Review of Economic Studies (2000) 67, 359–378 0034-6527/00/00180359$02.00

Incentives, Information, and Organizational Form

ERIC MASKIN
Harvard University

YINGYI QIAN
Stanford University

and

CHENGGANG XU
London School of Economics

First version received June 1997; *final version accepted April* 1999 (*Eds.*)

We model an organization as a hierarchy of managers erected on top of a technology (here consisting of a collection of plants). In our framework, the role of a manager is to take steps to reduce the adverse consequences of shocks that affect the plants beneath him. We argue that different organizational forms give rise to different information about managers' performance and therefore differ according to how effective incentives can be in encouraging a good performance. In particular, we show that, under certain assumptions, the M-form (multi-divisional form) is likely to provide better incentives than the U-form (unitary form) because it promotes yardstick competition (*i.e.* relative performance evaluation) more effectively. We conclude by presenting evidence that the assumptions on which this comparison rests are satisfied for Chinese data.

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 an *organization* we mean the hierarchy of managers built on *top* of technology, *e.g.* the way a corporation is subdivided into different divisions, and the way a planned economy (such as China or the former Soviet Union) is divided into different functional or regional governing bodies. In this paper we show how organizational form affects the quality of incentive schemes that can be given to managers.

Of course, 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, modelled 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.

Specifically, we show that different organizational forms give rise to different information about managers' performance. Therefore, we argue, they differ according to how effective incentives can be in encouraging good performance.

We focus on the comparison between two organizational forms: the M-form (multi-divisional form) and the U-form (unitary form). Both structures have figured prominently in corporate history (see Chandler (1962), Williamson (1975)). 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 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 economies of the former Soviet Union and 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.[1] Since 1958, however, the Chinese economy has more closely resembled an "M-form"; it comprises a large number of reasonably self-sufficient regions (*e.g.* provinces, prefectures, *etc.*).

The potential benefits from the U-form—mainly exploitation of scale economies—have been discussed at length in the literature on the Soviet economy (*e.g.* Kornai (1992)). What are the countervailing advantages of the M-form? We argue that one such benefit may be better incentives, deriving from the familiar principle of yardstick competition (see, for example, Lazear and Rosen (1981), Holmstrom (1982), Nalebuff and Stiglitz (1983), and Shleifer (1985)). Indeed, relative performance evaluation appears to be widespread in China: provinces, cities, counties, townships, and villages are continually ranked by their performance in growth, output, foreign investment, *etc.*[2] Interestingly, there did not appear to be such competition between the specialized ministries of the Soviet Union. The question is, why not? 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 produce more-or-less the same array of goods. But on what is this intuition founded?

One answer could be 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 so, yardstick competition between two regions will be more effective in providing incentives than that between two ministries, and thus an M-form will dominate from the standpoint of providing managerial incentives. Of course, this comes down in the end to a matter of empirics. But here our analysis of data from 520 Chinese state-owned enterprises seems to support the hypothesis that it is "easier" to compare different regions than different industries. In any case, the more general lesson that we are trying to draw is that different organizational forms give a rise to different *information* on which *incentives* can be based.

We proceed as follows. In Section 2, we lay out the model. In Section 3, we present our theoretical results. Proposition 2 shows that the M-form provides better incentives

1. The current Russian economy, including privatized firms, is still deeply affected by their U-form legacy. Data from field work show that Russian firms are still strongly influenced by industrial ministries (Earle and Ross (1996)).

2. The Chinese central government has pursued an explicit policy during reform to stimulate regional competition, such as encouraging 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, some county governments use the annual ranking of townships (by profit rate on total capital) as a primary criterion to evaluate township government officials (Chapter 2, Whiting (1995)). Moreover, 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 or regional statistical books publish national or regional rankings of provinces, cities and/or counties every year.

for middle-level managers provided that there is "less variation" in interregional performance than in interindustry performance. The comparison is independent of utility functions of both the principal and agents. Proposition 1 establishes that it is *only* at the middle 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.

Our empirical work is reported in Section 4, where we argue that there is indeed higher "variation" in performance across industries than across regions. We also offer systematic evidence to show the use of yardstick competition in the Chinese economy. We make a few concluding remarks in Section 5.

2. THE MODEL

Consider an economy with two *regions*, A and B; two *industries*, 1 and 2; and four *plants*, one plant 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 ($r = \text{A}, \text{B}$).[3] There are three kinds of *shocks*: shocks η hit all plants in the economy; shocks θ_i hit just plants in industry i, $i = 1, 2$; shocks δ_r hit region r, $r = \text{A}, \text{B}$. We assume that shocks are jointly normally distributed.[4]

A shock has two effects: (i) to increase the variance of output of those plants it hits, and (ii) to *potentially* decrease the mean of their output. We emphasize the qualification "potentially" because it is the job of a *manager* "assigned" that shock to take steps (*i.e.* to exert effort) before the shock hits to prevent the mean from falling too far: the higher the effort, the smaller the fall (we suppose however, that a manager cannot affect the variance of the shock). Hence the economy-wide (top) manager is assigned η; an industry i manager ($i = 1, 2$) is assigned θ_i; and a region r manager ($r = \text{A}, \text{B}$) is assigned δ_r.

For example, imagine that η corresponds to a potential increase in the world price of oil. The top manager could attempt to limit the effect of a possible such increase by investing in the development of machinery running on liquified coal. Similarly, suppose that δ_{A} corresponds to the consequences of a possible flood besetting region A. The region A manager might prepare for this eventuality by working out a contingent plan for reinforcing embankments along the river. Finally, assume that industry 1 is agriculture and that θ_i corresponds to the effects of a potential locust invasion. The agriculture manager could respond by stocking-up on suitable crop sprays.[5]

In this paper, we will focus on two particular but widely employed hierarchies of managers: the U-form and the M-form.

If the economy is set up as a U-form, then it is organized along industrial (ministerial) lines. Beneath the top manager, who is allocated shock η, there is a manager (minister) for each of the two industries (ministries). Then within each industry, there are managers

3. With this specification, we rule out the possibility that all plants from a given industry be located in the same region, although this was roughly the case for some industries in the Soviet economy. The reason for ruling it out is that it implies that organization by region is identical to organization by industry, whereas we are interested in the contrast between the two.

