

Decentralization, Collusion and Coalmine Deaths*

Ruixue Jia[†]

Huihua Nie[‡]

IIES, Stockholm University

Renmin University of China

December, 2011

Abstract

We propose a political-economics model of decentralization (by the central government) and collusion (between the local government and firms) to explain high death rates in the coalmine industry in China, where the local government allows the firms to choose dangerous but high-payoff production technology when collusion is feasible. Our model predicts that coalmine death rates are increasing in decentralization and collusion but decreasing in media exposure. We collect a provincial level panel dataset from 1995 to 2005 and compare the decentralization period (1998 to 2000) with the adjacent centralization periods. Consistent with the model, we find that decentralization increases death rates and that media exposure (proxied by newspapers published per 1000 individuals) decreases death rates. Assuming that the transaction costs of collusion are lower for native Safety governors, we find that given decentralization, death rates are about 100% higher in provinces with native Safety governors.

Keywords: decentralization, collusion, public bads

*Ruixue Jia is indebted to Torsten Persson for his guidance. We thank Davide Cantoni, James Fenske, Masa Kudamatsu, James Robinson, Tomas Sjöström, David Strömberg and participants at IIES Brown Bag, SITE Lunch Seminar, Tsinghua University, SSE Econ Lunch Seminar, SSE CERC Asia workshop for their comments.

[†]Institute for International Economic Studies, Stockholm University. Email: ruixue.jia@iies.su.se.

[‡]School of Economics, Renmin University of China. Email: niehuihua@gmail.com

1 Introduction

Over the last few decades, China has experienced rapid industrialization and economic growth. Meanwhile, China has become internationally notorious for weak occupational health and safety protection. Fatalities from coalmine accidents are among the most important health and safety issues in China, where coal makes up about 70% of the total energy consumption. Coal mining accounts for less than 4 percent of the broadly defined industrial workforce, but over 45 percent of industrial fatalities (Wright 2004).

China's coal mining industry is the deadliest in the world in terms of human safety where thousands of people die every year in the coal pits. Coalmine fatalities in China account for 80 percent of the world total¹. In 2007, China produced 41.1% and United States produced 18.7% of coal in the world. At least 3598 people died in coal mine mishap in China, compared with 34 in the United States. In fact, the death rate in China is not only much higher than that of developed countries such as the US, Japan and Germany but also more than ten times higher than that of many other developing economies such as India and African countries. As mentioned in Wright (2004), the head of China's safety bureaucracy admitted in 2001 that China's coalmine death rate was 11 times higher than that in Russia, and 15 times higher than that in India.

China's appalling record on coal-mine safety has drawn worldwide attention, especially in the media (The Economist 2002; The Economist 2005; The Economist 2007). Some hypotheses about causes have been put forward, such as insufficient safety input (Tang and Guo 2006) and lack of safety education (Asian Development Bank 2007), insecure property rights (Tan and Tang 2004) and possible collusion between governments and coalmines (Nie and Jiang 2011). However, to the best of our knowledge, all these existing hypotheses are mainly based on correlations and lack a clear theoretical framework as well as convincing empirical strategy. In this paper, we propose a model of decentralization and collusion to explain high coalmine death rates. The local government allows firms to choose dangerous but high-payoff technologies when collusion between the local government and the firms is feasible. As discussed below, our assumptions on technology and collusion are based on existing investigations of deadly coalmine mishaps.² Our model predicts that coalmine death rates are increasing in decentralization and collusion but decreasing in media exposure.

To bring our theory to data, we collect provincial-level panel data from 1995 to 2005 and explore an exogenous institutional change of the management power of key state coalmines.

¹“Coal Mine Fatalities in China Account for 80 percent of World Total,” BBC, Nov. 13, 2004. See:http://news.bbc.co.uk/chinese/simp/hi/newsid_4000000/newsid_4009000/4009043.stm.

²There are so many anecdotal evidences on collusion that a phrase was coined for this phenomenon "*guan-meì goujie*", which literally means "official-coal collusion".

In March 1998, the management of the key state coalmines was delegated to provincial government and this made collusion possible, or at least much easier. This decentralization period lasted until February 2001, when the State Administration of Coal Mine Safety was established and re-centralized coalmine safety supervision. Hence, we have one decentralization period (1998-2000) when collusion is possible and two centralization periods (1995-1997 and 2001-2005) as comparison.

We find that death rates were increased dramatically during the decentralized period. A competing hypothesis is that death rate is an increasing function of production, and hence, decentralization leads to higher death rates as might be correlated with higher level of production. We include the production level in the estimation and do not find this hypothesis true. In fact the correlation coefficient between death rates and production is negative (about -0.3). Thus, the competing hypothesis of convex safety costs is not supported.

To test the impact of media exposure, we use the number of newspapers published per 1000 population as a proxy and find that higher exposure decreases death rates. Since transaction costs of collusion are not observable, we proxy them by exploring biographical information on all the 71 Safety governors between 1995 and 2005. The Safety governors are one of the vice governors responsible for industrial production safety including coalmine safety. We focus on the Safety governors rather than the main governors because the Safety governors are directly in charge of coal production permissions.³ Assuming that being a native of the province lowers the transaction costs of collusion, we find that the death rate is about 100% higher in provinces with native Safety governors within the decentralized period.

Although the decentralization decision is a national policy and exogenous to difference provinces, there still might be a concern that native Safety governors are appointed to provinces with higher death rates. We do not find that this is true in our data. However, as a robustness check, we allow for more flexible specifications including the interactions between the dummy for being a native and year dummies. This was, we get the effects of being a native year by year. The results suggest that there is no pre-trend. We also control for Safety governor fixed effects and explore within-governor variation. Further, we limit our sample by dropping a subset of governors that might have been endogenously appointed.

Death rate is an intuitive measure of safety that is often discussed in the media. We use it as our dependent variable in our baseline estimations. For the concern of measurement, we also employ alternative measures including the number of deaths and the number of accidents. Since both are count variables, we employ fixed-effects Poisson estimations.

Our study contributes to a few lines of literature. Many fruitful theories in the or-

³We control for characteristics of main governors in our estimations.

ganizational literature have focused the costs and benefits of decentralization/delegation. Mookherjee (2006) discusses the existing theoretical literature and some empirical studies of industrial organizations. Delegation theory has been widely applied to the organizational forms of governments, especially to the trade-off between decentralization and centralization. Delegation and collusion have been paid a lot of attention in the theoretical literature (Tirole 1986; Kofman and Lawarree 1993; Baliga and Sjostrom 1998; Laffont and Martimort 1998; Mookherjee and Tsumagari 2004), which often stresses the cost side of decentralization when the principal has less information than the supervisor. Similar to this line of literature, our model emphasizes how decentralization makes the collusion between the supervisor (the Local government) and the agent (the Firm) more likely. As a result, the principal (the Central government) faces a trade-off between high payoff and high probability of accidents.

Although the theoretical incentive-based literature focuses more on the cost side of decentralization, existing empirical evidence has been celebrating the virtues of decentralization in terms of economic performance and government accountability. Bardhan and Mookherjee (2006) provide an excellent survey of existing research along this line. Within the context of China, fiscal decentralization since 1994 is often taken as one of the driving forces of China's growth miracle (Jin, Qian, and Weingast 2005; Lin and Liu 2000), despite sometimes opposite findings (Zhang and Zou 1998). A possible reason for the more positive findings in existing empirical studies is that the cost side is often less visible than economic growth rates and expenditure on public goods. We focus on coalmine mishap as it is one of the most visible health and safety problems. Our findings on the downside impacts of decentralization share some flavor with the findings on regional protectionism in Young (2000), who argues that increased autonomy and incentives induced local governments to engage in provincial protectionism. Our finding of bad outcomes due to the combination of decentralization and collusion remind of the views in Blanchard and Shleifer (2000), who argue that political centralization and economic decentralization together reduce the risk of local capture.

Our study is also related to the literature evaluating political connections. This line of literature has documented the benefits captured by those with stronger political connections (Fisman 2001; Faccio 2006). It is conceivable that these political benefits might be captured at the costs of some public interest. Our study provides new evidences on the cost side of political connections, and the measures are closely related to descriptions of Chinese networks ("*guanxi*") in political science (Guo 2001). The three main types of networks in China are locality of origin, kinship and job-related colleagues (Luo 2007). We focus on locality of origin in this paper, which might be the most relevant in this context. We do not focus on work experience because almost all of the safety governors have some work experience in a place before coming to the office.

Moreover, our study contributes to an emerging strand of literature which examines the impacts of media in nondemocratic countries (Reinikka and Svensson 2005; Egorov, Guriev, and Sonin 2009). Little is known about the role of media in China where media is owned and operated by the government. Our finding on the effect of newspapers supports the logic in Egorov, Guriev, and Sonin (2009), who argue that media allows dictators to monitor bureaucrats. However, this finding is very preliminary as we only have a very rough measure of media exposure.

