THE WILLIAM DAVIDSON INSTITUTE
AT THE UNIVERSITY OF MICHIGAN BUSINESS SCHOOL

Incentives, Scale Economies, and Organizational Form

by Eric Maskin, Yingyi Qian, and Chenggang Xu

Working Paper Number 51
May 1997

Comments Welcome

Presented by Chenggang Xu in June 1997 at the ***Davidson Institute Research Workshop on the Economics of Transition.***

Incentives, Scale Economies, and Organizational Form[1]

Eric Maskin
Department of Economics
Harvard University

Yingyi Qian
Department of Economics
Stanford University

Chenggang Xu
Department of Economics
London School of Economics

May 1997

Abstract

We model organization as the command-and-communication network of managers erected on top of technology (which is modeled as a collection of plants). In our framework, the role of a manager is to deal with shocks that affect the plants that he oversees directly or indirectly. Organizational form is then an instrument for (a) economizing on managerial costs, and (b) providing managerial incentives. We show that two particular organizational forms, the M-form (multi-divisional form) and the U-form (unitary form), are the optimal structures when shocks are sufficiently "big." We argue however that, under certain empirical assumptions, the M-form is likely to be strictly preferable once incentives are taken into account. We conclude by showing that the empirical hypotheses on which this comparison rests are satisfied for Chinese data.

Key Words: Organizational Form, Scale Economies, Incentives, Yardstick Competition, China

---

[1] We are grateful to Takeshi Amemiya, Masahiko Aoki, Patrick Bolton, Gary Chamberlain, Javier Hidalgo, Bengt Holmstrom, Jan Magnus, and Frank Wolak for helpful discussions, and to participants in seminars and conferences in Gerzensee, CalTech, Harvard-MIT, LSE, CEPR (Brussels), and CASS-OECD-UNDP (Beijing) for useful comments. We thank Juzhong Zhuang for his help with the data; and Nancy Hearst for her help with the literature. Angela Lee and Ying Qian provided research assistance. Maskin's research is supported by the NSF, Qian's research by a McNamara Fellowship at Stanford, and Xu's research by STICERD and CEP at London School of Economics (CEP is sponsored by the ESRC).

Incentives, Scale Economies, and Organizational Form

Eric Maskin, Yingyi Qian, and Chenggang Xu

1. Introduction

A central theoretical question is how organization makes a difference to economic performance. Obviously, technology will have a great bearing on the way a firm or economy performs. But, by organization we mean the command-and-communication structure built on top of technology, e.g., the way a corporation is subdivided into different divisions and subdivisions, and the way a planned economy (such as China or the Soviet Union) or a multinational bureaucracy (such as the World Bank or the European Union) is subdivided into different functional or regional governing bodies. In this paper we propose a simple theory through which the effects of different organizational forms on performance can be compared.

In reality, the choice of productive technique and that of organizational structure may not be altogether independent decisions: to some extent, the former may dictate the latter and vice versa. But to focus on the effect of organization, we abstract from this interaction and assume that technology, modeled as a collection of plants, is fixed. In this way, we can explore the implications of alternative organizational forms erected on top of these plants.

In our framework, an organization is a network of managers who oversee a set of plants directly or indirectly. Any theory of organization must articulate what it is that managers do. For example, in Mirrlees (1976), Calvo and Wellisz (1978), and Qian (1994), they monitor the efforts of other managers or of plant workers. In our setting, managers handle shocks that affect a plant's performance or the value of its output. Such shocks include shifts in input supply or output demand, changes in the weather, or technological advances. A manager's job is to determine the operational implications of these shocks for the plants under his charge and to communicate these implications to his subordinates.

We assume that managers are costly to hire, and so the first issue that arises is whether or not it is desirable to assign a manager at all to a given shock. In our model, managers are more valuable, the bigger the shock (as measured by its variance). And so, our prediction is that managers will be installed for sufficiently volatile shocks.

The next issue is whether organization itself is necessary: Why isn't each plant a fully autonomous entity? Mirrlees (1976), Calvo and Wellisz (1978), and Qian (1994) do not address this issue. They simply assume that all plants are "controlled" by the some central agency, but do not examine why this should be so, e.g., why some of the plants should not be spun off into separate organizations.

One answer is that there may be an organizational economy of scale to a more integrated structure.[2] In our setting, there is a natural source for such a scale economy: a shock may affect more than one plant. Thus, if managers are costly to hire, having one high-level manager handle the shock for all affected plants can be more efficient than having the same shock handled repeatedly by separate managers in each of these plants.

The implication of this is that, in a setting in which there are large shocks hitting multiple plants, the organization that is optimal in the sense of maximizing net economic value will be highly integrated. This leads us naturally to the M-form (multi-divisional form) and U-form (unitary form), which in our model are the two possible fully-integrated organizations.

Both structures have figured prominently in corporate history (see Chandler, 1962). A classic example of the U-form was the Ford Motor Company before the Second World War. In those days, Ford was organized into a number of functionally specialized departments: production, sales, purchasing, and so on. In other words, the various departments carried out complementary tasks; none was independent of the others. By contrast, General Motors under Alfred Sloan became the prototypical M-form; GM comprised (and still comprises) a collection of fairly self-contained

[2] Radner (1993) also emphasizes economies of scale, but in his case they are computational. Other recent explanations for integration include Grossman and Hart (1986) and Hart and Moore (1990), who examine how different ownership patterns affect the incentives to invest. Our approach is silent on the issue of who owns the various plants. Bolton and Farrell (1990), Bolton and Dewatripont (1994), and Qian, Roland, and Xu (1997) concentrate on the coordination advantages of integration.

divisions, e.g., Chevrolet, Pontiac, and Oldsmobile.

The terms 'M-form' and 'U-form' have been applied primarily to corporations. Recently, however, they have been brought into the study of comparative economic systems. In particular, Qian and Xu (1993) observed that an important difference between the economy of the former Soviet Union and that of China lies in their respective organizational structures. The Soviet economy was, in effect, a gigantic U-form; it consisted of approximately sixty specialized ministries, e.g., steel or mining. Since 1958, however, the Chinese economy has more closely resembled an 'M-form;' it comprises thirty reasonably self-sufficient provinces or regions.

The Soviet economy turned out to be a disaster, whereas the growth rate in China remains strong. Could the difference in organizational form help explain this contrast? We argue that an M-form seems likely to dominate from the standpoint of providing managerial incentives. Suppose that managers -- regardless of the organizational form -- must be provided with incentives to act in the organization's interest. One way to do this is to reward them on the basis of performance. But, in view of the shocks we have been discussing, performance will not be perfectly correlated with managerial effort. Thus, if the manager of region A shows a poor performance, he may try to blame the outcome on bad luck rather than on lack of effort. This defense will not be so persuasive, however, if other regions are prospering. Thus, it will, in general, be desirable to make the manager's reward depend not only on absolute performance but also on performance relative to that in other regions.

This is, of course, a familiar idea -- the principle of yardstick competition (see, for example Lazear and Rosen, 1981, Holmstrom, 1982, Nalebuff and Stiglitz, 1883, and Shleifer 1985). But the question arises: why can we not do the same thing in a U-form? After all, in theory, we could compare the steel minister's performance with that of the mining minister. Admittedly, this seems intuitively more difficult than comparing regions that producing more-or-less the same array of goods. But on what is this intuition founded?

We argue that one possible foundation is the idea that the 'variation' between the performances of two regions producing similar outputs is likely to be lower (in the appropriate statistical sense) than that between the performances of two production ministries. If this is, in fact, so, then it may shed light on the Chinese success story. Of course, this comes down in the end to a matter of empirics. But here our analysis from 520 Chinese state-owned enterprises data seems to support the hypothesis that it is "easier" to compare different regions than different industries.

The more general lesson that our Chinese example illustrates is that different organizational forms give rise to different information on which incentives can be based. Thus in the end, our theory of organizational form consists of three guiding principles: (1) an organization should exist when shocks are big enough to warrant the expense of managers; (2) it should exploit managerial scale economies through integration; and (3) it should take into account the informational structure it induces and, in particular this structure's effect on incentives.

We proceed as follows. In section 2, we lay out the model. In section 3, we present our theoretical results. Proposition 1 establishes that if all shocks are sufficiently big, then a fully-integrated organization -- an M-form or U-form -- is the optimal organizational form. Proposition 3 shows that the M-form provides better incentives for middle-level managers provided that there is "less variation" in interregional performance than in interindustry performance. But Proposition 2 establishes that it is only at this level that organizational form has any bearing on incentive issues: both top- and bottom-level managers' incentives turn out to be independent of whether the M-form or the U-form is employed.

Then, in section 4, we develop the test statistics needed to analyze our Chinese data set. Our empirical work is reported in section 5, where we argue that there is indeed higher "variation" in performance across industries than across regions. We also offer some evidence for the use of yardstick competition in the Chinese economy We make a few concluding remarks in section 6.

2. The Model

Consider an economy with two regions, A and B; two industries, 1 and 2; and four plants, one for each region-industry combination: 1A, 1B, 2A, and 2B, where plant ir produces industry i output ($i=1,2$) and is located in region r ($r=A,B$). There are three kinds of shocks s: shock η hits all plants in the economy; shock θ_i hits just plants in industry i, $i=1,2$; shock δ_r hits region r, $r=A,B$.[3] We assume that the shocks are jointly normally distributed with zero mean. Each shock may (but need not) be assigned to one or more managers.[4]

A chain of command is a sequence of managers, who directly or indirectly, oversee a given plant, where each manager "reports" to the next higher manager (if any), and the highest manager in the sequence reports to no one. We assume that each manager can report to at most one other manager in the model.[5]

A manager who is assigned shock s can exert effort $e \geq 0$ to deal with it. The effect of the manager's presence depends on the magnitude of the shock as measured by its variance σ_s^2. The effect takes the form of an increase in the value of the output of any plant that (i) is hit by the shock that he is assigned and (ii) is connected to him by a chain of command. Let $e + f(\sigma_s^2)$ be that increase in output,[6] where f is differentiable and $f(0)=0$. Assume that there exists a constant $k>0$

---

[3] There is another -- and perhaps more "standard" -- interpretation of our model. Instead of an entire economy, think of a corporation, say, an automobile manufacturer. The "regions" would then correspond to two different car models, whereas the two "industries" would become two different specialized departments, e.g., production and purchasing. Shocks to "regions" (models) could then be interpreted as shifts in demand for these models, whereas shocks to "industries" (departments) might reflect changes in the cost of labor or parts.

[4] We assume that a single manager cannot efficiently handle more than one shock. Implicit here is the idea that to deal with a shock properly a manager must take a significant fixed investment of time and knowledge (indeed, we can interpret $\underline{U}$ as that investment).

[5] The justification for the restriction that a manager reports to only one direct superior (i.e., each manager has only one "boss") is that bosses give orders and that, with multiple bosses, one boss's orders might fail to be informed by another boss's information. To give a stylized example, suppose that industry 1 is "agriculture" and that region A is "Hunan Province." Assume that the provincial minister (governor) oversees the agricultural minister in Hunan. Imagine that the governor detects a demographic trend that, as applied to agriculture, calls for growing more soy beans. He therefore issues a directive to that effect to the agriculture minister. Given this directive, the agriculture minister then investigates how shocks may have affected supplies of the principal input into soy production -- say, a certain kind of fertilizer. Accordingly, he gives orders to ensure that soy bean growers get sufficient quantities of this fertilizer. Notice that if he had not received the governor's directive, he would not have known to concentrate particularly on soy beans in his evaluation of agricultural shocks. Therefore, an arrangement in which the agriculture minister is independent of the governor either would no work at all or else would require that the agricultural minister issue complicated contingent orders.

[6] We are normalizing output so that zero corresponds to expected output in the absence of a manager (since means are unimportant in our analysis, we can normalize them any way we wish).

such that

(1) $$df/d(\sigma_s^2) > k.$$

Formula (1) says that the bigger the expected magnitude of a shock (as measured by its variance), the bigger is the expected impact that a manager's presence has. We assume that the impact takes the form of a fixed effect f that rises with variance plus a variable effect that depends on the manager's effort.

Hence, if there is chain of command for plant ir, consisting of a manager overseeing shock θ_i, who reports to a manager overseeing shock δ_r, who in turn reports to a manager overseeing shock η, then output is:

$$x_{ir} = e_i + f(\sigma_i^2) + e_r + f(\sigma_r^2) + e_\eta + f(\sigma_\eta^2) + \theta_i + \delta_r + \eta,$$

where e_i, e_r and e_η are the efforts corresponding to shocks θ_i, δ_r and η respectively. The cost of effort e is C(e) where C(0)=0, and

(2) $$dC/de > 0, \text{ and } d^2C/de^2 > 0.$$

A manager's utility is given by

$$U(t) - C(e),$$

where t is the manager's (monetary) payment and U is his von Neumann-Morgenstern utility function. Let $\underline{U}$ be the manager's reservation utility, where $\underline{U}>0$. We shall suppose that:

(3) $$\underline{U} > \max_e (e-C(e)).$$

That is, unless a shock is of sufficient magnitude, it is not worthwhile hiring a manager to handle it.