4. For simplicity, we are limiting our attention to normal distributions. However, using the methods of Kim (1995) we could obtain extensions of Propositions 1 and 2 to general distributions.

5. There is another—and perhaps more "standard"—interpretation of our model. Instead of an entire economy, think of the organization as 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 labour or parts.

for the two regional shocks (and for the two plants). Hence, a U-form is illustrated by Figure 1.

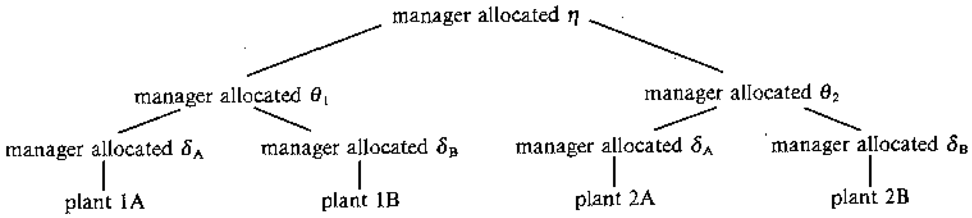

FIGURE 1
A U-form organization

If the economy is configured as an M-form, then it is organized along regional lines. In this case there is again a top manager, who is allocated shock η. Then, at the next level down, there is a manager (governor) in charge of each of the two regions. Next, within a region, there is a manager for each of the industrial shocks. The M-form is depicted in Figure 2.

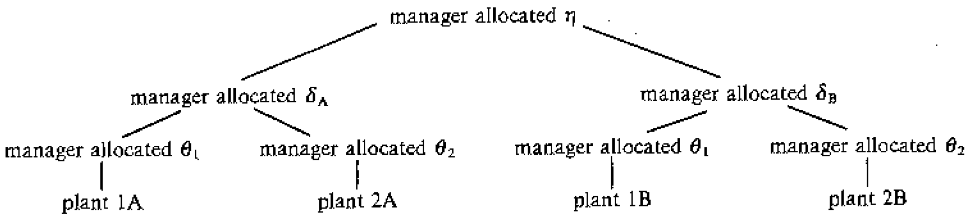

FIGURE 2
An M-form organization

Consider a manager who is allocated shock δ_r. Suppose that in the absence of his taking any corrective action, the shock induces an expected decrease of d in the output of each plant under him in the hierarchy. In the flood example above, d might correspond the average fall in output if there were no embankment fortification.[6] Let e_r measure the extent to which the manager prepares for the flood (it corresponds to his *effort*): expected output is raised by e_r in all plants under him.

Now, because absolute levels play no role in our analysis, we can take the value of d, as well the means of all shocks, to be zero. Hence the output of a plant in industry i and region r, when the managers allocated the shocks θ_i, δ_r, and η hitting that plant exert efforts e_i, e_r, and e_η respectively, is

$$x_{ir} = e_i + e_r + e_\eta + \theta_i + \delta_r + \eta.$$

The cost of effort e is $C(e)$ where $C(0) = 0$, and

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

6. We have been speaking of shocks as though they are necessarily a bad thing. But, in the case of a favourable shock, we can reinterpret d as the maximum possible average increase in output that the shock permits. Thus, if the manager does nothing in response to the shock, average output is lower by d relative to what it would have been had the manager taken full advantage of the shock.

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.

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 and a set of reward schemes $t_j(\cdot)$ for each manager j so as to maximize the expected value of net output

$$\sum_i \sum_r x_{ir} - \sum_{j \in M} t_j(\cdot),$$

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

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

3. INFORMATION AND INCENTIVES

The central point of our paper is that different organizational forms give rise to different information on which incentives can be based. We are particularly interested in comparing managers' incentives in M-form and U-form organizations.

We will argue that if there is less "variation" (in the appropriate sense) in shocks across regions than 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

$$x_i = e_i + \varepsilon_i,$$

where e_i is the effort of the manager in charge of shock ε_i, and $(\varepsilon_1, \varepsilon_2)$ 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

$$x_r = e_r + \varepsilon_r,$$

e_r is the effort of the manager in charge of shock ε_r, and $(\varepsilon_A, \varepsilon_B)$ 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

$$\mathrm{Var}(\varepsilon_A | \varepsilon_B) < \mathrm{Var}(\varepsilon_1 | \varepsilon_2), \tag{1}$$

then manager A can be given better incentives than manager 1.[7] Moreover, if both

$$\min\{\mathrm{Var}(\varepsilon_A | \varepsilon_B), \mathrm{Var}(\varepsilon_B | \varepsilon_A)\} < \min\{\mathrm{Var}(\varepsilon_1 | \varepsilon_2), \mathrm{Var}(\varepsilon_2 | \varepsilon_1)\}, \tag{2}$$

7. Here, as in the remainder of the paper, we must impose an exogenous upper bound on penalties or else, as a referee pointed out, there would exist incentive schemes approximating the first-best arbitrarily closely (*à la* Mirrlees (1974)).

and

$$\max\{\mathrm{Var}(\varepsilon_A|\varepsilon_B), \mathrm{Var}(\varepsilon_B|\varepsilon_A)\} < \max\{\mathrm{Var}(\varepsilon_1|\varepsilon_2), \mathrm{Var}(\varepsilon_2|\varepsilon_1)\}, \tag{3}$$

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

The less noisy performance is as a function of effort, the easier it is to provide a manager with the incentive to supply effort. Condition (1) 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 (1) holds, manager A can be provided with better incentives than manager 1, fix an effort level e' for manager 2 and assume that managers choose effort levels noncooperatively. Suppose that $t_1(\cdot,\cdot)$ 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 (1) holds, we can find a transfer scheme $t_A(\cdot,\cdot)$ as a function of (x_A, x_B) such that, if manager B exerts effort e', the scheme $t_A(\cdot,\cdot)$ is equivalent to $t_1(\cdot,\cdot)$. To see this, note that (1) is equivalent to

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

where $\sigma_r^2 = \mathrm{Var}(\varepsilon_r)$, $r = A, B$; $\sigma_{AB} = \mathrm{Cov}(\varepsilon_A, \varepsilon_B)$; $\sigma_i^2 = \mathrm{Var}(\varepsilon_i)$, $i = 1, 2$; and $\sigma_{12} = \mathrm{Cov}(\varepsilon_1, \varepsilon_2)$.