Although we focus on a very specific example of coalmine deaths, our theory has external validity to other "*public bads*" such as ecology destruction and air pollution. Our theory is consistent with the current dilemma of China: dramatic growth but overwhelming problems with workplace safety and the environment. Moreover, our study also serves as the first attempt to provide both a theory and empirical evidences on high coalmine death rates in China. To explore an exogenous change in our identification, we only focus on key state coalmines in this paper. This perspective also sheds possible light on the death rates in smaller coalmines. For example, Wright (2004) mentions that township and village coalmines are closely related to local governments in a nexus of "local state corporatism" (Oi 1999).

The organization of the paper is as follows. Section 2 provides a model of collusion and decentralization and the predictions derived from this model. Section 3 describes the data. Section 4 presents the main estimation results, where the predictions are confronted with the data. Section 5 provides various robustness checks to deal with concerns of endogeneity, measurement and heterogeneity. Section 6 discusses alternative explanations. Section 7 concludes.

2 A Model of Decentralization and Collusion

Our model is based on the three-tier principal-supervisor-agent (PSA) framework (Tirole 1986), where the Central government is the principal, the Local government serves as the supervisor and the Firm is the agent. Decentralization is characterized by the share of resources assigned to the Local government in a grand contract. Collusion is characterized by a contract signed by the Local government and the Firm. Below we describe the environment, the timing of the game, motivating evidences for our assumptions and the solution.

2.1 Environment

The centralization or decentralization decision of the Central government (C) is exogenously given. In a more complete model, the Central government would generate utility from out-

put and safety and would face a trade-off between high output and dangerous technology. Its choice of decentralization would depend on the relative utility of output and safety concerns. However, we do not attempt to endogenize its choice of decentralization in this paper. Instead we will bring the predictions of our theory to data, given the decentralization, or centralization, decisions of the Central government.

The Firm (F) cares about output. To produce y units of coal, F can use a *good* technology or a *bad* technology. The good way costs $\bar{c}y^2$ and there is no accident. The bad way costs $\underline{c}y^2$ and an accident A happens with probability p . Parameters \underline{c} and \bar{c} are known to F and S but cannot be observed by C . To be able to employ a *bad* technology, F needs to bribe the Local government (L) to get permission.

Production y is taxed at the rate of t and income ty is divided between the Central government and the Local government. The Local government gets a share of s , which measures the extent of decentralization.

The Local government (L) cares about sy as well as the bribes from F . This is a short-cut to think about the incentives of the Local government. In the real world, sy is related to power. Meanwhile, L cares about the private payoff from collusion with F .

When an accident happens and is detected, the local government (L) and the firm (F) get punished. We assume that both of them have limited liability. Without loss of generality, assume that L loses his future expected payoff R and F gets a fine of f in the case of punishment. The probability of being punished is mp , where m denotes the detection probability; especially, this is positively related to media exposure.

However, L and F can collude by signing a contract⁴, although they face some transaction costs to agree with this arrangement. For simplicity, we assume transaction costs are proportional to output, i.e., the costs are $\phi y(\underline{c})$. L gets a bribe b if he allows F to choose \underline{c} , where b is proportional to the payoff of F , i.e., $b = k(1 - t)(1 - \phi)y(\underline{c})$. This way, k can be thought as a result of bargaining between L and F . In our solution, we take k as a choice variable of L that satisfies F 's participation constraint in the side contract.

2.2 Timing

First, C decides the policy of centralization or decentralization characterized by s . C also specifies the punishment of discovering an accident. The punishment is known to F and L , and denoted by $(-R, f)$.

Note that our model takes the decentralization decision of the Central government as

⁴In collusion, F chooses \underline{c} . Note that we assume away the case that F chooses \bar{c} and there is no collusion. This is because the information is symmetric between F and L .

given rather than provide an optimal contract from the perspective of the Central government. Given the policies of C , the timing of the game for F and L is as follows.

1. Given $(s, -R, f)$, F and L decide to collude or not by signing a side contract. The side contract specifies the share k , a choice of L that satisfies F 's participation constraint in the side contract. There are some transaction costs in this agreement given by ϕy .
2. Given $(s, -R, f)$ and k , F chooses \underline{c} or \bar{c} . \underline{c} or \bar{c} is not observed by C .
3. The side contract, if it exists, is implemented.
4. Output (y) is realized. Accidents are observed, if they are detected. The punishment schedule gets implemented in the case of detection.

2.3 Discussion of assumptions

Before turning to the solution of the game, we present some qualitative evidences in support of the model's assumptions. Specially, we clarify what the technologies are, how collusion works and why there might be transaction costs of collusion.

The technology in the model broadly refers to coal production technology and safety monitoring technology. We focus on the short-run decisions of firms such as where to mine, how many hours to work and how closely to monitor the mine safety. In the real world, most coalmine accidents occur because of these short-run decisions. For instance, one of the most common types of mining accidents in China is gas explosion. Some coalmine managers disregard the monitoring system for the sake of production. After 32 miners died in a gas explosion in Henan Province on October 18, 2010, it was revealed that gas emissions had exceeded the standard 22 hours prior to the disaster but this warning was disregarded by the mine management.⁵ Similarly, 108 miners died of gas explosion in Heilongjiang Province on November 21, 2010 because the managers decided to mine at places that did not have a qualified gas drainage system.⁶ Because short-run decisions of firms are essential in affecting coalmine safety, centralization or decentralization of management and supervision may play an important role.

As for collusion, government officials allow mine managers to employ dangerous technology in exchange for money or shares in the enterprises. When accidents occur, officials also help the owners cover them up as they will be punished together. For instance, when 81 miners died of mine flooding in Nandan County (Guangxi Province) on July 17, 2001, local officials quickly teamed up with the coalmine owner to cover up the accident. Once the catastrophe was disclosed to the public, it was revealed that the head of Nandang County

⁵Source: <http://society.people.com.cn/GB/12986786.html>

⁶Source: <http://society.people.com.cn/GB/8217/174547/index.html>

had received 3.21 million yuan (\$410,000) in bribes during the prior two years (Tu 2007). Wright (2008) provides more information on rent seeking in the coal industry.

Since collusion is illegal, trust is essential in the agreement of collusion. Li and Wu (2010) provide some qualitative evidences on the transactions between bribers and bribees in China and the Philippines. They show that personal networks play an important role in facilitating corruption, especially in a relation-based governance system such as the Chinese case. Empirically, we will argue that a common background and a long period of interaction facilitates trust-based collusion.

2.4 Solution

We solve the game by backward induction.

First, consider the firm's choice between the production technologies characterized by \bar{c} and \underline{c} , given $(s, -R, f)$ and k . If it chooses the good \bar{c} -technology, its problem is,

$$\max_y (1-t)y - \frac{\bar{c}}{2}y^2.$$

The solution gives $y(\bar{c}) = \frac{1-t}{\bar{c}}$ and $\pi(\bar{c}) = \frac{(1-t)^2}{2\bar{c}}$. Instead, if it chooses the bad \underline{c} -technology, the firm's problem becomes,

$$\max_y (1-t)(1-\phi)y - \frac{\underline{c}}{2}y^2 - k(1-t)(1-\phi)y - mpf.$$

Now, we get $y(\underline{c}) = \frac{(1-t)(1-k)(1-\phi)}{\underline{c}}$ and $\pi(\underline{c}) = \frac{(1-t)^2(1-k)^2(1-\phi)^2}{2\underline{c}} - mpf$.

Therefore, it is optimal to choose the \underline{c} -technology iff $\pi(\underline{c}) \geq \pi(\bar{c})$, i.e. the participation constraint below as follows,

$$\frac{(1-t)^2(1-k)^2(1-\phi)^2}{2\underline{c}} - \frac{(1-t)^2}{2\bar{c}} \geq mpf. \quad (\text{PC}_F)$$

Now consider the local government's choice of k . The local government cares about the revenues from coal production as well as the private payment from collusion.⁷ Thus, its problem is as follows,

$$\max_k [(1-\phi)sty(\underline{c}) + k(1-t)(1-\phi)y(\underline{c})] + (1-mp)R$$

⁷Note that we do not differ the local governor from the local government. It can be thought that the local governor faces the maximization problem. He cares about the revenues from production, because it is positively related to his power. Meanwhile, he also cares about the private payoff from collusion with firms.

s.t. its participation constraint,

$$[(1 - \phi)sty(\underline{c}) + k(1 - t)(1 - \phi)y(\underline{c})] + (1 - mp)R \geq sty(\bar{c}) + R \quad (\text{PC}_L)$$

where $y(\underline{c}) = \frac{(1-t)(1-k)(1-\phi)}{\underline{c}}$, as solved above.

We temporarily ignore the participation constraint of L and solve its maximization problem. Then we check when the participation constraint is satisfied.

