

Does Competition Encourage Corporate Profit Misreporting? Evidence from Chinese Firms*

Hongbin Cai [†] Qiao Liu [‡] Geng Xiao [§]

Abstract

This paper investigates whether market competition enhances firms' incentives to hide profits. We develop a theoretical model of firms' profit hiding behavior in competitive environments and derive several testable hypotheses. We then test the model using a database that covers more than 20,000 large-and-medium-sized industrial firms in China during the period 1995-2002. Our findings show that firms in more competitive market environments — as well as firms in relatively disadvantageous positions — hide a larger share of their profits. This suggests that policies intended to promote competition should be accompanied by policies aiming at strengthening institutional infrastructure and at leveling playing fields.

JEL Classification: L10, D21, H26, G30

Keywords: Competition, firm behavior, tax evasion, Chinese economy

*Part of the statistical analysis of this paper was implemented at the National Bureau of Statistics of China (NBS) under arrangements to maintain legal confidentiality requirements. We thank Junling Xing and Xiaoyun Yu at the NBS for technical support. The views expressed here are those of the authors and do not reflect official positions of the NBS. We thank Paul Devereux, Amar Hamoudi, Ginger Z. Jin, Jiang Luo, Jean-Laurent Rosenthal, Alan Siu, C. Y. Tse, Keith Wong, and seminar participants at HKUST for helpful comments and Zhengge Tu for excellent research assistance. Liu acknowledges support by a grant from the University Grants Committee of Hong Kong (Project No. AOE/H-05/99).

[†]Department of Economics, UCLA, Los Angeles, CA 90095-1477. Contact:(310)794-6495, cai@econ.ucla.edu.

[‡]School of Economics and Finance, University of Hong Kong, Pokfulam, Hong Kong. Contact: (852)2859-1059, qliu@hku.hk.

[§]School of Economics and Finance, University of Hong Kong, Pokfulam, Hong Kong.

1 Introduction

Traditionally in economics, competition is believed to improve productive efficiency of firms and increase social welfare.¹ However, it has become well recognized that in environments where the standard assumptions do not apply, competition may achieve neither of these ends. For example, Milgrom and Roberts (1992) argue that competition pressure exacerbated the moral hazard problems in the savings and loan (S&L) industry in the U.S. by forcing S&L executives to gamble on risky investments in order to survive. In a recent article, Shleifer (2004) argues that competition encourages the spread of a wide range of unethical behavior such as employment of child labor, corruption, excessive executive pay, and corporate earnings manipulation. These claims clearly illustrate that *the effects of competition critically depend on the instruments firms use in order to compete*. If firms use unethical or illegal, socially unproductive means to gain competitive advantage, then competition may not lead to socially desirable outcomes. While this is theoretically plausible, empirically there is no study that bears on the relationship between competition and unethical behavior.

In this paper, using a large dataset of Chinese large and medium sized industrial firms, we examine empirically how product market competition affects firms' tendency to hide profit. We focus on profit hiding for two reasons. First, as a way to reduce tax, profit hiding can potentially be used by firms as a cost-saving device to gain competitive advantage. However, profit hiding is likely to be socially unproductive.² Second, profit hiding is a common phenomenon around the world and causes serious economic problems in many economies, even though its severity is likely to vary across countries.³ We study corporate profit hiding of Chinese firms, because

¹This view has been expressed in the classic writings of Adam Smith (1976) and Hicks (1935), and many others. For more recent analysis of the effects of competition, see, e.g., Leibenstein(1966) and Machlup (1967). The available empirical evidence is weak, but in general supports the view that competition improves firm efficiency, see, for example, Porter (1990), Nickell (1996), and Fee and Hadlock (2000).

²Some forms of profit hiding can be legal (e.g., taking advantage of loopholes in tax laws), others are illegal (e.g., failing to report revenue). For our purpose, there is no need to distinguish them, since both are generally considered to be socially wasteful activities. If tax rates are excessively high, it might be possible that profit hiding reduces distortion. However, this caveat is not central to our analysis.

³For example, the U.S. Internal Revenue Service estimated that about 17% of income tax liability is not paid (Slemrod and Yitzhaki, 2000), and according to the U.S General Accounting Office, about "one-third of large U.S. corporations reported zero tax liability" from 1996-2000 (reported in Desai, 2005a). In China, the National Auditing Office uncovered 13.39 billion yuan (\$1.6 billion) in unpaid or underpaid tax in 2002, and 11.89 billion

(i) we have access to a comprehensive dataset of a large number of Chinese industrial firms, which allows us to assess the degree of profit hiding; (ii) there is broad variation in terms of both competitiveness and profit reporting practices in China; and (iii) recently the Chinese economy has become increasingly market-oriented, and taken on an increasingly important role in the world economy. Thus, lessons learned here are relevant to other parts of the world.

The dataset we use is maintained by the National Bureau of Statistics of China (NBS) and contains firm-level information based on the annual accounting briefing reports filed by all large- and medium-sized industrial firms in China from 1995 to 2002. We develop a novel approach to test how profit hiding is affected by competition intensity and firm characteristics. Our main empirical findings can be summarized as follows.

- Firms in more competitive industries tend to hide more profits, all else equal.
- Firms positioned unfavorably in competitive environments, such as firms facing higher corporate tax rates, firms facing more severe financing constraints, smaller firms, and private/collective firms, display stronger propensities to hide profits.

We also find that these results are robust to various measures of competition intensity, to different market (or industry) definitions, and to various choices of estimators and model specifications.

More specifically, we develop a simple model in which a representative firm with a certain amount of realized profit decides how much profit to report to the government—which determines its tax liability — and then invests the retained profit to strengthen its competitive position in the marketplace. A firm’s expected future profit depends on its investment and its competitors’ investments. The more competitive the industry is, the more future profit opportunities a firm will lose if its investment lags behind its competitors’. In equilibrium, firms under-report profits, and the equilibrium amount of profit reported by a firm is a linear function of its true profit. We show that all firms will hide more profits when the market becomes more competitive. This is because as the market becomes more competitive, firms lose more if their investments lag behind

yuan in 2003 based on a four-month, nationwide investigation of 788 companies selected at random in 17 provinces and cities (*The Asian Wall Street Journal*, A2, September 20, 2004). It is safe to say that these cases of uncovered tax evasion represent only a tiny fraction of the tax evasion in China. Using an innovative approach, Fisman and Wei (2004) find evidence of tariff evasion in China. Johnson, Kaufmann, McMillan and Woodruff (2000) conduct a cross-country comparison of the sizes of hidden “unofficial” economies, which can be considered as an extreme form of profit hiding and tax evasion.

their competitors'. Thus, firms hide more profit in order to have more funds available in order to protect their competitive positions. Our model also predicts that a firm will hide more profit when it faces a higher tax rate or tighter financial constraints. In such cases marginal returns from hiding profits are higher. Furthermore, we find that within an industry, firms in disadvantageous competitive positions (e.g., greater market entry barriers or unfavorable treatment in government procurements) have stronger propensities to hide profits than other firms.

A main challenge for our empirical analysis is that firms' true accounting profits are not observable. We overcome this difficulty by computing corporate profit based the national income account system — that is, by deducting intermediate inputs from gross output. This measure of corporate profit can legitimately differ from a firm's true accounting profit based on the General Accounting and Auditing Principles (GAAP) because of differences in the revenue and expense recognition rules of the two systems.⁴ However, since both measures of corporate profits reflect a firm's economic fundamentals, they should be positively correlated. We assume that the technical relationship between the two profit measures is not affected by the competitiveness of the market.⁵ Subject to this assumption, our theoretical predictions lead to testable hypotheses regarding the relationships between the variables of interest and the correlation between reported and imputed profits.

The theoretical predictions of our model are all confirmed by our empirical results. Specifically, we find strong evidence indicating that competitiveness in the market enhances firms' profit hiding behavior. The estimated effect on profit hiding has the predicted sign and is statistically significant for several measures of competitiveness (the number of firms, concentration, or industry average profit margin) and alternative definitions of industries (2-digit or 3-digit industry codes) and markets (national or regional). The competition effect is also economically significant. Based on our estimation of the baseline model, a representative firm in an industry that is one standard deviation *more competitive* than the average reports (up to) 18% *less* profit than an

⁴The GAAP accounting system was adopted and implemented in China before the beginning of our sample period 1995. The national income account system and the GAAP accounting system can legitimately differ for several reasons. For example, not all gross output in the current year necessarily converts into firm revenue in the same year. Asset depreciation rules can be different. They can also differ because of accounting accruals and earning management by firms.

⁵In Section 6, we conduct extensive robustness checks and provide strong evidence that this assumption holds in our context.

identical firm in the industry with the average level of competitiveness (Section 5.1). Our main empirical results are quite robust to alternative specifications. Overall, the evidence is strong that competition encourages profit hiding in our sample.

We also find that after controlling for other characteristics, firms facing higher tax rates or tighter financial constraints hide more profits. The estimated effects of these factors have the predicted signs and are statistically and economically significant. Based on one estimation of the baseline model, an increase of one standard deviation in tax rate reduces reported profits by about 10% relative to imputed profits (Section 5.2); and an increase of one standard deviation in our measure of accessibility to capital markets increases the share of profit which is hidden by about 2.6% (Section 5.3). Furthermore, after controlling for tax rates and financial constraints and other characteristics, firms that are competitively disadvantageous in other dimensions have a higher propensity to hide profits. In all cases estimated effects have the predicted signs and are statistically and economically significant. Based on the same estimation, an increase of one standard deviation in firm employment size increases the share of profit which is hidden by about 4.8% (Section 5.4); and private and collective firms report 18.5% *less* reported profits than other types of firms (Section 5.4). Although the magnitudes of these effects vary with the regression specification, their economic significance is consistently large.

Governments in developing and post-socialist countries are often advised to implement market oriented reforms aiming to promote competition. China has implemented such reforms extensively —loosening control over prices, giving more discretion to state firms, opening the market to foreign direct investment, and easing regulations to allow entry in most industries. To be clear, we are not trying to argue against competition or against policies promoting competition. Rather, we argue that promoting competition is not sufficient to obtain socially desirable outcomes. At the same time of promoting competition, it is important to improve the institutional infrastructure in the economy so that firms do not easily use socially wasteful instruments to gain competitive advantage. Moreover, policies that help equalize opportunities for all market participants should be important components of any reform strategy. While it might be useful to give preferential treatment to some firms (such as tax breaks for foreign invested firms), discrimination can be quite harmful over the long run. Such policy discrimination not only directly reduces the efficiency benefits of competition, but also causes long run deterioration of the institutional infrastructure of the economy because firms that are discriminated against try to

compensate their inherent disadvantages with illegal and socially wasteful means.

Our empirical approach is similar to that of Desai (2003, 2005a,b). In these studies, Desai examines the discrepancy in U.S. corporations' book income and tax income. For a given (unobservable) true profit, a corporation can legitimately report different book income and tax income due to the different requirements of the two reporting systems. However, after controlling for accounting accruals and earning management activity, Desai shows that there is still an ever-widening divergence between book income and tax income, which he attributes to firms' growing tax sheltering activities. Like Desai (2003, 2005a, b), we also use the relationship between firms' reports to two different systems (in our case the tax system and the National Income Account system) to infer possible tax avoidance. We then investigate how competition affects firms' incentives to misreport.

Following Becker's (1957) classical study of discrimination, Shleifer (2004) argues that if firms treat honesty as a normal good, then their demand for honesty will be lower in more competitive environments since competition reduces profits. Our theoretical model generates very similar conclusions to Shleifer's. However, our analysis is different in that we do not rely on income effects. We explicitly model firms' strategic use of profit hiding in competitive environments and derive a rich set of implications on how profit hiding is affected by competitiveness and firm characteristics from the equilibrium analysis of the model. Of course, we do not argue that income effects are irrelevant. In another related paper, Cummins and Nyman (2004) illustrate a different dark side of competition — competition makes firms (e.g., investment bankers) reluctant to act on private information that is unpopular with consumers, resulting in socially under-use of valuable information (see also Harris, 1998). Our paper is also related to the literature that studies the effects of product market competition on managerial incentives and corporate performance, e.g., Hart (1983), Nalebuff and Stiglitz (1983), Scharfstein (1988), Hermalin (1992), and Schmidt (1997). While Hart (1983) and Nalebuff and Stiglitz (1983) show that competition strengthens managerial incentives to maximize firm value in standard moral hazard models, Scharfstein (1988), Hermalin (1992) and Schmidt (1997) demonstrate that there can be countervailing effects so that the net effect of competition on efficiency is ambiguous. The focus of our paper is different, i.e., on the effect of competition on firms' incentives to engage in profit hiding. Moreover, while these papers are all theoretical, the main contribution of our paper is empirical.

The rest of the paper proceeds as follows. Section 2 presents the theoretical model. Section

3 then discusses our empirical methodology and develops empirical hypotheses. We describe the dataset and our empirical strategy in Section 4, and then present the empirical results in Section 5. Section 6 examines robustness issues of our empirical approach. Concluding remarks are in Section 7.