Choose scalars

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

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 $[\mathrm{Var}(\varepsilon_1|\varepsilon_2) - \mathrm{Var}(\varepsilon_A|\varepsilon_B)]$. 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

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

$t_A(\cdot,\cdot)$ will be equivalent to $t_1(\cdot,\cdot)$. 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:

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

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

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

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

$$\begin{aligned}\mathrm{Var}(x_A - \alpha x_B + z) &= \sigma_A^2 - 2\alpha\sigma_{AB} + \alpha^2\sigma_B^2 + [\mathrm{Var}(\varepsilon_1|\varepsilon_2) - \mathrm{Var}(\varepsilon_A|\varepsilon_B)]\\ &= \sigma_A^2 - \sigma_{AB}^2/\sigma_B^2 + \sigma_{12}^2/\sigma_2^2 + \sigma_1^2 - \sigma_{12}^2/\sigma_2^2 - (\sigma_A^2 - \sigma_{AB}^2/\sigma_B^2) = \sigma_1^2 = \mathrm{Var}(x_1),\end{aligned}$$

as claimed.

We have been taking e' as fixed for managers 2 and B. But if (2) and (3) 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 setup that is simpler than the model that we are really interested in. Let us return, thereafter, to the model of Section 2. 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 have argued that, at least in the stripped-down model, the M-form provides better incentives than the U-form for *middle-level* managers (those one level down from top management), provided that conditional variances for regions are smaller than those for industries. It may appear at first that the comparison should go exactly the other way, once we move down to bottom-level managers. After all, the bottom-level managers are industrial in the M-form and regional in the U-form. Moreover, were the comparison to flip, we would get no clear-cut answer about the M-form versus the U-form. However, it turns out that the incentives for bottom-level managers do not depend on whether an M-form or U-form is employed (nor do they for top-level managers). Thus it suffices to consider only the incentives of middle-level managers:

Proposition 1. *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 $t_{ir}(\cdot)$ 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

$$\begin{aligned} t'_{A1}(x_{1A}, x_{2A}, x_{1B}, x_{2B}) = t_{1A}(&x_{1A} + e^*_A + e^*_\eta - e^{**}_1 - e^{**}_\eta, x_{2A} + e^*_A + e^*_\eta \\ &- e^{**}_2 - e^{**}_\eta + e^*_{2A} - e^{**}_{A2}, x_{1B} + e^*_B + e^*_\eta - e^{**}_1 - e^{**}_\eta \\ &+ e^*_{1B} - e^{**}_{B1}, x_{2B} + e^*_B + e^*_\eta - e^{**}_2 - e^{**}_\eta + e^*_{2B} - e^{**}_{B2}). \end{aligned}$$

It is then straightforward to verify that, for any effort choice e by managers A1 or 1A, the random variables $t'_{A1}(\cdot, \cdot, \cdot, \cdot)$ and $t_{1A}(\cdot, \cdot, \cdot, \cdot)$ are the same. The argument for top managers is similar. ||

Proposition 1 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 framework already 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 the 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 framework, 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 model of Section 2, 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.

For $i = 1, 2$ and $r = A, B$, we denote

$$\varepsilon_{ir} = \theta_i + \delta_r.$$

Let λ_A solve

$$\min_\lambda \operatorname{Var}(\lambda\varepsilon_{1A} + (1-\lambda)\varepsilon_{2A} | \varepsilon_{1B}, \varepsilon_{2B}), \tag{4}$$

and let λ_1 solve

$$\min_\lambda \operatorname{Var}(\lambda\varepsilon_{1A} + (1-\lambda)\varepsilon_{1B} | \varepsilon_{2A}, \varepsilon_{2B}). \tag{5}$$

Define λ_B and λ_2 analogously. Let

$$\varepsilon_r = \lambda_r\varepsilon_{1r} + (1-\lambda_r)\varepsilon_{2r}, \tag{6}$$

for $r = A, B$, and

$$\varepsilon_i = \lambda_i\varepsilon_{iA} + (1-\lambda_i)\varepsilon_{iB}, \tag{7}$$

for $i = 1, 2$. We 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.

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.

Proof. See Appendix. ||

Lemma 1 can be understood from standard linear regression theory. Vector (x_1, x_1, x_2, x_2) is the fitted regression vector under the "best linear unbiased estimation" procedure for $(x_{1A}, x_{1B}, x_{2A}, x_{2B})$. Therefore, the residual vector (u_1, u_2, u_3, u_4) is uncorrelated (as well as independent due to normality) to (x_1, x_1, x_2, x_2). This decomposition of $(x_{1A}, x_{1B}, x_{2A}, x_{2B})$ essentially makes (x_1, x_2) a sufficient statistic for $(x_{1A}, x_{1B}, x_{2A}, x_{2B})$

under the U-form, and therefore, (x_1, x_2) becomes an appropriate aggregation from the point of view of providing incentives (Holmstrom (1982)).

With the help of Lemma 1, we can establish the following lemma, which is the counterpart to our analysis of the stripped-down framework:

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,

$$\mathrm{Prob}\,(t_A(x^*_{1A}, x^*_{2A}, x^*_{1B}, x^*_{2B}) = \tau | e) = \mathrm{Prob}\,(t_1(x_{1A}, x_{1B}, x_{2A}, x_{2B}) = \tau | e),$$

if and only if

$$\mathrm{Var}\,(\varepsilon_A | \varepsilon_B) \leqq \mathrm{Var}\,(\varepsilon_1 | \varepsilon_2).$$

Proof. See Appendix. ||

Finally, because the labels "1", "2", "A", and "B" are arbitrary, applying Lemma 2, we can compare the M-form and U-form straightforwardly as follows:

Proposition 2. *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*

$$\max\,\{\mathrm{Var}\,(\varepsilon_A | \varepsilon_B), \mathrm{Var}\,(\varepsilon_B | \varepsilon_A)\} \leqq \max\,\{\mathrm{Var}\,(\varepsilon_1 | \varepsilon_2), \mathrm{Var}\,(\varepsilon_2 | \varepsilon_1)\}, \tag{8}$$

and

$$\min\,\{\mathrm{Var}\,(\varepsilon_A | \varepsilon_B), \mathrm{Var}\,(\varepsilon_B | \varepsilon_A)\} \leqq \min\,\{\mathrm{Var}\,(\varepsilon_1 | \varepsilon_2), \mathrm{Var}\,(\varepsilon_2 | \varepsilon_1)\}, \tag{9}$$

where $\varepsilon_B, \varepsilon_A, \varepsilon_1$, and ε_2 are given by (6) *and* (7).

