-839.

[5] P. Bolton and D. Scharfstein, Corporate Finance, the Theory of the Firm, and Organization, *Journal of Economic Perspectives*, **12** 1998, 95-114.

[6] M. Buble, Transition Processes of Organizational Structures in Large Enterprises, *Management*, **1** (1), 1996, 1-16.

[7] T. Burch and V. Nanda, Divisional Diversity and Conglomerate Discount: Evidence from Spinoffs, *Journal of Financial Economics*, **70** 2003, 69-98.

[8] G. Calvo and S. Wellisz, Supervision, Loss of Control and the Optimal Size of the Firm, *Journal of Political Economy*, **86** 1978, 943-952.

[9] E. Chou, Flattened Resource Allocations, Separation of Ownership and Control, and Diversification of the Firm, in *Proceedings of the 2002 North American Summer Meetings of the Econometric Society, 2002.*

[10] J. Graham, M. Lemmon, and J. Wolf, Does Corporate Diversification Destroy Value? *Journal of Finance*, **57** (1) 2002, 965-720.

[11] M. Hammer and J. Champy, *Reengineering the Corporation*, New York: Harper Collins, 1993.

[12] M. Harris and A. Raviv, Capital Budgeting and Delegation, *Journal of Financial Economics*, **50** 1998, 259-289.

[13] O. Hart, *Firms, Contract, and Financial Structure*, Oxford University Press, 1995.

[14] O. Hart and J. Moore, On the Design of Hierarchies: Coordination and Specialization, NEBR working paper 7388, 1999.

[15] H. Itoh, Corporate Restructuring in Japan, Part I: Can M-form Organization Manage Diverse Businesses? *Japanese Economic Review*, **54** 2003, 49-73.

[16] L. Lang and R. Stulz, Tobin's q, Corporation Diversification, and Firm Performance, *Journal of Political Economy,* **102** 1994, 1248-1280.

[17] E. Maskin, Y. Quian, and C. Xu, Incentives, Information, and Organizational Form, *Review of Economic Studies*, **67** 2000, 359-378.

[18] J. A. Mirrless, The Optimal Structure of Incentives and Authority within an Organization, *Bell Journal of Economics*, **7** 1976, 105-131.

[19] D. Mookherjee and S. Reichelstein, Incentives and Coordination in Hierarchies, *Advances in Theoretical Economics*, **1** (1) 2001, article 4.

[20] R. Rajan, H. Servaes, and L. Zingales, The Cost of Diversity: The Diversification Discount and Inefficient Investment, *Journal of Finance*, **55** (1) 2000, 35-80.

[21] R. Rajan and J. Wulf, The Flattening Firm: Evidence from Panel Data on the Changing Nature of Corporate Hierarchies, NEBR working paper w9633, 2003.

[22] R. Rajan and L. Zingales, The Firm as a Dedicated: A theory of the Origins and Growth of Firms, *Quarterly Journal of Economics*, **116** (3) 2001, 805-851.

[23] D. Scharfstein, The Dark Side of Internal Capital Markets II: Evidence from Diversified Conglomerate, NEBR working paper 6352, 1998.

[24] D. Scharfstein and J. Stein, The Dark Side of Internal Capital Markets: Divisional Rent Seeking and Inefficient Investment, *Journal of Finance*, **55** (6) 2000, 2537-64.

[25] H. Shin and R. Stulz, Are Internal Capital Markets Efficient? *Quarterly Journal of Economics*, **113** 1998, 531-553.

[26] L. Stole and J. Zwiebel, Organizational Design and Technology Choice under Intrafirm Bargaining, *American Economic Review*, **86** (1) 1996, 195-222.

[27] J. Stein, Information Production and Capital Allocation: Decentralized versus Hierarchical Firms, , *Journal of Finance*, **57** (5) 2002, 1891-1921.

[28] T. Whited, Is it Inefficient Investment that Causes the Diversification Discount? *Journal of Finance*, **56** 2000, 1667-92.

[29] O. Williamson, Hierarchical Control and Optimal Firm Size, *Journal of Political Economy*, **75** (2) 1967, 123-138.

[30] O. Williamson, *Markets and Hierarchies: Analysis and Antitrust Implications*, New York: Free Press, 1975.