2 A Theoretical Model

In market j there are $n + m$ firms, of which n firms are “competitively advantageous” and $m \geq n$ are “competitively disadvantageous” (to be defined below). To allow for firm heterogeneity while keeping the analysis tractable, we suppose that all n competitively advantageous firms are identical to each other, as are the m competitively disadvantageous firms. At the end of any given year t , firm i has a realized profit of $\pi_{i,t}$. It faces a tax rate of $\tau_{i,t}$, and chooses to report a profit of $\hat{\pi}_{i,t}$, resulting in an after-tax profit of $\pi_{i,t} - \tau_{i,t}\hat{\pi}_{i,t}$. Misreporting profit is costly to firms (otherwise they would always report zero profit), because (i) they have to invest resources (e.g., hiring additional accountants) and time to take advantage of loopholes in tax laws; and (ii) they may have to change their accounting and business practices to hide profit; and (iii) they face financial penalties and legal punishments if caught by government auditing. For simplicity, we suppose that the cost of hiding profit is a quadratic function of the amount of profit hidden, that is, $C = 0.5\gamma(\pi_{i,t} - \hat{\pi}_{i,t})^2$, where γ is a positive parameter.⁶

For simplicity we assume that firms re-invest all their retained profits.⁷ Thus, at the beginning of year $t + 1$, firm i ’s available resources are given by

$$k_{i,t+1} = \pi_{i,t} - \tau_{i,t}\hat{\pi}_{i,t}$$

Firm i invests $k_{i,t+1}$ to compete with other firms in the market, trying to maintain and expand its market share. Such investments can take many forms, such as R&D, advertising and other marketing expenses, discounts and promotions, or expenses to build relationships with clients

⁶By assuming that γ is exogenous, we assume that tax enforcement is exogenous, and in particular, is independent of competitiveness of the industry a firm is in. This seems reasonable in our context, because tax authorities in China are not as experienced and sophisticated as their counterparts in developed economies. Our empirical findings can potentially be useful in identifying types of firms as more likely suspects of profit hiding (e.g., those in more competitive industries), and thus may help tax authorities improve their auditing strategies in the future.

⁷This assumption can be relaxed. In particular, our analysis will be almost unchanged if firms pay out a proportion of their retained profits as dividends.

and government officials. Firm i 's expected future profit depends on its own investment and those of its competitors as follows:⁸

$$\pi_{i,t+1} = f(k_{i,t+1}, k_{-i,t+1})$$

where $k_{-i,t+1}$ is the vector of investments by all firms other than i , and f is the firm i 's expected future earnings. Naturally, f is increasing in $k_{i,t+1}$ and decreasing in $k_{-i,t+1}$. For simplicity, we assume f takes the following form:

$$f(k_{i,t+1}, k_{-i,t+1}) = a_{i,t} + b_{i,t}k_{i,t+1} - 0.5c_{i,t}k_{i,t+1}^2 - g(k_{-i,t+1}) - 0.5\mu_j(\bar{k}_{-i,t+1} - k_{i,t+1})^2$$

where $a_{i,t}$, $b_{i,t}$, $c_{i,t}$, and μ_j are all positive parameters, g is an increasing function, and $\bar{k}_{-i,t+1} = \sum_{l \neq i} k_{l,t+1}$ is the aggregate investment of firm i 's competitors. The parameter $a_{i,t}$ represents firm i 's expected future profit that can be achieved without any additional investment. The parameter $b_{i,t}$ represents the base marginal product of firm i 's retained profit. The parameter $c_{i,t}$ represents how fast the marginal product of retained profit decreases as firm i 's available funds increase. The last quadratic term captures the idea that in a competitive environment, the further a firm's investment lags behind that of its competitors, the further behind it falls in terms of market share. In this formulation, the parameter μ_j is a measure of the competitiveness of market j : the larger μ_j , the more market share firm i loses when it lags behind its competitors. The number of firms in the industry, $n + m$, represents another measure of competitiveness: the more firms, the more competitors firm i faces. In the above formulation, this implies that $\bar{k}_{-i,t+1}$ will be larger, thus the more market share firm i loses when it lags behind its competitors. Therefore, the last term is the "competition effect."

Aside from this competition effect, the overall marginal product of retained profit for firm i is $b_{i,t} - c_{i,t}k_{i,t+1}$. For firms that have better access to the capital market and hence are less constrained by liquidity, we expect that their marginal returns of retained profits should be lower, i.e., they should have smaller $b_{i,t}$ and larger $c_{i,t}$. It is well established in development economics that access to credit is very important for firm performance and economic growth in developing countries (e.g, Rajan and Zingales, 1998, Banerjee and Duflo, 2004). China has experienced rapid

⁸Here the firm's next period profit should be interpreted as its whole future profit stream with proper discounting. One way to think about this is that we are studying a stationary equilibrium of a dynamic game in which firms make profit hiding decisions over time.

economic growth for more than two decades, but still has a very ineffective banking sector and an ill-functioning stock market. Thus, accessibility to the credit market constitutes an important competitive advantage in China.

In our model, we suppose that accessibility to the capital market is the main factor that differentiates competitively advantageous and disadvantageous firms. Specifically, all n competitively advantageous firms have the same parameters $\{\tau_t^s, a_t^s, b_t^s, c_t^s, \pi_t^s\}$ and all m competitively disadvantageous firms have the same parameters $\{\tau_t^w, a_t^w, b_t^w, c_t^w, \pi_t^w\}$. *Compared with competitively disadvantageous firms, competitively advantageous firms have smaller marginal returns of retained profits, i.e., $b_t^s < b_t^w$ and $c_t^s > c_t^w$.* Their advantages come primarily from better access to the capital market, but also from favorable initial conditions in other respects. For example, competitively advantageous firms may have better market entry conditions (so investments needed to enter markets are smaller and less important), or they may receive better treatment by regulators and government, including for example better protection of property and contractual rights (so that expenses needed to build relationships with government agencies are smaller and less important). Because of these advantages, competitively advantageous firms are likely to have higher profit than competitively disadvantageous firms (i.e., $\pi_t^s \geq \pi_t^w$) and to be less threatened by competition than competitively disadvantageous firms (i.e., $a_s \geq a_w$). However, these differences are not essential to our analysis. In a broad sense, competitively advantageous firms in China likely have better tax treatments and face lower tax rates than competitively disadvantageous firms, i.e., $\tau_t^s \leq \tau_t^w$. However, since we can separate out and control for the effects of tax rates, we will focus on the other advantages of competitively advantageous firms.

Firm i chooses an optimal level of profit to report, $\hat{\pi}_{i,t}$, to maximize its expected total payoff

$$\max U_i = \pi_{i,t+1} - C = f(k_{i,t+1}, k_{-i,t+1}) - 0.5\gamma(\pi_{i,t} - \hat{\pi}_{i,t})^2$$

From the first order conditions, we get

$$\left[\gamma + (\mu_j + c_{i,t})\tau_{i,t}^2 \right] \hat{\pi}_{i,t} = [\gamma + (\mu_j + c_{i,t})\tau_{i,t}] \pi_{i,t} - \tau_{i,t} b_{i,t} - \mu_j \tau_{i,t} \bar{k}_{-i,t+1}$$

We focus on the symmetric equilibrium of the model. Specifically, all n competitively advantageous firms choose the same $\hat{\pi}_t^s$ and all m competitively disadvantageous firms choose the same $\hat{\pi}_t^w$. Then the first order conditions can be rewritten as

$$\left[\gamma + c_t^s (\tau_t^s)^2 - \mu_j (\tau_t^s)^2 (n - 2) \right] \hat{\pi}_t^s = [\gamma + c_t^s \tau_t^s - \mu_j \tau_t^s (n - 2)] \pi_t^s - \tau_t^s b_t^s - \mu_j \tau_t^s m (\pi_t^w - \tau_t^w \hat{\pi}_t^w) \quad (1)$$

$$\left[\gamma + c_t^w (\tau_t^w)^2 - \mu_j (\tau_t^w)^2 (m - 2) \right] \hat{\pi}_t^w = [\gamma + c_t^w \tau_t^w - \mu_j \tau_t^w (m - 2)] \pi_t^w - \tau_t^w b_t^w - \mu_j \tau_t^w n (\pi_t^s - \tau_t^s \hat{\pi}_t^s) \quad (2)$$

Solving these equations we get

$$\hat{\pi}_t^s = d_t^s \pi_t^s + e_t^s \quad \text{and} \quad \hat{\pi}_t^w = d_t^w \pi_t^w + e_t^w \quad (3)$$

where for $i = s, w$, d_t^i and e_t^i are functions of $\{\mu_j, n, m, \tau_t^i, b_t^i, c_t^i\}$. Thus, the amount of profit a firm reports is a linear function of its true accounting profit. If $d_t^i = 1$ and $e_t^i = 0$ for $i = s, w$, then all firms report *truthfully*. In general firms under-report profits, that is, $d_t^i < 1$ and $e_t^i \leq 0$ for $i = s, w$ (see Equations 11 – 14 in the Appendix). Clearly, *the amount of hidden profit*, $\pi_t^i - \hat{\pi}_t^i$, is decreasing in d_t^i and e_t^i . In other words, when d_t^i and e_t^i are greater, firm i hides less profit. For reasons that will become clear later, we focus on the comparative statics of d_t^s and d_t^w . These parameters measure the degree of profit hiding on the margin (i.e., firms hide $1 - d_t$ yuan of profit more if the true profit increases by one yuan). Henceforth, the phrase “profit hiding” will be used to refer to profit hiding at the margin.

We can prove the following results (all proofs are in the Appendix).

Proposition 1 *For $i = s, w$, d_t^i is decreasing in μ_j , n , and m . Thus, all else equal, profit hiding is increasing in the degree of competitiveness of the market.*

In our model, a greater μ_j or the total number of firms means a higher degree of competitiveness in industry j . Proposition 1 shows that as the market becomes more competitive, all firms report a smaller share of their profits. Intuitively, we would expect that as competition heats up, each firm would hide more profit in order to avoid losing market share. Besides this direct effect, there is also a feedback effect. When all other firms hide more profits and compete more aggressively, each firm has additional incentives to hide profit so as to meet the challenges presented by competition.⁹

⁹Technically, this game features strategic complementarities: the marginal benefit of profit hiding for one firm is increasing in the amount of profit hiding by its competitors.

For the next result, we assume that n and m are sufficiently large so that $(n - 2) \mu_j \geq c_t^s$ and $(m - 2) \mu_j \geq c_t^w$.

Proposition 2 *For $i = s, w$, d_t^i is decreasing in τ_t^i and increasing in c_t^i . Thus, all else equal, profit hiding is increasing in tax rates and in the marginal returns of retained profits.*

Proposition 2 shows that firms will report less profits if the tax rates they face are higher or if the marginal returns of their retained profits are higher. With higher tax rates, one yuan of hidden profit saves more taxes, hence profit hiding is more profitable. With higher marginal returns of retained profits, one yuan of saved tax will generate more future profit. In either case, firms will tend to hide more profits.

Proposition 3 *Suppose $\tau_t^s = \tau_t^w$. Then $d_t^s \geq d_t^w$. That is, all else equal, competitively advantageous firms hide less profits than competitively disadvantageous firms.*

Proposition 3 says that a one yuan increase of true profit leads to more reported profit from a competitively advantageous firm than from a competitively disadvantageous firm.¹⁰ The reason is that with better access to the capital market, competitively advantageous firms are able to compete more aggressively without relying heavily on (costly) profit hiding. Note that Propositions 2 and 3 have different empirical implications. Proposition 2 is about comparative statics with regards to firms' own characteristics, while Proposition 3 focuses on firms' relative positions in the industry.

3 Empirical Methodology and Testable Hypotheses

If all variables were observable, then a straightforward test of our model would be simply to estimate equation (3) as follows:

$$\hat{\pi}_{i,t} = d_{i,t} \pi_{i,t} + e_{i,t} + \epsilon_{i,t} \tag{4}$$

¹⁰The reversed causality, i.e., that firms become competitively advantageous because they pay more taxes and thus are better treated by governments, is not likely. Government agencies at different levels (from central to provincial to city) that affect firms' competitive positions (e.g., procurement, regulations) are separate from tax collection agencies. It is unlikely they coordinate to reward firms who pay more taxes.

with some appropriately chosen functional forms for $d_{i,t}(\mu_j, n, m, \tau_t^i, b_t^i, c_t^i)$ and $e_{i,t}(\mu_j, n, m, \tau_t^i, b_t^i, c_t^i)$, and with the standard assumption that $\epsilon_{i,t}$ is uncorrelated with either $d_{i,t}$ or $e_{i,t}$. The estimation results would then allow us to directly test the comparative statics results of Propositions 1-3. However, the main challenge for our empirical analysis is that firms' true profits $\pi_{i,t}$ are not observable. To overcome this difficulty, we adopt the following approach.

Using the NBS database, which we will detail in Section 4, we compute firm i 's corporate profit $PRO_{i,t}$ in year t according to the national income accounting system as follows:

$$PRO_{i,t} = Y_{i,t} - MED_{i,t} - FC_{i,t} - WAGE_{i,t} - CURRD_{i,t} \quad (5)$$

where $Y_{i,t}$ is the firm's gross output; $MED_{i,t}$ measures its intermediate inputs excluding financial charges; $FC_{i,t}$ is its financial charges (mainly interest payments); $WAGE_{i,t}$ is the firm's total wage bill; and $CURRD_{i,t}$ is the amount of current depreciation.

