Does China overinvest? Evidence from a panel of Chinese firms

Sai Ding[+]*

(*University of Glasgow*)

Alessandra Guariglia

(*Durham University*)

and

John Knight

(*University of Oxford*)

Abstract

This paper addresses the hotly-debated question: do Chinese firms overinvest? A firm-level dataset of 100,000 firms over the period of 2000-07 is employed for this purpose. We initially calculate measures of investment efficiency, which is typically negatively associated with overinvestment. Despite wide disparities across various ownership groups, industries and regions, we find that corporate investment in China has become increasingly efficient over time. However, based on direct measures of overinvestment that we subsequently calculate, we find evidence of overinvestment for all types of firms, even in the most efficient and most profitable private sector. We find that the free cash flow hypothesis provides a good explanation for China's overinvestment, especially for the private sector, while in the state sector, overinvestment is attributable to the poor screening and monitoring of enterprises by banks.

JEL classification: G31; O16; O53

Keywords: Overinvestment; Investment efficiency; Free cash flow; Debt; China

[+] *Corresponding author*: Sai Ding, Economics, Business School, Adam Smith Building, University of Glasgow, Glasgow, UK, G12 8RT. Email: sai.ding@glasgow.ac.uk.

*The authors thank the ESRC for providing financial support for this research (RES-000-22-3140). The research was conducted while Sai Ding was in the Department of Economics, University of Oxford.

1. Introduction

China has experienced an investment boom in recent years. Gross fixed capital formation has averaged 33 percent of GDP since economic reform commenced in 1978, and 39 percent over the last decade. A wide range of arguments is offered by way of explanation. For instance, Gong and Lin (2008) argue that the vast surplus of labour in rural areas, and the easy and cheap credit provided by the government via its state banking system, are the preconditions for the high investment rate. According to Aziz and Dunaway (2007), it is the attractive returns on investment that provide Chinese firms with strong incentives to invest. They point out that low bank lending rates and abundant retained earnings have kept the cost of investment funds low. Barnett and Brooks (2006) provide evidence that the non-state sector has been the driver in the recent investment surge, and that it has been funded mainly by 'self-raised' funds emanating from the growth of company profits. Knight and Ding (2010) stress the high growth expectations and investment confidence that flow from China's 'development state'.

Nevertheless, whether or not China overinvests is a matter of controversy. Although investment and investment-generated improvements in productivity are important drivers of China’s rapid economic growth, the high investment rate may also be an important source of macroeconomic imbalances. Concern has been expressed that too much investment may create industrial overcapacity, generate inefficiency, and threaten profits and employment.

This paper aims to assess the extent to which Chinese firms overinvest and to analyze the determinants of their overinvestment. To this end, we use a firm-level dataset of 100,112 Chinese firms over the period 2000-2007. We first look at overinvestment indirectly, constructing firm-level measures of investment efficiency, and provide descriptive statistics so as to shed light on the linkages between investment efficiency and ownership, industry, time, and geographic location. A GMM estimator is used to examine the determinants of investment efficiency. We then proceed to measure firm-level overinvestment and free cash flow directly by employing approaches pioneered by Richardson (2006) and Bates (2005). Where overinvestment does exist, we examine whether it can be explained by the free cash flow hypothesis (Jensen, 1986) or the absence of disciplinary role of debt (Stulz, 1990). To the best of our knowledge, none of the existing literature has attempted to use both such a comprehensive dataset and systematic empirical methods to explore the overinvestment issue in relation to Chinese firms.

We find that, despite significant differences across ownership groups, industries and regions, the general investment efficiency of Chinese firms has increased over time. Debt contributes positively to the investment efficiency of private firms, but not that of state-owned enterprises (SOEs). This implies that preferential lending to the state sector by the banking system remains problematic. Evidence of overinvestment is found for all types of firms. In the state sector, overinvestment is attributable to the poor screening and monitoring of SOEs by banks, whereas in the private sector, it can be explained by abundant cash flow generated from high profits.

The remainder of the paper is organized as follows. Section 2 briefly reviews the relevant theories and empirical evidence on overinvestment, both in general and in the context of China. Section 3 describes the data. Section 4 examines the investment efficiency of Chinese firms using both descriptive statistics and regression methods. Section 5 employs various methodologies to attempt to answer the question: do Chinese firms overinvest? Section 6 draws conclusions.

2. Literature review

2.1 Overinvestment – general literature

According to the free cash flow hypothesis (Jensen, 1986), managers have incentives to expand their firms beyond the optimal size. The underlying reason is that growth strengthens managers' power by increasing the resources under their control: as a firm becomes larger, more opportunities exist for managers to indulge their desires for pecuniary and non-pecuniary (power and prestige) benefits. Hence, there exist conflicts of interest between shareholders and managers over dividend policies. The conflict is especially severe when firm have 'free cash flow', i.e. cash flow in excess of that required to fund all projects that have positive net present values (NPV) when discounted at the relevant cost of capital. Managers have to be monitored in order to prevent them from investing their free cash flow at below the cost of capital or wasting it on organizational inefficiencies. Jensen (1986) argues that, by serving this monitoring role, external capital markets in general, and debt in particular, could and should discipline managerial use of funds and prevent overinvestment.