We will suppose that managers' efforts cannot be directly monitored. Hence, a manager's reward t will depend only on the observable outputs $\{x_{ir}\}$. The organizational problem is to choose a set M of managers, a corresponding set of chains of command, and a set of reward schemes $t_j(.)$ for each manager j so as to maximize the expected value of net output

$$\Sigma_i \Sigma_r x_{ir} - \Sigma_{j \in M} t_j(.),$$

subject to the constraints that each manager get at least his reservation utility and that he choose an effort level e* that maximizes his own net expected utility:

$$E[U(t(.)) - C(e)].$$

It may be useful to give some illustrations of possible organizational forms. At one extreme is the case in which none of the shocks is handled by a manager, i.e., there is no organization at all. In this case, each plant operates independently. Another instance of independent plants is the case in which each plant has a different manager who oversees and handles the industrial shock hitting it (see Figure 1). However, such a configuration involves considerable redundancy, and since managers are costly, it is dominated by one in which each manager oversees two plants (see Figure 2).[7] In that case, there are, in effect, two organizations.

Figure 1

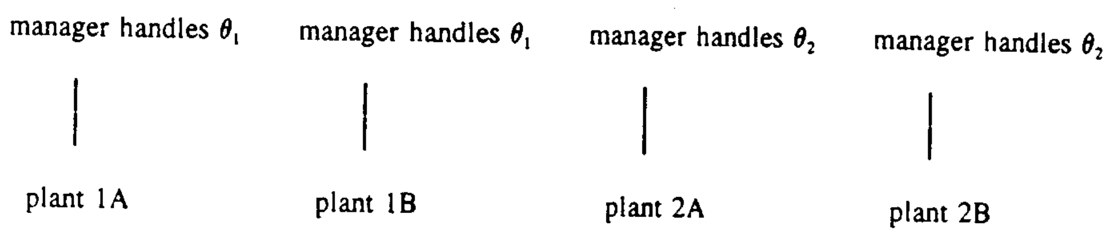
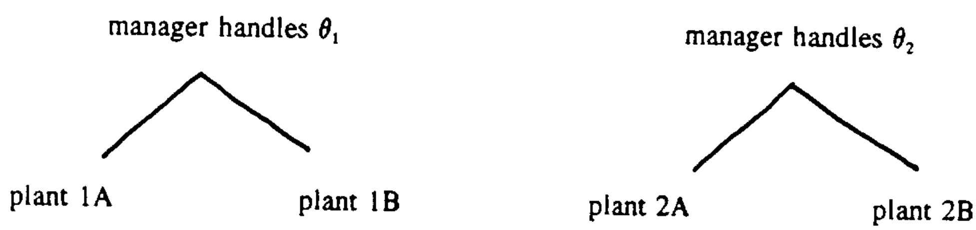


Figure 2



Next let us examine organizations in which both regional and industrial shocks are handled by

[7] As we will see in the proof of Proposition 1, this conclusion relies on our assumption (3).

managers. The configurations illustrated by Figures 3 and 4 represent the cases, respectively, of two independent industries and two independent regions. Note that both entail some duplication of effort -- in Figure 3, there are two managers each handling shocks δ_r, whereas in Figure 4 there are two managers each handles shocks θ_i. However, given our assumption that no manager can report to two bosses, some duplication is unavoidable if all regional and industrial shocks are to be covered. One aspect of optimal organizational design will thus be how to best minimize the cost of such duplication.

Figure 3

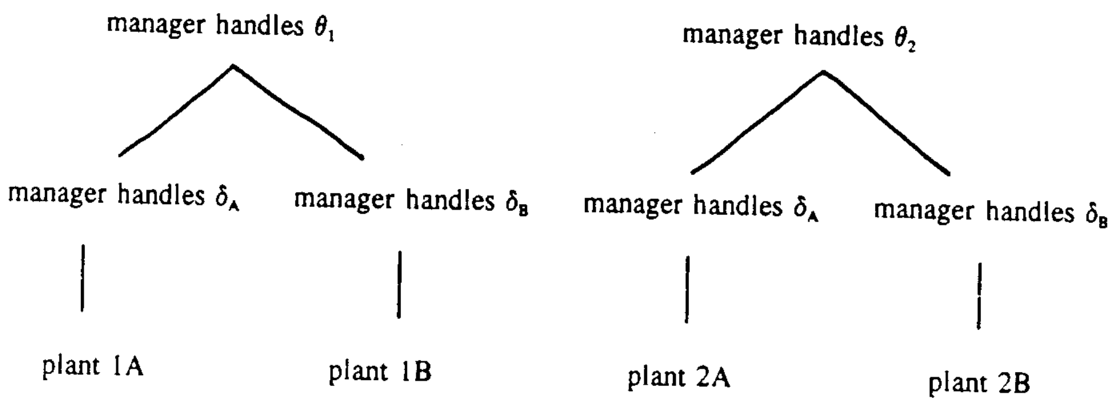
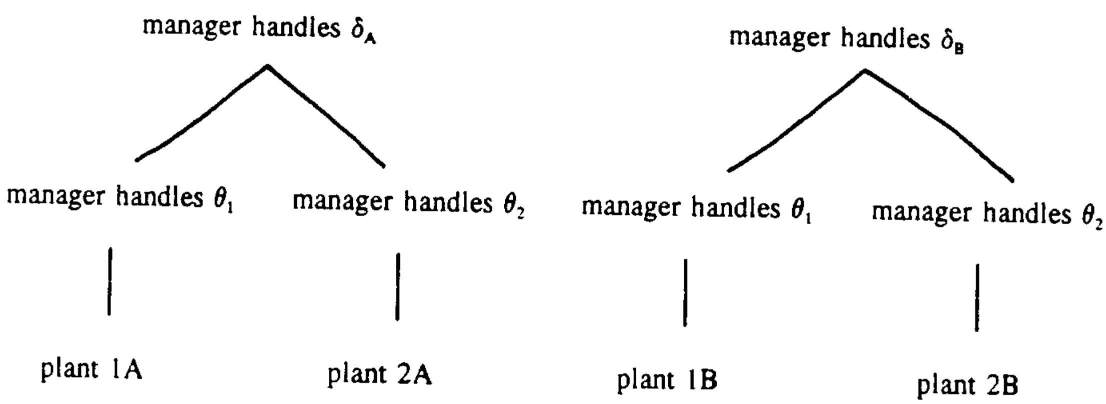


Figure 4



If the shock η hitting plants is big enough, then a completely integrated organization may be desirable. These are the U-form and M-form, illustrated in Figures 5 and 6, respectively.

Figure 5

U-form

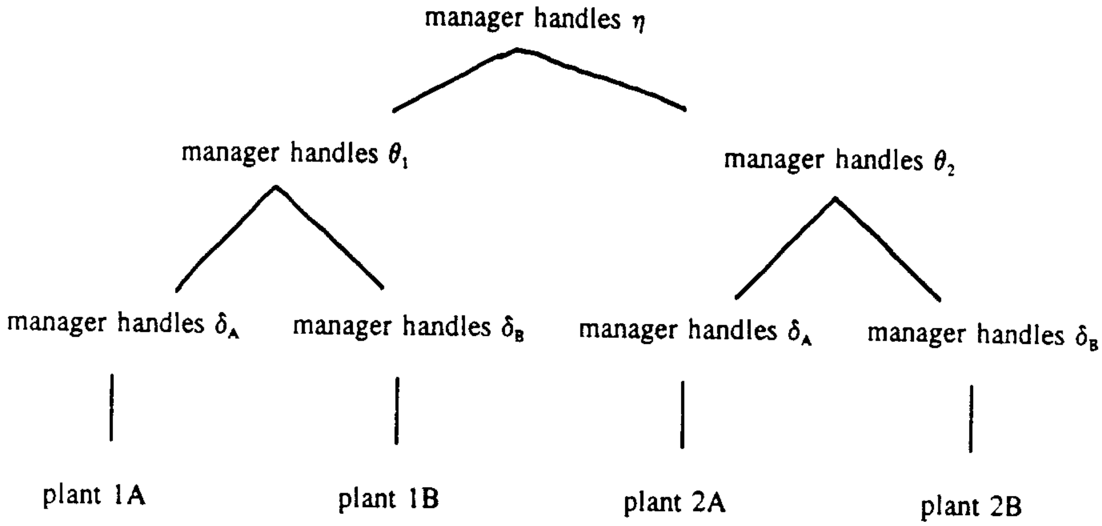
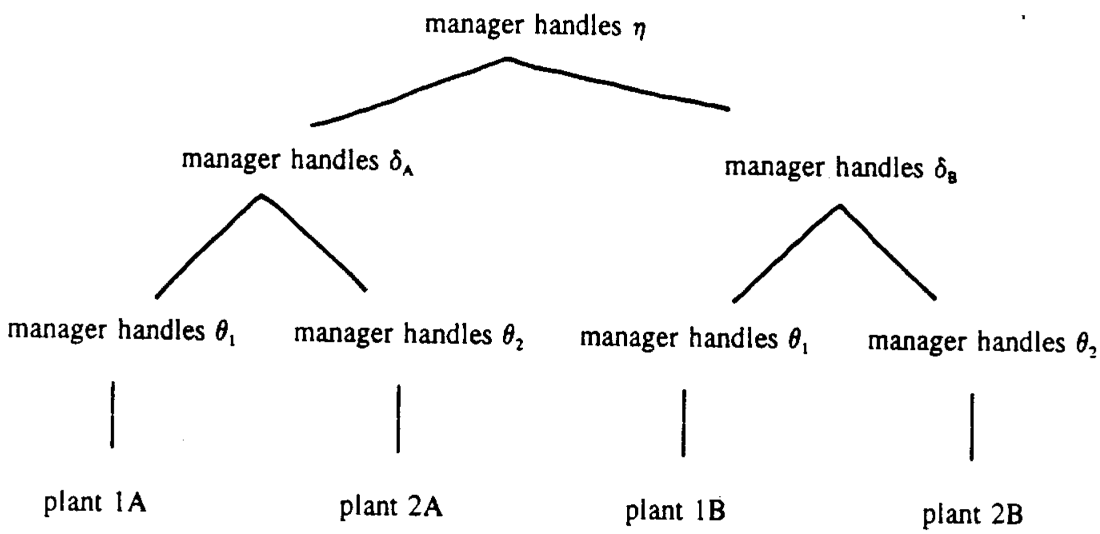


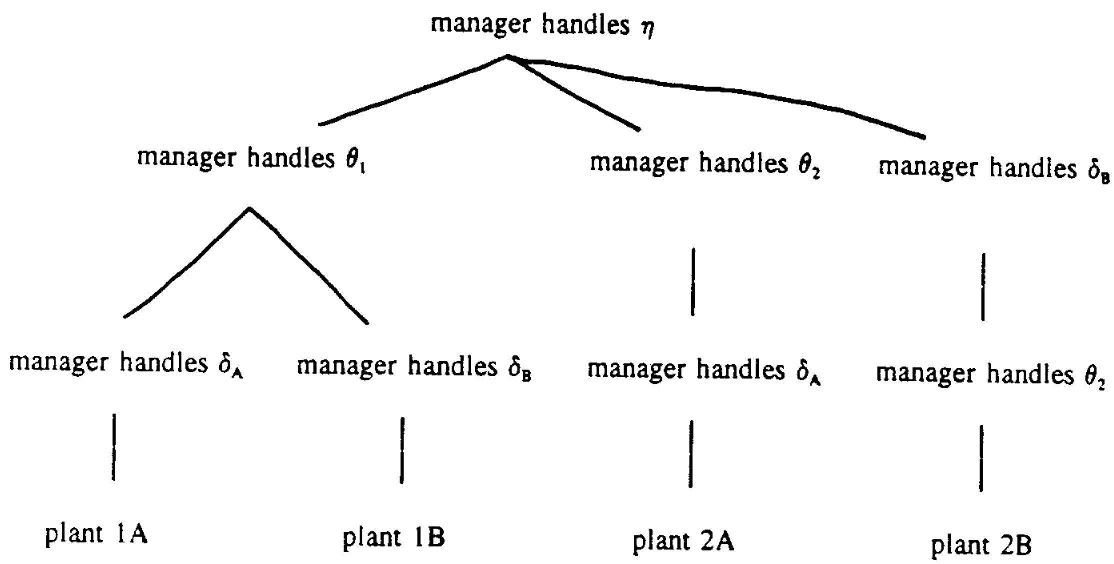
Figure 6

M-form



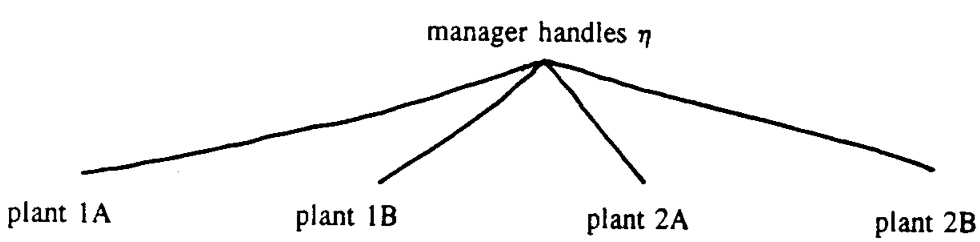
There are, of course, other possibilities too (see, for example, Figure 7).