The variable $PRO_{i,t}$ defined here is not a firm's *true accounting profit* $\pi_{i,t}$, the profit calculated truthfully according to the general accounting principles. They differ by the timing of when revenues and expenses are recognized into income, because outputs and costs in the current year do not necessarily convert into revenues and expenses in the same year. Differences inherent in revenue and expense recognition rules between the two systems account for discrepancies between $\pi_{i,t}$ and $PRO_{i,t}$. Given the exogenous differences in the rules of the two systems, we suppose that with a linear approximation, $PRO_{i,t}$ and $\pi_{i,t}$ are related in the following way:

$$\pi_{i,t} = \delta_{i,t} PRO_{i,t} + \eta_{i,t} \quad (6)$$

where $\delta_{i,t}$ and $\eta_{i,t}$ are firm-specific parameters that depend on firm production cycles, demand seasonal shocks, equipment life cycles, firm locations (land valuation fluctuations), etc. We assume that $\delta_{i,t} > 0$, since one expects that firms' earning fundamentals generally move $PRO_{i,t}$ and $\pi_{i,t}$ in the same directions and thus cause them to be positively correlated. On the other hand, the sign of $\eta_{i,t}$ is hard to determine *a priori*.

Since we do not observe $\delta_{i,t}$ and $\eta_{i,t}$, we still do not know $\pi_{i,t}$. However, by substituting equation (6) into equation (4), we derive

$$RPRO_{i,t} = D_{i,t} PRO_{i,t} + E_{i,t} + \epsilon_{i,t} \quad (7)$$

where $RPRO_{i,t}$ replaces $\hat{\pi}_{i,t}$ as the amount of profit reported by firm i in year t , $D_{i,t} = d_{i,t}\delta_{i,t}$ and $E_{i,t} = d_{i,t}\eta_{i,t} + \epsilon_{i,t}$. Using linear approximations, we propose the following econometric specification for $D_{i,j,t}$ (j denotes the industry of firm i):

$$D_{i,j,t} = \beta_0 + \beta_1 * Compet_{j,t} + \beta_2 * Tax_{i,t} + \beta_3 * Finance_{i,t} + \beta_4 * Position_{i,t} + \beta_5 * X_{i,t} + \epsilon_{i,t} \quad (8)$$

where $Compet_{j,t}$ is a variable that captures the level of competitiveness in industry j (corresponding to μ_j , n , and m in the model) ; $Tax_{i,t}$ (corresponding to $\tau_{i,t}$ in the model) is firm i 's tax rate in year t ; $Finance_{i,t}$ (corresponding to $c_{i,t}$ in the model) is a measure of how easily firm i can access the capital market; $Position_{i,t}$ is a set of variables that proxy for firm i 's relative competitive position in the marketplace; and $X_{i,t}$ is a set of control variables that includes firm characteristics, time fixed effect, and location fixed effects.

In order to test Propositions 1-3, we need to make the following assumption.

Assumption 1 $\delta_{i,t}$ is not affected by industry competition intensity.

Assumption (1) implies that if competition affects $D_{i,t}$, it does so only through $d_{i,t}$, but not $\delta_{i,t}$. This assumption is critical to our empirical strategy. *A priori*, one may wonder whether competition may affect $\delta_{i,t}$ by affecting the timing of revenues (expenses) recognition. Specifically, it might be possible that firms in a more competitive industry would face greater difficulties in selling their products and converting $PRO_{i,t}$ into $\pi_{i,t}$, which leads to a technical correlation between $\delta_{i,t}$ and competition. As a basic control for this concern, we include in all our regressions a variable called $RSALE_{i,t}$, which is defined as the ratio of firm i 's sales to its total output ($Y_{i,t}$) in year t . Since $RSALE_{i,t}$ measures how effectively firm i converts final outputs into revenues, effects of competition on $D_{i,t}$ after controlling for $RSALE_{i,t}$ should be attributed to $d_{i,t}$, not $\delta_{i,t}$.

Moreover, we investigate whether Assumption (1) holds in our context later in Section 6. We will provide evidence that in our dataset, all major factors responsible for discrepancies between $PRO_{i,t}$ and $\pi_{i,t}$, such as changes in inventories, and current liabilities, and depreciation, are not correlated with the competition intensity of the industry firm i operates in. Controlling for the impact of these variables does not change our baseline results either. Moreover, we will show that our main empirical results still hold after controlling for PRO in previous year. These robustness checks indicate that the parameters $\delta_{i,t}$ and $\eta_{i,t}$ capturing the technical relationship between $\pi_{i,t}$ and $PRO_{i,t}$ are unlikely to be affected by competition.

Under Assumption (1), $d_{i,t}$ is decreasing in competition intensity if and only if β_3 is negative. Therefore, Proposition 1 leads to the following testable hypothesis.

Hypothesis 1 $\beta_1 < 0$, *i.e.*, a firm's incentives to hide profit are positively correlated with the degree of product market competitiveness.

Since equation (6) represents the technical relationship between profit measures from the two different accounting systems, we expect that $\delta_{i,t}$ is unaffected by firms' tax rates, their access to credit market, or their relative positions in the industry.¹¹ This is because Equation (6) is about a firm's true profit, profit that is calculated strictly according to the accounting rules of the two systems without strategic manipulation. The reported profits contain possible manipulations, which our model predicts will respond to competition and firm characteristics in a systematic way. Therefore, we have the following hypotheses from Propositions 2-3:

Hypothesis 2 $\beta_2 < 0$, *i.e.*, a firm's incentives to hide profits are positively correlated with its tax rate.

Hypothesis 3 *Firms' incentives to hide profits are negatively correlated with its accessibility to capital market.*

For proxies of competitively disadvantageous firm characteristics $Position_{i,t}$, Proposition 3 implies that their effects on profit hiding are positive, and hence the coefficient estimators in the profit reporting equation should be negative. For such proxies, we have the following hypothesis:

Hypothesis 4 *Firms with disadvantageous market positions tend to have stronger incentives to hide profits.*

In the empirical implementation, we have a specification for $E_{i,t}$ similar to that for $D_{i,t}$ as in Equation (8). However, since we cannot determine the sign of $\eta_{i,t}$ in Equation (6) *a priori*, and since $E_{i,t} = d_{i,t}\eta_{i,t} + e_{i,t}$, our model has no prediction about how $E_{i,t}$ will be affected by competition or other variables. Thus, we do not have predictions about the signs of the coefficients in the estimation of $E_{i,t}$.

¹¹Recall that we include the ratio of sales to total output $RSALE_{i,t}$ in all our regressions. After controlling for $RSALE_{i,t}$, $\delta_{i,t}$ should be very unlikely to be affected by those variables.

4 Data and Variable Definitions

4.1 Dataset

Our main data source is the NBS database, which is compiled based on the annual accounting briefing reports filed by all large and medium sized Chinese industrial firms with the NBS during the period from 1995 to 2002. The Data Appendix details how this database was created, structured, and cleaned.

The NBS database covers more than 20,000 firms in 37 two-digit manufacturing industries, from 28 provinces or province-equivalent municipal cities.¹² As shown in Table A1, over the sample period, the total value added for all of our sample firms ranges from RMB 958 billion to 2013 billion, which account for 33.3% to 43.3% of the total industrial value added in China and 14.4% to 19.2% of China's GDP. Our sample firms hired between 26 and 38 million employees during 1995-2002, or about 10% to 20% of total urban employment over the sample years.

The NBS assigns each firm covered by this database a unique legal identification number. A firm may leave, enter, or re-enter the database when its operation scale fluctuates around the classification criterion of large and medium sized firms set by the NBS. However, we are unable to track an individual firm if it leaves or re-enters the sample. For example, if a firm covered in 1995 did not appear in 1996, it could have gone bankrupt, or been acquired by another firm, or reclassified as a small firm, or simply changed its ownership (e.g., privatization). This implies that we have an unbalanced panel dataset.¹³

4.2 Profit Measures

The NBS requires all above scale firms in China to report their accounting data on an annual basis. The NBS database contains the pre-tax accounting profit reported by each firm, which gives the dependent variable in our regressions, *RPRO*. The pre-tax accounting profits in our dataset should be the same as those reported to the tax authority, since cross-checking is trivial and is regularly carried out. Our dataset is used by the NBS to calculate the Gross Domestic

¹²We combine the three adjacent provinces QingHai, NingXia, and Tibet, because of their economic similarities and the small sizes of their economies. ChongQing is included as a part of SiChuan province, since it was separated from SiChuan only recently.

¹³Focusing on the subsample of observations with all 8 year data does not qualitatively change our basic empirical results. However, it reduces the sample size substantially.

Product, and contains inputs and outputs information for all the sample firms. This allows us to compute PRO , profit from the national income account system, as in equation (5). The inputs and outputs information in our dataset should be quite accurate, because the Statistical Bureau has implemented standard procedures in calculating the national income account since 1995 and has strict double checking procedures for large and medium sized firms.¹⁴

In our analysis, we scale both PRO and $RPRO$ with firms' total assets (TA). After the scaling, the sample mean of $RPRO$ is 0.003 and that of PRO is 0.042 (Table 2). On average, PRO is almost 12 times $RPRO$. However, as we argued before, the difference between the two in itself is not evidence of profit hiding. It could simply reflect the exogenous differences between the accounting system and the national income account system, as postulated in equation (6).

4.3 Competition Variables, Industry and Market Definitions

Following the standard practice in the Industrial Organization literature, we construct four variables to measure competition intensity in product markets. The first variable is the number of above-scale firms operating in an industry, N , which is collected from the China Statistical Year Books, 1995 - 2002. We use its natural logarithm, $LOGN$, in our analysis. The variable $LOGN$ corresponds to $n + m$ in our model, and correlates positively with competitiveness.

The second measure is the industry Herfindal index, $H - Index$, which is the sum of squares of the market shares (by sales) by the ten largest firms in an given industry.¹⁵ As another way to measure concentration, we also compute the total market share accounted for by the four largest firms in an industry (by sales) and name it $CONCEN$. Both $H - Index$ and $CONCEN$ are negatively correlated with competitiveness.

A fourth measure of competitiveness we use is the industry average profit margin, $PMARGIN$, which is the ratio of total pre-tax profit to total sales in an industry. As competitiveness increases,

¹⁴Moreover, firms do not have clear incentives to misreport their inputs and outputs information, because the imputed profits from the national income account can legitimately differ from true accounting profits and cannot be used by governments to directly detect profit hiding. Misreporting of statistical data was commonly suspected for some time in China, but the most likely misreported information was provincial GDP growth data provided by provincial governments. National income accounting of large and medium firms is done by the central Statistical Bureau, hence is less subject to misreporting. In addition, such misreporting was mostly over-reporting, which would bias against our results.

¹⁵As a robustness check, we also calculate the Herfindal index by total assets, which yields similar results.

one may expect that firm profit on average will fall, thus *PMARGIN* should be negatively correlated with competitiveness. As a measure of competitiveness, *PMARGIN* is probably more controversial than the others. We include it in our analysis to show that our main empirical results are robust to different measures of competitiveness.

Table 1 presents these competition measures for the thirty-seven two-digit manufacturing industries in China averaged over 1995-2002. All measures show substantial variations across industries. We report the bivariate correlations among the four competitiveness measures in Table A2. As measures of industry concentration, not surprisingly, *H-Index* and *CONCEN* are highly correlated; the correlation coefficient is 0.847. *LOGN* negatively correlates with both *H-Index* and *CONCEN*; the coefficients are -0.544 and -0.624 , respectively. *PMARGIN* negatively correlates with *LOGN* (correlation coefficient of -0.233), and positively correlates with both *H-Index* and *CONCEN* but the correlation coefficients are quite small (0.053 and 0.090 , respectively). Overall, it appears that the four variables measure the degree of competitiveness quite consistently, yet offer somewhat differentiated perspectives.

Besides the measurement of competitiveness, specifying the appropriate scope of a product market is crucial in gauging the competitive pressure firms face. In addition to the two-digit industry codes, we also define an industry according to the three-digit industry codes specified by the NBS. There are in total 138 three-digit industries in the manufacturing sectors in China. Table A3 lists all of them. We calculate all competition measures except *LOGN* for every three-digit industry, because we do not have the information on the number of above scale firms at the three-digit industry level.

In defining a product market at either two-digit or three-digit industry level, one assumes that firms operate in the national market. Clearly this does not hold for all firms. In particular, regional protectionism commonly practiced in China may limit a firm's reach in the national marketplace. One remedy might be to specify a product market as a two or three-digit industry confined to one single province or province-equivalent municipal city, which assumes that these firms only operate locally. However, given their size and importance, the large and medium sized industrial firms in our data are more likely to operate beyond the province level. Therefore, we define a product market as a two-digit industry in one of China's eight economic regions specified by the State Council of China (see Table A4 for details). We compute *H-Index*, *CONCEN*, and *PMARGIN* based on this market definition.

We report the summary statistics of the competition variables based on the three different market definitions in Table 2. The two concentration measures, *H – Index* and *CONCEN*, are very sensitive to market definition. As the market becomes smaller (from 2-digit industry to 3-digit industry, to 2-digit industry/region), as one would expect, concentration increases. On the other hand, *PMARGIN* is stable with respect to the first two market definitions, but is quite different when the 2-digit industry/region definition is used. This is likely due to the great disparity in regional development in China.

4.4 Other Independent Variables

From our dataset, we construct the following variables of firm characteristics and report their summary statistics in Table 2.