Stulz (1990) develops a theoretical model of the relationship between the source of financing and agency costs of managerial discretion over investment funds. Given poor investment opportunities, the likelihood that management invests in negative NPV projects increases in the level of managerial discretion over investment funds. It is shown that debt reduces such overinvestment by forcing managers to pay out when cash flows accrue. Thus, firms with poor investment opportunities benefit from higher leverage because increased capital market monitoring and discipline reduce the overinvestment problem. In other words, debt financing pre-commits management to pay out free cash flow rather than to waste it when positive NPV investment opportunities are exhausted.

Aghion *et al*. (1999) argue similarly that debt instruments reduce the agency costs of free cash flow by reducing the cash available for spending at the discretion of managers. In their theoretical model, this not only mitigates managerial slack but also accelerates the rate at which managers adopt new technologies and thus fosters growth.

An alternative explanation for overinvestment can be found in the literature on financial constraints. According to Myers and Majluf (1984), information asymmetries increase the cost of capital for firms forced to raise external finance, therefore reducing investment. Yet, financial constraints are eased by the existence of abundant internally generated funds, which creates a tendency for overinvestment.

Much empirical work has been conducted in this field. Using either US or Canadian data, Lang *et al*. (1996), Aivazian *et al*. (2005) and Ahn *et al*. (2006) find a negative relationship between investment and leverage. The correlation is much stronger for firms with low growth. This is consistent with the hypothesis of Stulz (1990) that leverage inhibits managers of low-growth firms from investing in non-profitable capital expenditures.

Using US data of 400 sales of subsidiaries in the 1990s, Bates (2005) relates the use of proceeds from asset sales to overinvestment. If retained proceeds enable firms to bypass external capital markets in financing any remaining positive NPV projects, a positive relationship should appear between post-sale capital investment and the likelihood of retention, and growth would be enhanced. However, if managerial discretion were to result in the financing of negative NPV projects, inefficiencies would be generated. In the presence of managerial discretion and in the absence of measures to align the incentives of managers and shareholders, retention decisions would bear little relation to the firm's growth opportunities. Bates (2005) finds that retaining firms systematically overinvest relative to an industry benchmark, and that their retention probabilities are increasing with their contemporaneous growth opportunities. These results suggest that there exists a trade-off between the efficiency benefits and managerial agency costs of retained proceeds.

Richardson (2006) adopts an accounting-based framework to measure overinvestment and free cash flow. Excess investment is defined as the (positive) residuals from a regression of new investment on a group of explanatory variables, and free cash flow is defined as the amount of cash flow that is not encumbered by the need to maintain the existing assets of the firm. By regressing his measure of excess investment on his measure of free cash flow, he finds that overinvestment is concentrated in firms with the highest levels of free cash flow.

D'Mello and Miranda (2010) investigate the impact of long-term debt on the overinvestment decisions of firms. Employing a sample of 366 debt issues in the US over the period 1968-2001, they find that managers of unleveraged firms retain excessive liquidity, and that the issue of debt leads to a dramatic reduction in cash ratios and in abnormal capital expenditures. The results provide support for the hypothesis that debt reduces overinvestment.

2.2 Overinvestment in China – macroeconomic evidence

Several recent studies have used aggregate data to explore the question of whether China invests too much. They have obtained contradictory results. Zhang (2003) finds that China's ratio of fixed assets investment to GDP was not on a rising trend over the period 1978-2000, once the nominal investment and GDP were converted to real terms using the price index for capital goods and the GDP deflator respectively. His calculated incremental capital-output ratio shows a downward pattern, suggesting an improvement of investment efficiency at the aggregate level. He argues that this may have been due to rural industrialization and the proliferation of small firms, which could have improved allocative efficiency.

Bai *et al*. (2006) derive estimates of the real rate of return on capital in the economy as a whole over the reform period. They find that the return to capital in China has been remarkably high despite the high investment rate. Possible explanations are suggested in terms of rapid growth in total factor productivity (TFP) and a trend towards more capital-intensive industries. The authors also show that there were considerable differences in the return on capital across provinces, the return being highest in the east and lowest in the west of China, but that the provincial dispersion fell over time, suggesting that the misallocation of investment decreased over time.

In contrast, Rawski (2002) holds a negative view of China's investment performance, suggesting that there were low investment returns and widespread excess capacity across many industries throughout the 1990s. There is in fact direct evidence of underutilization of capital in certain industries, particularly heavy industries dominated by the state (European Chamber, 2009). For instance, in 2005 the percentage rate of excess capacity was reported to be 34, 46, 73, 84, and 88 percent[1] respectively in the steel, aluminium, calcium carbide, ferroalloy, and container industries.