Figure 7



Finally, it is possible, of course, that shock η may be big, but that neither the regional nor the industrial shocks are significant enough to warrant hiring managers to handle them. In that case, we might expect an organizational form as in Figure 8.

Figure 8



3. Theoretical Findings

A. Shocks and Scale Economies

Given that the only role of managers is to handle shocks and that managers are costly, assumption (3) implies that sufficiently small shocks should not be assigned to managers. On the other hand, assumption (1) tells us that big shocks should have managers. Our first result establishes that when all shocks are big enough, either the M-form or the U-form is the optimal organizational form.

Proposition 1: Under assumptions (1)-(3), there exists $v>0$ such that, if for all shocks s, $\sigma_s^2>v$, the optimal organization form is either the U-form (see Figure 5) or the M-form (see Figure 6).

Proof: From (1), we can choose v such that

(4) $$f(v) > \underline{U}.$$

Hence, if, for all shocks s, $\sigma_s^2>v$, then it is worthwhile assigning every shock to a manager. Now conceivably, a manager handling a given shock s could be given better incentives if there were another manager handling the same shock with whom his performance could be compared. But (3) ensures that any such benefit would be dwarfed by the expense of the additional manager. Hence it is preferable for each shock to be handled by as few managers as possible. One can readily check that the U-form and M-form are the two structures that minimize that total number of managers (seven), subject to the constraint that all shocks are handled. Q.E.D.

Because we are particularly interested in highly integrated organizations, we shall henceforth assume that all shocks exceed v as given by Proposition 1.

B. Information and Incentives

Although Proposition 1 elevates the M-form and U-form above all other organizational forms, it offers no clue about which will perform better. We will argue that there is if there is less "variation" (in the appropriate sense) in shocks across regions than in shocks across industries, the M-form dominates the U-form from the standpoint of incentives.

To get a feel for the issues involved, let us consider an even simpler framework than that of our model. Suppose that there are two industries, 1 and 2, and that output in industry i is given by

$$x_i = e_i + \epsilon_i,$$

where e_i is the effort of the manager in charge of shock ϵ_i, and (ϵ_1, ϵ_2) are jointly normally distributed. Let us compare this with the case of two regions, A and B, where output in region r is given by

$$x_r = e_r + \epsilon_r,$$

e_r is the effort of the manager in charge of shock ϵ_r, and (ϵ_A, ϵ_B) are jointly normal. All managers have preferences given by

$$U(t) - C(e),$$

where t is a transfer that in the industrial case can depend on (x_1, x_2), and in the regional case on (x_A, x_B).

In which scenario can better incentives be provided? It turns out that a comparison of conditional variances is the key. If

(5) $$\mathrm{Var}(\epsilon_A \mid \epsilon_B) < \mathrm{Var}(\epsilon_1 \mid \epsilon_2),$$

then manager A can be given better incentives than manager 1. Moreover, if both

(6) $$\min \{ \mathrm{Var}(\epsilon_A \mid \epsilon_B), \mathrm{Var}(\epsilon_B \mid \epsilon_A) \} < \min \{ \mathrm{Var}(\epsilon_1 \mid \epsilon_2), \mathrm{Var}(\epsilon_2 \mid \epsilon_1) \}$$

and

(7) $$\max \{ \mathrm{Var}(\epsilon_A \mid \epsilon_B), \mathrm{Var}(\epsilon_B \mid \epsilon_A) \} < \max \{ \mathrm{Var}(\epsilon_1 \mid \epsilon_2), \mathrm{Var}(\epsilon_2 \mid \epsilon_1) \},$$

then both managers A and B can be given better incentives than manager 1 and 2.

The less noisy performance is as a measure of effort, the easier it is to provide a manager with the incentive to supply effort. Condition (5) says that the residual noise that remains in manager

A's performance after it is compared with that of manager B is smaller than the residual noise that remains in manager 1's performance after it is compared with that of manager 2.

To see that if (5) holds, manager A can be predicted with better incentives than manager 1, fix an effort level e' for manager 2 and assume that managers choose effort levels noncooperatively. Suppose that $t_1(\bullet,\bullet)$ is an incentive scheme for manager 1 such that $t_1(x_1, x_2)$ is his transfer conditional on outputs (x_1, x_2). We will show that, if (5) holds, we can find a transfer scheme $t_A(\bullet,\bullet)$ as a function of (x_A, x_B) such that, if manager B exerts effort e', the scheme $t_A(\bullet,\bullet)$ is equivalent to $t_1(\bullet,\bullet)$. To see this, note that (5) is equivalent to

$$\sigma_A^2 - (\sigma_{AB})^2/\sigma_B^2 \leq \sigma_1^2 - (\sigma_{12})^2/\sigma_2^2,$$

where $\sigma_r^2 = \mathrm{Var}(\epsilon_r)$, r=A,B; $\sigma_{AB} = \mathrm{Cov}(\epsilon_A, \epsilon_B)$; $\sigma_i^2 = \mathrm{Var}(\epsilon_i)$, i=1,2; and $\sigma_{12} = \mathrm{Cov}(\epsilon_1, \epsilon_2)$.

Choose scalars

$$\alpha = \sigma_{AB}/\sigma_B^2 - \sigma_{12}/(\sigma_2^2\sigma_B^2)^{1/2},$$

$$\beta = (\sigma_2^2/\sigma_B^2)^{1/2}, \text{ and}$$

$$\gamma = (1-\beta)e'.$$

Also let z be a normally distributed random variable, independent of x_A and x_B, with mean $\alpha e'$ and variance $[\mathrm{Var}(\epsilon_1 | \epsilon_2) - \mathrm{Var}(\epsilon_A | \epsilon_B)]$. We claim that if managers 2 and B choose effort e', then for any choice of effort e by manager 1 and A, the two pairs of random variable (x_1, x_2) and $(x_A - \alpha x_B + z, \beta x_B + \gamma)$ have the same distributions. Hence, if we take

$$t_A(x_A, x_B) = t_1(x_A - \alpha x_B + z, \beta x_B + \gamma),$$

$t_A(\bullet,\bullet)$ will be equivalent to $t_1(\bullet,\bullet)$. But because all random variables are normal, it suffices to show that the two pairs have the same mean and the same covariance matrix for all e. In fact:

$E(x_A - \alpha x_B + z) = e - \alpha e' + \alpha e' = e = Ex_1;$

$E(\beta x_B + \gamma) = \beta e' + (1-\beta)e' = e' = Ex_2;$

$\mathrm{Var}(\beta x_B + \gamma) = \beta^2\,\mathrm{Var}(x_B) = \sigma_2^2 = \mathrm{Var}(x_2);$

$\mathrm{Cov}(x_A - \alpha x_B + z, \beta x_B + \gamma) = \beta\sigma_{AB}^2 - \alpha\beta\sigma_B^2 = \sigma_{12}^2 = \mathrm{Cov}(x_1, x_2);$

$$\begin{aligned}\mathrm{Var}(x_A - \alpha x_B + z) &= \sigma_A^2 - 2\alpha\sigma_{AB} + \alpha^2\sigma_B^2 + [\mathrm{Var}(\epsilon_1 | \epsilon_2) - \mathrm{Var}(\epsilon_A | \epsilon_B)] \\ &= \sigma_A^2 - \sigma_{AB}^2/\sigma_B^2 + \sigma_{12}^2/\sigma_2^2 + \sigma_1^2 - \sigma_{12}^2/\sigma_2^2 - (\sigma_A^2 - \sigma_{AB}^2/\sigma_B^2) = \sigma_1^2 = \mathrm{Var}(x_1),\end{aligned}$$

as claimed.

We have been taking e' as fixed for managers 2 and B. But if (6) and (7) hold, a similar argument shows that manager B can be induced to choose the same effort level as manager 2.

So far we have been examining a set-up that is simpler than the model that we are really interested in. Let us return, thereafter, to the model of section 2. Because the terms $f(\sigma_s^2)$ do not affect incentives, let us henceforth set them equal to zero:

(8) $$f(\bullet) \equiv 0.$$

As in the stripped-down framework, let us suppose that managers' effort cannot be directly monitored, so that their rewards can be based only on the vector of outputs

$$(x_{1A}, x_{2A}, x_{1B}, x_{2B}).$$

Let us also continue to assume that managers choose their effort levels noncooperatively. We first observe that in comparing the M-form and U-form, it suffices to consider the incentives of only the middle-level managers (regional managers in the M-form, industrial managers in the U-form); those of top- and bottom-level managers are the same for either organization.

Proposition 2: Given any incentive scheme $t_\eta(x_{1A}, x_{2A}, x_{1B}, x_{2B})$ for the top manager (the one handling η) in the M-form, there exists an equivalent scheme $t_\eta'(x_{1A}, x_{2A}, x_{1B}, x_{2B})$ for the top manager in the U-form (in the sense that it induces the same effort level and gives the managers the same expected payoff), and vice versa. Similarly, given any incentive scheme $t_{ir}(\bullet)$ for the industry i manager under the region r manager in the M-form, there exists an equivalent scheme t_{ri}' for the region r manager under the industry i manager in the U-form, and vice versa.

Proof: Suppose that the industry 1 manager in region A (manager 1A) in the M-form faces incentive scheme $t_{1A}(x_{1A}, x_{2A}, x_{1B}, x_{2B})$. Moreover, suppose that, given their incentive schemes, the other bottom-level managers are induced to choose levels e_{2A}^*, e_{1B}^*, e_{2B}^* (where e_{ir}^* is the effort level of manager ir), the middle-level managers are induced to choose levels e_A^* and e_B^*, and the top manager level e_η^*.

Now consider the U-form and suppose that the bottom-level managers other than A1 (the region A manager in industry 1) have incentive schemes that induce them to choose levels e_{A2}^{**}, e_{B1}^{**}, e_{B2}^{**}, the middle-level managers e_1^{**} and e_2^{**}, and the top-level manager e_η^{**}. Endow manager A1 with transfer function

$$t_{A1}'(x_{1A}, x_{2A}, x_{1B}, x_{2B}) = t_{1A}(x_{1A}+e_A^*+e_\eta^*-e_1^{**}-e_\eta^{**},\ x_{2A}+e_A^*+e_\eta^*-e_2^{**}-e_\eta^{**}+e_{2A}^*-e_{A2}^{**},$$
$$x_{1B}+e_B^*+e_\eta^*-e_1^{**}-e_\eta^{**}+e_{1B}^*-e_{B1}^{**},\ x_{2B}+e_B^*+e_\eta^*-e_2^{**}-e_\eta^{**}+e_{2B}^*-e_{B2}^{**}).$$

It is then straightforward to verify that, for any effort choice e by managers A1 or 1A, the random variables $t_{A1}'(\bullet,\bullet,\bullet,\bullet)$ and $t_{1A}(\bullet,\bullet,\bullet,\bullet)$ are the same. The argument for top managers is similar. Q.E.D.

The proof of Proposition 2 relies on a simple idea: the information available on which to base incentives is the same across organizational forms for both top- and bottom-level managers. However, as our stripped-down model above suggests, the same is not true of middle-level managers. Indeed, a major theme of this paper is that an important respect in which organizational forms differ is precisely in the information that they give rise to.

In both the M-form and U-form, incentive schemes can depend on $(x_{1A}, x_{2A}, x_{1B}, x_{2B})$. However, the way this set is partitioned into spheres of influence of the two middle-level managers differs. In the M-form, the region A and B managers affect (x_{1A}, x_{2A}) and (x_{1B}, x_{2B}) respectively, whereas in the U-form, the industry 1 and 2 managers affect (x_{1A}, x_{1B}) and (x_{2A}, x_{2B}) respectively. In our stripped-down model, the M-form dominated the U-form from the standpoint of incentives if the M-form's associated conditional variances were smaller than those of the U-form. Now, in the full-blown model, we must compare pairs of random variables, which may seem more complicated than the stripped-down analysis. But it turns out that the comparisons can be reduced to one dimension. Specifically, let λ_A solve

(9) $$\min_\lambda \mathrm{Var}\,(\lambda\epsilon_{1A} + (1-\lambda)\epsilon_{2A} \mid \epsilon_{1B}, \epsilon_{2B}),$$

and let λ_1 solve

(10) $$\min_\lambda \mathrm{Var}\,(\lambda\epsilon_{1A} + (1-\lambda)\epsilon_{1B} \mid \epsilon_{2A}, \epsilon_{2B}).$$

Define λ_B and λ_2 analogously. We will show that the M-form generates better incentives than the U-form provided that

$$\min \{ \mathrm{Var}(\epsilon_A \mid \epsilon_B), \mathrm{Var}(\epsilon_B \mid \epsilon_A) \} < \min \{ \mathrm{Var}(\epsilon_1 \mid \epsilon_2), \mathrm{Var}(\epsilon_2 \mid \epsilon_1) \}$$

and

$$\max \{ \mathrm{Var}(\epsilon_A \mid \epsilon_B), \mathrm{Var}(\epsilon_B \mid \epsilon_A) \} < \max \{ \mathrm{Var}(\epsilon_1 \mid \epsilon_2), \mathrm{Var}(\epsilon_2 \mid \epsilon_1) \},$$

where, for r = A, B,

$$\epsilon_r = \lambda_r \epsilon_{1r} + (1-\lambda_r)\epsilon_{2r}$$

and, for i = 1, 2,

$$\epsilon_i = \lambda_i \epsilon_{iA} + (1-\lambda_i)\epsilon_{iB}.$$

(The "max" and "min" operators reflect the fact that the labels "1", "2", "A", and "B" are arbitrary).