We construct a variable *TAX* from the ratio of actual corporate income tax paid by a firm to its reported pre-tax profit, and set it to zero for loss-making firms. Although the standard corporate income tax rate is 33% in China, the Chinese government gives various preferential tax treatments (e.g., tax reduction for a certain period of time) to various kinds of firms (e.g., foreign firms, high-tech firms, joint ventures) over the sample period. Local governments also grant tax holidays and rebates to various types of companies in order to promote local economic development. Furthermore, tax collection and enforcement is quite discretionary, leaving large room for distortion and bribery in exchange of tax reduction. As a result, there is substantial variation in the effective tax rates across firms. From Table 2, the variable *TAX* has a sample mean of 18.2% with a standard deviation of 26.8%. Note that a small fraction of firms have very high effective tax rates (the sample max is 87.2%), partly because they are seriously mistreated and partly because there are tax carryovers from previous years.

One may wonder whether including *TAX* in the regressions causes endogeneity problems as *TAX* is the ratio of actual taxes paid and the dependent variable: the reported profits. However, this worry is not warranted, because this calculation really is an identity equation: taxes collected equals tax rates times reported profits. For a fixed tax rate, the error terms in the reported profits enter proportionally into the taxes paid, thus not affecting *TAX*.

It is difficult to measure a firm’s access to credit markets. However, as the Chinese economy has been growing at a very fast pace, Chinese firms’ demand for credit has grown, but the banking sector and stock market have not developed quickly enough to keep pace with this growing

demand. Thus, the actual amount of debt a firm has reflects mostly how much it manages to borrow, not its endogenously chosen optimal capital structure. Consequently, we expect firms' access to credit and their debt to equity ratios to be positively correlated. Note that banks in China have little discretion over interest rates they charge borrowing firms. Therefore, interest payments on loans reflect how much a firm is able to borrow. Therefore, we compute the ratio of total financial charges to total assets for each firm, which is called *FINANCE*, and use it as a proxy for the firm's access to credit markets.¹⁶ From Table 2, *FINANCE* has a sample mean of 2.46% and a standard deviation of 2.32%.

We use two proxies for firms' relative competitive positions. One proxy is firm size measured by the logarithm of the number of employees, *LNLABOR*.¹⁷ Large firms are likely to be the "competitively advantageous" ones in our model for several reasons. First, large firms have more resources to compete in the marketplace, and hence they should rely less on retained profits and have smaller marginal returns from retained profits (i.e., smaller $b_{i,t}$ and larger $c_{i,t}$ in the model). Secondly, large firms have better access to the capital market, which is an important competitive advantage given China's poorly developed financial markets. To the extent that the variable *FINANCE* does not perfectly capture a firm's accessibility to the capital market, firm size may reflect some of the effect of financial constraints. Thirdly, large firms may also get better tax treatments, and thus firm size may pick up some tax effects that are not fully captured by the variable *TAX*. Finally, large firms in China usually enjoy better protections of property and contractual rights, better regulatory treatments (such as in international trade), and have lesser market entry restrictions. The variable *LNLABOR*'s sample mean is 6.572 (corresponding to 715 employees) with a standard deviation of 1.06.

Another proxy for relative competitive positions we use is firm ownership, whereby we create a dummy variable *OWN* that takes the value of 1 if a firm is either private or collective and 0 otherwise. Over our sample period, private and collective firms accounted for about 13.1% of

¹⁶Without tight financial constraints and fixed interest rates, higher *FINANCE* may indicate *lower* future borrowing ability because banks are reluctant to lend to firms with high levels of debts and will charge higher interest rates on firms with lower credit ratings. As further evidence supporting our use of *FINANCE* as a measure of a firm's borrowing capability in our setting, we compute the autocorrelation of *FINANCE* for each firm. We find that their mean is significantly positive, suggesting that firms with high *FINANCE* in the current year are more likely to borrow large amounts in the future.

¹⁷Using the logarithm of total assets as a measure of firm size yields similar results.

the sample (the rest consisted of state owned enterprises, mixed ownership firms, foreign firms, and Hong Kong/Taiwan firms). For obvious reasons, private and collective firms in China are at a disadvantage. Since private and collective firms in China typically do not have access to the state banking system and are often subject to higher tax rates, indicators of private and collective ownership will pick up some of the tax and financial constraint effects on profit hiding which are not fully captured by the variables TAX and $FINANCE$. Moreover, private and collective firms in China have long been subject to insecure property rights, have been discriminated against by various government policies and regulations, and have faced much higher hurdles in entering new markets. For them, marginal returns of retained profits should be quite high.

As mentioned before, we include $RSALE$, the ratio of sales to total output, in our regressions. Its sample mean is 0.991 and standard deviation is 0.524 (Table 2). Also in Table 2, one can see that on average, our sample firm has total assets (TA) of 341 million yuan. Finally, we create twenty-eight location dummies, and eight year dummies to capture the geographical and time-varying effects.

5 Main Empirical Results

Based on Equations (7) and (8), we estimate the following regression model:¹⁸

$$\begin{aligned}
RPRO_{i,t} = & (\beta_0 + \beta_1 Compet + \beta_2 TAX + \beta_3 FINANCE + \beta_4 LNLABOR + \beta_5 OWN \\
& + \beta_6 RSALE + \sum_{i=year} \beta_i d^{year} + \sum_{j=loc} \beta_j d^{loc}) * PRO_{i,t} + \alpha_1 Compet + \alpha_2 TAX \\
& + \alpha_3 FINANCE + \alpha_4 LNLABOR + \alpha_5 OWN + \alpha_6 RSALE + \sum_{i=year} \alpha_i d^{year} \\
& + \sum_{j=loc} \alpha_j d^{loc} + \epsilon_{i,t}
\end{aligned} \tag{9}$$

where d^{year} is a set of year dummies, and d^{loc} is a set of location dummies. Thus, our estimation controls for time-specific, and location-specific effects on profit reporting behavior in our sample. $RSALE$ controls for any shifts in the technical relationship between PRO and the true accounting profit π .

¹⁸Since we include OWN and location dummies in the constant term, a firm fixed effect model is not applicable here. However, when we exclude OWN and location dummies and add firm fixed effect in the constant term, we find qualitatively similar results.

We estimate equation (9) using four alternative measures of competitiveness of each of the three alternative definitions of the market. *In the above specifications, the signs and magnitudes of the coefficients of the interactive terms (i.e., β 's) indicate whether and to what extent firms misreport their profits.*¹⁹ We report the regression results in Table 3. Columns 1-4 report results when the market is defined by two-digit industry; Columns 5-7, by 3-digit industry; and Columns 8-10, by 2-digit industry and region. In each column, the heading identifies the measure of competitiveness which is used in the regression. To save space, only the estimated coefficients of interest are reported.

One may have the following concern about our empirical specification (9). Since profits should be lower in more competitive industries, thus the estimated coefficients of β 's may simply capture such a relationship perhaps with some non-linearity. This concern is addressed implicitly in two ways. First, if competition affects firm profits but does not affect profit-reporting by firms, then both *RPRO* and *PRO* should vary with competition but *their relationship does not depend on competition*, namely, the estimated coefficients of β 's should be insignificant. Second, in our empirical specification (9), the effect of competition on firm profits is directly captured by the coefficient on *Compet*, α_1 .

We report t-statistics in brackets, computed from robust standard errors. While computing the robust standard errors, we also control for the clustering at the industry (or industry/region) level to account for the common random shocks within an industry (or an industry/region), as suggested in Bertrand, Duflo, and Mullainathan (2004).²⁰

¹⁹One issue arises when reported/actual profits both are in the negative territory, in which case higher sensitivity between reported profits and actual profits may imply over-reporting rather than under-reporting. We address this issue in two ways. First, we note that both *RPRO* and *PRO* in our sample have positive means (Table 2), and most of industries in most of years report positive average profits. In fact, only about 16% of our sample observations have both negative *PRO* and *RPRO*. Thus, the influence of such observations is limited. Secondly, we have also run these regressions in the sub-sample where both reported profits and actual profits are positive. This experiment, although potentially introducing a selection bias problem, does not change our results. This indicates that our results are not driven by the potential "over-reporting" of observations with negative reported and actual profits.

²⁰We control for the clustering at the industry level throughout our regression analysis.

5.1 Does Competition Enhance Incentives to Hide Profits?

The main objective of the paper is to investigate whether competition encourages profit hiding. Our Hypothesis 1 says that all else equal, firms hide more profit (report less profits) as the market becomes more competitive. As is evident from Table 3, this hypothesis is strongly supported by our regression results. The estimated coefficient for $Competition \times PRO$, β_1 in Equation (9), is negative when $LOGN$ is used as the measure of competition intensity, and is positive in all other cases (higher $H - Index$, $CONCEN$ and $PMARGIN$ all indicate less competition). In all the regressions, the estimated β_1 is statistically very significant. The evidence here shows that firms tend to hide more profits in industries that are less concentrated, have more firms, or have lower average profit margin. Therefore, for each measure of competitiveness and each definition of the market, the empirical results are consistent with the hypothesis that competition enhances firms' incentives to hide profits.

Using the regression results in Table 3, we gauge the magnitude of the competition effect on profit hiding to give estimates of its economic significance. We set all of the independent variables at their means, and estimate how much the responsiveness of $RPRO$ to PRO would change when the competition measure used in the regression changes by one standard deviation.

For example, consider the regression in Column 4 where $CONCEN$ is used to measure competitiveness. The responsiveness level of $RPRO$ relative to PRO , when all independent variables take their mean values, is 0.287. It is calculated as the slope of the profit reporting equation (9) using the estimated coefficients from Column 4 and the means of all independent variables. If $CONCEN$ increases from its mean, 7.1%, by one standard deviation, to 12.7%, then the responsiveness of $RPRO$ to PRO increases by 0.019, representing a 6.6% increase from its previous level. This says that a firm's profit hiding propensity decreases by 6.6% when its industry concentration measured by $CONCEN$ increases by one standard deviation. Similarly, from Column 2, the economic significance $PMARGIN$ is quite large too — a one standard deviation decrease in profit margin from its mean level, 13.8%, to 9.5%, will lead to a 0.026 decrease in the responsiveness of $RPRO$ to PRO , representing a 9.2% decrease from its previous level. From Column 3, if $LOGN$ increases by one standard deviation 1.75, the responsiveness of $RPRO$ to PRO decreases by 0.037, representing a 12.8% decrease from its sample average level.

The economic significance of other measures of competitiveness is similar. Take $PMARGIN$ based on the 3-digit industry as an example (Column 6). All else equal, a one standard deviation

decrease in $PMARGIN$ will lead to a 0.038 decrease in β_1 , which represents a 17.9% decrease in the responsiveness of $RPRO$ to PRO (the average responsiveness level of $RPRO$ to PRO in this specification is 0.214). Take $H - Index$ based on the 2-digit industry/region as an example (Column 8). A one standard deviation decrease in $H - Index$ will lead to a 0.021 decrease in β_1 , representing a 7.4% decrease in the responsiveness of $RPRO$ to PRO in this specification.

5.2 How Does Disparate Tax Rate Affect Firms' Profit-Hiding Incentives?

Our theoretical model predicts that a firm's incentives to hide profits are positively correlated with its tax rate (Hypothesis 2). From the specification of Equation (9), this means β_2 should be negative. Table 3 shows that in all regressions, the estimated coefficient of $TAX \times PRO$ is negative and statistically significant. Thus, higher tax rates reduce the sensitivity of $RPRO$ to PRO , or in other words, firms facing higher tax rates report less profits (i.e., hide more profits).

To estimate the economic significance of the tax effect on profit hiding, we use the result in Column 4 of Table 3 as an example. When TAX increases by one standard deviation (0.268), a firm will report 0.0279 yuan less of profit. Considering that the average responsiveness of $RPRO$ to PRO — when all independent variables take their mean values — is 0.287, a 0.0279 yuan decrease in reported profit represent a reduction of 9.7% from its mean level. Thus, the tax effect on profit hiding is substantial. Note that the estimates of the coefficient of $TAX \times PRO$ are very close in all the regressions in Table 3, thus the magnitude of the tax effect on profit hiding should be close if using other regression results.

5.3 Do Financing Constraints Matter?

Our theoretical model predicts that all else equal, a firm's incentives to hide profits are negatively correlated with its access to the capital market (Hypothesis 3). Since higher $FINANCE$ means better access to the capital market, we expect that β_3 should be positive. It is evident from Table 3 that this hypothesis is strongly supported by our regression results. In all the regressions, the estimated coefficient of $FINANCE \times PRO$ is positive and statistically significant. In fact, the estimates of β_3 are quite stable in all regressions. Thus, in our sample, firms with better access to the capital market report more profits.

To estimate the economic significance of $FINANCE$ on profit hiding, we use the result in Column 4 of Table 3. When $FINANCE$ increases by one standard deviation (0.023), a firm will

report 0.008 yuan more of profit, representing a 2.6% reduction of its profit hiding propensity from its mean level (0.287).

5.4 Do Competitively Disadvantageous Firms Have Stronger Incentives to Hide Profits?

Our model also predicts that all else equal, firms with disadvantageous market positions tend to have greater incentives to hide profits (Hypothesis 4). We use two measures of a firm’s relative market position, its size $LNLABOR$ indicating competitively advantageous market position, and private and collective ownership OWN indicating competitively disadvantageous market position. Hence, we expect that the estimated coefficient of $LNLABOR \times PRO$, β_4 , to be positive, and the estimated coefficient of $OWN \times PRO$, β_5 , to be negative.