[1] These numbers were taken from the National Development and Reform Commission, Guofa No. 38 (2009).

Barnett and Brooks (2006) claim that the increase in investment over the period 1990-2005 led to a rise in the capital-output ratio and a fall in the marginal product of capital. They infer that the efficiency of capital declined, suggesting overinvestment. They also point out that it is the rising profits in both the state and non-state sectors that funded the investment boom since the late 1990s.

Qin and Song (2009) attempt to measure the extent of overinvestment in China using provincial data for the period 1989-2004. By estimating a production function, they are able to predict the profit-maximizing level of investment. Defining overinvestment as actual minus profit-maximizing investment, they find that there was widespread overinvestment. Moreover, the coastal provinces, being more reformed and more prosperous, were technically more efficient but allocatively less efficient, i.e., they tended to show a greater degree of overinvestment.

Gong and Lin (2008) point to the investment boom of the early 1990s, ignited by Deng Xiaoping's Southern Tour in the Spring of 1992. They claim that the immediate impact of the investment boom was to accelerate inflation, but when the government conducted an anti-inflationary policy via deflationary demand management, overinvestment became apparent. The authors postulate that investment fluctuations of this sort can account for the business cycle in China.

In brief, studies based on aggregate time-series data do not reach a consensus regarding the presence of overinvestment in the Chinese economy: microeconomic evidence is needed.

2.3 Overinvestment in China – microeconomic evidence

At the microeconomic level, there is some consensus on the profitability of Chinese firms but heterogeneous evidence on their investment efficiency. For instance, Liang (2006) claims that the reported investment-to-GDP ratio in China is significantly overstated owing to an overestimation of investment spending and an underestimation of consumption and GDP. His firm-level data show that the return on investment has been high and rising since the late 1990s, as a result of the declining share of investment undertaken by SOEs. He claims that the so-called overinvestment problem merely reflects data quality issues and that China's investment remains profitable and sustainable.

Making the assumption that in equilibrium the rate of return required by managers is equal to the cost of capital, Liu and Siu (2006) derive a measure of investment efficiency from a dynamic model of value optimization by firms. They find that the implied cost of capital is much lower in the state than in the non-state sector. Owing to soft budget constraints, managers of SOEs perceive a cost of capital that is inefficiently low, and therefore tend to overinvest from society's perspective. They estimate that the welfare loss resulting from the misallocation of capital amounts to 8% of GDP.

Dollar and Wei (2007) examine the efficiency of capital allocation in China using a sample of over 12,000 firms in 120 cities for the years 2002-04. They find that state ownership is systematically associated with easier access to formal finance and also lower returns to capital. After more than two decades of economic reforms, the bias in favour of SOEs remains intact. In contrast, non-state firms with higher returns to capital are held back by financial constraints. The authors conclude that the immature financial system has imposed costs on the economy in the form of investment misallocation towards inefficient SOEs equal to about 5% of GDP,

Lu *et al*. (2008) use profitability as a measure of investment efficiency of China's industrial firms. Their various indicators show a falling trend in the profit rate from 1978 to the mid-1990s, and a subsequent rise from 1998 onwards. This pattern holds for both SOEs and private firms, but there exists a sharp contrast in the sectoral composition of their respective profit sources. The most profitable SOEs concentrate in several highly monopolized sectors including electricity production and supply, petroleum and gas extraction, and tobacco production. Hence, the rising profitability of SOEs might in part reflect the outcome of government regulatory policies. In contrast, the most profitable industries for private firms are highly competitive ones such as textiles, machinery, and food.

Firth *et al*. (2008) test the relationship between leverage and investment in 1,200 Chinese listed firms over the period 1991-2004. They find that there is a negative link between the two variables, and that the relation is weaker in firms with low growth opportunities, poor operating performance, and high degree of state shareholding. They conclude that the state-owned banks in China impose few restrictions on the capital expenditures of slowly growing and poorly performing firms, as well as firms with greater state ownership, with consequent bias towards overinvestment.

Hsieh and Klenow (2009) use firm-level data to examine the impact of resource misallocation on aggregate manufacturing productivity in China (and India) in comparison with the US. They find evidence of greater distortions in resource allocation in China than in the US. They show that if capital and labour were hypothetically reallocated to equalize marginal products to the extent observed in the US, manufacturing TFP would rise by 30-50% in China (and by even more in India).

Chen *et al*. (2010) examine the connection between government intervention and investment efficiency using a panel of listed firms in China from 2001 to 2006. Instead of directly measuring firms' investment efficiency, they argue that the smaller sensitivity of investment expenditure to investment opportunities for SOEs compared to non-SOEs is evidence for lower investment efficiency in the state sector. They conclude that government intervention through majority state ownership or the appointment of politically connected managers distorts SOEs' investment behaviour and harms investment efficiency, particularly in those SOEs controlled by local governments.