To establish this result, we first establish that appropriately aggregated information is equivalent to disaggregated information for incentive purposes. Because the shock η plays no role in the subsequent analysis, we henceforth ignore it.

We take

$$\xi = (\theta_1, \theta_2, \delta_A, \delta_B)', \ \Sigma = \mathrm{var}(\xi),$$

$$\epsilon_u = (\epsilon_{1A}, \epsilon_{1B}, \epsilon_{2A}, \epsilon_{2B})', \ \Sigma_u = \mathrm{var}(\epsilon_u),$$

$$\epsilon_m = (\epsilon_{1A}, \epsilon_{2A}, \epsilon_{1B}, \epsilon_{2B})', \ \Sigma_m = \mathrm{var}(\epsilon_m),$$

where

$$(\epsilon_{1A}, \epsilon_{1B}, \epsilon_{2A}, \epsilon_{2B})' = A_u\xi, \ \ \Sigma_u = A_u \Sigma A_u',$$

$$(\epsilon_{1A}, \epsilon_{2A}, \epsilon_{1B}, \epsilon_{2B})' = A_m\xi, \ \ \Sigma_m = A_m \Sigma A_m'$$

and

$$A_u = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \qquad A_m = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}.$$

Note that both A_u and A_m are singular, and so are Σ_u and Σ_m. However, one can verify that $\mathrm{Rank}(A_u)=\mathrm{Rank}(A_m)=3$, and $A_u'R=0$ and $A_m'R=0$ for $R=(1,-1,-1,1)'$.

Lemma 1: If $(x_{1A}, x_{1B}, x_{2A}, x_{2B})$ and $(x_{1A}^*, x_{2A}^*, x_{1B}^*, x_{2B}^*)$ are the outputs in the U-form and M-form respectively, we can express

$$(x_{1A}, x_{1B}, x_{2A}, x_{2B}) = (x_1, x_1, x_2, x_2) + (u_1, u_2, u_3, u_4)$$

and
$$(x_{1A}^*, x_{2A}^*, x_{1B}^*, x_{2B}^*) = (x_A^*, x_A^*, x_B^*, x_B^*) + (v_1, v_2, v_3, v_4),$$

where (x_1, x_1, x_2, x_2) and (u_1, u_2, u_3, u_4) are uncorrelated, $(x_A^*, x_A^*, x_B^*, x_B^*)$ and (v_1, v_2, v_3, v_4) are uncorrelated, and

$$(x_1, x_2)' = (C_u'(C_{ul}\Sigma C_{ul}')^{-1} C_u)^{-1} C_u'(C_{ul}\Sigma C_{ul}')^{-1} Q_u'(x_{1A}, x_{1B}, x_{2A}, x_{2B})'$$

$$(x_A^*, x_B^*)' = (C_m'(C_{ml}\Sigma C_{ml}')^{-1} C_m)^{-1} C_m'(C_{ml}\Sigma C_{ml}')^{-1} Q_m'(x_{1A}^*, x_{2A}^*, x_{1B}^*, x_{2B}^*)'$$

$$C_u = Q_u'A,\ C_{ul} = Q_u'A_u,\ C_m = Q_m'A,\ C_{ml} = Q_m'A_m,$$

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \qquad Q_u = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{bmatrix} \qquad Q_m = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix},$$

and $(C_{ul}\Sigma C_{ul}')$ and $(C_{ml}\Sigma C_{ml}')$ are non-singular 3x3 matrices.

Proof: We prove Lemma 1 for the U-form (the M-form argument is similar). Let $x = (x_{1A}, x_{1B}, x_{2A}, x_{2B})'$ and $\beta=(e_1,e_2)'$. Then under the U-form:

$$x = A\beta + A_u\xi.$$

Let $\underline{x} = A(x_1, x_2)' = (x_1, x_1, x_2, x_2)'$ and $u = (u_1, u_2, u_3, u_4) = x-\underline{x}$. Because $Eu=Ex-E\underline{x}=A\beta-A\beta=0$, to show $\underline{x}$ and u are uncorrelated, we need only show that $E\underline{x}u'=0$. In fact,

$E\underline{x}u'$

$= E\underline{x}(x-\underline{x})'$

$= E\,\{A(C_u'(C_{ul}\Sigma C_{ul}')^{-1} C_u)^{-1} C_u'(C_{ul}\Sigma C_{ul}')^{-1} Q_u'(A\beta+A_u\xi)\}$
$\quad\{(I-A(C_u'(C_{ul}\Sigma C_{ul}')^{-1} C_u)^{-1} C_u'(C_{ul}\Sigma C_{ul}')^{-1} Q_u')(A\beta+A_u\xi)\}'$

$= E\,\{A(C_u'(C_{ul}\Sigma C_{ul}')^{-1} C_u)^{-1} C_u'(C_{ul}\Sigma C_{ul}')^{-1} Q_u'(A\beta+A_u\xi)\}$
$\quad(A_u\xi)'\{I-A(C_u'(C_{ul}\Sigma C_{ul}')^{-1} C_u)^{-1} C_u'(C_{ul}\Sigma C_{ul}')^{-1} Q_u'\}'$

$= E\,\{A\beta + A(C_u'(C_{ul}\Sigma C_{ul}')^{-1} C_u)^{-1} C_u'(C_{ul}\Sigma C_{ul}')^{-1} C_{ul}\ \xi\}$
$\quad\xi'A_u'\{I-Q_u'(C_{ul}\Sigma C_{ul}')^{-1} C_u\ (C_u'(C_{ul}\Sigma C_{ul}')^{-1} C_u)^{-1} A'\}$

$= A(C_u'(C_{ul}\Sigma C_{ul}')^{-1} C_u)^{-1} C_u'(C_{ul}\Sigma C_{ul}')^{-1} C_{ul} E(\xi\xi') A_u'$

$- A(C_u'(C_{ul}\Sigma C_{ul}')^{-1} C_u)^{-1} C_u'(C_{ul}\Sigma C_{ul}')^{-1} C_{ul} E(\xi\xi')C_{ul}'(C_{ul}\Sigma C_{ul}')^{-1} C_u (C_u'(C_{ul}\Sigma C_{ul}')^{-1} C_u)^{-1} A'$

$= A(C_u'(C_{ul}\Sigma C_{ul}')^{-1} C_u)^{-1} C_u'(C_{ul}\Sigma C_{ul}')^{-1} C_{ul} \Sigma A_u' - A(C_u'(C_{ul}\Sigma C_{ul}')^{-1} C_u)^{-1} A'$

We multiply $E\underline{x}u'$ from the right a non-singular matrix $[Q_u, R]$, we have,

$[A(C_u'(C_{ul}\Sigma C_{ul}')^{-1} C_u)^{-1} C_u'(C_{ul}\Sigma C_{ul}')^{-1} C_{ul} \Sigma A_u' - A(C_u'(C_{ul}\Sigma C_{ul}')^{-1} C_u)^{-1} A']Q_u$

$= A(C_u'(C_{ul}\Sigma C_{ul}')^{-1} C_u)^{-1} C_u' - A(C_u'(C_{ul}\Sigma C_{ul}')^{-1} C_u)^{-1} C_u'$

$= 0.$

We also have

$$[A(C_u'(C_{ul}\Sigma C_{ul}')^{-1} C_u)^{-1} C_u'(C_{ul}\Sigma C_{ul}')^{-1} C_{ul} \Sigma A_u' - A(C_u'(C_{ul}\Sigma C_{ul}')^{-1} C_u)^{-1} A']R = 0$$

because $A_u'R=0$ and $A'R = 0$.

Therefore, $E\underline{x}u' = 0$, that is, $\underline{x}$ and u are uncorrelated and $x = \underline{x} + u$. Q.E.D.

Remark: One can show x_1 and x_A as defined in Lemma 1 are just $\lambda_1 x_{1A} + (1-\lambda_1)x_{1B}$ and $\lambda_A x_{1A}^* + (1-\lambda_A)x_{2A}^*$, where λ_1 and λ_A are given by (9) and (10), respectively. Similarly for x_2 and x_B.

Lemma 2: Let $t_1(x_{1A}, x_{1B}, x_{2A}, x_{2B})$ be any transfer scheme for manager 1 in the U-form. Fix the effort levels at e' for all managers but manager A in the M-form and manager 1 in the U-form. There exists an equivalent transfer scheme for manager A in the M-form, i.e., a scheme $t_A(x_{1A}^*, x_{2A}^*, x_{1B}^*, x_{2B}^*)$ such that for all transfer values τ and all effort levels e by manager A or manager 1,

$$\text{Prob}(\, t_A(x_{1A}^*, x_{2A}^*, x_{1B}^*, x_{2B}^*) = \tau \mid e) = \text{Prob}(\, t_1(x_{1A}, x_{1B}, x_{2A}, x_{2B}) = \tau \mid e),$$

if and only if

$$\text{Var}(\epsilon_A \mid \epsilon_B) \leq \text{Var}(\epsilon_1 \mid \epsilon_2),$$

where

$$(\epsilon_1, \epsilon_2)' = (C_u'(C_{ul}\Sigma C_{ul}')^{-1} C_u)^{-1} C_u'(C_{ul}\Sigma C_{ul}')^{-1} C_{ul}\, \xi, \text{ and}$$

$$(\epsilon_A, \epsilon_B)' = (C_m'(C_{ml}\Sigma C_{ml}')^{-1} C_m)^{-1} C_m'(C_{ml}\Sigma C_{ml}')^{-1} C_{ml}\, \xi.$$

Proof: From our analysis of the stripped-down model, Var $(\epsilon_A \mid \epsilon_B) \leq \mathrm{Var}(\epsilon_1 \mid \epsilon_2)$ implies that there exist constant α, β, γ and random noise z uncorrelated with (x_A, x_B) such that for all $e_1 = e_A$,

$$(x_1, x_1, x_2, x_2) = (x_A^* - \alpha x_B^*, x_A^* - \alpha x_B^*, \beta x_B^*, \beta x_B^*) + (z, z, \gamma, \gamma) \text{ in distribution.}$$

By Lemma 1, we can choose a random vector (w_1, w_2, w_3, w_4) such that

(i) $\mathrm{Var}(w_1, w_2, w_3, w_4) = \mathrm{Var}(u_1, u_2, u_3, u_4) = \mathrm{Var}(x_{1A}, x_{1B}, x_{2A}, x_{2B}) - \mathrm{Var}(x_1, x_1, x_2, x_2)$; and

(ii) (w_1, w_2, w_3, w_4) is independent of (x_1, x_2), (x_A^*, x_B^*), and z.

Then we obtain,

$$\mathrm{Var}(x_{1A}, x_{1B}, x_{2A}, x_{2B})$$
$$= \mathrm{Var}(x_1, x_1, x_2, x_2) + [\mathrm{Var}(x_{1A}, x_{1B}, x_{2A}, x_{2B}) - \mathrm{Var}(x_1, x_1, x_2, x_2)]$$
$$= \mathrm{Var}(x_A^* - \alpha x_B^*, x_A^* - \alpha x_B^*, \beta x_B^*, \beta x_B^*) + \mathrm{Var}(z, z, \gamma, \gamma) + \mathrm{Var}(w_1, w_2, w_3, w_4)$$
$$= \mathrm{Var}(x_A^* - \alpha x_B^* + z + w_1, x_A^* - \alpha x_B^* + z + w_2, \beta x_B^* + \gamma + w_3, \beta x_B^* + \gamma + w_4).$$

Furthermore,

$$E\,(x_{1A}, x_{1B}, x_{2A}, x_{2B})$$
$$= E\,(x_1, x_1, x_2, x_2)$$
$$= E\,(x_A^* - \alpha x_B^* + z + w_1, x_A^* - \alpha x_B^* + z + w_2, \beta x_B^* + \gamma + w_3, \beta x_B^* + \gamma + w_4).$$

Therefore we obtain

$$(x_{1A}, x_{1B}, x_{2A}, x_{2B}) = (x_A^*-\alpha x_B^*+z+w_1, x_A^*-\alpha x_B^*+z+w_2, \beta x_B^*+\gamma+w_3, \beta x_B^*+\gamma+w_4)$$

in distribution.