Table 3 shows that in all regressions, the estimated coefficient of $LNLABOR \times PRO$ is positive and statistically significant. The estimates are in a close range from 0.012 and 0.021. Thus, in our sample, larger firms hide less profits, consistent with Hypothesis 4. We use the results in Column 4 to gauge the economic significance of firm size on a firm’s profit hiding propensity. Since β_4 , the coefficient of $LNLABOR \times PRO$ is 0.013, a one standard deviation increase in $LNLABOR$ by 1.06 can increase a firm’s reported profit by 0.0138, which represents a 4.8% reduction of the profit hiding propensity at the mean of 0.287. The firm size effect on profit hiding is substantial.

From Table 3, it is clear that the estimated coefficient of $OWN \times PRO$ is negative and statistically significant in all regressions. The estimates are very close across regressions, ranging from -0.57 to -0.48 . The results suggest that private and collective firms demonstrate higher profit hiding propensity than other types of firms. The economic magnitude is also substantial. Take the result from Column 4 of Table 3 as an example. Since β_5 , the coefficient of $OWN \times PRO$, is -0.053 , all else equal, a private or collective firm tends to report 0.052 yuan less of profit. Since the average responsiveness of $RPRO$ to PRO is 0.287, a private or collective firm’s profit hiding propensity is 18.5% higher, all else equal.

6 Robustness Checks and Extensions

In this section, we address the issue of whether the level of competition is endogenous in our setting. We then check the robustness of a critical assumption of our empirical strategy —

competition does not affect $\delta_{i,t}$, the parameter that captures the technical relationship between $PRO_{i,t}$ and true accounting profit $\pi_{i,t}$. We also examine whether the empirical results identified in Section 5 are robust to alternative specifications.

6.1 Is Competition Endogenous?

In general equilibrium, competition intensity in an industry is endogenously determined by firm behavior within the industry as well as behavior of other market participants in the economy. This may cause the endogeneity problem in regressions. To the best of our knowledge, the existing IO literature does not offer any good solution to this problem due to the lack of appropriate instrumental variables.

We believe that the potential endogeneity problem with respect to competition is not likely to be serious in our context. One reason is that the number of firms in most industries is quite large (Table A1), thus a single firm should have a very small effect on the industry competition intensity. Moreover, since the effects of an individual firm on the industry outcome differ across different measures of competition (number of the firm, concentration, profit margin) and across markets, the degree of the potential endogeneity problem should also vary. However, we obtain similar results using different measures of competition intensity and different market definitions. This suggests that the potential endogeneity problem is likely to be quite limited.

To further address the endogeneity of competition, we redo our empirical analysis with firm fixed effects. Specifically, we run regressions according to Equation (9), except that firm fixed effects replace the location and ownership dummies (since they are firm-specific and time-invariant). The results from the firm fixed effects models are very similar to those with the location and ownership dummies. This suggests that the location and ownership dummies can pick up unobserved factors that may affect the level of competition, thus the endogeneity of competition is not likely to be a serious concern.

As a final check on the endogeneity of competition, we examine how our measures of competition intensity change over time. From Table A5, we find that the number of the above scale firms in each two-digit industries has been surprisingly stable over our sample period. Its cross-year standard deviation is normally one or two magnitudes smaller than the average number of the above scale firms in an industry. Take Coal Mining industry (06) as an example. Its mean is 2,826, but its standard deviation is only 177. Timber Logging (12) and Garments (18) are the

only two industries that have eye-witnessed larger changes in number of firms. We also find that the time series distributions of other competition measures: *H – Index*, *CONCEN*, and *PMARGIN*, are all relatively stable across time. This suggests that the level of competition in almost all industries has been relatively stable. Thus, for a given firm, the competition intensity of its industry is more likely to be exogenous.

6.2 Does Competition Affect $\delta_{i,t}$?

The discrepancies between $\pi_{i,t}$ and $PRO_{i,t}$ are driven by the timing of when revenue and expense are recognized into income. Similar to Dechow, Kothari, and Watts’ (1998) analysis of the relation between cash flows and accounting earnings, the discrepancies between $\pi_{i,t}$ and $PRO_{i,t}$ in our context are likely to be accounted for by working capital accruals — especially changes in inventories, receivables, current liabilities, and depreciation. Thus, if competition affects $\delta_{i,t}$, it must do so through these intermediate variables.

However, we find no evidence that competitiveness is related to any of these variables. We compute each of these variables using the information from the NBS dataset. We define $DINV_{i,t}$ as the change in inventories scaled by total assets for firm i in year t , $DCL_{i,t}$ as the change of current liabilities (liquid liabilities in the NBS database) scaled by total assets, and $DCURRD$ as the ratio of depreciation in the current year to total assets. Aside from these measures of working capital accrual, we also define $DIADA$ as the ratio of intangible assets and deferred assets to total sales, and finally $EUP_{i,t}$ as the ratio of (un)employment insurance premium to total sales.

We compute the correlations between these variables and the various competition measures, and report the results in Panel A of Table 4.²¹ As shown in Panel A, all the correlation coefficients are extremely small—below 0.0065 in absolute value. Furthermore, most of them are not statistically significant. This indicates that competitiveness does not affect working capital accruals. Only $RCURRD$ correlates with two competition variables, *H – Index* and *LOGN*, and $DIADA$ correlated with *LOGN*, in the statistically significant sense. However, as we show below, adding these variables to our baseline model has no effect on our results.

As a more direct test, we add the above five variables and their interactions with *PRO* to the

²¹For brevity, we only report the correlations between these variables and the competition variables based on the 2-digit industry codes. Using other market definitions yields similar results.

baseline model (9). We report the results of using the four competition measures based on the 2-digit industry codes in Panel B of Table 4. We find that these variables do not enter regressions significantly for most specifications. Most importantly, as shown in Panel B, controlling for these variables has almost no effect on our basic results. All estimates of the coefficients of interest have the same predicted signs and are statistically significant. Moreover, the point estimates do not change much — in quite a few cases they do not change at all (e.g. $LNLABOR * PRO$).

Lastly, we note that in all results reported earlier, we have included an interaction between $RSALE$ and PRO to control for differences in the timing of revenues (expenses) recognition. The competition effect is clearly not driven by $RSALE$. We hence conclude that in our data, competition affects firms' profit hiding incentives through $d_{i,t}$, not $\delta_{i,t}$.

6.3 Using Differenced PRO and RPRO

Another concern about our empirical approach is that some unobserved time-invariant firm or industry factors other than the ones we have identified might also affect the responsiveness of $RPRO_{i,t}$ to $PRO_{i,t}$. To address this concern, we replace the profit measures in Equation (9) with their first difference counterparts, that is, $DRPRO_{i,t} = RPRO_{i,t} - RPRO_{i,t-1}$ and $DPRO_{i,t} = PRO_{i,t} - PRO_{i,t-1}$. This specification can also serve as an additional way to control for the timing difference between $PRO_{i,t}$ and $\pi_{i,t}$.

Table 5 reports the regression results, where we use competition variables based on the three different market definitions. Using the differenced profit measures yields similar competition effects. For example, the estimated coefficients of $Compet \times DPRO$ in all regressions are statistically significant and have expected signs. That is, as competitiveness increases, the sensitivities of the changes in reported profits, $DRPRO$, to the changes in PRO , $DPRO$, become smaller. The result provides further empirical support for our hypothesis that competition enhances firms' profit hiding incentives. We also find that $TAX \times PRO$, $FINANCE \times PRO$, and $OWN \times PRO$ all have the signs and significance levels consistent with our model predictions. The estimates of $LNLABOR \times PRO$ are not significant for some specifications, but have expected signs.

6.4 Including Lagged PRO As A Control Variable

Another possible factor behind the timing difference between $PRO_{i,t}$ and $\pi_{i,t}$ is that firms may try to smooth income over time. To control for this possibility, we allow $RPRO_{i,t}$ to be responsive to

both $PRO_{i,t}$ and $PRO_{i,t-1}$. Specifically, we include the lagged PRO as an additional independent variable in the baseline model (9). We report the results in Panel A of Table 6. The coefficients of the lagged PRO are statistically significant in all regressions, suggesting that firms' reported profits do depend on their PRO in the previous year. However, all of our previous results hold after controlling for lagged PRO . In all cases, the estimated coefficients have the predicted signs and are statistically significant.

6.5 RPRO Being Responsive to PROs in Various Years

As our last robustness check, we suppose that firms' reported profits respond to both last year's and this year's fundamental earnings, PRO . We specify the model as follows:

$$\begin{aligned}
RPRO_{i,t} = & (\beta_0 + \beta_1 Compet + \beta_2 TAX + \beta_3 FINANCE + \beta_4 LNLABOR + \beta_5 OWN + \beta_6 RSALE \\
& + \sum_{i=year} \beta_i d^{year} + \sum_{j=loc} \beta_j d^{loc}) * PRO_{i,t} + (\lambda_0 + \lambda_1 Compet + \lambda_2 TAX + \lambda_3 FINANCE \\
& + \lambda_4 LNLABOR + \lambda_5 OWN + \lambda_6 RSALE + \sum_{i=year} \lambda_i d^{year} + \sum_{j=loc} \lambda_j d^{loc}) * PRO_{i,t-1} \\
& + \alpha_1 Compet + \alpha_2 TAX + \alpha_3 FINANCE + \alpha_4 LNLABOR + \alpha_5 OWN + \alpha_6 RSALE \\
& + \sum_{i=year} \alpha_i d^{year} + \sum_{j=loc} \alpha_j d^{loc} + \epsilon_{i,t}, \tag{10}
\end{aligned}$$

where the sum of β and λ measures the true sensitivities of $RPRO$ to PRO .²² The previous test where lagged PRO is used as a control variable is a special case of Equation (10) in that all λ 's are set to zero except λ_0 .

The regressions results are reported in Panel B of Table 6, where we suppress the estimated coefficients of other variables and only report $\beta_0 + \lambda_0$, $\beta_1 + \lambda_1$, $\beta_2 + \lambda_2$, $\beta_3 + \lambda_3$, $\beta_4 + \lambda_4$, and $\beta_5 + \lambda_5$. We also test whether the sums are significantly different from zero. As shown in Panel B, $\beta_1 + \lambda_1$ is significantly different from zero in all regressions and have the expected signs. Thus, competition not only affects the responsiveness of $RPRP_{i,t}$ to $PRO_{i,t}$, but also affect its responsiveness to $PRO_{i,t-1}$. This implies that even if firms smooth income over time, competition pressure enhances firms' incentives to hide profits. From Table 6, it is clear that all our other

²²We also allow $RPRO_{i,t}$ to respond to $PRO_{i,t-2}$, or $PRO_{i,t+1}$. These experiments greatly reduce our sample size, but yield qualitatively similar results.

results also continue to hold: the signs and levels of significance of other variables are consistent with our model predictions.

7 Conclusion

In this paper, we have shown that competition pressure drives Chinese industrial firms to hide more profits. Our study provides evidence that in a market environment with poor institutional infrastructure, competition may very well encourage socially wasteful activities as firms use all possible instruments to gain competitive advantage. Thus, policies intended to promote competition in developing and transition economies must be accompanied by reforms which improve the institutional infrastructure. Such reforms would include improved tax enforcement, strengthened financial market regulation, and a more developed legal system to protect property rights and enforce contracts.

We also find strong evidence that firms that are competitively disadvantaged (e.g., smaller firms, firms facing high tax rates, financially constrained firms, and private or collective firms) hide a larger share of their profits. Our findings suggest that in order for competition to deliver desirable social outcomes, it is crucial that all market participants have the same opportunities. This lesson is especially relevant for countries like China, where government policies and regulations routinely favor some firms over others. Such discriminatory practices not only result in allocation inefficiency, but also force firms that are discriminated against to find ways to compensate their disadvantages. These compensating behaviors often include socially unproductive behaviors like profit hiding or bribery of government officials, which can be very harmful in the long run, further deteriorating already weak institutions.

8 Proofs of Propositions

Proof of Proposition 1: Let us define

$$\Delta = [\gamma + c_t^s(\tau_t^s)^2 - \mu_j(\tau_t^s)^2(n-2)][\gamma + c_t^w(\tau_t^w)^2 - \mu_j(\tau_t^w)^2(m-2)] - nm\mu_j^2(\tau_t^s\tau_t^w)^2$$

$$\begin{aligned}
x_t^s &= \mu_j \{ (n-2)\gamma + [2(m+n-2)\mu_j + c_t^w(n-2)](\tau_t^w)^2 \} - c_t^s [\gamma + c_t^w(\tau_t^w)^2 - \mu_j(\tau_t^w)^2(m-2)] \\
y_t^s &= \tau_t^s b_t^s [\gamma + c_t^w(\tau_t^w)^2 - \mu_j(\tau_t^w)^2(m-2)] + m\mu_j \tau_t^s (\tau_t^w)^2 b_t^w + \mu_j \tau_t^s (1 - \tau_t^w) m \gamma \pi_t^w \\
x_t^w &= \mu_j \{ (m-2)\gamma + [2(m+n-2)\mu_j + c_t^s(m-2)](\tau_t^s)^2 \} - c_t^m [\gamma + c_t^s(\tau_t^s)^2 - \mu_j(\tau_t^s)^2(n-2)] \\
y_t^w &= \tau_t^w b_t^w [\gamma + c_t^s(\tau_t^s)^2 - \mu_j(\tau_t^s)^2(n-2)] + n\mu_j \tau_t^w (\tau_t^s)^2 b_t^s + \mu_j \tau_t^w (1 - \tau_t^s) n \gamma \pi_t^s
\end{aligned}$$

From Equations (1) and (2), we obtain, for $i = s$,

$$d_t^s = 1 - \tau_t^s (1 - \tau_t^s) \frac{x_t^s}{\Delta} \quad (11)$$

$$e_t^s = -\frac{y_t^s}{\Delta} \quad (12)$$

and for $i = w$,

$$d_t^w = 1 - \tau_t^w (1 - \tau_t^w) \frac{x_t^w}{\Delta} \quad (13)$$

$$e_t^w = -\frac{y_t^w}{\Delta} \quad (14)$$

Since $\gamma > \max\{\mu_j(\tau_t^s)^2(n-2) - c_t^s(\tau_t^s)^2, \mu_j(\tau_t^w)^2(m-2) - c_t^w(\tau_t^w)^2\}$ (otherwise the model has no solution), Δ is decreasing in μ_j , n and m , while x_t^s is clearly increasing in μ_j , n and m .