The first-order condition with respect to k gives:

$$k^* = \frac{1 - t - st}{2(1 - t)}.$$

Clearly,

$$\frac{\partial k^*}{\partial s} = -\frac{t}{2(1 - t)} < 0.$$

Substituting k^* into the firm's participation constraint (PC_F) for collusion, we get the following equivalent condition (PC'_F):

$$\frac{(1 - t)^2(1 - k^*)^2(1 - \phi)^2}{2\underline{c}} - \frac{(1 - t)^2}{2\bar{c}} - mpf \geq 0. \quad (\text{PC}'_F)$$

First, we observe that the left hand side LSH_F is decreasing in k^* and hence increasing in the decentralization measure s . Second, a necessary condition for the condition to hold is that the difference between the two types of technology should be large enough. Otherwise, the left condition will be close to zero or even negative. The condition to make sure the first part of LSH_F ($\frac{(1-t)^2(1-k^*)^2(1-\phi)^2}{2\underline{c}} - \frac{(1-t)^2}{2\bar{c}}$) positive is as follows:

$$\bar{c} \geq \frac{2}{(1 - \phi)^2}.$$

We assume that this condition holds in the following discussions. Otherwise, collusion will not be profitable for the firm.

Now consider local government's participation constraint (PC_L):

$$[(1 - \phi)sty(\underline{c}) + k^*(1 - \phi)y(\underline{c})] - sty(\bar{c}) - mpR \geq 0 \quad (\text{PC}'_L)$$

Obviously, higher m, p and R make the local government less willing to participate in the collusion. Moreover, given that $\bar{c} \geq \frac{2}{(1-\phi)^2}$, higher s makes the local government more willing to participate in the collusion. This is proved in the following lemma. The intuition is that higher s makes the expected gains from the collusion higher.

Lemma 1 Let LHS_L denote the left-hand side of PC'_L . Given that $\frac{\bar{c}}{c} \geq \frac{2}{(1-\phi)^2}$, $\frac{\partial LHS_L}{\partial s} \geq 0$. In particular, if $s > 0, t \in (0, 1)$, then $\frac{\partial LHS_L}{\partial s} > 0$.

Proof.

$$\begin{aligned} LHS_L &= (1-\phi)y(\underline{c})(st + \frac{1-t-st}{2}) - st\frac{1-t}{\bar{c}} - mpR \\ &= \frac{(1-\phi)^2(1-t+st)^2(1-t)}{4} \frac{1-t}{\underline{c}} - st\frac{1-t}{\bar{c}} - mpR \end{aligned}$$

Given that $\frac{\bar{c}}{c} \geq \frac{2}{(1-\phi)^2}$, we have:

$$\frac{\partial LHS_L}{\partial s} = \frac{(1-\phi)^2 t(1-t+st)}{2\underline{c}} - t\frac{1-t}{\bar{c}} \geq \frac{t(1-t+st)}{\bar{c}} - t\frac{1-t}{\bar{c}} \geq 0.$$

In particular, if $s > 0, t \in (0, 1)$, then $\frac{\partial LHS_L}{\partial s} > 0$. ■

Given the lemma and (PC'_F) , we can characterize the likelihood of collusion (*Collusion*) by the following proposition.

Proposition 1 *Collusion is more likely for higher decentralization (s) and less likely for higher media exposure (m): $\frac{\partial Collusion}{\partial s} > 0$, $\frac{\partial Collusion}{\partial m} < 0$. Besides, higher transaction costs coefficient (ϕ) makes the effect of decentralization small: $\frac{\partial^2 Collusion}{\partial s \partial \phi} < 0$. The predictions for death rates have the same signs: $\frac{\partial Deathrate}{\partial s} > 0$, $\frac{\partial Deathrate}{\partial m} < 0$, $\frac{\partial^2 Deathrate}{\partial s \partial \phi} < 0$.*

Proof. Consider the comparative statics with respect to the left hand side (LHS_F) of (PC'_F) and the left hand side (LHS_L) of (PC'_L) .

First, consider the firm's participation constraint PC'_F .

$$\begin{aligned} \frac{\partial LHS_F}{\partial s} &= -\frac{(1-t)^2(1-k^*)(1-\phi)^2}{\underline{c}} \frac{\partial k^*}{\partial s} \\ &= \frac{(1-t)^2(1-k^*)(1-\phi)^2}{\underline{c}} \frac{t}{2(1-t)} > 0 \end{aligned}$$

$$\frac{\partial LHS_F}{\partial m} = -fp < 0$$

$$\frac{\partial^2 LHS_F}{\partial s \partial \phi} = -\frac{(1-t)^2(1-k^*)(1-\phi)}{\underline{c}} \frac{t}{(1-t)} < 0$$

Similarly, consider the local government's participation constraint:

$$\frac{\partial LHS_L}{\partial s} > 0,$$

if $s > 0$. This is shown by Lemma 1.

$$\frac{\partial LHS_L}{\partial m} = -pR < 0$$

$$\frac{\partial^2 LHS_L}{\partial s \partial \phi} = -\frac{(1-\phi)t(1-t+st)}{\underline{c}} < 0$$

Therefore, the impacts of s, m , and ϕ are of the same direction for the firm and the local government. Given that *Collusion* is more likely for a higher *LHS*, we get $\frac{\partial Collusion}{\partial s} > 0$, $\frac{\partial Collusion}{\partial m} < 0$ and $\frac{\partial^2 Collusion}{\partial s \partial \phi} < 0$. Since the \underline{c} -technology is chosen under *Collusion*, we will also get $\frac{\partial Deathrate}{\partial s} > 0$, $\frac{\partial Deathrate}{\partial m} < 0$ and $\frac{\partial^2 Deathrate}{\partial s \partial \phi} < 0$. ■

In terms of empirical predictions, we can observe s since the centralization or decentralization policy of the Central government is public. We can also observe m to some extent. For example, number of newspapers would be a reasonable measure for m . But we cannot observe the technology device by the coalmining firms (\bar{c} or \underline{c}) nor can we observe the transaction cost coefficient ϕ directly. Instead, we consider some proxies for ϕ . For instance, being a native (N) and having more local network connections will lower the transaction costs of collusion.

Hence, we have the following testable predictions.

(1) Higher decentralization leads to higher death rates: $\frac{\partial Deathrate}{\partial s} > 0$;

(2) More media supervision leads to lower death rates: $\frac{\partial Deathrate}{\partial m} < 0$;

(3) Having a native local leader increases the effect of decentralization on death rates:

$$\frac{\partial^2 Deathrate}{\partial s \partial N} > 0.$$

3 Data

We collect panel data on key state coalmines, governors in charge of coalmine safety, as well as provincial characteristics for 22 provinces across China between 1995 and 2005. Among China's remaining provinces, Tianjin, Shanghai, Hainan and Tibet do not produce any coals, while there are no key coalmines in Fujian, Hubei, Gungdong, Guangxi and Qinghai. Figure 1 maps the distribution of key coalmines across China in terms of production. Table 1 presents some summary statistics.

[Figure 1 about here]

[Table 1 about here]

3.1 Decentralization/delegation of key state coalmines

According to their ownership, Chinese coalmines can be divided into three types: key state coalmines, local state coalmines and township and village coalmines. There are 94 key state coalmines, which in 2003 produced 47.81% of the total coals in China whereas local state coalmines produced 16.94% and township and village coalmines produced 35.25% (The State Administration of Coal Mine Safety 2004).

In this paper, we focus on the management of the key state coalmines. Before 1998, the coal mining industry was overseen by the Ministry of Coal Industry in the Central government. Due to many policy burdens for state-owned enterprises and the competition of small coalmines, the profitability of the key state mines was negative in the 1990s. To provide more incentives for profitability, the management of the key state mines was shifted to provincial governments in 1998 (The State Council 1998). This delegation decision was passed at China's Ninth National People's Congress in March 1998. This delegation was also associated with 206 enterprises affiliated with these coalmines, an asset of 237.9 billion yuan (\$30 billion) and 4.35 million employees. After delegation, the taxes from the key state mines were handed over to the provincial government and the profits were left to the coalmines. Meanwhile, the powers of safety supervision were also shifted to the provincial governments.

This decentralization period of key state coalmines lasted until February 2001, when the State Administration of Coal Mine Safety (SACMS) was established and re-centralized safety supervision powers. The centralization was further increased in 2003, when SACMS became part of the general offices of the State Council.

Hence, we take the period between 1998 and 2000 as the decentralized period when collusion became possible or at least much easier, whereas 1995-1997 and 2001-2005 are centralized comparison periods. Note that the two centralized periods differ slightly. Before 1998, the key state coalmines were standard state-owned enterprises, whose management power and supervision power all belonged to the central government. These powers were all delegated to the provincial governments in 1998. Only the supervision power got centralized since 2001. We group the two centralized period together in our baseline estimations and separate them as a robustness check.