Finally, we define

$$t_A(x_{1A}^*, x_{2A}^*, x_{1B}^*, x_{2B}^*)$$
$$= t_1(x_A^*-\alpha x_B^*+z+w_1, x_A^*-\alpha x_B^*+z+w_2, \beta x_B^*+\gamma+w_3, \beta x_B^*+\gamma+w_4),$$

which in distribution is the same as $t_1(x_{1A}, x_{1B}, x_{2A}, x_{2B})$. Q.E.D.

Applying Lemma 2, we can compare the M-form and U-form straightforwardly as follows:

Proposition 3: Incentives under the M-form are at least as good as those under the U-form (in the sense that any U-form incentive scheme can be replicated by an M-form incentive scheme) provided that

$$\max \{ \operatorname{Var} (\epsilon_A \mid \epsilon_B), \operatorname{Var} (\epsilon_B \mid \epsilon_A) \}$$
$$\leq \max \{ \operatorname{Var} (\epsilon_1 \mid \epsilon_2), \operatorname{Var} (\epsilon_2 \mid \epsilon_1) \},$$

and

$$\min \{ \operatorname{Var} (\epsilon_A \mid \epsilon_B), \operatorname{Var} (\epsilon_B \mid \epsilon_A) \}$$
$$\leq \min \{ \operatorname{Var} (\epsilon_1 \mid \epsilon_2), \operatorname{Var} (\epsilon_2 \mid \epsilon_1) \},$$

where ϵ_B, ϵ_A, ϵ_1, and ϵ_2 are given by Lemma 2.

When there is symmetry across regions and across industries and no correlation between industrial and regional shocks, the formulas of Proposition 3 simplify into the following condition:

Corollary: Assume $\operatorname{Var}(\delta_A) = \operatorname{Var}(\delta_B) = V_R^2$, $\operatorname{Var}(\theta_1) = \operatorname{Var}(\theta_2) = V_I^2$, $\operatorname{Cov}(\theta_i, \delta_r) = 0$ for $i=1,2$ and $r=A,B$. Let $V_{12} = \operatorname{Cov}(\theta_1, \theta_2)$ and $V_{AB} = \operatorname{Cov}(\delta_A, \delta_B)$. Then, incentives under the M-form are at least as good as those under the U-form if and only if $V_R^2 - V_{AB} \leq V_I^2 - V_{12}$.

4. Test Statistics

We wish to test the inequalities in Proposition 3 using Chinese data. To do so, we need to derive the test statistics for the conditional variances under the M-form and U-form organizations.

Suppose we have sample industry-specific and region-specific shocks

$$\xi_t' = (\theta_{1t}, \theta_{2t}, \delta_{At}, \delta_{Bt})$$

$t=1,\ldots,T$, which is drawn from a population $N(0,\Sigma)$, where

$$\Sigma = \begin{bmatrix} \sigma_{11}, & \sigma_{12}, & \sigma_{1A}, & \sigma_{1B} \\ \sigma_{21}, & \sigma_{22}, & \sigma_{2A}, & \sigma_{2B} \\ \sigma_{A1}, & \sigma_{A2}, & \sigma_{AA}, & \sigma_{AB} \\ \sigma_{B1}, & \sigma_{B2}, & \sigma_{BA}, & \sigma_{BB} \end{bmatrix}$$

Let sample mean and sample covariance be, respectively,

$\xi_T = (1/T)\, \Sigma_{t=1}^{T}\, \xi_t,$ and

$$\hat{\Sigma} = (1/T)\, \Sigma_{t=1}^{T}\, (\xi_t - \xi_T)(\xi_t - \xi_T)' = \begin{bmatrix} \hat{\sigma}_{11}, & \hat{\sigma}_{12}, & \hat{\sigma}_{1A}, & \hat{\sigma}_{1B} \\ \hat{\sigma}_{21}, & \hat{\sigma}_{22}, & \hat{\sigma}_{2A}, & \hat{\sigma}_{2B} \\ \hat{\sigma}_{A1}, & \hat{\sigma}_{A2}, & \hat{\sigma}_{AA}, & \hat{\sigma}_{AB} \\ \hat{\sigma}_{B1}, & \hat{\sigma}_{B2}, & \hat{\sigma}_{BA}, & \hat{\sigma}_{BB} \end{bmatrix}.$$

Vectorizing Σ and $\hat{\Sigma}$, we have

$$T^{1/2}\, (\mathrm{vec}(\hat{\Sigma}) - \mathrm{vec}(\Sigma)) \dashrightarrow N(0, B(\Sigma)), \text{ as } T \dashrightarrow \infty,$$

where $B(\Sigma)$ is a 16x16 matrix with elements $b_{ij,kl} = \sigma_{ik}\sigma_{jl} + \sigma_{il}\sigma_{jk}$, $i,j,k,l = 1,2,A,B$ (Theorem 3.4.4 from Anderson, 1984).

We want to test

$$\mathrm{Var}\, (\epsilon_A \mid \epsilon_B) \leq \mathrm{Var}(\epsilon_1 \mid \epsilon_2) \text{ and } \mathrm{Var}\, (\epsilon_B \mid \epsilon_A) \leq \mathrm{Var}(\epsilon_2 \mid \epsilon_1),$$

where

$$(\epsilon_1, \epsilon_2)' = (C_u'(C_{ul}\Sigma C_{ul}')^{-1}\, C_u)^{-1}\, C_u'(C_{ul}\Sigma C_{ul}')^{-1}\, C_{ul}\, \xi'$$

$$(\epsilon_A, \epsilon_B)' = (C_m'(C_{ml}\Sigma C_{ml}')^{-1}\, C_m)^{-1}\, C_m'(C_{ml}\Sigma C_{ml}')^{-1}\, C_{ml}\, \xi'$$

Write $\epsilon = (\epsilon_1, \epsilon_2, \epsilon_A, \epsilon_B)'$, and

$$\Phi(\Sigma) = \begin{bmatrix} \Phi_u(\Sigma) \\ \Phi_m(\Sigma) \end{bmatrix} = \begin{bmatrix} (C_u'(C_{ul}\Sigma C_{ul}')^{-1}\, C_u)^{-1}\, C_u'(C_{ul}\Sigma C_{ul}')^{-1}\, C_{ul} \\ (C_m'(C_{ml}\Sigma C_{ml}')^{-1}\, C_m)^{-1}\, C_m'(C_{ml}\Sigma C_{ml}')^{-1}\, C_{ml} \end{bmatrix}.$$

Then

$$\epsilon = \Phi(\Sigma)\, \xi \sim N(0, \Omega),$$

where

$$\Omega = \Phi(\Sigma)\, \Sigma\, \Phi(\Sigma)' = \begin{bmatrix} \omega_{11}, & \omega_{12}, & \omega_{1A}, & \omega_{1B} \\ \omega_{21}, & \omega_{22}, & \omega_{2A}, & \omega_{2B} \\ \omega_{A1}, & \omega_{A2}, & \omega_{AA}, & \omega_{AB} \\ \omega_{B1}, & \omega_{B2}, & \omega_{BA}, & \omega_{BB} \end{bmatrix}$$

Define

$$h_1(\Omega) = V(\epsilon_A \mid \epsilon_B) - V(\epsilon_1 \mid \epsilon_2) = [\omega_{AA} - \omega_{AB}\omega_{BB}^{-1}\omega_{BA}] - [\omega_{11} - \omega_{12}\omega_{22}^{-1}\omega_{21}],$$

$$h_2(\Omega) = V(\epsilon_B \mid \epsilon_A) - V(\epsilon_2 \mid \epsilon_1) = [\omega_{BB} - \omega_{BA}\omega_{AA}^{-1}\omega_{AB}] - [\omega_{22} - \omega_{21}\omega_{11}^{-1}\omega_{12}], \text{ and}$$

$$h(\Omega)' = (h_1(\Omega), h_2(\Omega)).$$

We want to test:

$$H_0\text{: } h(\Omega) \leq 0 \text{ vs. } H_1\text{: } h(\Omega) \in R^2$$

Using matrix differential calculus (Magnus and Neudecker, 1988, pp.27-31, 46-48, 94-97, 147-149, and 173-184), we derive the following

<u>Lemma 3:</u> We have

$$T^{1/2} (h(\Omega(\hat{\Sigma})) - h(\Omega(\Sigma))) \dashrightarrow N(0, D(\Sigma)),$$

with $D(\Sigma) = (dh(\Omega)/d(vec(\Omega))')\ (dvec(\Omega)/d(vec(\Sigma))')\ B(\Sigma)\ (dvec(\Omega)/d(vec(\Sigma))')'\ (dh(\Sigma)/d(vec(\Sigma))')'$,

$$dh(\Omega)/d(vec(\Omega))' = \begin{bmatrix} dh_1(\Omega)/d(vec(\Omega))' \\ dh_2(\Omega)/d(vec(\Omega))' \end{bmatrix}$$

$$= \begin{bmatrix} -1, & \omega_{22}^{-1}\omega_{21}, & 0, & 0, & \omega_{12}\omega_{22}^{-1}, & -\omega_{12}\omega_{22}^{-2}\omega_{21}, & 0, & 0, & 0, & 0, & 1, & -\omega_{BB}^{-1}\omega_{BA}, & 0, & 0, & -\omega_{AB}\omega_{BB}^{-1}, & \omega_{AB}\omega_{BB}^{-2}\omega_{BA} \\ -\omega_{21}\omega_{11}^{-2}\omega_{12}, & \omega_{21}\omega_{11}^{-1}, & 0, & 0, & \omega_{11}^{-1}\omega_{12}, & -1, & 0, & 0, & 0, & 0, & \omega_{BA}\omega_{AA}^{-2}\omega_{AB}, & -\omega_{BA}\omega_{AA}^{-1}, & 0, & 0, & -\omega_{AA}^{-1}\omega_{AB}, & 1 \end{bmatrix},$$

$$dvec(\Omega)/d(vec(\Sigma))' = \Phi(\Sigma)\otimes\Phi(\Sigma) + (K+I)(I\otimes\Phi(\Sigma)\Sigma)V(\Sigma),$$

where $\otimes$ is the Kronecker product, K is the commutation matrix, I is the identity matrix, and $V(\Sigma)$ is defined below.[8]

<u>Proof:</u> Using matrix differential calculus we obtain

$$d\Omega = d(\Phi(\Sigma)\Sigma\Phi(\Sigma)') = d\Phi(\Sigma)\ \Sigma\Phi(\Sigma)' + \Phi(\Sigma)\ d\Sigma\ \Phi(\Sigma)' + \Phi(\Sigma)\Sigma\ d\Phi(\Sigma)'. \tag{3}$$

Using matrix differential calculus again, for i = u,m, we obtain:

---

[8] Let U be an $m \times n$ matrix and W be an $p \times q$ matrix. The Kronecker product of U and W ($U \otimes W$) is defined by the following $mp \times nq$ matrix

$$\begin{bmatrix} u_{11}W \ldots u_{1n}W \\ \\ u_{m1}W \ldots u_{mn}W \end{bmatrix}$$

(Magnus and Neudecker, 1988, p.27).