Although the studies surveyed above present intuitively appealing results, the datasets used in most of these studies cover either a fairly small number of firms or a relatively short time period, making the representativeness of the findings questionable. Moreover, each single method of measuring investment efficiency, or overinvestment, inevitably involves strong assumptions. We contribute to this literature by using a more comprehensive dataset and by adopting several methodologies to measure investment efficiency and overinvestment, with the objective of finding a robust answer to our question of whether or not Chinese firms overinvest.

3. Data and sample

Firm-level data offer several advantages for the study of investment behaviour: the problem of aggregation over firms is eliminated in estimation, and heterogeneity among various types of firms can be taken into account (Bond and Van Reenen, 2007). This is particularly important for China owing to the institutional differences between state and non-state enterprises.

We use data drawn from the annual accounting reports filed by industrial firms with the National Bureau of Statistics (NBS) over the period of 2000-2007. This dataset includes all SOEs and other types of enterprises with annual sales of five million yuan (about $650,000) or more. These firms operate in the manufacturing and mining sectors and in 31 provinces or province-equivalent municipal cities. We deleted observations with negative values of: sales; total assets minus total fixed assets; total assets minus liquid assets; and accumulated depreciation minus current depreciation. Firms that did not have complete records of our main regression variables were also dropped. To control for the potential influence of outliers, we excluded observations in the one percent tails of each regression variable. Finally, we removed all firms with less than five years of consecutive observations.

Our final dataset covers 100,112 (mainly unlisted) firms, which corresponds to 639,382 firm-year observations[2]. Our sample is unbalanced, and its structure can be observed in Table A1 in the Appendix. The number of observations ranges from a minimum of 49,639 in 2000 to a maximum of 93,330 in 2003. There was entry and exit of firms during our sample period: less than 30 percent of firms have the full 8-year accounting information. The active entry and exit of firms is the consequence of enterprise restructuring, which began in earnest in the mid-1990s. It can be viewed as a source of dynamism in this sector (see, for instance, Brandt *et al*., 2009).

The NBS data contain a continuous measure of ownership, which is based on the fraction of paid-in-capital contributed by six different types of investors, namely the state; foreign investors (excluding those from Hong Kong, Macao, and Taiwan); investors from Hong Kong, Macao, and Taiwan; legal entities; individuals; and collective investors. The rationale for dividing foreign investors into those from Hong Kong, Macao, and Taiwan, and those from other parts of the world is that the former capture the so-called 'round-tripping' foreign direct investment, whereby domestic firms may register as foreign invested firms from nearby regions to take advantage of the benefits (such as tax and legal benefits) granted to foreign invested firms (Huang, 2003). Ownership by legal entities is a mixture of ownership by state and private legal entities: it represents a form of corporate ownership[3]. Finally, collective firms are typically owned collectively by communities in urban or rural areas (they are known as 'township and village enterprises' or TVEs).

We group all foreign firms (from Hong Kong, Macao, Taiwan, and other parts of the world) into a single category (which are labelled *foreign*); and all firms owned by legal entities and individuals into a single category (labelled *private*)[4]. Thus our firms fall into four broad categories - state-owned, collective, private, and foreign - based on the shares of paid-in-capital contributed by the four types of investors each year.

[2] The NBS dataset does not allow separate identification of publicly listed companies in China. Specifically, it is difficult to track these companies as their legal identification numbers were changed when they went public (Liu and Xiao, 2004). Over the period considered, there were slightly more than 1000 listed companies operating in the manufacturing and mining sectors - amounting to less than 0.3% of the total number of firms in our sample.

[3] Legal entities represent a mix of various domestic institutions, such as industrial enterprises, construction and real estate development companies, transportation and power companies, security companies, trust and investment companies, foundations and funds, banks, technology and research institutions etc.

[4] Within this category, firms owned by individuals represent about two thirds of the total. As firms owned by legal entities include firms owned by state legal entities, one could question their inclusion in the *private* category. One reason for including them is that while the state's primary interest is political (i.e. aimed at maintaining employment levels or control over certain strategic industries), legal entities are profit-oriented (Wei *et al*., 2005). Since our dataset does not allow us to discriminate between state and non-state legal entities, we are unable to exclude the former from our *private* category. However, all our results are robust to excluding all firms owned by legal entities from the *private* category.

We adopt two methods to classify firms by ownership. First, we group firms according to the majority average ownership shares. For instance, if the average share of capital paid-in by private investors over the period 2000-2007 is greater than 50%, then the firm is classified as privately owned. One potential problem with this method is that the size of the private ownership group is likely to be exaggerated. According to Haggard and Huang (2008), defining China's private sector is difficult, as genuinely private domestic firms are different from government-controlled firms. They argue that the former group has remained relatively small and subject to many controls and permissions, for instance with regard to the provision of finance and the requirement of official approval of investment projects above a certain size. To take account of this phenomenon, our second approach to classification is based on a 100% rule. For instance, a firm is classified as privately-owned when all the paid-in-capital in each year is contributed by private investors. This method allows us to focus on the *de jure* private firms which are more likely to represent the true private sector. The cost of the second approach is that a significant number of firms are left in a residual category. This is referred to as the mixed ownership group, in which firms do not have a dominant investor (by the majority rule) or a single-type investor (by the 100% rule).