3.2 Transaction costs of collusion and media exposure

As mentioned above, we gauge the transaction costs of collusion (ϕ in the model) by examining the biographical information of the Safety governor. There is one main governor and four to six vice governors in charge of economic-related activities in a province. The Safety governor is one of the vice governors and has a tenure of at most 10 years. Besides

coalmine safety, the vice governor is also in charge of the safety of other industries such as construction-site safety and road-traffic safety. We trace the careers of all the 71 Safety governors across 22 provinces between 1995 and 2005. These data come from yearly provincial government reports and the CVs can be found on People’s Daily online. The average length in office is very close for native governors and non-native governors, which is around 4 years.

We proxy the transaction costs of collusion by looking at whether the governor in charge of coalmine safety is a native, i.e. born in the same province that he governs. Since all of the governor work some years before coming to the office, being native naturally implies that they have longer experience in a province. About 43% of the vice governors in charge of coalmine safety are native. This share is very similar to the share of natives among the major provincial governors. Figure 2 shows the share of native governments over time. We do not see any change along with decentralization and only a slight increase after 2001.

[Figure 2 about here]

We realize that these measures might also capture something else besides the possibility of collusion. For example, being a native might care more about safety of the local people as well as make it easier to under-report death numbers. However, these concerns would only bias our findings in the opposite direction from that of our main hypotheses.

We measure media exposure (m in the model) by newspapers published per 1000 individuals between 1996 and 2005. Numbers of newspapers published come from China Statistical Data of Press and Publication published since 1997. We divide them by provincial population sizes from China Statistical Yearbook. The number of newspapers published per 1000 individuals varies from 0.36 to 8.58.

3.3 Death rates

Number of deaths and the yearly production of key state coalmines come from China Coal Industry Yearbook and China Statistical Yearbook. Death rates are measured by number of deaths per million tons of coal production. As shown in Table 1, the mean of death rate is about 2.5 people per million tons of production.

As for measurement error in deaths, the death rates of key state coalmines are believed to be more reliable than those of smaller coalmines. This is another reason why we focus on them in this paper. On top of this, suppose the provincial governments have incentives to under-report deaths, they would have better ability of under-reporting in the decentralized period. This should again bias our findings in the opposite direction from the predictions of our model.

Figure 3 plots the changes of death rates over time by whether the governor in charge of coalmine safety is a native. This figure serves as a motivational picture. Clearly, the death rate jumped dramatically between 1998 and 2000, especially for the provinces with native observations. However, it is worthwhile pointing out the picture has some limitations. First, there are compositional changes in the two groups, because governors may vary from year to year. There are very few provinces whose governors are always native or always non-native. Second, there is a jump in 1998 because the death rate in Zhejiang Province is exceptionally high in that year. In our estimations in Section 5.1, we will drop this outlier in Zhejiang Province as a robustness check. We will also visualize the estimation effects year by year in Figure 4 below, to compare with Figure 3 from the raw data.

[Figure 3 about here]

As an alternative to measure safety, we also use number of deaths and the number of accidents in our estimations below. The data source is the same as the information on death rates. However, the information on number of accidents is only available since 1997 and is not complete for each year.

3.4 Other controls

We also control for characteristics of coalmines, other information on the governors and other provincial characteristics as well as their interactions with decentralization.

Data on coalmines includes average yearly wages and size of fixed assets. The average yearly wage is a proxy for labor quality, whereas size of fixed assets is a proxy for the scale and technological inputs of coalmines. We also include the production level. These data also come from China Coal Industry Yearbook and China Statistical Yearbook.

Other information on the governors in charge of coalmines includes their age, how many years they have been in office and whether they have some experience of safety management before coming to office. We use these variables to control for the career concerns and experience of the governors. Like the proxies for transaction costs of collusion, these data come from yearly provincial government reports.

Provincial characteristics include real GDP per capita, whether the main provincial governor is a native, and the distance of the provincial capital to Beijing. GDP per capita comes from China Statistical Yearbook. The biographical information of the main governors is from yearly provincial government reports. The distance of the provincial capital to Beijing is calculated with ArcGIS and works as a proxy for the supervision intensity by the Central government. In practice, we use $\log(1+\text{distance to Beijing})$ to take account of the 0 observations of Beijing.

4 Baseline Estimates

In our baseline estimation, we explore within-province variation by running a fixed-effect model including province fixed effects and year fixed effects. To test our basic prediction, we run the following specification:

$$\text{Deathrate}_{it} = \lambda_i + \beta D_{it} + \sum_{i=2}^{22} \delta_i \text{prov}_i * \text{year}_t + \varepsilon_{it}, \quad (1)$$

where λ_i denotes province fixed effects. D_{it} is the decentralization dummy for year 1998-2000. In this case, we cannot control for year fixed effects, as their effects are absorbed by D_{it} . We also control for provincial specific trends by including the interaction of provincial dummies and year, $\sum_{i=2}^{22} \delta_i \text{prov}_i * \text{year}_t$.

Our theory predicts that $\beta > 0$. The results presented in Table 2 are consistent with our theory. Column (1) shows the results from OLS, column (2) and (3) are the results from fixed effects regression with and without province-specific trends, and column (4) also include the production level. The magnitude of the decentralization effect is about half the mean and 25% of the standard deviation of the dependent variable.

[Table 2 about here]

To test predictions (2) and (3), we run the following specification:

$$\begin{aligned} \text{Deathrate}_{it} = & \lambda_i + \gamma_t + \beta_1 N_{it} * D_{it} + \beta_2 \text{News}_{it} + \beta_N N_{it} + \gamma' X_{it} * D_{it} + v' X_{it} \\ & + \sum_{i=2}^{22} \delta_i \text{prov}_i * \text{year}_t + \varepsilon_{it}, \end{aligned} \quad (2)$$

where N_{it} is a binary indicator whether the governor is a native in province i and year t . News_{it} is the number of newspapers published per 1000 individuals. We do not include the dummy of D_{it} itself as we have already controlled for year fixed effects γ_t . Including it does not change our results.

X_{it} is a vector of controls, including some other characteristics of the governor in charge of coalmine safety, such as his age, how many years he has been in office and whether he has work experience in safety supervision departments. We control for coalmine characteristics including the log amount of fixed assets and log of the average yearly wage of mine workers. Moreover, we also control for other provincial characteristics such as GDP per capita, whether the major provincial governor is a native and the log distance of the provincial capital to

Beijing. We also control for the variables mentioned above as well as their interactions with decentralization.

To be cautious about autocorrelation concerns, we report AR (1) standard errors for all the specifications to follow.⁸

The results are presented in Table 3. Column (1) and (2) are the results from the fixed effects model with and without controls. Column (3) presents the results including provincial specific trends. Similarly, Column (4)-(6) are the results after considering the data of newspapers per 1000 individuals.

Our theory predicts that $\beta_1 > 0$ and $\beta_2 < 0$. The results are in line with these predictions and stable over different specifications. In the decentralized period, the mean of the dependent variable is about 3.3 and its standard deviation is about 6.6. Within the decentralized period, the death rate for the province with native Safety governor is higher by the magnitude of 4, which is about 100% of the mean of dependent variable. On the other hand, a standard deviation of newspaper per 1000 individuals (1.36) decreases the death rate by the magnitude of 3.

Among the controls, the ones that have significant impacts are production level and tenure of the Safety governors. The interaction between the production level and decentralization has negative impact on death rates, which is again different from the hypothesis that death rates are increase with production. Tenure is a dummy to measure whether the Safety governor has been in office longer than average (4 years). The interaction between the tenure dummy and decentralization has positive impacts on death rates. Although being in office longer captures different dimensions such as network and learning experience, this findings tends to support our hypothesis of collusion, as being in office longer would make collusion easier.

[Table 3 about here]

5 Robustness Checks

In this section, we conduct a couple of robustness checks to deal with concerns of measurement, endogeneity as well as possible heterogeneity.

First, we conduct two robustness checks about measurement. In our baseline estimations, we look at the average impact of collusion likelihood in the whole decentralized period. As a robustness check, we allow for more flexible specifications by looking at the interactions

⁸We also tried with other ways of dealing with standard errors. FGLS (xtgls) gives us similar magnitudes and higher significance. Clustering at the provincial level gives the same magnitudes and lower significance levels (around 10% in the baseline estimations) as we only have 22 provinces in our sample.

of having a native Safety governor and year dummies. We also exclude the outlier in 1998. This will generate a more convincing picture, as a comparison to the motivational picture in Figure 3. Another measurement check is to use alternative measures as the dependent variable, including number of deaths and accidents.

We also conduct two robustness checks related to endogeneity concerns. Although the national decentralization decision is exogenous to the local governments, there still might be a concern that natives are appointed to provinces with higher death rates. We do not find this true in our data. We also allow for flexible specifications by looking at the effects of the interactions between the Native dummy and year dummies. This way, we can check whether there are pre-trend before the decentralization. As another robustness check for endogeneity, we also control for governor fixed effects and explore within-governor variation. Besides, we limit our sample by excluding possibly endogenous governors.