Proposition 2 implies an incomplete ranking of the M-form and U-form in terms of managerial incentives. If both (8) and (9) hold, the M-form is at least as good as the U-form; if both fail, the U-form is at least as good as the M-form; and one of them is satisfied and another fails, the result is inclusive.

Notice that our method of argument is to compare two probability distributions of output signals under the alternative organizational forms. The managerial incentive schemes are entirely based on these probability distributions, regardless of particular form of the utility functions of managers. Therefore, our comparison of organizational forms is independent of the utility functions of the managers. Furthermore, it is also independent of the solution concept of the (non-cooperative) game played by the managers, such as Nash or dominant strategy equilibrium, or others.

When there is symmetry across regions and across industries and no correlation between industrial and regional shocks, the formulas of Proposition 2 can be simplified into the following intuitive conditions:

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

$$V_R^2 - V_{AB} \leq V_1^2 - V_{12}.$$

The corollary demonstrates a linear tradeoff between variances and covariances for the purpose of incentives.

4. 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.[8]

It is well documented that enterprises in EEFSU were grouped by industry, each of which was supervised by a ministry (Gregory and Stuart (1989)). 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 between 1953 and 1957, which was formulated with the help of the Soviets. 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 administration: 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.[9]

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

8. 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.

9. China's M-form economy is not mere decentralization at the national level due to its large size. Compare Hungary and Guangdong province. The former was organized in a U-form hierarchy with specialized ministries managing all firms, while the latter itself is also organized in an M-form with multiple regions consisting of prefectures, counties, townships and villages, all of them being self-contained economic units.

a 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.[10] 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 neighbouring 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^I_{it} and D^R_{it}. The coefficients of these dummies serve as proxies for the industry-specific and region-specific shocks in the given period. Formally, we have

$$
\begin{aligned}
E(y \mid L, k)
&= (\beta + \textstyle\sum_{i=1}^{I} \beta_i D^I_i)L + (\gamma + \sum_{i=1}^{I} \gamma_i D^I_i)K \\
&\quad + \textstyle\sum_{t=1}^{T} D_t \eta_t + \sum_{t=1}^{T} \sum_{r=1}^{R} D^R_{rt} \delta_{rt} + \sum_{t=1}^{T} \sum_{i=1}^{I} D^I_{it} \theta_{it} \\
&= (\beta + \textstyle\sum_{i=1}^{I} \beta_i D^I_i)L + (\gamma + \sum_{i=1}^{I} \gamma_i D^I_i)K \\
&\quad + \textstyle\sum_{t=1}^{T} D_t \eta_t + \sum_{t=1}^{T} [D^R_{Rt} \delta_{Rt} + \sum_{r=1}^{R-1} D^R_{rt} \delta_{rt}] + \sum_{t=1}^{T} [D^I_{It} \theta_{It} + \sum_{i=1}^{I-1} D^I_{it} \theta_{it}] \\
&= (\beta + \textstyle\sum_{i=1}^{I} \beta_i D^I_i)L + (\gamma + \sum_{i=1}^{I} \gamma_i D^I_i)K \\
&\quad + \textstyle\sum_{t=1}^{T} D_t \eta_t + \sum_{t=1}^{T} \{(D^T_t - \sum_{r=1}^{R-1} D^R_{rt})\delta_{Rt} + \sum_{r=1}^{R-1} D^R_{rt} \delta_{rt}\} \\
&\quad + \textstyle\sum_{t=1}^{T} \{(D^T_t - \sum_{i=1}^{I-1} D^I_{it})\theta_{It} + \sum_{i=1}^{I-1} D^I_{it} \theta_{it}\} \\
&= (\beta + \textstyle\sum_{i=1}^{I} \beta_i D^I_i)L + (\gamma + \sum_{i=1}^{I} \gamma_i D^I_i)K \\
&\quad + \textstyle\sum_{t=1}^{T} D^T_t \zeta_t + \sum_{t=1}^{T} \sum_{r=1}^{R-1} D^R_{rt} \delta'_{rt} + \sum_{t=1}^{T} \sum_{i=1}^{I-1} D^I_{it} \theta'_{it},
\end{aligned}
$$

where

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

$$\delta'_{rt} = \delta_{rt} - \delta_{Rt},$$

10. The data were collected by the China System Reform Research Institute, Beijing, China.

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,[11] 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 $\xi_t = (\theta_{1t}, \theta_{2t}, \delta_{At}, \delta_{Bt})$. Then we treat these estimated shocks as if they were real shocks that are uncorrelated over time.

Given the limitation of the data, *i.e.* the too short time series ($T = 6$), we are not able to perform a formal test.[12] In the following, we compare the conditional variances under the M-form with those under the U-form in a descriptive way. 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) summarizes the comparison.

Of the 63 results in Table 1, there are 44 cases in which the estimated means of both conditional variances under the M-form are smaller than their counterparts under the U-form. In the remaining 29 cases where at least one of the conditional variances under the M-form is smaller than its counterparts under the U-form. There is no case where both conditional variances under the M-form are larger than their counterparts under the U-form. The results in Table 2 show that out of 36 possible pairs of comparisons, in 25 pairs both conditional variances under the M-form are smaller than their counterparts under the U-form. In the remaining 11 pairs of our test statistic at least one estimated mean conditional variance under the M-form is greater than its counterpart under the U-form. Again, there is no case where both conditional variances under the M-form are larger than their counterparts under the U-form. Therefore, in view of Proposition 2, these results suggest 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 4B 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.