Let A be an $m \times n$ matrix. The commutation matrix K is such that

$$K \text{ vec } A = \text{vec } A'$$

(Magnus and Neudecker, 1988, pp.46-48).

$$d\Phi_i(\Sigma)' = d\{C_{i1}'(C_{i1}\Sigma C_{i1})^{-1}C_i(C_i'(C_{i1}\Sigma C_{i1})^{-1}C_i)^{-1}\}$$

$$= -C_{i1}'(C_{i1}\Sigma C_{i1}')^{-1}C_{i1}(d\Sigma)\Phi_i(\Sigma)' + C_{i1}'(C_{i1}\Sigma C_{i1}')^{-1}C_i\Phi_i(\Sigma)(d\Sigma)\Phi_i(\Sigma)'.$$

Thus,

$$d\Phi(\Sigma)' = [-C_{u1}'(C_{u1}\Sigma C_{u1}')^{-1}C_{u1}(d\Sigma)\Phi_u(\Sigma)' + C_{u1}'(C_{u1}\Sigma C_{u1}')^{-1}C_u\Phi_u(\Sigma)(d\Sigma)\Phi_u(\Sigma)',$$
$$-C_{m1}'(C_{m1}\Sigma C_{m1}')^{-1}C_{m1}(d\Sigma)\Phi_m(\Sigma)' + C_{m1}'(C_{m1}\Sigma C_{m1}')^{-1}C_m\Phi_m(\Sigma)(d\Sigma)\Phi_m(\Sigma)']$$

and

$$\mathrm{vec}(d\Phi(\Sigma)') = \mathrm{vec}\ \{[-(\Phi_u(\Sigma)\otimes C_{u1}'(C_{u1}\Sigma C_{u1}')^{-1}C_{u1}) + (\Phi_u(\Sigma)\otimes C_{u1}'(C_{u1}\Sigma C_{u1}')^{-1}C_u\Phi_u(\Sigma)]d(\mathrm{vec}(\Sigma)),$$
$$[-(\Phi_m(\Sigma)\otimes C_{m1}'(C_{m1}\Sigma C_{m1}')^{-1}C_{m1}) + (\Phi_m(\Sigma)\otimes C_{m1}'(C_{m1}\Sigma C_{m1}')^{-1}C_m\Phi_m(\Sigma)]d(\mathrm{vec}(\Sigma))\}$$
$$= \begin{bmatrix} -\Phi_u(\Sigma)\otimes C_{u1}'(C_{u1}\Sigma C_{u1}')^{-1}C_{u1} + \Phi_u(\Sigma)\otimes C_{u1}'(C_{u1}\Sigma C_{u1}')^{-1}C_u\Phi_u(\Sigma) \\ -\Phi_m(\Sigma)\otimes C_{m1}'(C_{m1}\Sigma C_{m1}')^{-1}C_{m1} + \Phi_m(\Sigma)\otimes C_{m1}'(C_{m1}\Sigma C_{m1}')^{-1}C_m\Phi_m(\Sigma) \end{bmatrix} d(\mathrm{vec}(\Sigma)).$$

Write

$$V(\Sigma) = \begin{bmatrix} -\Phi_u(\Sigma)\otimes C_{u1}'(C_{u1}\Sigma C_{u1}')^{-1}C_{u1} + \Phi_u(\Sigma)\otimes C_{u1}'(C_{u1}\Sigma C_{u1}')^{-1}C_u\Phi_u(\Sigma) \\ -\Phi_m(\Sigma)\otimes C_{m1}'(C_{m1}\Sigma C_{m1}')^{-1}C_{m1} + \Phi_m(\Sigma)\otimes C_{m1}'(C_{m1}\Sigma C_{m1}')^{-1}C_m\Phi_m(\Sigma) \end{bmatrix}.$$

Then

$$\mathrm{vec}(d\Phi(\Sigma)') = V(\Sigma)\ d(\mathrm{vec}(\Sigma)), \text{ and}$$

$$\mathrm{vec}(d\Phi(\Sigma)) = \mathrm{vec}(d\Phi(\Sigma)')' = K\ \mathrm{vec}(d\Phi(\Sigma)')$$

$$= K\ V(\Sigma)\ d(\mathrm{vec}(\Sigma)).$$

Therefore, using (3), we get

$$d\mathrm{vec}(\Omega) = \mathrm{vec}(d\Omega)$$

$$= (\Phi(\Sigma)\Sigma\otimes I)\mathrm{vec}(d\Phi(\Sigma)) + (\Phi(\Sigma)\otimes\Phi(\Sigma))\mathrm{vec}(d\Sigma) + (I\otimes\Phi(\Sigma)\Sigma)\mathrm{vec}(d\Phi(\Sigma)')$$

$$= (\Phi(\Sigma)\Sigma\otimes I)KV(\Sigma)d(\mathrm{vec}(\Sigma)) + (\Phi(\Sigma)\otimes\Phi(\Sigma))\mathrm{vec}(d\Sigma) + (I\otimes\Phi(\Sigma)\Sigma)V(\Sigma)d(\mathrm{vec}(\Sigma))$$

$$= [\Phi(\Sigma)\otimes\Phi(\Sigma) + (K+I)(I\otimes\Phi(\Sigma)\Sigma)V(\Sigma)]d(\mathrm{vec}(\Sigma)).$$

Hence,

$$d\mathrm{vec}(\Omega)/d(\mathrm{vec}(\Sigma))' = \Phi(\Sigma)\otimes\Phi(\Sigma) + (K+I)(I\otimes\Phi(\Sigma)\Sigma)V(\Sigma).\ \text{Q.E.D.}$$

We now consider the problem

$$h(\Omega(\hat{\Sigma})) = h(\Omega(\Sigma)) + T^{-1/2}\ v,$$

where $v \sim N(0, D(\Sigma))$. We test $h(\Omega(\Sigma)) \leq 0$ vs. $h(\Omega(\Sigma)) \epsilon R^2$ with the test statistic

$$(11) \qquad TS_b = \min \{ T (h(\Omega(\hat{\Sigma})) - x)' D(\hat{\Sigma})^{-1} (h(\Omega(\hat{\Sigma})) - x), \text{ s.t. } x \leq 0\},$$

which has the asymptotic distribution of a weighted sum of χ_1^2 and χ_2^2 distributions and the weights are (½) and w respectively, where $0 \leq w = (½)\pi^{-1} \arccos(\rho_{12}) \leq ½$ (ρ_{12} is the correlation coefficient associated with D) (Wolak, 1987).

Define c such that

$$(12) \qquad (½) \text{ prob } (\chi_1^2 \geq c) + (w) \text{ prob } (\chi_2^2 \geq c) = 0.01\alpha.$$

Then if $TS_b \leq c$, H_0 is accepted; and if $TS_b > c$, H_0 is rejected (at α% significance level).

An easier method without using the weight w is to find lower bound c_ℓ and upper bound c_u which are solutions to:

$$(1/2) \text{ prob } (\chi_1^2 \geq c_\ell) = 0.01\alpha$$

$$(13) \qquad (1/2) \text{ prob } (\chi_1^2 \geq c_u) + (1/2) \text{ prob } (\chi_2^2 \geq c_u) = 0.01\alpha.$$

Obviously, $c_\ell < c < c_u$. We calculate from the χ^2 distribution table that $c_\ell = 2.7$ and $c_u = 5.2$ at 5% significance level. If $TS_b \leq c_\ell$, H_0 is accepted; if $TS_b \geq c_u$, H_0 is rejected; and if $c_\ell < TS_b < c_u$, it is inconclusive (at α% significance level). However, when it is inconclusive, an exact test (with the above c) has to be done and therefore (13) has to be solved.

5. An Application to China

A. The M-form Economy of China

Chandler (1966) and Williamson (1975) characterized the two predominant organizational forms of business corporations: the U-form and the M-form. The U-form corporation has a unitary structure and is organized along functional lines. It was popular in the late 1800s and early 1900s. The M-form corporation, by contrast, consists of reasonably self-contained divisions and emerged in the 1920s. Recently, Qian and Xu (1993) proposed comparing the transition paths of economies in Eastern Europe and the former Soviet Union (EEFSU) with that of China from the standpoint of organizational structures. They observed that the economies of EEFSU resembled U-forms (also known as "branch organizations"), whereas the Chinese hierarchy has taken an M-form structure, in which divisions correspond to regions.[9]

It is well documented that enterprises in EEFSU were grouped by industry, each of which was supervised by a ministry (Gregory and Stuart 1989).[10] In order to fully exploit scale economies and avoid conflicting operations, there was little overlap of functions across ministries. Enterprises were highly specialized. Because of the strong interdependence between enterprises in different regions, comprehensive planning and administrative coordination between ministries at the top level of government were crucial for the normal operation of the economy.

China's planning system began by imitating the U-form Soviet model in its first five-year plan (1953-57), which was formulated with the help of the Soviets. As serious economic problems emerged, however, China started to deviate from the Soviet scheme and moved toward an M-form economy in the late 1950s. In the process, "blocks" (*kuaikuai*), i.e., regions, replaced "branches" (*tiaotiao*), i.e., specialized ministries, as the foundation of the planning system. In fact, there are now six regional levels for adminstration: central, provincial, prefecture, county, township and village (a municipality can have the rank of province, prefecture or county). Regions at the county level and above are relatively self-contained; indeed, they are nearly self-sufficient in function. Hence, the Chinese M-form is "deep" and differs from the U-form of the Soviet Union and Eastern Europe in a thorough-going way.

B. Evidence on Conditional Variances of Industrial and Regional Shocks

We now investigate whether the conditional variance condition of Proposition 3 holds empirically. Implicitly, we are comparing the Chinese organizational form (M-form) with a

---

[9] Qian and Xu (1993) discussed the overall costs and benefits of U-forms and M-forms in terms of scale economies, incentives, and coordination, and also the implications of these costs and benefits for alternative approaches to reform.

[10] Khrushchev tried to change the Soviet economy from a U-form to an M-form by abolished the ministries all together and introduced 105 Regional Economic Councils (*Sovnarkhozy*) in 1957. But his endeavor soon failed (Gregory and Stuart, 1989).

hypothetical U-form. In this U-form, all firms would be organized into hypothetical industrial ministries (although some industrial ministries actually exist in China, most state firms are under the control of regional governments). We will compare conditional variances of regional and industrial shocks under M-form and U-form arrangements.

Our data set consists of 520 Chinese state-owned enterprise from 1986 to 1991.[11] The enterprises sampled are drawn from more than thirty manufacturing industries, located in major cities in 20 different provinces. The data set contains industry classification codes and location codes for each firm.

In our regressions, we group the data by region and by industry so that a proper sample size is maintained. Moreover, as much as possible, we try to reflect actual organization. For industries, we group the data into units similar to Eastern European-style ministries, with headings such as "machinery," "chemicals," and "textiles." Indeed, because of data limitations, we concentrate on these three industries in particular, since they have the largest sample sizes. Because sample sizes in individual cities are too small, our regional exercises are carried out in two ways. In the first scheme (Table 1), the cities are grouped into provinces. We select the five provinces with the largest sample sizes. These are Liaoning, Hubei, Hunan, Jiangsu (which includes Shanghai), and Hebei (which includes Beijing and Tianjin). In the second scheme (Table 2), we organize cities into "large regions," where each region contains three to six neighboring provinces. We choose the four regions with the largest sample sizes. These are "East" (Jiangsu, Anhui, Zhejiang, and Shanghai), "North" (Hebei, Henan, Shandong, Shanxi, Beijing, Tianjin), "Northeast" (Heilongjiang, Jilin, Liaoning), and "Central South" (Hubei, Hunan, Guangdong, Guangxi, and Fujian), which comprise a total of 18 provinces.

We use the log-linear Cobb-Douglas production function as our regression model to estimate industry-specific shocks (θ) and region-specific shocks (δ). For every industry i, region r, and period t, we include dummy variables D_{it}^{I} and D_{rt}^{R}. The coefficients of these dummies serve as proxies for

[11] The data were collected by the China System Reform Research Institute, Beijing, China.

the industry-specific and region-specific shocks in the given period. Formally, we have

$$
\begin{aligned}
& E(y \mid L, k) \\
&= (\beta + \Sigma_{i=1}^{I} \beta_i D_i^I)L + (\gamma + \Sigma_{i=1}^{I} \gamma_i D_i^I)K + \Sigma_{t=1}^{T} D_t \, \eta_t + \Sigma_{t=1}^{T} \Sigma_{r=1}^{R} D_{rt}^{R} \, \delta_{rt} + \Sigma_{t=1}^{T} \Sigma_{i=1}^{I} D_{it}^{I} \, \theta_{it} \\
&= (\beta + \Sigma_{i=1}^{I} \beta_i D_i^I)L + (\gamma + \Sigma_{i=1}^{I} \gamma_i D_i^I)K + \Sigma_{t=1}^{T} D_t \, \eta_t \\
&\quad + \Sigma_{t=1}^{T} [D_{Rt}^{R} \, \delta_{Rt} + \Sigma_{r=1}^{R-1} D_{rt}^{R} \, \delta_{rt}] + \Sigma_{t=1}^{T} [D_{It}^{I} \theta_{It} + \Sigma_{i=1}^{I-1} D_{it}^{I} \, \theta_{it}] \\
&= (\beta + \Sigma_{i=1}^{I} \beta_i D_i^I)L + (\gamma + \Sigma_{i=1}^{I} \gamma_i D_i^I)K + \Sigma_{t=1}^{T} D_t \, \eta_t \\
&\quad + \Sigma_{t=1}^{T} \{(D_t^T - \Sigma_{r=1}^{R-1} D_{rt}^{R})\delta_{Rt} + \Sigma_{r=1}^{R-1} D_{rt}^{R} \, \delta_{rt}\} + \Sigma_{t=1}^{T} \{(D_t^T - \Sigma_{i=1}^{I-1} D_{it}^{I})\theta_{It} + \Sigma_{i=1}^{I-1} D_{it}^{I} \, \theta_{it}\} \\
&= (\beta + \Sigma_{i=1}^{I} \beta_i D_i^I)L + (\gamma + \Sigma_{i=1}^{I} \gamma_i D_i^I)K + \Sigma_{t=1}^{T} D_t^T \, \zeta_t + \Sigma_{t=1}^{T} \Sigma_{r=1}^{R-1} D_{rt}^{R} \, \delta_{rt}' + \Sigma_{t=1}^{T} \Sigma_{i=1}^{I-1} D_{it}^{I} \, \theta_{it}'
\end{aligned}
$$

where

$$\zeta_t = \eta_t + \delta_{Rt} + \theta_{It},$$

$$\delta_{rt}' = \delta_{rt} - \delta_{Rt}, \text{ and}$$

$$\theta_{it}' = \theta_{it} - \theta_{It},$$

for $t = 1, \ldots, T$; $r = 1,2, \ldots, R-1$; and $i = 1,2, \ldots, I-1$.