Therefore, d_t^s is decreasing in μ_j , n and m . Similarly, d_t^w is decreasing in μ_j , n and m . *Q.E.D.*

Proof of Proposition 2: Note that for $\tau_t^s \leq 0.5$, $\tau_t^s(1 - \tau_t^s)$ is increasing in τ_t^s . In addition, x_t^s is independent of τ_t^s . When $c_t^s \leq \mu_j(n-2)$, Δ is decreasing in τ_t^s . From Equation (11), d_t^s is decreasing in τ_t^s . Since x_t^s is decreasing in c_t^s and Δ is increasing in c_t^s , then d_t^s must be increasing in c_t^s . Similar conclusions hold for d_t^w . *Q.E.D.*

Proof of Proposition 3: From Equations (11) and (13), $d_t^s > d_t^w$ when $\tau_t^s(1 - \tau_t^s)x_t^s < \tau_t^w(1 - \tau_t^w)x_t^w$, or when $x_t^s < x_t^w$ (since $\tau_t^s = \tau_t^w$). It is easy to verify that

$$x_t^w - x_t^s = \mu_j \gamma [(m-n) + c_t^s - c_t^w]$$

Since $m > n$ and $c_t^s > c_t^w$, we have $d_t^s > d_t^w$. *Q.E.D.*

9 Data Appendix

Data for this study are primarily from the annual accounting briefing reports filed by all industrial firms with the National Bureau of Statistics of China (NBS) during the 1995 - 2002 period. Before

1995, due to changes in accounting rules and collection methods, firm-level information collected by the NBS was fragmented and inconsistent. In 1995, China conducted its third nationwide industrial census. The NBS introduced a more rigorous and internally consistent statistical reporting system in preparation for the 1995 industrial census. As a result, the quality of data collection and database management has improved substantially. The NBS database compiles firm level information on large and medium-sized industrial firms annually to calculate the Gross Domestic Product (GDP). The database covers more than 20,000 firms annually in thirty-seven two-digit industries and 28 provinces or province-equivalent municipal cities.

To ensure the reliability of our analysis, we screened the original firm-level data and deleted problematic observations. Specifically, we deleted those observations whose information on critical parameters (such as total assets, the number of employees, gross value of industrial output, net value of fixed assets, or sales) was missing and those misclassified observations whose operation scales were clearly much smaller than the classification standards of large and medium-sized firms. The latter type of deleted observations includes firms whose operation scales measured by one of the following: (i) the value of fixed assets, (ii) the total value of intermediate inputs, (iii) the total value of industrial output, (iv) the total sales, or (v) the total assets, is less than RMB 100,000. It also includes firms who have fewer than 30 employees, or have one of the following variables at a negative value: (i) the total assets minus liquid assets, (ii) the total assets minus total fixed assets, (iii) the total assets minus net value of fixed assets, or (iv) the accumulated depreciation minus current depreciation is negative.

Historical factors may underlie the misclassifications. The classification criteria for industrial firms were established in April 1988 by several government agencies, and were based on the measurement of quantity rather than value. These criteria, a legacy of the centrally planned economy, are being phased out. However, the coexistence of different selection criteria may lead to some misclassifications.

Based on the above selection criteria, we deleted, from year to year, between 2% and 4.8% of observations from the original data source. We did not observe any significant cross-ownership, cross-industry, or geographical patterns in the probability of an observation being dropped, which implies that the “bad data” problem has been random.

After the screening, we have more than 20,000 firms for each year from 1995 to 2002. To ensure that a few outlier observations do not determine our results, we delete firm-year observations with

variable values either below the 1% level or above the 99% level. Our final sample consists of 163,618 firm-year observations.

References

- [1] Banerjee, Abjhit, and Esther Duflo, 2004, Do Firms Want to Borrow More? Testing Credit Constraints Using a Directed Lending Program, MIT Working Paper.
- [2] Becker, Gary. 1957, The economics of discrimination, *Chicago: University of Chicago press*.
- [3] Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, 2004, How Much Should We Trust Differences-in-Differences Estimates? *Quarterly Journal of Economics*, 119, 249-75.
- [4] Cummins, Jason G., and Ingmar Nyman, 2004, The dark side of competitive pressure, *Rand Journal of Economics*, forthcoming.
- [5] Dechow, P., S. P. Kothari, and R. L. Watts, 1998, The relation between earnings and cash flows, *Journal of Accounting and Economics*, 25, 133-168.
- [6] Desai, Mihir, A., 2003, The divergence between book and tax income”, in James Poterba, ed., *Tax Policy and the Economy (17)*, MIT Press.
- [7] Desai, Mihir, A., 2005(a), Corporate tax avoidance and high powered incentives, forthcoming, *Journal of Financial Economics*.
- [8] Desai, Mihir, A., 2005(b), The degradation of corporate profit reporting, forthcoming, *Journal of Economic Perspectives*.
- [9] Fee, E., and C., Hadlock, 2000, Management turnover and product market competition: empirical evidence from the U.S. newspaper industry, *Journal of Business*, 73, 205-243.
- [10] Fisman, Raymond, and Shang-Jin Wei, 2004, Tax rates and tax evasion: evidence from “missing imports” in China, *Journal of Political Economy*, 112(2), 471-496.
- [11] Harris, Mary Stanford, 1998, The association between competition and manager’s business segment reporting decisions, *Journal of Accounting Research*, 36, 111-128.

- [12] Hart, Oliver, 1983, The market mechanism as an incentive scheme, *Bell Journal of Economics*, 14, 366-82.
- [13] Hermalin, Benjami, 1992, The effects of competition on executive behavior, *Rand Journal of Economics*, 23, 350-365.
- [14] Johnson, Simon, Daniel Kaufmann, John McMillan, and Christopher Woodruff, 2000, Why do firms hide? bribes and unofficial activity after communism, *Journal of Public Economics*, 76:495-520.
- [15] Leibenstein, Harvey, 1966, Allocative efficiency vs. 'X-efficiency', *American Economic Review*, 56, 392-415.
- [16] Machlup, Fritz, 1967, Theories of the firm: marginalist, behavioral, managerial, *American Economic Review*, 57, 1-33.
- [17] Milgrom, Paul, and John Roberts, 1992, Economics, organization, and management, *Upper Saddle River: Prentice-Hall*.
- [18] Nalebuff, Berry, and Joseph Stiglitz, 1983, Information, competition, and markets, *American Economics Review*, 73: 278-283.
- [19] Nickell, Stephen, 1996, Competition and corporate performance, *Journal of Political Economy*, 104, 724-746.
- [20] Porter, Michael E., 1990, The competitive advantage of nations, London, Macmillan Press.
- [21] Rajan, Raghuram, and Luigi Zingales, 1998, Financial dependence and growth, *American Economic Review*, 88, 559-586.
- [22] Schmidt, Klaus M., 1997, Managerial incentives and product market competition, *Review of Economics Studies*, 64, 191-213.
- [23] Shleifer, Andrei, 2004, Does competition destroy ethical behavior? *American Economic Review*, 94, 414-418.
- [24] Slemrod, Joel, and Shlomo Yitzhaki, 2000, Tax avoidance, evasion, and administration, NBER Working Paper No. W7473.

Table 1: Competition intensity across the 2-digit industries in China (1995-2002)

Industry	N	LOGN	PMARGIN	H-Index	CONCEN
[06]Coal Mining	2826	7.945	0.220	0.007	0.129
[07]Petroleum Extraction	82	4.399	0.324	0.134	0.596
[08]Ferrous Mining	614	6.419	0.179	0.009	0.164
[09]Nonferrous Mining	1386	7.234	0.255	0.004	0.097
[10]Nonmetal Mining	1787	7.488	0.208	0.001	0.057
[12]Timber Logging	516	6.236	0.309	0.003	0.085
[13]Food Processing	10978	9.303	0.069	0.000	0.035
[14]Food Production	4858	8.487	0.121	0.001	0.062
[15]Beverage	3498	8.159	0.154	0.003	0.088
[16]Tobacco	334	5.809	0.145	0.026	0.264
[17]Textile	11561	9.354	0.097	0.000	0.024
[18]Garments	7459	8.912	0.116	0.000	0.033
[19]Leather	3396	8.128	0.092	0.000	0.035
[20]Timber	2651	7.880	0.094	0.001	0.062
[21]Furniture	1563	7.353	0.145	0.001	0.040
[22]Papermaking	4856	8.487	0.156	0.001	0.045
[23]Printing	3778	8.237	0.224	0.001	0.042
[24]Cultural	1967	7.580	0.125	0.001	0.056
[25]Petroleum Processing	1038	6.944	0.079	0.025	0.232
[26]Raw Chemical	11210	9.344	0.140	0.002	0.078
[27]Medical	3399	8.130	0.202	0.002	0.070
[28]Chemical Fiber	841	6.734	0.127	0.027	0.249
[29]Rubber	1793	7.492	0.108	0.005	0.119
[30]Plastic	6534	8.782	0.125	0.000	0.026
[31]Nonmetal Products	14650	9.592	0.151	0.000	0.020
[32]Pressing Ferrous	3151	8.055	0.121	0.011	0.182
[33]Pressing of Nonferrous	2609	7.865	0.132	0.002	0.072
[34]Metal Products	8741	9.073	0.121	0.000	0.022
[35]Ordinary Machinery	9641	9.173	0.175	0.001	0.040
[36]Special Equipment	6492	8.778	0.143	0.002	0.069
[37]Transport Equipment	6962	8.848	0.136	0.009	0.161
[40]Electric Equipment	8146	9.003	0.132	0.003	0.087
[41]Electronic and Telecom	4613	8.433	0.120	0.004	0.107
[42]Instruments	1920	7.559	0.149	0.004	0.103
[44]Electric Power	4915	8.500	0.153	0.002	0.077
[45]Gas Production	307	5.724	-0.078	0.030	0.252
[46]Tap Water	2398	7.782	0.192	0.010	0.158

NOTE. N is the number of the above-scale firms in a given industry (source: *the China Statistical Yearbook*, various issues for 1995-2002). LOGN is the natural logarithm of N. PMARGIN is the total before-tax profits divided by total sales in a given industry. H-Index is the Herfindahl concentration index, which is the sum of squares of the market shares (by sales) of the ten largest firms in an industry. CONCEN is calculated as the total market share (by sales) accounted for by the four largest firms in an industry. The numbers in brackets before industry names are the 2-digit industry codes designated by the National Bureau of Statistics of China (NBS). The sample period is 1995-2002.

Table 2: Summary statistics of key variables

Variable	# of Obs.	Mean	Std. Dev.	Min.	Max.
Competition measures					
<u>2-digit industry</u>					
H-Index	163,618	0.003	0.007	0	0.175
CONCEN	163,618	0.071	0.056	0	0.621
LOGN	163,618	8.694	1.751	4.317	9.634
PMARGIN	163,618	0.138	0.043	-0.135	0.511
<u>3-digit industry</u>					
H-Index	163,618	0.008	0.017	0	0.773
CONCEN	163,618	0.122	0.089	0	0.842
PMARGIN	163,618	0.141	0.052	-0.142	0.517
<u>2-digit industry/ region</u>					
H-Index	163,618	0.050	0.068	0.001	0.831
CONCEN	163,618	0.310	0.182	0.048	0.917
PMARGIN	163,618	0.012	0.027	-0.171	0.532
Other Variables					
RPRO	163,618	0.003	0.068	-0.687	0.861
PRO	163,618	0.042	0.139	-0.816	1.076
RSALE	163,618	0.991	0.524	0.138	1.172
FINANCE	163,618	0.024	0.023	0	0.287
LNLABOR	163,618	6.572	1.061	3.453	12.012
TAX	163,618	0.182	0.268	0	0.872
OWN	163,618	0.131	0.119	0	1
TA (RMB million)	163,618	341.23	156.09	3.48	67,841

NOTE. This table reports the summary statistics of the key variables used in our regressions for our final sample of 163,618 firm-year observations. We specify a market in three different ways: a 2-digit industry, a 3-digit industry, and a 2-digit industry in one of China's eight economic regions (see Table A.4 for details about economic regions). Based on the three market definitions, we compute H-Index, PMARGIN, and CONCEN accordingly. Since *the China Statistical Yearbooks* do not have information on the number of above-scale firms at the 3-digit industry or region level, we cannot calculate LOGN for them. PRO is the imputed corporate profit from the national income account divided by total assets. RPRO is a firm's reported profit scaled by total assets. RSALE is a firm's total sales divided by total output in monetary terms. FINANCE is the total amount of financial charges divided by total assets. TAX is a firm's income tax actually paid divided by its pre-tax profit, and is set to zero for loss-making firms. OWN is a dummy variable taking the value of 1 if a firm is either a private or collective firm. TA is total assets in million RMB. LNLABOR is the logarithm of the number of employees in a firm.