Given that we have two periods of centralization (1995-1997 and 2001-2005) and one period of decentralization (1998-2000), we can test our theory separately to see whether the effects are heterogeneous over time. We find the effects are generally larger for the period between 1995 and 2000.

5.1 Measurement check I: more flexible specifications

To check whether there were already some different trends for the provinces with native vice governors before the decentralization decision in 1998, we allow for more flexible specifications. In particular, we run a regression as follows

$$\text{Deathrate}_{it} = \lambda_i + \gamma_t + \sum_{\tau=1996}^{2005} \beta_{\tau} N_{i\tau} * \text{year}_{\tau} + \sum_{\tau=1996}^{2005} \gamma' X_{it} * \text{year}_{\tau} + v' X_{it} + \varepsilon_{it},$$

where $\sum_{\tau=1996}^{2005} N_{i\tau} * \text{year}_{\tau}$ are the interactions between the dummy for being native and different year dummies and $\sum_{\tau=1996}^{2005} X_{it} * \text{year}_{\tau}$ are the interactions between different controls mentioned above and year dummies.

The results are presented in Table 4. Column 1 presents the results from the fixed effects model. Column 2 shows the results after dropping the outlier in 1998. Column 3 further are the results after dropping the outlier and controlling for other variables as well as their interactions with year dummies. Similar to the baseline specification, we can also control for province-specific trends by including $\sum_{i=2}^{22} \text{prov}_i * \text{year}_t$. The results are shown in column 4.

These results imply that having a native vice governor did not matter before the decentralization. The effect became significant during the decentralized period: the year of 1998, 1999 and 2000. The significance of the effect disappears again in the second centralized period since 2001.

[Table 4 about here]

Figure 4 visualizes the results from column 2. The main message is consistent with that of Figure 3. However, Figure 4 better deals with the compositional concern and the outlier problem.

[Figure 4 about here]

5.2 Measurement check II: using number of deaths and accidents

In our baseline estimations, we use number of deaths per million tons of coal production as an indicator of safety. It is an intuitive measure that is discussed in the media. A potential concern is whether the results are driven by a couple of extreme values. As mentioned above, the results are robust to dropping outliers. As another robustness check, we can also use the number of deaths and the number of accidents as the dependent, while controlling for the amount of coal production.

Since both the number of deaths and the number of accidents are count variables, we employ Poisson fixed effects. The specification is as follows,

$$Y_{it} = \exp(\lambda_i + \gamma_t + \beta_1 N_{it} * D_{it} + \beta_N N_{it} + \beta_2 News_{it} + \gamma' X_{it} * D_{it} + \nu' X_{it} + \varepsilon_{it})$$

whereas Y_{it} can be the number of deaths and the number of accidents. Controls included in X_{it} are the same as those in the baseline estimations.

Panel A of Table 5 presents the results for the number of deaths. Columns 1-3 report the coefficients, which can be interpreted as elasticity. We also report the marginal effects in column 4-6, which measure the effects of the discrete change from 0 to 1 for $N_{it} * D_{it}$. Thus, when the interaction of $N_{it} * D_{it}$ switches from 0 to 1, the number of deaths increase by about 27, which is close to the mean of deaths (30).

[Table 5 about here]

Similarly, Panel B presents the results for the number of accidents. As mentioned above, the information on the number of accidents is not complete. Although the signs are consistent with our model predictions, the estimation results are not precise.

5.3 Endogeneity check I: including governor fixed effects

Since we have information for each governor, we can also include governor fixed effects as a robustness check. This way, we look at the interaction effect of decentralization and being native within the same governor. Analogous to the baseline estimations, we run fixed-effect models after controlling for governor fixed effects. The controls are the same as in the baseline estimations.

The estimation results are presented in Table 6. The impact of newspapers is still negative but not significant in this specification. This is because there is not enough variation in terms of this variable within governors. The rest of the results are very similar to what we find in the within-province estimation. This provides further evidence on decentralization and collusion.

[Table 6 about here]

5.4 Endogeneity check II: excluding a subgroup of governors

We trace the career paths of the governors and define a subset of governors whose predecessor was politically demoted as possibly endogenous governors. The motivation is that this subset of governors might be more likely to be endogenously assigned as the previous ones are demoted for some reason.

In our data, 25% of this subset are native whereas 45% of the rest are native. This information does not suggest that natives are more likely to be appointed to provinces where previous ones are demoted.

As a robustness check, we drop the subset of governors whose predecessors were demoted, which is about 20% of our observations. We then run the same regression as (2) and the results are presented in Table 7. The results from column (1) to column (6) are close to our baseline estimates in magnitude.

[Table 7 about here]

5.5 Heterogeneity check: comparing 1995-2000 and 1998-2005

We know from government official documents that the decentralization decision in 1998 was due to the consideration of profitability (The State Council 1998). However, we do not know the precise motives behind the centralization decision in 2001. Besides, the centralization refers to the supervision power particularly. Our model predicts that we should see an effect, both when decentralization is introduced and when it is taken away. To check whether the

effects are heterogenous between the two decisions, we can estimate separate regressions for subperiods.

Now we replicate the same regressions in (1) and (2) but separate the time into two sample periods: 1995-2000 and 1998-2005. In the first subsample 1995-2000, 1995-1997 is the centralization period whereas 1998-2000 is the decentralization period. In the second subsample 1998-2005, 1998-2000 is the decentralization period where 2001-2005 is the centralization period.

The results are presented in Table 8. Panel A presents the results for the period 1995-2000. Column 1 and 2 are the results for the impact of decentralization. Columns 3-6 are the results for the interaction effects. Similarly, Panel B presents the results for the period 1998-2005. We find the effects of the decentralization decision are generally somewhat larger in the first subsample. But the conclusions from the baseline estimations hold for both periods.

[Table 8 about here]

6 Alternative explanations

As mentioned in the introduction, both media and academia have provided quite a few hypotheses to explain China's high coalmine death rate. We categorize the alternative hypotheses into three groups. The first hypothesis is insufficient safety input (Tang and Guo 2006). This argues that insufficient inputs are the main reason behind the high accident rates in China, especially for small local coalmines. For example, gas drainage systems have not been set up in 40% of the coalmines.

A second hypothesis is insecure property right (Tan and Tang 2004). This would argue that the ownership of mines is not stable and often subject to policies of the Central government. The Central government sometimes closes coalmines without noticing owners of the mines when safety becomes politically sensitive. Insecure property rights makes owners short-sighted. Hence, the owners under-invests in safety facilities.

A third hypothesis is low quality of coalmine workers, which emphasizes the demand side of safety inputs. This would argue that the coalmine workers are usually uneducated, lack knowledge on safety and do not have strong demands of improvements in their working conditions (Whyte 1999).

Insufficient safety inputs no doubtedly matter for the high death rates. But the next question to answer is why the safety inputs are not sufficient, which is what we try to do in this paper. The property right hypothesis may provide a possible explanation. However, this

hypothesis would generate opposite results from our findings, since decentralization should provide better property rights protection.

It is not easy to measure the demand for workplace safety. We can use yearly wages of mine workers as proxies for labor quality. This proxy does not matter after controlling for province fixed effects and year fixed effects, as shown in Table 3.

Our theory provides a political-economics explanation about the insufficient supply of safety inputs. Our study contributes to the understanding of high coalmine death rates by providing empirical evidence in support of this theory.

7 Conclusion

We propose a theory of decentralization and collusion to explain China's abysmal record on coalmine safety. In this theory, the local government allows firms to choose dangerous but high payoff production technology when collusion is feasible. We explore an institutional change in the management of key coalmines to test our theory, using a provincial level panel dataset from 1995 and 2005.

We might be the first to write down a model of collusion and provide empirical evidence for it. However, as mentioned in the introduction, this insight about collusion between governments and firms has been well recognized in the media. We might be the first to write down a model of collusion and provide empirical evidence for it. However, as mentioned in the introduction, this insight about collusion between governments and firms has been well recognized in the media. This study focuses on a specific industry of coalmine. But the political logic can be generalized to other industries and other type of problems.

Our theory is consistent with the current dilemma of China: dramatic growth and overwhelming problems with workplace safety and the environment. While decentralization has led to growth achievements, it has also brought many unexpected problems. More careful pondering is needed to evaluate the consequences of public-sector decentralization.

We use aggregate data at the provincial level in this paper. Another perspective to investigate the death rates in coalmines would be to study data at the level of individual coalmines. Coalmine level data may also help us better analyze the welfare implications to coalmine workers. This task is left for future work.