Table A2 in the Appendix presents the distribution of observations by ownership using both methods. Our sample is dominated by private firms: 62% of firms are classified as privately-owned by the majority rule and 38% by the 100% rule. SOEs, collective firms and foreign firms represent 8%, 8% and 18% of our sample respectively, based on the majority rule, and 4%, 3% and 10% respectively using the 100% rule. The second approach leads to a decrease of the number of firms in all four types of ownership groups, and an increase of firms in the mixed ownership group (46% of our observations are classified as mixed ownership firms). Since the composition of investors in this residual group is unclear, the second method involves a significant loss of observations despite its clearer identification of private ownership. We therefore rely on the majority classification rule.

Table A2 also shows an interesting pattern of the evolution of ownership over the eight-year period. Taking the majority classification rule as an example, we find that the proportion of SOEs in our sample declined dramatically, from 12% in 2000 to 5% in 2007. A similar pattern holds for collective firms, whose share declines from 11% to 7%. In contrast, the share of private firms climbed from 52% to 66%. The share of foreign firms remained roughly stable at between 17 and 19%. Privatization of small SOEs and TVEs became significant after 1998 (Haggard and Huang, 2008). Our dataset reflects the restructuring process involved in the shrinkage of the state and collective sectors and the expansion of the private sector.

4. Investment efficiency of China

4.1 Methodology

We first adopt the method proposed by Dollar and Wei (2007) to measure the investment efficiency of Chinese firms. Investment efficiency can be seen as an indirect measure of overinvestment: firms that overinvest are likely to exhibit low levels of investment efficiency. In Dollar and Wei's (2007) simple model, the profit-maximizing firm i faces the following problem:

$$Max\ \pi_i = p_i Y_i - r_i K_i - w_i L_i \quad , \tag{1}$$

where π_i is the firm's profit, p_i is the output price, Y_i is output, r_i is the rental cost of capital, K_i is capital usage, w_i is the wage rate, and L_i is the firm's labour usage. The firm subscript i reflects the fact that distortions in the output and factor markets can be firm-specific and make the firm's effective output price and input costs deviate from the market prices. The production function is assumed to take the form: $Y_i = A_i K_i^{\alpha} L_i^{1-\alpha}$, where A_i is firm-specific TFP, and α, the capital share in output, is assumed to be the same for all firms in each industry.

The first-order condition for profit maximization is that the marginal revenue product of capital ($MRPK$) should equal the firm-specific interest rate, $MRPK_i = p_i A_i f_k^{'}(K_i, L_i) = r_i$. Since not all distortions faced by the firm are observable, $MRPK$ is difficult to calculate. By virtue of the Cobb-Douglas production function, $MRPK$ is proportional to an observable variable, the average revenue product of capital ($ARPK$), $ARPK_i = \frac{p_i Y_i}{K_i} = \frac{1}{\alpha} MRPK_i$. In our panel data context, we define $ARPK$ as the ratio of value added to capital, i.e.:

$$ARPK_{i,t} = \frac{VA_{i,t}}{K_{i,t}} \quad , \tag{2}$$

where $VA_{i,t}$ is the real value added of firm i in period t, which is equal to the sum of pre-tax profit income and labour compensation, deflated by the provincial ex-factory producer price index; and $K_{i,t}$ is real tangible fixed assets, deflated using a fixed capital formation deflator.

Instead of inferring MRPK from the estimated ARPK, an alternative method is to compute it from the rate of profit on capital:

$$MRPK_{i,t} = \frac{VA_{i,t} - w_{i,t}L_{i,t}}{K_{i,t}} \ , \tag{3}$$

where $w_{i,t}L_{i,t}$ is the total wage bill of firm i at time t. Despite the very strong and contestable assumptions involved in such approximation[5], this alternative approach has the advantage of not relying on the Cobb-Douglas production function specification, based on the capital share being the same across firms. Since $ARPK$ and $MRPK$, as measures of firm efficiency, inevitably involve respective sets of assumptions, we make use of both proxies in order to combine the strength of both and to circumvent the limitations of each.