Table 3: Estimates of profit reporting equations for 1995-2002: the effects of competition, tax, financing constraints, firm size, and ownership

Covariates	2-digit industry				3-digit industry			2-digit industry/ region		
	H-Index	PMARGIN	LOGN	CONCEN	H-Index	PMARGIN	CONCEN	H-Index	PMARGIN	CONCEN
PRO	0.150*** [39.21]	0.069*** [16.76]	0.254*** [20.43]	0.151*** [29.27]	0.158*** [31.31]	0.089*** [11.33]	0.177*** [29.56]	0.274*** [34.43]	0.215*** [36.93]	0.312*** [22.51]
Competition * PRO	1.231*** [13.74]	0.611*** [28.77]	-0.021*** [16.17]	0.337*** [25.84]	1.112*** [12.15]	0.714*** [39.69]	0.241*** [22.61]	0.314*** [14.62]	0.511*** [23.06]	0.112*** [12.19]
TAX*PRO	-0.105*** [14.22]	-0.098*** [16.84]	-0.108*** [16.43]	-0.104*** [17.48]	-0.097*** [15.56]	-0.107*** [18.10]	-0.107*** [18.01]	-0.091*** [16.51]	-0.104*** [14.56]	-0.097*** [14.98]
FINANCE*PRO	0.306*** [7.41]	0.325*** [7.40]	0.356*** [6.91]	0.331*** [7.11]	0.371*** [8.11]	0.353*** [8.23]	0.344*** [7.87]	0.412*** [6.56]	0.375*** [6.45]	0.387*** [6.67]
LNLABOR*PRO	0.013*** [10.88]	0.012*** [11.01]	0.013*** [11.13]	0.013*** [11.53]	0.014*** [12.09]	0.013*** [12.51]	0.013*** [11.49]	0.016*** [12.91]	0.020*** [11.38]	0.021*** [12.19]
OWN*PRO	-0.052*** [11.17]	-0.048*** [10.36]	-0.058*** [11.99]	-0.053*** [9.32]	-0.053*** [11.34]	-0.051*** [10.11]	-0.049*** [9.86]	-0.054*** [10.98]	-0.052*** [10.44]	-0.057*** [10.01]
RSALE * PRO	0.155*** [41.54]	0.155*** [50.65]	0.154*** [41.11]	0.153*** [49.71]	0.154*** [41.48]	0.157*** [51.48]	0.156*** [42.29]	0.161*** [52.81]	0.148*** [38.49]	0.146*** [46.33]
Observations	163,618	163,618	163,618	163,618	163,618	163,618	163,618	163,618	163,618	163,618
Adj. R-squared	0.39	0.40	0.39	0.39	0.40	0.40	0.39	0.38	0.38	0.38

*, **, *** - significant at the 10%, 5%, and 1% levels, respectively. The absolute values of t-statistics based on robust standard errors appear in brackets. When computing the robust standard errors, we control for the clustering at the industry level to account for common random shocks within an industry (or an industry/region).

NOTE. The dependent variable is RPRO, the reported profits scaled by total assets. We specify the basic regression model as

$$RPRO_{it} = \left(\beta_0 + \beta_1 Compet + \beta_2 TAX + \beta_3 FINANCE + \beta_4 LNLABOR + \beta_5 OWN + \beta_6 RSALE + \sum_{i=year} \beta_i d^{year} + \sum_{j=loc} \beta_j d^{loc} \right) PRO_{i,t} + \alpha_1 Compet + \alpha_2 TAX + \alpha_3 FINANCE + \alpha_4 LNLABOR + \alpha_5 OWN + \alpha_6 RSALE + \sum_{i=year} \alpha_i d^{year} + \sum_{j=loc} \alpha_j d^{loc} + \varepsilon_{i,t}$$

In our regressions, we specify a market in three different ways (2-digit industry, 3-digit industry, and 2-digit industry by economic region), and calculate the competition measures accordingly. Since the *China Statistical Yearbook* does not have information on the number of above-scale firms at the 3-digit industry and region levels, LOGN is not available for the last two market definitions. In the interest of brevity, we do not report the estimates for the constant, competition variables, TAX, FINANCE, LNLABOR, OWN, RSALE, year dummies, location dummies, year dummies*PRO, and location dummies* PRO. For variable definitions, see Table 2.

Table 4: Robustness check – does competition affect $\delta_{i,t}$?
Panel A Pearson correlation coefficient matrix

	<u>Competition variables based on 2-digit codes</u>			
	H-Index	PMARGIN	LOGN	CONCEN
Change in inventories from t-1 to t scaled by total assets (DINV)	0.0046	-0.0028	0.0029	0.0042
Change in current liabilities from t-1 to t scaled by total assets (DCL)	-0.0037	0.0023	-0.0035	-0.0034
Depreciation in the current year / total assets (RCURRD)	0.0054*	-0.0046	-0.0063**	0.0037
Change in intangible assets and deferred assets / total sales (DIADA)	-0.0041	-0.0034	0.0064**	-0.0023
(Un)employment insurance premium / total sales (EUP)	0.0012	0.0018	-0.0031	0.0028

*, **, *** - significant at the 10%, 5%, and 1% levels, respectively.

Panel B Estimates of profit reporting equations while controlling for variables that may affect $\delta_{i,t}$

	<u>Competition variables based on the 2-digit codes</u>			
Covariates	H-Index	PMARGIN	LOGN	CONCEN
Competition*PRO	1.191*** [15.63]	0.625*** [26.78]	-0.025*** [15.21]	0.341*** [23.48]
TAX*PRO	-0.105*** [14.691]	-0.102*** [17.72]	-0.099*** [16.36]	-0.094*** [18.21]
FINANCE*PRO	0.298*** [5.46]	0.274*** [6.53]	0.371*** [8.41]	0.343*** [6.77]
LNLABOR*PRO	0.013*** [11.07]	0.013*** [10.94]	0.012*** [9.99]	0.013*** [12.34]
OWN*PRO	-0.051*** [15.94]	-0.052*** [13.95]	-0.058*** [11.23]	-0.054*** [12.18]
Observations	117,736	117,736	117,736	117,736
Adj. R-squared	0.41	0.42	0.42	0.42

*, **, *** - significant at the 10%, 5%, and 1% levels, respectively. The absolute values of t-statistics based on robust standard errors appear in brackets. When computing the robust standard errors, we control for the clustering at the industry level to account for common random shocks within an industry (or an industry/region).

NOTE. We create a set of five variables to capture factors that may affect the technical relationship between RPRO and PRO. DINV is the change in inventories from t-1 to t scaled by total assets. DCL is the change in current liabilities from t-1 to t scaled by total assets. RCURRD is the ratio of the depreciation in the current year to total assets. DIADA is the change in intangible and deferred assets scaled by total sales. Finally, EUP is the ratio of unemployment insurance premium to total sales. Panel A reports the Pearson correlation coefficients between the above five variables and the competition variables based on the 2-digit industry codes. Panel B reports the regression results of the profit reporting equation when we add the five variables, and their interactive terms with PRO to the baseline model specified in Table 3. For our purpose, we only report the estimates of the interactions of Competition, TAX, FINANCE, LNLABOR, and OWN with PRO. Also, we only report the results of using the competition variables defined based on the 2-digits codes. Using other competition variables yields similar results.

Table 5: Estimates of profit reporting equations using differenced profits for 1995-2002

Covariates ^a	<u>2-digit industry</u>				<u>3-digit industry</u>			<u>2-digit industry/ region</u>		
	H-Index	PMARGIN	LOGN	CONCEN	H-Index	PMARGIN	CONCEN	H-Index	PMARGIN	CONCEN
DPRO	0.113*** [18.14]	0.074*** [5.63]	0.001*** [2.92]	0.113*** [16.98]	0.116 [15.95]	0.075*** [4.73]	0.074*** [5.64]	0.073*** [5.27]	0.071*** [5.42]	0.041** [2.30]
Competition * DPRO	0.027* [1.92]	0.172*** [6.43]	-4e-4** [2.38]	0.124*** [4.17]	0.167** [2.31]	0.243*** [8.07]	0.054*** [3.87]	0.019* [1.89]	0.034** [2.13]	0.027*** [2.91]
TAX * DPRO	-0.076*** [13.75]	-0.076*** [13.73]	-0.081*** [14.40]	-0.076*** [13.29]	-0.075*** [13.73]	-0.074*** [13.74]	-0.071*** [13.78]	-0.076*** [13.82]	-0.076*** [13.89]	-0.075*** [13.74]
FINANCE*DPRO	0.832*** [14.40]	0.831*** [14.37]	0.831*** [14.37]	0.813*** [14.05]	0.826*** [14.29]	0.867*** [14.85]	0.823*** [14.25]	0.867*** [13.18]	0.813*** [13.45]	0.826*** [14.31]
LNLABOR*DPRO	1.5e-4 ** [2.08]	1.5e-4 [1.06]	1.6e-4 * [1.96]	1.5e-4 [1.44]	1.6e-4 ** [2.32]	1.5e-4 [1.14]	1.5e-4 [1.18]	1.6e-4 * [1.92]	1.5e-4 [0.86]	1.5e-4 [0.92]
OWN*DPRO	-0.045*** [6.84]	-0.038*** [8.07]	-0.031*** [6.47]	-0.049*** [10.98]	-0.046*** [6.34]	-0.037*** [7.72]	-0.042*** [9.96]	-0.039*** [6.11]	-0.04*** [7.92]	-0.038*** [8.73]
Observations	120,046	120,046	120,046	120,046	120,046	120,046	120,046	120,046	120,046	120,046
R-squared	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11

*, **, *** - significant at the 10%, 5%, and 1% levels, respectively. The absolute values of t-statistics based on robust standard errors appear in brackets. . When computing the robust standard errors, we control for the clustering at the industry level to account for common random shocks within an industry (or an industry/region).

NOTE. The dependent variable is $(RPRO_{it} - RPRO_{it-1})$, the change of reported profit from year t-1 to year t. We specify the basic regression model as

$$(RPRO_{it} - RPRO_{it-1}) = \left(\beta_0 + \beta_1 Compet + \beta_2 TAX + \beta_3 FINANCE + \beta_4 LNLABOR + \beta_5 OWN + \beta_6 RSALE + \sum_{i=year} \beta_i d^{year} + \sum_{j=loc} \beta_j d^{loc} \right) (PRO_{i,t} - PRO_{i,t-1}) + \alpha_1 Compet + \alpha_2 TAX + \alpha_3 FINANCE + \alpha_4 LNLABOR + \alpha_5 OWN + \alpha_6 RSALE + \sum_{i=year} \alpha_i d^{year} + \sum_{j=loc} \alpha_j d^{loc} + \varepsilon_{i,t}$$

We specify a market in three different ways (2-digit industry, 3-digit industry, and 2-digit industry by economic region), and calculate the competition measures accordingly..

Table 6: Robustness analysis of firms' profit reporting behavior 1995-2002

Covariates ^a	<u>2-digit industry</u>				<u>3-digit industry</u>			<u>2-digit industry/ region</u>		
	H-Index	PMARGIN	LOGN	CONCEN	H-Index	PMARGIN	CONCEN	H-Index	PMARGIN	CONCEN
<u>Panel A</u>										
PRO	0.101*** [29.31]	0.026*** [4.53]	0.136*** [15.43]	0.081*** [22.37]	0.065*** [23.72]	0.028*** [10.49]	0.078*** [28.18]	0.161*** [13.08]	0.145*** [9.40]	0.136*** [7.70]
Competition * PRO	1.312*** [11.47]	0.771*** [29.49]	-0.038*** [18.47]	0.404*** [21.41]	0.466*** [11.64]	0.778*** [35.36]	0.248*** [29.74]	0.210*** [12.47]	0.9697*** [24.58]	0.569*** [19.38]
TAX*PRO	-0.038*** [7.20]	-0.041*** [7.07]	-0.031*** [5.05]	-0.039*** [7.12]	-0.031*** [6.71]	-0.041*** [7.75]	-0.034*** [6.47]	-0.033*** [8.12]	-0.036*** [6.78]	-0.031*** [4.12]
FINANCE*PRO	0.283*** [5.17]	0.279*** [5.91]	0.334*** [6.71]	0.286*** [5.44]	0.294*** [7.12]	0.273*** [6.83]	0.305*** [7.39]	0.299*** [5.88]	0.245*** [4.38]	0.291*** [6.76]
LNLABOR*PRO	0.007*** [5.86]	0.007*** [6.71]	0.008*** [6.17]	0.007*** [5.92]	0.007*** [6.37]	0.007*** [5.76]	0.007*** [6.58]	0.007*** [5.22]	0.007*** [6.31]	0.007*** [7.21]
OWN*PRO	-0.067*** [9.29]	-0.063*** [9.93]	-0.067*** [9.23]	-0.065*** [9.01]	-0.064*** [9.39]	-0.063*** [8.92]	-0.061*** [9.08]	-0.067*** [9.67]	-0.061*** [9.81]	-0.066*** [8.92]
Lagged PRO	0.095*** [23.33]	0.047*** [23.42]	0.113*** [26.45]	0.093*** [23.32]	0.083*** [25.98]	0.052*** [26.21]	0.109*** [25.91]	0.098*** [26.45]	0.064*** [26.49]	0.112*** [26.45]
<u>Panel B</u>										
PRO ($\beta_0 + \lambda_0$)	0.207***	0.131***	0.223***	0.214***	0.233***	0.118***	0.214***	0.235***	0.248***	0.217***
Competition*PRO ($\beta_1 + \lambda_1$)	1.567***	0.923***	-0.044***	0.641***	1.627***	1.144***	0.389***	2.460***	1.111***	0.651***
TAX*PRO ($\beta_2 + \lambda_2$)	-0.116***	-0.109***	-0.117***	-0.113***	-0.108***	-0.115***	-0.115***	-0.102***	-0.115***	-0.109***