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

11. Dummy variables here have the following property:

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

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

12. We have derived a formal asymptotic test statistics which could allow us to perform a rigorous test if we had a better data set (see Maskin, Qian and Xu (1997)).

TABLE I

Comparing industrial and regional variance and conditional variance (by province)

	(1) $V(\varepsilon_A \mid \varepsilon_B)$	(2) $V(\varepsilon_B \mid \varepsilon_A)$	(3) $V(\varepsilon_1 \mid \varepsilon_2)$	(4) $V(\varepsilon_2 \mid \varepsilon_1)$	(5) Comparison
PR11	0·0008187	0·0007571	0·0030142	0·0024633	Both $V^M < V^U$
PR12	0·0009656	0·0013551	0·0028583	0·0100515	Both $V^M < V^U$
PR13	0·0009623	0·0021445	0·0040004	0·0016873	One $V^M < V^U$
PR14	0·0003254	0·0009305	0·0038165	0·0011762	Both $V^M < V^U$
PR15	0·0007978	0·0003191	0·0045005	0·0016317	Both $V^M < V^U$
PR16	0·0015251	0·0019038	0·0028962	0·0022384	Both $V^M < V^U$
PR17	0·0019566	0·0010847	0·0024202	0·0015482	Both $V^M < V^U$
PR18	0·0006965	0·0005134	0·0019411	0·0011535	Both $V^M < V^U$
PR19	0·0036261	0·0010941	0·0027306	0·0027949	One $V^M < V^U$
PR21	0·0035188	0·0075737	0·004532	0·0014271	One $V^M < V^U$
PR22	0·0125974	0·0033652	0·0064482	0·0027321	One $V^M < V^U$
PR23	0·0007243	0·0005763	0·0032571	0·0019396	Both $V^M < V^U$
PR24	0·0035512	0·0017348	0·0083718	0·0076624	Both $V^M < V^U$
PR25	0·0008053	0·0056084	0·0031696	0·0041095	One $V^M < V^U$
PR26	0·0011198	0·0012982	0·0280702	0·0043337	Both $V^M < V^U$
PR27	0·0035219	0·0014758	0·0014134	0·0045301	One $V^M < V^U$
PR28	0·0032017	0·0036066	0·0022232	0·0039406	One $V^M < V^U$
PR29	0·0009339	0·0006265	0·0059682	0·0121066	Both $V^M < V^U$
PR31	0·0041727	0·0043189	0·0086372	0·0023456	One $V^M < V^U$
PR32	0·0116553	0·00325	0·0114781	0·0047748	Both $V^M < V^U$
PR33	0·0003516	0·0002477	0·0032018	0·0012271	Both $V^M < V^U$
PR34	0·0022851	0·0012984	0·0096523	0·0044628	Both $V^M < V^U$
PR35	0·0014829	0·0069914	0·0063509	0·0042792	One $V^M < V^U$
PR36	0·001434	0·0013309	0·0381816	0·0041133	Both $V^M < V^U$
PR37	0·0017801	0·0006927	0·0026892	0·0043122	Both $V^M < V^U$
PR38	0·0039548	0·0055769	0·0033635	0·0040625	One $V^M < V^U$
PR39	0·0002523	0·000138	0·0012346	0·0083788	Both $V^M < V^U$
PR41	0·0022107	0·0042633	0·0063843	0·0044466	Both $V^M < V^U$
PR42	0·0045825	0·0036649	0·009163	0·0031834	One $V^M < V^U$
PR43	0·0042087	0·0078392	0·0067047	0·0037866	One $V^M < V^U$
PR44	0·0026567	0·0021724	0·0101624	0·0078117	Both $V^M < V^U$
PR45	0·0048027	0·0033586	0·0066146	0·0599119	Both $V^M < V^U$
PR46	0·0052495	0·0025457	0·0174984	0·0066967	Both $V^M < V^U$
PR47	0·0028687	0·002188	0·0051838	0·0071966	Both $V^M < V^U$
PR48	0·0044801	0·0044515	0·0031598	0·0106808	One $V^M < V^U$
PR49	0·0047402	0·003174	0·0051034	0·0113329	Both $V^M < V^U$
PR51	0·0019198	0·00453	0·0024876	0·0018453	One $V^M < V^U$
PR52	0·002041	0·0014745	0·00241	0·0018914	Both $V^M < V^U$
PR53	0·0005682	0·0008727	0·0025407	0·0024924	Both $V^M < V^U$
PR54	0·0040096	0·001339	0·003392	0·0067175	One $V^M < V^U$
PR55	0·0006512	0·0012565	0·0028805	0·0143283	Both $V^M < V^U$
PR56	0·0008774	0·001117	0·0029163	0·0067734	Both $V^M < V^U$
PR57	0·0019528	0·0024209	0·0020744	0·0072606	Both $V^M < V^U$
PR58	0·0011385	0·0013931	0·0018707	0·0102239	Both $V^M < V^U$
PR59	0·0006934	0·0008906	0·0039924	0·0069858	Both $V^M < V^U$
PR61	0·0021139	0·0058389	0·0051549	0·0041415	One $V^M < V^U$
PR62	0·0038292	0·0020893	0·0047473	0·0045516	Both $V^M < V^U$
PR63	0·0011121	0·0016206	0·0049029	0·0046627	Both $V^M < V^U$
PR64	0·0021749	0·0014442	0·0044179	0·0027746	Both $V^M < V^U$
PR65	0·0018646	0·0028379	0·0040928	0·012857	Both $V^M < V^U$
PR66	0·001113	0·0010564	0·0062856	0·0031007	Both $V^M < V^U$
PR67	0·0030414	0·0009696	0·0041567	0·0027338	Both $V^M < V^U$
PR68	0·0022042	0·0028437	0·0041358	0·0069915	Both $V^M < V^U$
PR69	0·001927	0·0008932	0·0063637	0·002802	Both $V^M < V^U$