We first compare $ARPK$ and $MRPK$ across various categories of firms using simple summary statistics. Then formal regression analysis is adopted to examine not only the disparity of investment efficiency but also its determinants. We initially estimate the following equation:

$$IE_{i,t} = \beta_0 + \beta_1 cash\ flow_{i,t-1} + \beta_2 sales\ growth_{i,t-1} + \beta_3 leverage_{i,t-1} + \beta_4 firm\ age_{i,t} + \beta_5 firm\ size_{i,t-1} + \beta_6 ownership\ dummies_{i,t} + v_i + v_t + v_j + v_r + \varepsilon_{i,t} \ , \tag{4}$$

where $IE_{i,t}$ is the investment efficiency of firm i at time t, measured in turn as $ARPK$ and $MRPK$. $cash\ flow_{i,t-1}$ is the ratio of cash flow over total tangible fixed assets of firm i at time $t-1$, where cash flow is defined as the sum of the firm's net income and depreciation. $sales\ growth_{i,t-1}$ is firm i's rate of growth of real sales; $leverage_{i,t-1}$ is the ratio of total debt divided by total assets. Firm $age_{i,t}$ is the natural logarithm of firm age; and $firm\ size_{i,t-1}$ is the natural logarithm of the book value of the firm’s real total assets. $ownership\ dummies_{i,t}$ include three dummy variables equal to 1 if the firm is owned respectively by the state, collective, or foreign agents, and 0 otherwise. The private ownership group is set as the benchmark group. A deflator for capital stock is used to deflate tangible fixed assets, and the provincial ex-factory producer price indices taken from various issues of *China Statistical Yearbook* are used to convert other variables from nominal to real terms.

The error term in equation (4) comprises five components. v_i is a firm-specific time-invariant component, encompassing all time-invariant firm characteristics likely to influence investment efficiency, as well as the time-invariant component of the measurement error affecting any of the regression variables. v_t is a time-specific component accounting for

[5] In aggregate models of perfect competition, in which a single good is produced and that good is used both in consumption and as a capital good, the marginal revenue product of capital equals its rate of return. This method is also based on the assumption that wage payment can be accurately observed.

possible business cycles; v_j is an industry-specific component reflecting industrial features associated with investment efficiency; v_r is a region-specific component which captures geographic factors that influence investment; and $\varepsilon_{i,t}$ is an idiosyncratic component. We control for v_i by estimating our equations in first-differences, and for v_t , v_j , and v_r by including year, industry and regional dummies in all our specifications.

We are particularly interested in the role of debt in limiting any overinvestment bias caused by the managerial agency problem, i.e. we wish to test the hypothesis that high leverage discourages management from undertaking non-profitable investments. In the case of China, it is interesting to examine the differential impact of debt on the investment efficiency of different ownership groups. We therefore interact the leverage term with the ownership dummies in our model specification, as follows:

$$\begin{aligned} IE_{i,t} = {} & \beta_0 + \beta_1 cash\,flow_{i,t-1} + \beta_2 sales\ growth_{i,t-1} + \beta_3 firm\ age_{i,t} + \beta_4 firm\ size_{i,t-1} + \\ & \beta_5 leverage_{i,t-1} * SOE_i + \beta_6 leverage_{i,t-1} * COL_i + \beta_7 leverage_{i,t-1} * PRIV_i + \\ & \beta_8 leverage_{i,t-1} * FOR_i + v_i + v_t + v_j + v_r + \varepsilon_{i,t} \ , \end{aligned} \tag{5}$$

where SOE_i, COL_i, $PRIV_i$ and FOR_i are dummy variables equal to 1 if firm i is classified as state-owned, collectively-owned, privately-owned and foreign-owned respectively, and 0 otherwise.

The system GMM estimator is used to estimate equations (4) and (5) in order to take into account unobserved firm heterogeneity and possible endogeneity and mismeasurement problems of the regressors. It combines the standard set of equations in first-differences with an additional set of equations in levels. By adding the original equation in levels to the system and exploiting these additional moment conditions, Arellano and Bover (1995) and Blundell and Bond (1998) found a dramatic improvement in efficiency and a significant reduction in finite sample bias compared with first-differenced GMM. Our initial instrument choice in the first-differenced equations consists in all regressors (except firm age) lagged twice or more, while the instruments in the level equations are first-differences of all regressors (except firm age) lagged once.

In assessing whether our instruments are legitimate and our model is correctly specified, the Hansen *J* test of over-identifying restrictions is employed to evaluate the overall validity of the set of instruments. In addition, we assess the presence of n^{th}-order serial correlation in the differenced residuals using the *m(n)* test, which is asymptotically distributed as a standard normal under the null of no n^{th}-order serial correlation of the differenced residuals. In the presence of serial correlation of order *n* in the differenced residuals, the instrument set needs to be restricted to lags *n*+1 and further. Since all our models generally fail the test for second-order autocorrelation of the differenced residuals, levels of all endogenous variables lagged three times or more are used as instruments in the first-differenced equations, and first-differenced variables lagged twice are used as additional instruments in the levels equations. By restricting the number of instruments used in each first-differenced equation, we alleviate the potential instrument proliferation problem (Bowsher, 2002; Roodman, 2009).