FINANCE*PRO ($\beta_3 + \lambda_3$)	0.118***	0.134***	0.156***	0.141***	0.172***	0.135***	0.136***	0.177***	0.184***	0.238***
LNLABOR*PRO ($\beta_4 + \lambda_4$)	0.017***	0.016***	0.017***	0.016***	0.017***	0.016***	0.017***	0.019***	0.021***	0.021***
OWN*PRO ($\beta_5 + \lambda_5$)	-0.054***	-0.051***	-0.061***	-0.052***	-0.054***	-0.057***	-0.052***	-0.053***	-0.057***	-0.059***

*, **, *** - significant at the 10%, 5%, and 1% levels respectively. The absolute values of t-statistics based on robust standard errors appear in brackets. . When computing the robust standard errors, we control for the clustering at the industry level to account for common random shocks within an industry (or an industry/region). NOTE. The dependent variable is RPRO, the reported profits scaled by total assets. In both panels, we choose to only report the estimated coefficients of the variables of interests. In Panel A, we add lagged PRO ($PRO_{i,t-1}$) in the baseline model defined in Tables 3 to control for the potential income smoothing over time. In Panel B, we specify the regression models as

$$RPRO_{i,t} = \left(\beta_0 + \beta_1 Compet + \beta_2 TAX + \beta_3 FINANCE + \beta_4 LNLABOR + \beta_5 OWN + \beta_6 RSALE + \sum_{i=year} \beta_i d^{year} + \sum_{j=loc} \beta_j d^{loc} \right) PRO_{i,t} + \alpha_1 Compet + \alpha_2 TAX + \alpha_3 FINANCE + \alpha_4 LNLABOR + \alpha_5 OWN + \alpha_6 RSALE + \sum_{i=year} \alpha_i d^{year} + \sum_{j=loc} \alpha_j d^{loc} + (\lambda_0 + \lambda_1 Compet + \lambda_2 TAX + \lambda_3 FINANCE + \lambda_4 LNLABOR + \lambda_5 OWN + \lambda_6 RSALE + \sum_{i=year} \lambda_i d^{year} + \sum_{j=loc} \lambda_j d^{loc}) PRO_{i,t-1}.$$

In Panel B, we report sums of betas and lambdas, and tests of whether they are significantly different from zero. The asterisks behind the coefficients indicate the significance level.

Table A1 - The significance of the sample firms in the Chinese economy

	1995	1996	1997	1998	1999	2000	2001	2002
(1) Number of firms in the sample	22,543	22,974	23,311	22,293	21,463	20,738	21,898	22,220
(2) Number of all industrial firms in China with annual sales above RMB 5 million	N.A.	N.A.	N.A.	165,080	162,033	162,885	171,256	181,557
(3) Reported value added for all firms in the sample (RMB billion)	958	1,017	1,080	1,131	1,289	1,521	1,742	2,013
(4) Total industrial value added in China (RMB billion)	2,472	2,908	3,241	3,339	3,509	3,905	4,238	4,654
(3)/(4) = Sample firms' value added / China industrial value added	38.8%	35%	33.3%	33.9%	36.7%	39%	41.1%	43.3%
(5) GDP (RMB billion)	5,848	6,789	7,446	7,835	8,207	8,947	9,731	10,479
(3)/(5) = Sample firms' value added / China GDP	16.4%	15%	14.5%	14.4%	15.7%	17%	17.9%	19.2%
(6) Number of employees for all firms in the sample (million)	38	38	37	34	31	28	27	26
(7) Urban employment in China (million)	191	198	202	216	224	232	239	248
(6)/(7) = Sample employment / China urban employment	20.0%	18.9%	18.1%	15.5%	13.7%	12.2%	11.3%	10.7%

Source: the NBS database; China Statistical Yearbook 1995-2002.

A2 Bivariate correlations of the 2-digit industry based competition measures

	LOGN	H-Index	PMARGIN	CONCEN
LOGN	1.0000 (0.000)			
H-Index	-0.5437 (0.000)	1.0000 (0.000)		
PMARGIN	-0.2332 (0.000)	0.0528 (0.001)	1.0000 (0.000)	
CONCEN	-0.6239 (0.000)	0.8472 (0.000)	0.0899 (0.001)	1.0000 (0.000)

NOTE. All four variables measure the competition intensity from different perspectives. For variable definitions, see Table 1.

A3 The list of manufacturing industries based on 3-digit codes in China

[061]Coal Mining	[252+251+253]Petroleum Refining
[062]Coal Processing	[257]Coking
[071+073]Petroleum Extraction	[261]Basic Raw Chemical
[072]Natural Gas Extraction	[262] Chemical Fertilizer
[081]Ferrous Mining	[263] Chemical Pesticide
[082] Other Ferrous Mining	[265]Organic Chemical products
[091] Heavy Metal Mining	[266]Synthetic Chemical Materials
[093]Light Metal Mining	[267]Special Chemical Products
[095]Costly Metal Mining	[268]Chemical Commodity Products
[096]Rare Metal Mining	[271]Chemical Medica
[101]Limestone Mining	[272]Medical Preparation
[102]Chemical Minerals Mining	[273]Chinese Medical
[103]Salt Mining	[275+275]Biological and Animal Medical
[109]Other minerals Mining	[281+285]Chemical Fiber
[121]Timber Logging and Transport	[282]Synthetic Fiber
[131]Food Processing	[291]Tire Products
[132]Vegetable Oil Processing	[296]Recycled Rubber
[133]Sugar Processing	[297~9+292~5]Other Rubber Products
[134]Meat Processing	[301]Film Plastic
[135]Fish Processing	[302]Panel Plastic
[139+136]Other Food Processing	[303]Silk Plastic
[141]Candy Production	[304]Foam Plastic
[142]Dairy Production	[309~5]Others Plastic products
[143]Canning Production	[311]Cement Products
[144]Barmy Production	[312]Finished Cement Products
[145]Condiment Production	[313]Tile Products
[149]Other food production	[314]Glass Products
[151]Alcohol production	[315]Chinaware
[152]Soft Beverage	[316]Fire-resistant Products
[159+155]Tea and Other beverage	[317]Carbon Products
[162]Tobacco Production	[318]Glass Fiber Products
[161+169]Other Tobacco	[319]Other Nonmetal Mineral Products
[171]Fiber Raw Material Processing	[321]Smelting of Ferrous
[172]Cotton Textile	[322]Smelting of Steel
[174]Wool Textile	[324]Refining of Steel
[176]Hemp Textile	[326]Refining of Ferrous
[177]Silk Textile	[331]Refining of Heavy Nonferrous
[178]Knit Fabric	[332]Refining of Light Nonferrous
[179]Others Textile	[336+333~4]Refining of Alloy Nonferrous
[181]Garments Products	[338]Pressing of Nonferrous
[189+182+183]Other Fiber Products	[343]Instrument Products
[191]Leather Making	[344]Metal Container Products
[192+193+195]Leather, Furs Products	[345]Tight Wire Products
[201+202+203+204]Lumber Processing	[346]Structural Metal Products
[211]Timber Furniture	[348]Metal Commodity
[219+212~4]Other Furniture	[349+341~2+347]Others Metal Products
[222+221+223]Papermaking	[351]Boiler Gas Engine
[231+232]Printing and Record Medium	[352]Processing Machinery
[241~5+249]Cultural Sports Products	[353]Ordinary Equipment

A3 The list of manufacturing industries based on the 3-digit codes in China (continued)

[354]Axletree Valv Products	[421]General Instruments Machinery
[356]Others Ordinary Accessory	[422~5+428~9]Special Instruments Machinery
[357~9]Other Ordinary Machinery	[426]Horologe
[361]Mining Equipment	[431+435+439]Craftwork Manufacturing
[362]Petrochemical Equipment	[441]Production of Electric Power
[363]Textile Equipment	[442]Supply of Electric Power
[364]Agricultural Equipment	[443]Steam and Hot Water
[365]Medical Equipment	[451]Gas Production
[367]Others Special Equipment	[452]Gas Supply
[368]Special Repairing Equipment	[461]Production of Tap Water
[371]Railway Transport Equipment	[462]Supply of Tap Water
[372]Automobile	
[373+374+375]Cycle Equipment	
[376]Shipping Equipment	
[377]Navigation Equipment	
[379~8]Other Transport Equipment	
[401]Electric Machine	
[402]Transformer Equipment	
[404]Electric Cable	
[406]Electric Commodity Equipment	
[407]Lighting Equipment	
[408~9]Others Electric Equipment	
[411]Telecom Equipment	
[412~3]Broadcasting Radar Equipment	
[414]Computer Equipment	
[415]Electronic Parts	
[416]Electronic Elements	
[417]Electronic Commodity	
[418~9]Other Electronic and Telecom	

NOTE. We combine various 3-digit industries mainly because there are no firms in some 3-digit industries.

A4 The list of the eight economic regions in China

Economic Regions	Province or province-equivalent cities included
Northern Coastal (NC)	ShangDong, HeBei, Beijing, Tianjin
Southern Coastal (SC)	GuangDong, FuJian, HaiNan
Eastern Coastal (EC)	ShangHai, JiangSu, ZheJiang
North East (NE)	LiaoNing, HeLongJiang, JiLin
Mid-Yangze Range (MYR)	HuNan, HuBei, JiangXi, AnHui
Mid-Huanghe Range (MHR)	ShaanXi, HeNan, ShanXi, Inner Mongolia
South West (SW)	GuangXi, YunNan, SiChuan, ChongQing, GuiZhou
North West (NW)	GanSu, Qinghai, NingXia, Tibet, XinJiang

A5 The number of firms for each two-digit industry over time, 1995 – 2002

2-digit industry	1995	1996	1997	1998	1999	2000	2001	2002
[06]Coal Mining	2859	2816	2816	3202	2795	2666	2603	2812
[07]Petroleum Extraction	79	81	81	76	75	82	90	84
[08]Ferrous Mining	595	604	619	577	578	598	646	696
[09]Nonferrous Mining	1411	1394	1377	1416	1428	1439	1310	1291
[10]Nonmetal Mining	1802	1793	1779	1849	1817	1770	1749	1711
[12]Timber Logging	545	526	502	634	564	498	431	383
[13]Food Processing	11152	11024	10922	11909	11231	10676	10381	10413
[14]Food Production	4949	4885	4840	5368	4963	4691	4563	4615
[15]Beverage	3561	3518	3480	3817	3579	3409	3307	3287
[16]Tobacco	344	340	331	352	352	343	320	287
[17]Textile	11267	11400	11708	11276	10981	10968	12065	13248
[18]Garments	7030	7198	7508	6768	6611	7064	8037	9061
[19]Leather	3285	3327	3428	3312	3192	3164	3539	3932
[20]Timber	2541	2585	2660	2487	2420	2552	2808	3033
[21]Furniture	1507	1527	1567	1470	1473	1498	1625	1767
[22]Papermaking	4755	4800	4881	4763	4657	4672	5027	5285
[23]Printing	3788	3772	3777	3863	3824	3703	3691	3806
[24]Cultural	1865	1892	1964	1785	1807	1879	2024	2327
[25]Petroleum Processing	1019	1020	1041	1052	988	993	1027	1144
[26]Raw Chemical	10932	11115	9473	11303	11337	11430	12031	12063
[27]Medical	3321	3349	3404	3280	3272	3301	3488	3681
[28]Chemical Fiber	824	834	847	803	803	834	885	909
[29]Rubber	1791	1789	1794	1785	1805	1783	1777	1822
[30]Plastic	6242	6349	6568	6016	6047	6230	6884	7665
[31]Nonmetal Products	14529	14558	14683	14496	14366	14540	14707	15305
[32]Pressing Ferrous	3118	3127	3162	3260	3042	2997	3176	3333
[33]Pressing of Nonferrous	2512	2564	2627	2405	2426	2538	2823	2942
[34]Metal Products	8407	8551	8799	8132	8176	8376	9274	10039
[35]Ordinary Machinery	9400	9504	9715	9282	9160	9338	10027	10767
[36]Special Equipment	6497	6479	6490	6638	6470	6406	6391	6546
[37]Transport Equipment	6840	6885	6983	6779	6701	6850	7114	7470
[40]Electric Equipment	7842	7981	8215	7544	7624	7845	8675	9385
[41]Electronic and Telecom	4399	4470	4612	4166	4289	4459	4824	5320
[42]Instruments	1864	1890	1932	1821	1817	1860	2018	2146
[44]Electric Power	4917	4910	4916	4994	4941	4825	4873	4946
[45]Gas Production	299	303	307	291	295	300	320	329
[46]Tap Water	2394	2395	2399	2363	2405	2408	2398	2420