TABLE 1—continued

(1) $V(\varepsilon_A \mid \varepsilon_B)$	(2) $V(\varepsilon_B \mid \varepsilon_A)$	(3) $V(\varepsilon_1 \mid \varepsilon_2)$	(4) $V(\varepsilon_2 \mid \varepsilon_1)$	(5) Comparison
PR71 0·0024754	0·0050791	0·0102891	0·0083058	Both $V^M < V^U$
PR72 0·0033104	0·002966	0·0083431	0·0060816	Both $V^M < V^U$
PR73 0·0033779	0·0083795	0·0154754	0·0061027	One $V^M < V^U$
PR74 0·002703	0·001844	0·0066952	0·0023344	Both $V^M < V^U$
PR75 0·0047986	0·003487	0·0113504	0·042088	Both $V^M < V^U$
PR76 0·0041169	0·0020171	0·0082085	0·0022948	Both $V^M < V^U$
PR77 0·0055198	0·0018183	0·0048365	0·0022892	One $V^M < V^U$
PR78 0·0043981	0·0047133	0·0047888	0·0067743	Both $V^M < V^U$
PR79 0·0058755	0·0018537	0·0043716	0·0023171	One $V^M < V^U$

In the comparison, "Both $V^M < V^U$" means that the estimated means of both conditional variances under the M-form are smaller than their counterparts under the U-form; and "One $V^M < V^U$" means that at least one of the estimated means conditional variances under the M-form is smaller than its counterparts under the U-form.
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 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.

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 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.[13]

13. 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-

TABLE 2

Comparing industrial and regional variance and conditional variance (by large region)

	(1) $V(\varepsilon_A\|\varepsilon_B)$	(2) $V(\varepsilon_B\|\varepsilon_A)$	(3) $V(\varepsilon_1\|\varepsilon_2)$	(4) $V(\varepsilon_2\|\varepsilon_1)$	(5) Comparison
LR11	0·0009876	0·000717	0·0025903	0·0014356	Both $V^M < V^U$
LR12	0·0008853	0·0005627	0·0061475	0·0015846	Both $V^M < V^U$
LR13	0·0020441	0·0007281	0·0026581	0·0040902	Both $V^M < V^U$
LR14	0·0008268	0·0007193	0·0024435	0·0027816	Both $V^M < V^U$
LR15	0·0006873	0·0011279	0·0092834	0·0012659	Both $V^M < V^U$
LR16	0·0007685	0·0011829	0·0033395	0·0015871	Both $V^M < V^U$
LR17	0·0005151	0·0011678	0·0016032	0·0024889	Both $V^M < V^U$
LR18	0·0007956	0·000601	0·0014715	0·0018291	Both $V^M < V^U$
LR19	0·0008224	0·0015445	0·0018787	0·0012828	One $V^M < V^U$
LR21	0·0005335	0·001472	0·000835	0·0095962	Both $V^M < V^U$
LR22	0·0016437	0·0005659	0·0014412	0·0060323	One $V^M < V^U$
LR23	0·0012722	0·0007375	0·007627	0·0068863	Both $V^M < V^U$
LR24	0·0009478	0·0012606	0·003262	0·0062783	Both $V^M < V^U$
LR25	0·0004361	0·0003698	0·0005611	0·0006254	Both $V^M < V^U$
LR26	0·000618	0·0006669	0·0020611	0·0013201	Both $V^M < V^U$
LR27	0·0015519	0·0010175	0·0072856	0·0062593	Both $V^M < V^U$
LR28	0·0002648	0·0016529	0·0072931	0·000737	One $V^M < V^U$
LR29	0·0002031	0·0021522	0·0058771	0·000439	One $V^M < V^U$
LR31	0·0005775	0·0012875	0·001453	0·0040046	Both $V^M < V^U$
LR32	0·0017089	0·0005776	0·0013437	0·0021527	One $V^M < V^U$
LR33	0·0011759	0·0008448	0·0035925	0·0021431	Both $V^M < V^U$
LR34	0·0007168	0·0007843	0·0016867	0·0047914	Both $V^M < V^U$
LR35	0·000916	0·0006686	0·0011856	0·0015126	Both $V^M < V^U$
LR36	0·0010723	0·0008347	0·0053094	0·0019045	Both $V^M < V^U$
LR37	0·0011079	0·000722	0·0025876	0·0029192	Both $V^M < V^U$
LR38	0·0007325	0·0034549	0·0022474	0·002	One $V^M < V^U$
LR39	0·0007093	0·0028758	0·0023074	0·0017212	One $V^M < V^U$
LR41	0·0010433	0·0040568	0·0016852	0·0015885	One $V^M < V^U$
LR42	0·0050836	0·0011803	0·0015829	0·0036032	One $V^M < V^U$
LR43	0·0008049	0·0006459	0·0027002	0·0015179	Both $V^M < V^U$
LR44	0·0012797	0·0016134	0·0015776	0·0027378	Both $V^M < V^U$
LR45	0·0005507	0·0011771	0·0016617	0·0014342	Both $V^M < V^U$
LR46	0·0004704	0·0009846	0·0167323	0·001396	Both $V^M < V^U$
LR47	0·0031506	0·0007645	0·0014811	0·0039824	One $V^M < V^U$
LR48	0·0006126	0·001314	0·0022435	0·0013026	One $V^M < V^U$
LR49	0·0010576	0·0007848	0·0040202	0·0029437	Both $V^M < V^U$

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):

$$\mathrm{PINDEX}_r = -0{\cdot}453 + \underset{(0{\cdot}246)}{1{\cdot}76}\ \mathrm{EINDEX}_r,\quad R^2 = 0{\cdot}671,$$

where

$$\mathrm{EINDEX}_r = 10^{*}\{(1/\mathrm{erank86}_r) - (1/\mathrm{erank76}_r) + (1/\mathrm{erank86}_r)^2\},$$

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.

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
Zheijiang	17	19	3	8

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

and

$$\text{PINDEX}_r = 10^*\{(1/\text{prank87}_r) - (1/\text{prank77}_r) + (1/\text{prank87}_r)^2\}.$$

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, $(1/\text{erank86}_r)^2\}$ and $(1/\text{prank87}_r)^2\}$ 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.[14]

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.

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

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. Cremer (1980) studies how activities should be optimally grouped into shops in a resource allocation problem. Aoki (1986) investigates how Japanese firms are organized differently from those in the U.S. and what implications these differences have for comparative performance. Holmstrom and Milgrom (1991, 1994) study how tasks should be allocated to firms and managers when managers may perform more than one task.