Because of an identification problem,[12] we cannot estimate $(\theta_{it}, \delta_{rt})$ directly. Instead, we drop the dummy variables of one region and one industry, and estimate the coefficients of the dummy variables for the remaining regions and industries. This can be interpreted as using the shocks in one region and one industry as a benchmark to estimate relative industry-specific and relative region-specific shocks $(\theta_{it}', \delta_{rt}')$.

For any three regions and three industries, $R = I = 3$, and $T = 6$, we take region 3 (or region C) and industry 3 as benchmarks. From the regressions we obtain a time series $(\theta_{1t}', \theta_{2t}', \delta_{At}', \delta_{Bt}')$, which, for notational simplicity, we denote by $\xi_t = (\theta_{1t}, \theta_{2t}, \delta_{At}, \delta_{Bt})$. In the test, we treat these estimated shocks as if they were real shocks that are uncorrelated over time.

We test the hypothesis that the conditional variances under the M-form are no greater than

[12] Dummy variables here have the following property:

$$D_t^T = \Sigma_{r=1}^{R} D_{rt}^{R} = \Sigma_{i=1}^{I} D_{it}^{I} \quad \begin{aligned} &= 1 \text{ in period } t \\ &= 0, \text{ otherwise,} \end{aligned}$$

that is, the sum of the regional dummies is the same as that of the industrial dummies creating a collinearity problem.

those under the U-form. The results are reported in Tables 1 and 2. Columns (1)-(4) report estimated conditional variances of regional shocks and industrial shocks, and column (5) reports the estimated test statistic TS_b.

Of the 63 results in Table 1, there are 44 cases in which $TS_b = 0$, that is, the estimated means of both conditional variances under the M-form are smaller than their counterparts under the U-form. In these cases, our hypothesis cannot be rejected at any significance level. In the remaining 29 cases where $TS_b > 0$, no value of TS_b is larger than 1.64 which is the lower bound at the 1% significance level. Thus, the hypothesis cannot be rejected at the 1% level. The results in Table 2 show that out of 36 possible pairs of comparisons, $TS_b = 0$ in 25 pairs. In the remaining 11 pairs our test statistic TS_b is positive, that is, at least one estimated mean conditional variance under the M-form is greater than its counterpart under the U-form. At the 5% significance level, we have the lower bound $c_\ell = 2.7$ and the upper bound $c_u = 5.2$. In 10 out of the 11 pairs $TS_b < 2.7$, and only one pair ($TS_b = 3.16$) falls into the inconclusive interval (2.7, 5.2). Therefore, except for one case out of 36, our hypothesis cannot be rejected at the 5% significance level. In view of Proposition 3, this test result suggests that, for the case of Chinese enterprises, the M-form provides better information than the U-form on relative performance.

C. Evidence on Regional Yardstick Competition

The findings of section 5B suggest that the M-form facilitates yardstick competition, but one may ask whether such relative performance evaluations are actually used in China. We now provide some evidence that they are.

Anecdotes

The Chinese central government has pursued an explicit policy during reform to encourage regions to "get rich first." Indeed, relative performance criteria are sometimes formally incorporated in the procedures for determining government officials' promotions and bonuses. For example, from her (1995) field work, Whiting reports that county governments use the annual ranking of townships (by profit rate on total capital) as a primary criterion to evaluate township government officials. As she notes, "Such a system also allowed county officials to compare the performance of township leaders across locales and helped identify the most competent cadres for promotion" (Chapter 2 in Whiting, 1995).

All regional governments lay great stress on relative performance in their public actions and pronouncements.[13] Government statistical reports and the mass media regularly publish rankings of regions in terms of their performances in growth, profit, foreign investment, etc. Most authoritative national statistical books, such as *the Almanac of China's Economy*, the *Finance Year Book of China*, *the Market Statistical Yearbook of China*, *the China Means of Production Market Statistical Yearbook*, and *the China Statistical Yearbook*, publish national rankings of provinces. Almost all of the 30 provincial statistical yearbooks report rankings of cities and/or counties within the provinces every year.

Systematic evidence

We next provide some evidence on promotions of regional government officials based on relative performance evaluation. The Chinese political system is still under one-party rule, and so the representation of a region in the Party Central Committee indicates the status and power of the regional government officials. Reflecting the increased importance of regions in government, regional representation in the Party's Congress and Central Committee as a whole has increased significantly over the reform period. For example, in the 14th Party Congress, more than 70% of delegates were from provinces, whereas only about 16% were from the central government and central Party organs (Saich, 1992).

We use a province's representation in the Party's Central Committee as a proxy for the promotion chances of officials in that province. We normalize the representation by the province's

[13] For instance, when Ningguo county of Anhui province was ranked in the top one hundred counties in China, the governor of the province made some highly publicized inspection visits to the county (*Statistic Yearbook of Anhui*, 1995).

population so as to use the "per capita number of Central Committee members" as an index. This is the ratio between the number of Central Committee members from that region and the region's population. We measure economic performance of a province by its growth rate in "national income" (the rough equivalent of GDP).

Table 3 lists the ranking of provincial per capita number of Central Committee members in the 11th Party Congress in 1977 ($prank77_r$) and in the 13th Party Congress in 1987 ($prank87_r$), and the ranking of provincial economic performance in growth rate one year before the Party Congress, that is, in 1976 ($erank76_r$) and in 1986 ($erank86_r$) respectively (data for Ningxia and Tibet are not available). The 11th Central Committee was formed before reform started, and at that time promotion criteria were mostly political. It could, therefore, be viewed as a benchmark. The 13th Central Committee was formed in 1987 when reform had been ongoing for almost a decade, and improving economic performance was officially stated as the central task of the Party. Table 3 shows that some provinces (e.g., Fujian, Jiangsu, Xinjiang, Zhejiang) improved their relative growth rankings, and their relative rankings of representation in the Central Committee also increased significantly. In contrast, the relative growth rankings of some provinces (e.g., Anhui, Guangxi, and Qinghai) deteriorated, and so did their rankings in representation in the Central Committee.[14]

To investigate the use of relative performance incentives, we focus on how the change of relative ranking in economic performance is related to the change of relative ranking in per capita number of Central Committee members. A simple regression model using the data in Table 3 shows the following result (standard error of the estimated coefficient is in parentheses):

$$PINDEX_r = -0.453 + \underset{(0.246)}{1.76}\, EINDEX_r, \quad R^2 = 0.671.$$

where

[14] There are of course important political factors that also had influence on the selection of the Central Committee members. Before reform, provinces such as Hunan, Hubei, and Jiangxi provinces were over-represented in the Central Committee because these were the home provinces of many revolutionary leaders (e.g. Mao Zedong was from Hunan), and other provinces such as Beijing were under-represented because of the purge in the Cultural Revolution, which ended just before the 11th Party Congress. Furthermore, some provinces such as Xinjiang have always been over-represented because of their political significance.

$$EINDEX_r = 10*\{(1/erank86_r) - (1/erank76_r) + (1/erank86_r)^2\},$$

and

$$PINDEX_r = 10*\{(1/prank87_r) - (1/prank77_r) + (1/prank87_r)^2\}.$$

For province r, $EINDEX_r$ is the index that measures the change in rank in economic performance between 1976 and 1986, while $PINDEX_r$ is the index that measures the change in rank in political position between 1977 and 1987. Note that we work with inverses. The third terms in EINDEX and PINDEX, $(1/erank86_r)^2\}$ and $(1/prank87_r)^2\}$ respectively, are incorporated into the indices of change in order to capture the feature that staying at the top requires more effort -- and thus requires greater reward -- than staying at the bottom.[15]

The significant positive correlation between the change of relative economic performance and the change of relative political position of a region suggests the use of regional yardstick competition.

6. Concluding Remarks

Our work is complementary to some other comparative studies of organizations. Arrow (1974) argues, as we do, that the information structures to which organizations give rise constitute an important characteristic by which they should be compared. Williamson (1975) suggests that in a U-form organization, the CEO may be overloaded with daily operational decisions, and therefore cannot concentrate on strategic decisions. An M-form organization helps to mitigate the overload by decentralizing decision-making. Milgrom and Roberts (1992) emphasize the advantage of the M-form corporation in coordinating finance and investment decisions. Aghion and Tirole (1995) compare the M-form and U-form from the standpoint of encouraging managerial initiative. Qian, Roland and Xu (1997) focus on organizational coordination issues, which they model as the problem of getting attributes suitably matched. They compare the M-form and U-form's efficacy in coordinating changes such as reform and innovation. Holmstrom and Milgrom (1991, 1994) study how tasks should be allocated to firms and managers when managers may perform more than one task. Aoki (1986)

[15] We have run many more regressions with alternative data sets and have obtained qualitatively similar results. Those results are available upon request.

investigates how Japanese firms are organized differently from those in the U.S. and what implications these differences have for comparative performance.

Our approach is potentially applicable to a range of economic and political issues beyond those treated in this paper. For example, it may form the basis of a theory of how nations respond to major emergencies, such as wars or natural disasters. If one thinks of such an emergency as a shock that is common to most aspects of a nation's operations, then our framework explains why a frequent reaction is to centralize authority (see also Bolton and Farrell, 1990 and Weitzman, 1974).

The framework may also be able to shed light on corporate mergers and spin-offs. Our theory has the testable implication that there should be more mergers in industries where there are strong common shocks. Moreover, it implies that one should normally see a greater correlation between the performances of two divisions within the same organization, say Cadillac and Oldsmobile within General Motors, than between those of divisions in separate corporations, say Cadillac in General Motors and Lincoln in Ford.

In closing, let us note that one hotly debated issue of current interest is how the European Union -- in particular, the EU central bank -- should be organized. One point of view favors organization according to specialization, another according to region. Our theory offers clear-cut advice: regional organization is preferable if and only if the conditional variances associated with regions are smaller than those associated with the bank's special tasks.

References

Aghion, Philippe and Jean Tirole, "Some Implications of Growth for Organizational Form and Ownership Structure," *European Economic Review*, 1995.

Anderson, T.W., *An Introduction to Multivariate Statistical Analysis* (2nd edition), New York: Wiley, 1984.

Aoki, Masahiko, "Horizontal vs. Vertical Information Structure of the Firm," *American Economic Review*, 76(5): 971-83, December 1986.

Arrow, Kenneth, *The Limits of Organization*, New York: Norton, 1974.

Bartke, Wolfgang, *Biographical Dictionary and Analysis of China's Party Leadership, 1922-1988*, NY: K.G.Saur, 1990.

Bolton, Patrick and Mathias Dewatripont, "Firms As Networks of Communication," *Quarterly Journal of Economics*, 1994.

Bolton, Patrick, and Joseph Farrell, "Decentralization, Duplication and Delay," *Journal of Political Economy*, 98(4), August, 1990.

Chandler, Alfred Jr., *Strategy and Structure*, New York: Doubleday & Company, Inc., 1962.

Calvo, Guillermo A. and Wellisz, Stanislaw, "Supervision, Loss of Control, and the Optimum Size of the Firm," *Journal of Political Economy*; 86(5): 943-52, October, 1978.

Galambos, Louis, and David Milobsky, "Organizing and Reorganizing the World Bank, 1946-1972: A Comparative Perspective," *Business History Review*, 69:156-190, 1995.

Gregory, Paul R., and Robert C. Stuart, *Soviet Economic Structure and Performance*, New York:Harper & Row, Publishers, Inc., 1981.

Grossman, Sanford J., and Oliver D. Hart, "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*; 94: 691-719, 1986.

Hart, Oliver D. and John Moore, "Property Rights and the Nature of the Firm," *Journal of Political Economy*; 98(6): 1119-1158, 1990.

Holmstrom, Bengt, "Moral Hazard in Teams," *Bell Journal of Economics*, 13:324-340, 1982.

Holmstrom, Bengt and Paul Milgrom, "Multi-task Principal-Agent Analysis: Incentive Contracts, Asset Ownership and Job Design," *Journal of Law, Economics, and Organization*, 7:24-52, 1991.

Holmstrom, Bengt and Paul Milgrom, "The Firm as an Incentive System," *American Economic Review*, 84:972-991, 1994.

Kornai, Janos, *The Socialist System*, Princeton University Press, 1992.

Lazear, Edward P. and Sherwin Rosen, "Rank-Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy*, 89(5), October, 1981.

Magnus, Jan R., and Heinz Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, John Wiley & Sons, 1988.

Milgrom, Paul, and John Roberts, *Economics, Organization, and Management*, Prentice Hall, 1992.