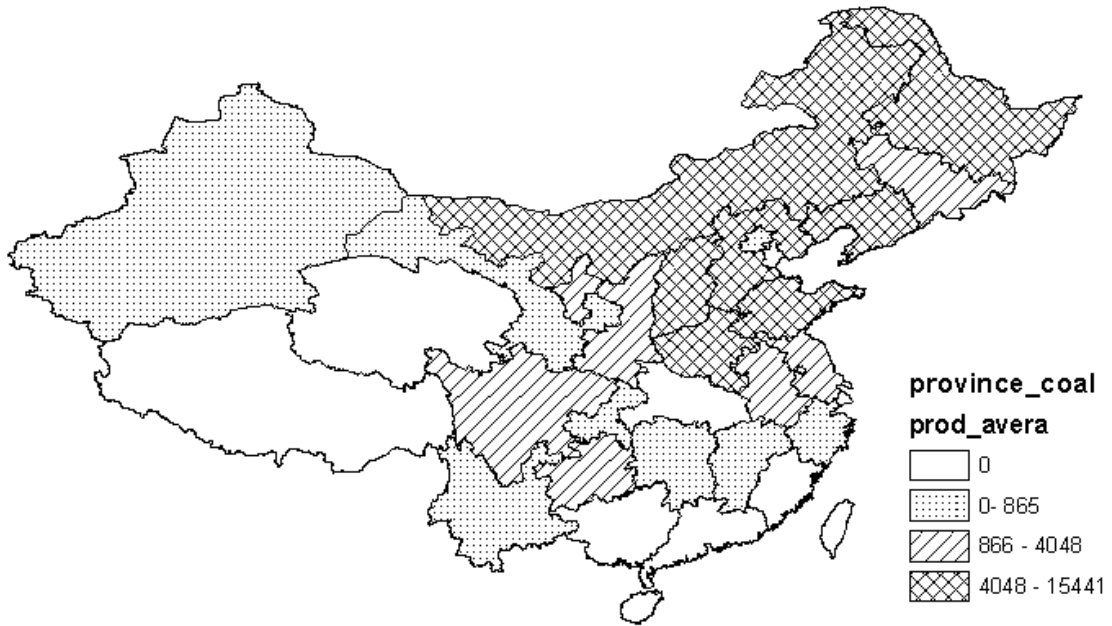
References

- Asian Development Bank (2007). People's Republic of China: Coal Mine Safety Study. www.adb.org/Documents/Reports/Consultant/39657-PRC/39657-PRC-TACR.pdf.
- Baliga, S. and T. Sjostrom (1998). Decentralization and Collusion. *Journal of Economic Theory* 83(2), 196–232.
- Bardhan, P. and D. Mookherjee (2006). Decentralization, Corruption and Government Accountability. *International Handbook on the Economics of Corruption*, 161–188.
- Blanchard, O. and A. Shleifer (2000). Federalism With and Without Political Centralization: China versus Russia. *NBER*.
- Egorov, G., S. Guriev, and K. Sonin (2009). Why Resource-Poor Dictators Allow Freer Media: A Theory and Evidence From Panel Data. *American Political Science Review* 103(04), 645–668.
- Faccio, M. (2006). Politically Connected Firms. *The American Economic Review* 96(1), 369–386.
- Fisman, R. (2001). Estimating the Value of Political Connections. *The American Economic Review* 91(4), 1095–1102.
- Guo, X. (2001). Dimensions of Guanxi in Chinese Elite Politics. *The China Journal* (46), 69–90.
- Jin, H., Y. Qian, and B. Weingast (2005). Regional Decentralization and Fiscal Incentives: Federalism, Chinese style. *Journal of Public Economics* 89(9-10), 1719–1742.
- Kofman, F. and J. Lawarree (1993). Collusion in Hierarchical Agency. *Econometrica* 61(3), 629–656.
- Laffont, J. and D. Martimort (1998). Collusion and Delegation. *The Rand Journal of Economics* 29(2), 280–305.
- Li, S. and J. Wu (2010). Why Some Countries Thrive Despite Corruption: The Role of Trust in the Corruption–Efficiency Relationship. *Review of International Political Economy* 17(1), 129–154.
- Lin, J. and Z. Liu (2000). Fiscal Decentralization and Economic Growth in China. *Economic Development and Cultural Change*, 1–21.
- Luo, Y. (2007). *Guanxi and Business*, Volume 5. World Scientific Pub Co Inc.
- Mookherjee, D. (2006). Decentralization, Hierarchies, and Incentives: A Mechanism Design Perspective. *Journal of Economic Literature* 44(2), 367–390.

- Mookherjee, D. and M. Tsumagari (2004). The Organization of Supplier Networks: Effects of Delegation and Intermediation. *Econometrica* 72(4), 1179–1219.
- Nie, H. and M. Jiang (2011). Zhengqi hemou yu kuangnan(Coal Mine Accidents and Collusion between Local Governments and Firms). *Jingjiyanjiu (Economic Research Journal)* 6.
- Oi, J. (1999). *Rural China Takes off: Institutional Foundations of Economic Reform*. University of California Press.
- Reinikka, R. and J. Svensson (2005). Fighting Corruption to Improve Schooling: Evidence from a Newspaper Campaign in Uganda. *Journal of the European Economic Association* 3(2-3), 259–267.
- Tan, M. and X. Tang (2004). Chanquan niuqu: kuangnandeshencengsikao (Property Rights and Coalmine Accidents). *Meitanxuebao (Journal of Coal Mine)* 29(6), 756–759.
- Tang, L. and X. Guo (2006). Woguo xianjieduan kuangnan pinfa chenyinjiquiduce: jiyu anquan touru de shjiao (The Reasons for Coalmine Mishap in China and Policy Implications). *Zhongguo gongye jingji (China Industrial Economy)* 12.
- The Economist (2002). Pits of Death: Counting the Cost of Coal. <http://www.economist.com/node/1284235>.
- The Economist (2005). Going Under. <http://www.economist.com/node/3966158>.
- The Economist (2007). Bottoming Out. <http://www.economist.com/node/9687898>.
- The State Administration of Coal Mine Safety (Ed.) (2004). *China Coal Industry Yearbook*. The Coal Industry Press.
- The State Council (1998). Guanyu guoyou zhongdian meikuang guanli tizhi youguan wenti de tongzhi (Notice on the Reform of the Key State Coalmines).
- Tirole, J. (1986). Hierarchies and Bureaucracies: On the Role of Collusion in Organizations. *Journal of Law, Economics, and Organization* 2(2), 181.
- Tu, J. (2007). Coal Mining Safety: China's achilles' heel. *China Security* 3, 36–63.
- Whyte, M. (1999). *The Paradox of China's Post-Mao Reforms*. Harvard Uni.
- Wright, T. (2004). The Political Economy of Coal Mine Disasters in China: Your Rice Bowl or Your Life. *The China Quarterly* 179(-1), 629–646.
- Wright, T. (2008). *Rents and Rent Seeking in the Coal Industry*. Taylor & Francis.

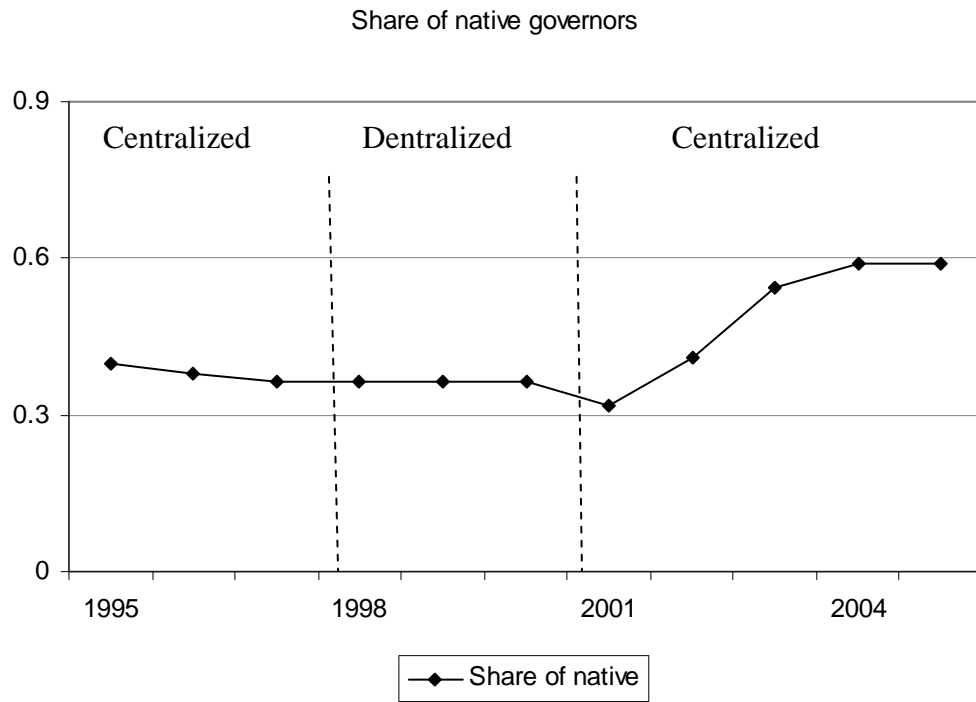
- Young, A. (2000). The Razor's Edge: Distortions and Incremental Reform in the People's Republic of China. *Quarterly Journal of Economics* 115(4), 1091–1135.
- Zhang, T. and H. Zou (1998). Fiscal Decentralization, Public Spending, and Economic Growth in China. *Journal of Public Economics* 67(2), 221–240.