4.2 Descriptive statistics

Table 1 presents descriptive statistics for the investment rate and investment efficiency proxies. Both means and medians are provided, as the latter are not influenced by outliers. We focus our discussion on means. When firms are classified by ownership (Panel A), we find that net fixed investment as a proportion of tangible fixed assets (I/K) is lowest for SOEs (4.1%). Private firms have the highest investment rate (10.6%), followed by foreign firms (9.9%). SOEs have the lowest investment efficiency as measured by both $ARPK$ (52.7%) and $MRPK$ (19.2%). Foreign firms have the highest $ARPK$ (110.8%), collective firms the highest $MRPK$ (54.1%), and private firms the second highest values of both. This initial evidence conflicts with the conventional view that SOEs are the main source of overinvestment: although they are the least efficient sector in their use of capital, they have accumulated capital less rapidly than other ownership groups.

In Panel B, we group firms into ten industries, i.e. metal and metal product; non-mental products and petroleum processing; chemicals and plastic products; machinery and equipment; electrical equipment; transport equipment; food and tobacco; textile; leather, timber and furniture; and mining and logging. Electronic equipment and transport equipment have the highest values for both the investment rate and the two investment efficiency measures. In contrast, food and tobacco, non-metal products, and petroleum processing have the lowest ratios for all three variables. Interestingly, the labour-intensive textile industry has a very high $ARPK$, perhaps reflecting the efficiency improvement associated with rapid expansion of textile exports and the profitability of exports. In summary, our results suggest that the industries that invest more are also those that are more efficient, and there is no evidence supporting the hypothesis that overinvestment occurs particularly in heavy industries.

Both $ARPK$ and $MRPK$ follow a strictly rising trend over the period 2000-07 (Panel C), suggesting a consistent improvement of firm-level investment efficiency. The investment rate was lowest in 2004 (5.2%), perhaps reflecting the tight monetary and fiscal policies implemented between August 2003 and October 2004 to reduce overheating in the economy. The co-existence of the highest investment rate (15.1%) and the lowest investment efficiency (as measured by both $ARPK$ and) in 2000 suggests the presence of overinvestment at the start of the period. Over time, investment efficiency gradually increased, implying that any overinvestment that existed initially had diminished by the end of the period.

Panel D shows that the coastal provinces have the highest investment rate (10.2%), the highest $ARPK$ (106%), and the highest $MRPK$ (48.3%), while the western provinces have the lowest ratios for all three variables. Capital accumulation was more rapid in the regions with more productive and more profitable capital.

Our initial descriptive statistics are not suggestive of much overinvestment. Firms with high investment rates (i.e., private and foreign firms, operating in electronic and transport equipment industries, located in the coastal region) are also those with high average and marginal revenue product of capital. The year 2000, with its high investment and low efficiency, might be the exception, but our proxies for investment efficiency increased consistently thereafter. It should be noted that the examination of firm efficiency without standardizing for firm-specific factors such as firm size, firm age, and growth opportunities may be misleading. We therefore next analyze the determinants of investment efficiency, making use of a regression analysis.

4.3 Regression analysis

i. Basic model of investment efficiency

The estimates of our basic model (equation 4) are reported in Table 2. In the $ARPK$ regression (column 1), we find that the cash flow ratio has a positive and significant coefficient, suggesting a positive relationship between investment efficiency and the abundance of internal funds. Cash flow, however, may be an imperfect proxy for changes in net worth, as it might contain information about expected future profitability or more in general, demand factors, which would be relevant to investment decisions even under the null hypothesis of perfect capital markets (Bond *et al*., 2003; Carpenter and Guariglia, 2008). To ensure that this does not happen, we include in the regression a distinct measure of investment opportunities. Since most firms in our sample are not listed in the stock market, we follow Konings *et al*. (2003) and Guariglia (2008) and use the annual growth rate of real sales to this end. Our results show that firms with higher investment opportunities tend to invest more efficiently.

The coefficient associated with the leverage term is significantly positive: a one percentage point rise in the debt to assets ratio raises the average revenue product of capital by 0.2 percentage points. This supports the arguments by Jensen (1986) and Stulz (1990) that debt reduces managerial discretion to invest in negative NPV projects, and thus improves firms' investment efficiency. In addition, we find that investment efficiency measured by $ARPK$ is higher for firms that are relatively old and small.

After accounting for firms' internal and external finance, investment opportunities, firm size, firm age and other factors, $ARPK$ is found to be statistically and substantially lower for SOEs than for domestic private firms. The difference is as much as 11.8 percentage points. The coefficient for collective firms is insignificant. Foreign firms have a higher $ARPK$ than private firms, and thus have the highest ratio among all ownership groups. These results are in line with our initial descriptive statistics.

In the case of $MRPK$ (column 2), we obtain similar results for the cash flow and sales growth variables. Yet, the leverage term is insignificant in determining $MRPK$. This might reflect the offsetting effects of debt among various ownership groups. Firm age and firm size also have poorly determined coefficients. Interestingly, the coefficient on the SOE dummy remains negative and highly significant, indicating a lower $MRPK$ than in the case of private firms. In line with the descriptive statistics, collective firms have the highest $MRPK$, and $MRPK$ is slightly lower for foreign than for private firms.