Mirrlees, James A., "The Optimal Structure of Incentives and Authority within an Organization," *Bell Journal of Economics*, 7(1):105-31, Spring, 1976.

Nalebuff, Barry, and Joseph Stiglitz, "Prizes and Incentives: Towards a General Theory of Compensation and Competition," *Bell Journal of Economics*, 1983.

Qian, Yingyi, "Incentives and Loss of Control in an Optimal Hierarchy," *Review of Economic Studies*, 61(3):527-544, July, 1994.

Qian, Yingyi, Gerard Roland, and Chenggang Xu, "Coordinating Changes in M-form and U-form Organizations," mimeo, Stanford University, ECARE, and London School of Economics, 1997.

Qian, Yingyi, and Chenggang Xu, "Why China's Economic Reform Differ: The M-form Hierarchy and Entry/Expansion of the Non-state Sector," *Economics of Transition*, 1(2), pp.135-170, June, 1993.

Radner, Roy, "The Organization of Decentralized Information Processing," *Econometrica*, 61(5), September, 1993.

Saich, Tony, "The Fourteenth Party Congress: A Program for Authoritarian Rule," *China Quarterly*, 132:1136-1160, December, 1992.

Shleifer, Andrei, "A Theory of Yardstick Competition," *RAND Journal of Economics*, 16(3):319-27, Autumn, 1985.

State Statistical Bureau, *Quanguo gesheng, zizhiqu, zhixiashi tongji ziliao huibian, 1949-1989* (*Collections of Historical Statistical Materials of All Provinces, Autonomous Regions, and Provincial Level Municipalities, 1949-1989*), Beijing: National Statistics Press, 1990.

Weitzman, Martin, "Price vs. Quantity," *Review of Economic Studies*, 1974.

Whiting, Susan, *The Micro-Foundations of Institutional Change in Reform China: Property Rights and Revenue Extraction in the Rural Industrial Sector*, Ph.D. dissertation, University of Michigan, 1995.

Williamson, Oliver, *Markets and Hierarchies*, New York: Free Press, 1975.

Wolak, Frank, "An Exact Test for Multiple Inequality and Equality Constraints in the Linear Regression Model," *Journal of the American Statistical Association*, 82(399), September 1987.

Table 1

Testing Industrial and Regional Variance and Conditional Variance (by Province)

	(1) $V(\epsilon_A \mid \epsilon_B)$	(2) $V(\epsilon_B \mid \epsilon_A)$	(3) $V(\epsilon_1 \mid \epsilon_2)$	(4) $V(\epsilon_2 \mid \epsilon_1)$	(5) TS_h
PR11	0.0008187	0.0007571	0.0030142	0.0024633	0
PR12	0.0009656	0.0013551	0.0028583	0.0100515	0
PR13	0.0009623	0.0021445	0.0040004	0.0016873	0.1606919
PR14	0.0003254	0.0009305	0.0038165	0.0011762	0
PR15	0.0007978	0.0003191	0.0045005	0.0016317	0
PR16	0.0015251	0.0019038	0.0028962	0.0022384	0
PR17	0.0019566	0.0010847	0.0024202	0.0015482	0
PR18	0.0006965	0.0005134	0.0019411	0.0011535	0
PR19	0.0036261	0.0010941	0.0027306	0.0027949	0.1239956
PR21	0.0035188	0.0075737	0.004532	0.0014271	1.2422196
PR22	0.0125974	0.0033652	0.0064482	0.0027321	1.1546736
PR23	0.0007243	0.0005763	0.0032571	0.0019396	0
PR24	0.0035512	0.0017348	0.0083718	0.0076624	0
PR25	0.0008053	0.0056084	0.0031696	0.0041095	0.1412761
PR26	0.0011198	0.0012982	0.0280702	0.0043337	0
PR27	0.0035219	0.0014758	0.0014134	0.0045301	1.4297282
PR28	0.0032017	0.0036066	0.0022232	0.0039406	0.4174463
PR29	0.0009339	0.0006265	0.0059682	0.0121066	0
PR31	0.0041727	0.0043189	0.0086372	0.0023456	0.5450603
PR32	0.0116553	0.00325	0.0114781	0.0047748	0
PR33	0.0003516	0.0002477	0.0032018	0.0012271	0
PR34	0.0022851	0.0012984	0.0096523	0.0044628	0
PR35	0.0014829	0.0069914	0.0063509	0.0042792	0.2695177
PR36	0.001434	0.0013309	0.0381816	0.0041133	0
PR37	0.0017801	0.0006927	0.0026892	0.0043122	0
PR38	0.0039548	0.0055769	0.0033635	0.0040625	0.1768892
PR39	0.0002523	0.000138	0.0012346	0.0083788	0

Each line of the Tables 1 and 2 corresponding to one set of results corresponding to a specific three regions and three industries with one of the them taken as a benchmark. All the 63 lines in Table 1 are divided into seven groups. The seven groups are the following: group 1: Jiangsu, Hebei, Liaoning; group 2: Jiangsu, Liaoning, Hubei; group 3: Jiangsu, Liaoning, Hunan; group 4: Hubei, Liaoning, Hunan; group 5: Hebei, Liaoning, Hubei; group 6: Hebei, Liaoning, Hunan; and group 7: Hubei, Jiangsu, Hunan. In Table 2, the 36 lines are divided into four groups: group 1: East, North, Northeast; group 2: East, North, Central South; group 3: Northeast, North, Central South; group 4: Northeast, East, Central South. Within each group, we have nine comparison results by rotating the benchmark region and the benchmark industry among the three regions and three industries within the group.

Table 1 (continued)

Testing Industrial and Regional Variance and Conditional Variance (by Province)

	(1) $V(\epsilon_A \mid \epsilon_B)$	(2) $V(\epsilon_B \mid \epsilon_A)$	(3) $V(\epsilon_1 \mid \epsilon_2)$	(4) $V(\epsilon_2 \mid \epsilon_1)$	(5) TS_h
PR41	0.0022107	0.0042633	0.0063843	0.0044466	0
PR42	0.0045825	0.0036649	0.009163	0.0031834	0.0330869
PR43	0.0042087	0.0078392	0.0067047	0.0037866	1.283058
PR44	0.0026567	0.0021724	0.0101624	0.0078117	0
PR45	0.0048027	0.0033586	0.0066146	0.0599119	0
PR46	0.0052495	0.0025457	0.0174984	0.0066967	0
PR47	0.0028687	0.002188	0.0051838	0.0071966	0
PR48	0.0044801	0.0044515	0.0031598	0.0106808	0.2720435
PR49	0.0047402	0.003174	0.0051034	0.0113329	0
PR51	0.0019198	0.00453	0.0024876	0.0018453	0.8994377
PR52	0.002041	0.0014745	0.00241	0.0018914	0
PR53	0.0005682	0.0008727	0.0025407	0.0024924	0
PR54	0.0040096	0.001339	0.003392	0.0067175	0.0483989
PR55	0.0006512	0.0012565	0.0028805	0.0143283	0
PR56	0.0008774	0.001117	0.0029163	0.0067734	0
PR57	0.0019528	0.0024209	0.0020744	0.0072606	0
PR58	0.0011385	0.0013931	0.0018707	0.0102239	0
PR59	0.0006934	0.0008906	0.0039924	0.0069858	0
PR61	0.0021139	0.0058389	0.0051549	0.0041415	0.2965862
PR62	0.0038292	0.0020893	0.0047473	0.0045516	0
PR63	0.0011121	0.0016206	0.0049029	0.0046627	0
PR64	0.0021749	0.0014442	0.0044179	0.0027746	0
PR65	0.0018646	0.0028379	0.0040928	0.012857	0
PR66	0.001113	0.0010564	0.0062856	0.0031007	0
PR67	0.0030414	0.0009696	0.0041567	0.0027338	0
PR68	0.0022042	0.0028437	0.0041358	0.0069915	0
PR69	0.001927	0.0008932	0.0063637	0.002802	0
PR71	0.0024754	0.0050791	0.0102891	0.0083058	0
PR72	0.0033104	0.002966	0.0083431	0.0060816	0
PR73	0.0033779	0.0083795	0.0154754	0.0061027	0.2567824
PR74	0.002703	0.001844	0.0066952	0.0023344	0
PR75	0.0047986	0.003487	0.0113504	0.0428088	0
PR76	0.0041169	0.0020171	0.0082085	0.0022948	0
PR77	0.0055198	0.0018183	0.0048365	0.0022892	0.0508461
PR78	0.0043981	0.0047133	0.0047888	0.0067743	0
PR79	0.0058755	0.0018537	0.0043716	0.0023171	0.2324766

Table 2

Testing Industrial and Regional Variance and Conditional Variance (by Large Region)

	(1) $V(\epsilon_A \mid \epsilon_B)$	(2) $V(\epsilon_B \mid \epsilon_A)$	(3) $V(\epsilon_1 \mid \epsilon_2)$	(4) $V(\epsilon_2 \mid \epsilon_1)$	(5) TS_h
LR11	0.0009876	0.000717	0.0025903	0.0014356	0
LR12	0.0008853	0.0005627	0.0061475	0.0015846	0
LR13	0.0020441	0.0007281	0.0026581	0.0040902	0
LR14	0.0008268	0.0007193	0.0024435	0.0027816	0
LR15	0.0006873	0.0011279	0.0092834	0.0012659	0
LR16	0.0007685	0.0011829	0.0033395	0.0015871	0
LR17	0.0005151	0.0011678	0.0016032	0.0024889	0
LR18	0.0007956	0.000601	0.0014715	0.0018291	0
LR19	0.0008224	0.0015445	0.0018787	0.0012828	0.0485825
LR21	0.0005335	0.001472	0.000835	0.0095962	0
LR22	0.0016437	0.0005659	0.0014412	0.0060323	0.0622105
LR23	0.0012722	0.0007375	0.007627	0.0068863	0
LR24	0.0009478	0.0012606	0.003262	0.0062783	0
LR25	0.0004361	0.0003698	0.0005611	0.0006254	0
LR26	0.000618	0.0006669	0.0020611	0.0013201	0
LR27	0.0015519	0.0010175	0.0072856	0.0062593	0
LR28	0.0002648	0.0016529	0.0072931	0.000737	0.1577814
LR29	0.0002031	0.0021522	0.0058771	0.000439	0.0807599
LR31	0.0005775	0.0012875	0.001453	0.0040046	0
LR32	0.0017089	0.0005776	0.0013437	0.0021527	0.5853984
LR33	0.0011759	0.0008448	0.0035925	0.0021431	0
LR34	0.0007168	0.0007843	0.0016867	0.0047914	0
LR35	0.000916	0.0006686	0.0011856	0.0015126	0
LR36	0.0010723	0.0008347	0.0053094	0.0019045	0
LR37	0.0011079	0.000722	0.0025876	0.0029192	0
LR38	0.0007325	0.0034549	0.0022474	0.002	0.3518275
LR39	0.0007093	0.0028758	0.0023074	0.0017212	0.0907792
LR41	0.0010433	0.0040568	0.0016852	0.0015885	0.9165561
LR42	0.0050836	0.0011803	0.0015829	0.0036032	3.1585944
LR43	0.0008049	0.0006459	0.0027002	0.0015179	0
LR44	0.0012797	0.0016134	0.0015776	0.0027378	0
LR45	0.0005507	0.0011771	0.0016617	0.0014342	0
LR46	0.0004704	0.0009846	0.0167323	0.001396	0
LR47	0.0031506	0.0007645	0.0014811	0.0039824	0.6695789
LR48	0.0006126	0.001314	0.0022435	0.0013026	0.0003513
LR49	0.0010576	0.0007848	0.0040202	0.0029437	0

Table 3

Provincial Ranking in Economic Performance and Political Position

	1976	1977	1986	1987
Province	Rank in Economic Growth[a] (erank76)	Rank in Party Central Committee Membership[b] (prank77)	Rank in Economic Growth[a] (erank86)	Rank in Party Central Committee Membership[b] (prank87)
Anhui	24	15	27	21
Beijing	1	27	1	1
Fujian	21	6	10	5
Gansu	8	23	20	15
Guangdong	12	21	12	9
Guangxi	11	16	25	26
Guizhou	27	24	24	22
Hebei	18	10	21	11
Heilongjiang	7	26	16	23
Henan	20	20	17	25
Hubei	22	5	14	17
Hunan	19	2	23	19
Jiangsu	16	12	7	4
Jiangxi	25	1	26	24
Jilin	14	22	18	10
Liaoning	4	17	6	7
NeiMongolia	9	14	15	14
Qinghai	3	9	5	27
Shaanxi	6	7	8	20
Shandong	10	13	11	6
Shanghai	2	4	2	2
Shanxi	23	3	19	13
Sichuan	26	18	22	18
Tianjin	5	11	4	12
Xinjiang	13	8	9	3
Yunnan	15	25	13	16
Zhejiang	17	19	3	8

Sources: (a) State Statistic Bureau, 1990; and (b) Bartke, 1990, p.374.