On the literature of the U-form vs. M-form, 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.

APPENDIX

Proof of Lemma 1.

We take

$$\xi = (\theta_1, \theta_2, \delta_A, \delta_B)', \qquad \Sigma = \mathrm{var}(\xi),$$
$$\varepsilon_u = (\varepsilon_{1A}, \varepsilon_{1B}, \varepsilon_{2A}, \varepsilon_{2B})', \qquad \Sigma_u = \mathrm{var}(\varepsilon_u),$$
$$\varepsilon_m = (\varepsilon_{1A}, \varepsilon_{2A}, \varepsilon_{1B}, \varepsilon_{2B})', \qquad \Sigma_m = \mathrm{var}(\varepsilon_m),$$

where

$$(\varepsilon_{1A}, \varepsilon_{1B}, \varepsilon_{2A}, \varepsilon_{2B})' = A_u\xi, \qquad \Sigma_u = A_u \Sigma A_u',$$
$$(\varepsilon_{1A}, \varepsilon_{2A}, \varepsilon_{1B}, \varepsilon_{2B})' = A_m\xi, \qquad \Sigma_m = A_m \Sigma A_m',$$

and

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

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

We prove the case for the U-form (the case for the M-form is similar). Let

$$(x_1, x_2)' = (C_u'(C_{u1}\Sigma C_{u1}')^{-1}C_u)^{-1}C_u'(C_{u1}\Sigma C_{u1}')^{-1}Q_u'(x_{1A}, x_{1B}, x_{2A}, x_{2B})',$$
$$(x_A^*, x_B^*)' = (C_m'(C_{m1}\Sigma C_{m1}')^{-1}C_m)^{-1}C_m'(C_{m1}\Sigma C_{m1}')^{-1}Q_m'(x_{1A}^*, x_{2A}^*, x_{1B}^*, x_{2B}^*)',$$
$$C_u = Q_u'A, \qquad C_{u1} = Q_u'A_u, \qquad C_m = Q_m'A, \qquad C_{m1} = Q_m'A_m,$$

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

and $(C_{u1}\Sigma C'_{u1})$ and $(C_{m1}\Sigma C'_{m1})$ are non-singular 3×3 matrices.

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,

$$\begin{aligned}
E\underline{x}u' &= E\underline{x}(x-\underline{x})' \\
&= E\{A(C'_u(C_{u1}\Sigma C'_{u1})^{-1}C_u)^{-1}C'_u(C_{u1}\Sigma C'_{u1})^{-1}Q'_u(A\beta+A_u\xi)\} \\
&\quad \times \{(I-A(C'_u(C_{u1}\Sigma C'_{u1})^{-1}C_u)^{-1}C'_u(C_{u1}\Sigma C'_{u1})^{-1}Q'_u)(A\beta+A_u\xi)\}' \\
&= E\{A(C'_u(C_{u1}\Sigma C'_{u1})^{-1}C_u)^{-1}C'_u(C_{u1}\Sigma C'_{u1})^{-1}Q'_u(A\beta+A_u\xi)\} \\
&\quad \times (A_u\xi)'\{I-A(C'_u(C_{u1}\Sigma C'_{u1})^{-1}C_u)^{-1}C'_u(C_{u1}\Sigma C'_{u1})^{-1}Q'_u\}' \\
&= E\{A\beta+A(C'_u(C_{u1}\Sigma C'_{u1})^{-1}C_u)^{-1}C'_u(C_{u1}\Sigma C'_{u1})^{-1}C_{u1}\xi\} \\
&\quad \times \xi'A'_u\{I-Q'_u(C_{u1}\Sigma C'_{u1})^{-1}C_u(C'_u(C_{u1}\Sigma C'_{u1})^{-1}C_u)^{-1}A'\} \\
&= A(C'_u(C_{u1}\Sigma C'_{u1})^{-1}C_u)^{-1}C'_u(C_{u1}\Sigma C'_{u1})^{-1}C_{u1}E(\xi\xi')A'_u \\
&\quad - A(C'_u(C_{u1}\Sigma C'_{u1})^{-1}C_u)^{-1}C'_u(C_{u1}\Sigma C'_{u1})^{-1}C_{u1}E(\xi\xi')C'_{u1}(C_{u1}\Sigma C'_{u1})^{-1}C_u(C'_u(C_{u1}\Sigma C'_{u1})^{-1}C_u)^{-1}A' \\
&= A(C'_u(C_{u1}\Sigma C'_{u1})^{-1}C_u)^{-1}C'_u(C_{u1}\Sigma C'_{u1})^{-1}C_{u1}\Sigma A'_u - A(C'_u(C_{u1}\Sigma C'_{u1})^{-1}C_u)^{-1}A'.
\end{aligned}$$

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

$$\begin{aligned}
&[A(C'_u(C_{u1}\Sigma C'_{u1})^{-1}C_u)^{-1}C'_u(C_{u1}\Sigma C'_{u1})^{-1}C_{u1}\Sigma A'_u - A(C'_u(C_{u1}\Sigma C'_{u1})^{-1}C_u)^{-1}A']Q_u \\
&\quad = A(C'_u(C_{u1}(C_{u1}\Sigma C'_{u1})^{-1}C_u)^{-1}C'_u - A(C'_u(C_{u1}\Sigma C'_{u1})^{-1}C_u)^{-1}C'_u \\
&\quad = 0.
\end{aligned}$$

We also have

$$[A(C'_u(C_{u1}\Sigma C'_{u1})^{-1}C_u)^{-1}C'_u(C_{u1}\Sigma C'_{u1})^{-1}C_{u1}\Sigma A'_u - A(C'_u(C_{u1}\Sigma C'_{u1})^{-1}C_u)^{-1}A']R = 0,$$

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

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

Finally, one can show x_1 and x_A as defined above 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 (4) and (5), respectively. Similarly for x_2 and x_B. ||

Proof of Lemma 2.