Figure 1: Average production of key state coalmines between 1995 and 2005
(million tons)



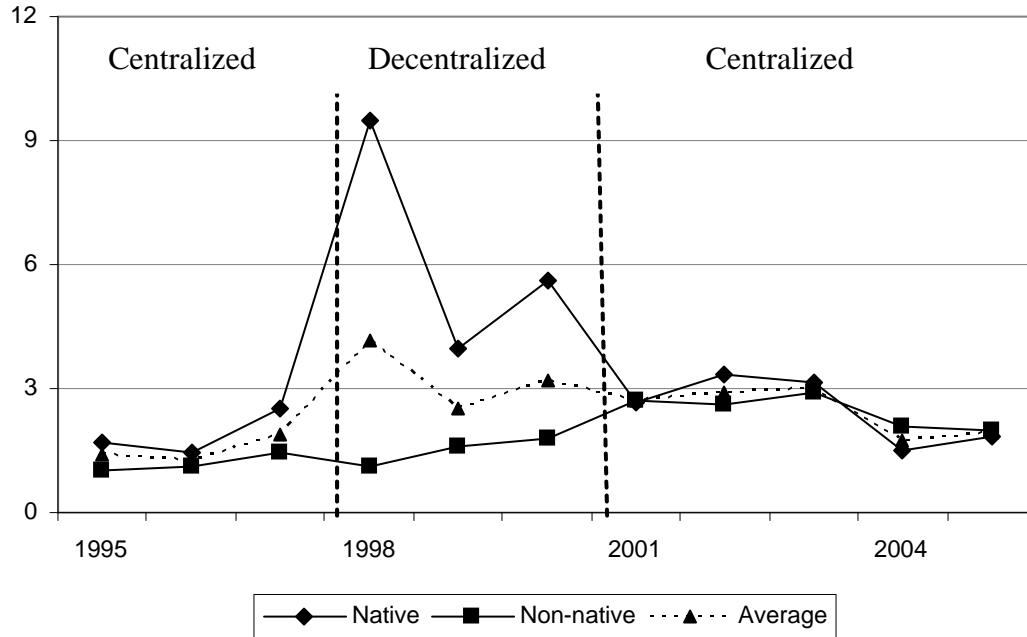
There are 22 provinces with key state coalmines.

Figure 2: Share of native governors in charge of coalmine safety between 1995 and 2005



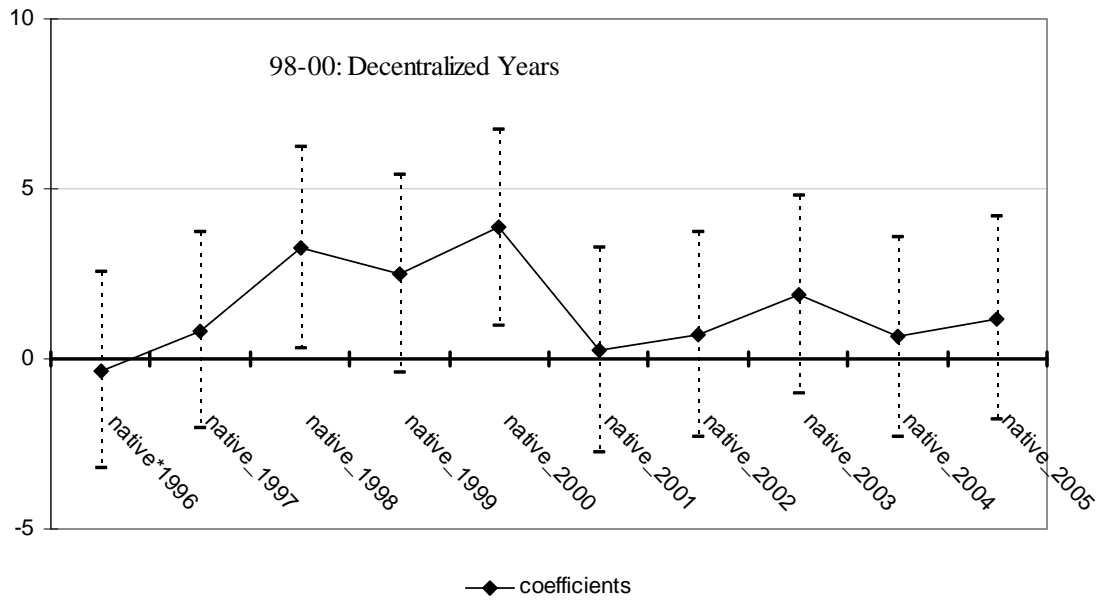
The figure shows that the share of native governors is not correlated with the decision of decentralization.

Figure 3: Raw data: coalmine death rates between 1995 and 2005



This figure is from the raw data. (1) There are compositional changes, as very few provinces whose governor are always native or non-native. (2) The death rate in Zhejiang whose governor is native in 1998 is particularly high. We will drop this observation as a robustness check.

Figure 4: Estimation Results



This figure visualizes the results from the fixed effects model.

Table 1. Summary statistics for the province-year data

Variables	Mean	S.D.	Min	Max	Number
Transaction costs of collusion and media exposure					
Native	0.43	0.50	0	1	239
Published newspapers per 1000 individuals	1.58	1.36	0.36	8.58	218
Coalmine characteristics					
Death rates (million tons)	2.45	4.10	0	46.25	233
Production (1,0000 tons)	2976	3815	42	25621	238
Number of Deaths	31.3	37.1	0	251	234
Number of Accidents	18.7	11.9	1	57	172
Yearly wage (RMB)	10292	4782	4321	25827	240
Fixed assets (100 million RMB)	36.99	59.69	0.02	329.84	238
Other leader's characteristics					
Whether has been in office more than 4 years	0.41	0.49	0	1	239
Years in office	3.6	2.3	1	10	239
Age	52.6	4.67	36	61	238
Safety Work Experience	0.31	0.46	0	1	242
Other provincial characteristics					
GDP per capita	7438	4926	1826	28050	241
Whether the main governor is native	0.42	0.49	0	1	242
Distance to Beijing (Km)	1075	674	0	2559	242

Table 2. Decentralization and Death Rates

Dependent Variable Model	Deaths per million tons of coal production			
	OLS	Fixed Effects	Fixed Effects	Fixed Effects
	(1)	(2)	(3)	(4)
Decentralization	1.182** (0.595)	1.043* (0.592)	1.227* (0.644)	1.285* (0.668)
Coal production				0.123 (0.364)
Province fixed effects		Y	Y	Y
Year fixed effects		Y	Y	Y
Province specific trends			Y	Y
Mean of the dependent var.	2.45	2.45	2.45	2.45
S.d. of the dependent var	(4.10)	(4.10)	(4.10)	(4.10)
Observations	233	211	211	211

Note: AR(1) standard error are reported. *** indicates $p < 0.01$, ** indicates $p < 0.05$, and * indicates $p < 0.1$.