Let

$$(\varepsilon_1, \varepsilon_2)' = (C'_u(C_{u1}\Sigma C'_{u1})^{-1}C_u)^{-1}C'_u(C_{u1}\Sigma C'_{u1})^{-1}C_{u1}\xi,$$

and

$$(\varepsilon_A, \varepsilon_B)' = (C'_m(C_{m1}\Sigma C'_{m1})^{-1}C_m)^{-1}C'_m(C_{m1}\Sigma C'_{m1})^{-1}C_{m1}\xi.$$

From our analysis of the stripped-down framework, $\text{Var}(\varepsilon_A|\varepsilon_B)\leqq\text{Var}(\varepsilon_1|\varepsilon_2)$ implies that there exist constant α, β, γ and random noise z uncorrelated with (x_A, x_B) such that for all $e_1=e_A$,

$$(x_1, x_1, x_2, x_2) = (x^*_A-\alpha x^*_B, x^*_A-\alpha x^*_B, \beta x^*_B, \beta x^*_B)+(z, z, \gamma, \gamma),$$

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

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

$$\begin{aligned}
&\text{Var}(x_{1A}, x_{1B}, x_{2A}, x_{2B}) \\
&\quad = \text{Var}(x_1, x_1, x_2, x_2)+[\text{Var}(x_{1A}, x_{1B}, x_{2A}, x_{2B})-\text{Var}(x_1, x_1, x_2, x_2)] \\
&\quad = \text{Var}(x^*_A-\alpha x^*_B, x^*_A-\alpha x^*_B, \beta x^*_B, \beta x^*_B)+\text{Var}(z, z, \gamma, \gamma)+\text{Var}(w_1, w_2, w_3, w_4) \\
&\quad = \text{Var}(x^*_A-\alpha x^*_B+z+w_1, x^*_A-\alpha x^*_B+z+w_2, \beta x^*_B+\gamma+w_3, \beta x^*_B+\gamma+w_4).
\end{aligned}$$

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 is the same as $t_1(x_{1A}, x_{1B}, x_{2A}, x_{2B})$ in distribution. ||

Acknowledgements. 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 three referees and participants in seminars and conferences in Gerzensee, California Institute of Technology, Harvard, MIT, LSE, CEPR (Brussels), and the Chinese Academy of Social Sciences (Beijing) for useful comments. We thank Juzhong Zhuang for his help with the data; Nancy Hearst for her help with the literature; Angela Lee and Ying Qian for their 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 LSE (CEP is sponsored by the ESRC).

REFERENCES

AGHION, P. and TIROLE, J. (1995), "Some Implications of Growth for Organizational Form and Ownership Structure", *European Economic Review*, **39**, 440–455.

AOKI, M. (1986), "Horizontal vs. Vertical Information Structure of the Firm", *American Economic Review*, **76**, 971–983.

ARROW, K. (1974) *The Limits of Organization* (New York: Norton).

BARTKE, W. (1990) *Biographical Dictionary and Analysis of China's Party Leadership, 1922–1988* (New York: K. G. Saur).

BOLTON, P. and FARRELL, J. (1990), "Decentralization, Duplication and Delay", *Journal of Political Economy*, **98**, 803–826.

CHANDLER, A. Jr. (1962) *Strategy and Structure* (New York: Doubleday & Company, Inc).

CREMER, J. (1980), "A Partial Theory of the Optimal Organization of Bureaucracy", *Bell Journal of Economics*, **11**, 683–693.

EARLE, J. and ROSS, R. (1996), "Ownership Transformation, Economic Behavior, and Political Attitudes in Russia", (Mimeo).

GREGORY, P. R. and STUART, R. C. (1981) *Soviet Economic Structure and Performance* (New York: Harper & Row, Publishers).

HOLMSTRON, B. (1982), "Moral Hazard in Teams", *Bell Journal of Economics*, **13**, 324–340.

HOLMSTROM, B. and MILGROM, P. (1991), "Multi-task Principal-Agent Analysis: Incentive Contracts, Asset Ownership and Job Design", *Journal of Law, Economics, and Organization*, **7**, 24–52.

HOLMSTROM, B. and MILGROM, P. (1994), "The Firm as an Incentive System", *American Economic Review*, **84**, 972–991.

KIM, S. K. (1995), "Efficiency of an Information System in an Agency Model", *Econometrica*, **63**, 89–102.

KORNAI, J. (1992) *The Socialist System* (Princeton, NJ: Princeton University Press).

LAZEAR, E. P. and ROSEN, S. (1981), "Rank-Order Tournaments as Optimum Labor Contracts", *Journal of Political Economy*, **89**, 841–864.

MASKIN, E., QIAN, Y. and XU, C. (1997), "Incentives, Information, and Organizational Form" (HIID Development Discussion Papers 618, Harvard University; CEP Discussion Paper 371, LSE).

MILGROM, P. and ROBERTS, J. (1992) *Economics, Organization, and Management* (Englewood Cliffs, NJ: Prentice Hall).

MIRRLEES, J. A. (1974), "Notes on Welfare Economics, Information and Uncertainty", in M. Balch, D. McFadden, and S. Wu, (eds.), *Essays on Economic Behavior under Uncertainty*, (Amsterdam: North Holland) pp. 243–258.

NALEBUFF, B. and STIGLITZ, J. (1983), "Prizes and Incentives: Towards a General Theory of Compensation and Competition", *Bell Journal of Economics*, **14**, 21–43.

QIAN, Y., ROLAND, G. and XU, C. (1997), "Coordinating Changes in M-form and U-form Organizations", (Mimeo, Stanford University, ECARE and LSE).

QIAN, Y. and XU, C. (1993), "Why China's Economic Reform Differ: The M-form Hierarchy and Entry/Expansion of the Non-state Sector", *Economics of Transition*, **1**, 135–170.

SAICH, T. (1992), "The Fourteenth Party Congress: A Program for Authoritarian Rule", *China Quarterly*, **132**, 1136–1160.

SHLEIFER, A. (1985), "A Theory of Yardstick Competition", *Rand Journal of Economics*, **16**, 319–327.

STATE STATISTICAL BUREAU (1990) *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).

WEITZMAN, M. (1974), "Price vs. Quantity", *Review of Economic Studies*, **41**, 477–491.

WHITING, S. (1995) *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).

WILLIAMSON, O. (1975) *Markets and Hierarchies* (New York: Free Press).