Table 3. Within-province estimations

	(1)	(2)	(3)	(4)	(5)	(6)
Native*Decentralization	4.216***	4.773***	5.130***	3.711***	4.317***	4.980***
	(1.280)	(1.539)	(1.715)	(1.354)	(1.628)	(1.763)
Native	-0.067	0.339	0.226	0.361	0.435	0.171
	(0.868)	(1.041)	(1.426)	(0.961)	(1.134)	(1.548)
Newspaper per 1000				-2.227**	-1.957*	-1.635
				(0.962)	(1.040)	(1.165)
Production*Decentralization		-0.610**	-0.625**		-0.548**	-0.659**
		(0.250)	(0.263)		(0.270)	(0.286)
Tenure*Decentralization		2.754**	2.621*		2.412	2.039
		(1.382)	(1.562)		(1.478)	(1.597)
Safety* Decentralization		-1.073	-0.964		-1.035	-0.835
		(1.608)	(1.795)		(1.696)	(1.880)
Age* Decentralization		0.291	0.341		0.273	0.352
		(0.221)	(0.249)		(0.233)	(0.263)
Fixedassets*Decentralization		-0.184	0.180		-0.079	0.413
		(0.578)	(0.673)		(0.600)	(0.667)
Wage*Decentralization		2.929	0.159		2.557	-1.393
		(3.282)	(3.584)		(3.458)	(3.757)
GDP per*Decentralization		1.133	1.427		0.157	0.999
		(1.714)	(1.846)		(1.861)	(1.985)
Native_main*Decentralization		0.688	0.365		1.107	0.796
		(1.499)	(1.609)		(1.601)	(1.669)
Distance_Beijing*Decentralization		0.555	0.443		0.235	-0.298
		(0.666)	(0.778)		(0.747)	(1.010)
Production		-0.134	-0.082		-0.066	-0.157
		(0.232)	(0.436)		(0.253)	(0.457)
Tenure		0.402	0.914		0.922	1.695
		(0.877)	(1.027)		(0.988)	(1.164)
Safety experience		-0.226	-0.727		-0.457	-1.659
		(1.307)	(1.693)		(1.457)	(1.809)
Age		0.082	0.034		0.075	0.048
		(0.084)	(0.102)		(0.094)	(0.127)
Fixedassets		-0.192	-0.453		-0.153	-0.468
		(0.423)	(0.489)		(0.454)	(0.499)
Wage		1.093	-0.255		1.906	0.562
		(3.857)	(5.051)		(4.197)	(5.297)
GDP per capital		-2.785	1.894		-3.583	-2.324
		(4.213)	(7.635)		(4.731)	(15.841)
Native main governor		-1.015	-0.352		-1.380	-1.019
		(0.917)	(1.189)		(1.029)	(1.210)
Province FE and Year FE	Y	Y	Y	Y	Y	Y
Province-specific trends			Y			Y
Observations	209	206	206	189	187	187

.Note: AR(1) standard error are reported.

Table 4. Measurement check I: more flexible specifications

	(1)	(2)	(3)	(4)
Native*Year1996	-0.376 (1.034)	-0.737 (1.169)	0.299 (2.782)	-0.993 (1.536)
Native*Year1997	0.663 (1.034)	0.398 (1.168)	1.974 (1.650)	0.099 (1.469)
Native*Year1998	2.395** (1.102)	3.154*** (1.190)	3.394** (1.529)	2.816** (1.400)
Native*Year1999	2.120** (1.034)	1.993* (1.158)	1.822 (1.452)	1.640 (1.284)
Native*Year2000	3.769*** (1.034)	3.605*** (1.141)	4.693*** (1.383)	3.448*** (1.232)
Native*Year2001	0.830 (1.102)	0.023 (1.155)	0.353 (1.522)	-0.140 (1.210)
Native*Year2002	1.480 (1.034)	0.102 (1.108)	-0.324 (1.753)	-0.090 (1.187)
Native*Year2003	1.284 (0.856)	1.466 (1.081)	3.051** (1.458)	0.807 (1.299)
Native*Year2004	-0.362 (0.825)	0.294 (1.099)	0.361 (1.509)	-0.453 (1.434)
Native*Year2005	0.013 (0.856)	0.808 (1.115)	-0.313 (1.526)	-0.232 (1.581)
Native	-0.376 (1.034)	-0.737 (1.169)	0.299 (2.782)	-0.993 (1.536)
Year Fixed Effects		Y	Y	Y
Province Fixed Effects		Y	Y	Y
Dropping the outlier in 1998	Y	Y	Y	Y
Control trends			Y	
Province specific trends				Y
Observations	230	208	206	208

Notes: The results are from the fixed effects models which allow flexible specifications: allowing the interactions of the dummy for being a native and year dummies. It shows that death rates increases for the natives during the decentralized period: the year of 1998, 1999 and 2000. The results from column (2) are visualized in Figure 4. AR(1) standard error are reported.

Table 5: Measurement check II: using alternative dependent variable

Panel A. Dependent Variable: Number of deaths						
	Coefficient			Marginal effect		
	(1)	(2)	(3)	(4)	(5)	(6)
Native*Decentralization	0.803*** (0.061)	0.828*** (0.075)	0.819*** (0.077)	25.56*** (2.62)	26.37*** (3.28)	27.26*** (3.49)
Native	-0.003 (0.038)	0.012 (0.051)	0.031 (0.054)	-0.06 (0.85)	0.26 (1.15)	0.74 (1.27)
Newspaper per 1000			-0.191*** (0.045)			-4.51*** (1.05)
Province fixed effects	Y	Y	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y	Y	Y
Controls		Y	Y		Y	Y
Mean of the depend.var.	31.3	31.3	31.3	31.3	31.3	31.3
S.d. of the depend. var.	(37.1)	(37.1)	(37.1)	(37.1)	(37.1)	(37.1)
Observations	232	228	209	232	228	209
Panel B. Dependent Variable: Number of Accidents						
	Coefficient			Marginal effect		
	(1)	(2)	(3)	(4)	(5)	(6)
Native*Decentralization	0.222** (0.094)	0.121 (0.115)	0.107 (0.115)	3.76** (1.72)	1.98 (1.96)	1.73 (1.94)
Native	-0.116* (0.064)	-0.096 (0.082)	-0.109 (0.082)	-1.79* (0.97)	-1.49 (1.27)	-1.69 (1.27)
Newspaper per 1000			-0.194*** (0.065)			-3.93*** (1.01)
Province fixed effects	Y	Y	Y	Y	Y	Y
Year Fixed Effects	Y	Y	Y	Y	Y	Y
Controls		Y	Y		Y	Y
Mean of the depend.var.	18.7	18.7	18.7	18.7	18.7	18.7
S.d. of the depend. var.	(11.9)	(11.9)	(11.9)	(11.9)	(11.9)	(11.9)
Observations	172	170	170	172	170	170

Note: The results are from fixed effects Poisson estimations.

Table 6. Endogeneity check I: including governor fixed effects

	(1)	(2)	(3)	(4)	(5)	(6)
Native*Decentralization	4.559*** (1.519)	5.072** (2.013)	5.663** (2.454)	4.627*** (1.681)	4.547** (2.002)	5.047 (4.272)
Native	0.342 (2.516)	1.456 (4.966)	-3.723 (6.846)	-0.668 (3.424)	11.033* (5.583)	-0.627 (5.535)
Newspaper per 1000				-3.164** (1.369)	-1.788 (1.158)	-1.214 (2.620)
Year Fixed Effects	Y	Y	Y	Y	Y	Y
Province Fixed Effects	Y	Y	Y	Y	Y	Y
Control trends		Y	Y		Y	Y
Province-specific trends			Y			Y
Observations	209	206	206	189	187	187

Notes. The results are from baseline regressions including governor fixed effects. AR(1) standard error are reported.

Table 7. Endogeneity check II: dropping possibly endogenous leaders

	(1)	(2)	(3)	(4)	(5)	(6)
Native*Decentralization	4.293**	4.015*	4.268*	3.618**	2.859	3.312
	(1.661)	(2.128)	(2.437)	(1.753)	(2.291)	(2.554)
Native	0.370	0.554	1.387	0.763	0.499	1.132
	(1.250)	(1.411)	(2.077)	(1.365)	(1.570)	(2.270)
Newspaper per 1000				-2.542**	-2.793**	-2.613*
				(1.121)	(1.217)	(1.425)
Year Fixed Effects	Y	Y	Y	Y	Y	Y
Province Fixed Effects	Y	Y	Y	Y	Y	Y
Control trends		Y	Y		Y	Y
Province-specific trends			Y			Y
Observations	209	206	206	189	187	187

Notes: This table shows the results after excluding a subgroup of governors whose predecessors were demoted. AR(1) standard error are reported.

Table 8. Heterogeneity check: comparing two periods

Panel A: 1995-2000						
	(1)	(2)	(3)	(4)	(5)	(6)
Decentralization	1.776** (0.876)	1.862** (0.790)				
Native*Decentralization			5.061*** (1.750)	7.108*** (2.666)	4.917** (2.116)	7.357* (3.686)
Native			-2.334 (1.959)	-1.998 (3.112)	-2.348 (2.488)	-1.099 (3.813)
Newspaper per 1000					-2.206 (2.396)	-2.152 (2.436)
Year fixed effects		Y	Y	Y	Y	Y
Province fixed effects		Y	Y	Y	Y	Y
Controls				Y		Y
Observations	125	125	123	121	103	102
Panel B: 1998-2005						
	(1)	(2)	(3)	(4)	(5)	(6)
Decentralization	0.852 (0.727)	0.665 (0.624)				
Native*Decentralization			4.024*** (1.432)	4.597*** (1.742)	3.716** (1.426)	4.264** (1.743)
Native			0.313 (1.064)	0.893 (1.297)	0.520 (1.059)	0.849 (1.289)
Newspaper per 1000					-1.882** (0.952)	-1.703 (1.049)
Year fixed effects		Y	Y	Y	Y	Y
Province fixed effects		Y	Y	Y	Y	Y
Controls				Y		Y
Observations	173	173	173	171	173	171

Note: This table provides the results for two separate periods: 1995 to 2000 and 1998 to 2005. The results show that the effects are larger for the first period.