The validity of the instrument sets is confirmed by the *m3* test. The *p* values of the Hansen *J* test is significant, which may result from the large size of our panel. The Monte Carlo evidence of Blundell *et al*. (2000) show that, when using system GMM on a large panel data to estimate a production function, the Sargan test tends to over-reject the null hypothesis of instrument validity. Consistent with this, Nickell and Nicolitsas (1999) report significant Sargan test statistics for all of their estimation results on UK firms, and Benito (2005), and Benito and Hernando (2007) for several of theirs. We are therefore inclined to pay little attention to the *J* test, as long as the test for third order autocorrelation of the differenced residuals is satisfactory.

In summary, our basic model shows that SOEs have much lower investment efficiency - in terms of both $ARPK$ and $MRPK$- than private firms. Having sufficient internal funds and more investment opportunities contributes positively to both measures of investment efficiency. The role of debt in alleviating overinvestment bias and in promoting investment efficiency is found for $ARPK$ but not $MRPK$.

ii. The effects of leverage on investment efficiency for different ownerships

Table 3 reports the estimates of equation (5). Our focus is on the interactions between leverage and the four ownership dummies. In the case of $ARPK$ (column 1), leverage reduces the investment efficiency of SOEs: a one percentage point rise in the debt to assets ratio in a state-owned firm reduces its average revenue product of capital by 0.5 percentage points. Furthermore, debt contributes positively and significantly to the investment efficiency of private firms and foreign firms, with bigger effects for the latter. Similar results hold for the $MRPK$ regression (column 2), except for the insignificant coefficient of the leverage term for foreign firms. The findings that leverage has a significantly negative effect on investment efficiency in SOEs but a significantly positive effect in private firms remain robust.

Table 3 also reports *p*-values associated with χ^2 tests aimed at assessing whether the impact of leverage on investment efficiency is equal across various ownerships. The hypothesis is rejected when SOEs are compared with private and foreign firms.

The differences in the impact of debt on investment efficiency among various ownership groups can be attributed to China's inefficient financial system. China has a 'repressed' financial system, and the degree of government intervention in bank lending decisions has been remarkably high (Riedel *et al*., 2007). Despite the 15-year reform of the banking sector, bank loans constitute a major share of investment financing only for SOEs, whereas private firms are generally discriminated against by the formal financial system and have to rely predominantly on internal funds to finance their investment (Allen *at al*., 2005; Guariglia *et al*., 2011). There is evidence that these problems have become less severe since 2000 (Guariglia and Poncet, 2008). Although the state-owned banks have become more profit-oriented over the decade, Haggard and Huang (2008) argue that private investment has remained financially constrained.

Our results provide further evidence that Chinese banks are generous in their lending to SOEs, but do not carry out effective monitoring. As a consequence, investment efficiency in the state sector is lower the higher the leverage. In contrast, because banks have incentives to impose disciplinary pressures on private and foreign firms, debt plays a significant role in enhancing their investment efficiency.

In summary, our regression analysis shows that there exist significant differences in investment efficiency among the four ownership groups after controlling for several firm-, industry-, time- and region-specific factors. SOEs are found to invest much less efficiently than their non-state counterparts, especially private and foreign firms. We also find evidence that bank lending to state firms creates a bias towards low investment efficiency and, hence, overinvestment in these firms. The disciplinary role of debt in alleviating the overinvestment bias (Jensen, 1986; Stulz, 1990) is found only in the private sector, which is both the driver of the investment boom and the engine of the rapid economic growth in China (Ding and Knight, 2008; 2010).

5. Does China overinvest?

5.1 Methodology

By measuring and explaining the investment efficiency of firms, Section 4 only provided indirect evidence of overinvestment. In this section, we attempt to provide a more direct answer to the question: do Chinese firms overinvest? This also enables us to test directly for the free cash flow hypothesis of overinvestment and to assess whether debt plays a disciplinary role in alleviating managerial discretion (Jensen, 1986; Stulz, 1990). Two approaches are adopted, each involving strong and different assumptions. Given that neither overinvestment nor free cash flow is directly observable, our use of two methods is intended to provide a more robust answer to our research question.

We first adopt the approach introduced by Richardson (2006), which uses an accounting-based framework to measure both overinvestment and free cash flow. According to Richardson (2006), overinvestment is defined as investment expenditure beyond that required to maintain assets in place and to finance expected new investment in positive NPV projects. The following identity holds:

$$I_{TOTAL,it} = I_{MAINTENANCE,it} + I_{NEW,it} \quad , \tag{6}$$

where total investment expenditure of firm i at time t ($I_{TOTAL,it}$) is decomposed into the required investment expenditure to maintain assets in place ($I_{MAINTENANCE,it}$) and new investment expenditure ($I_{NEW,it}$). The new investment expenditure can in turn be decomposed into the expected investment expenditure in new positive NPV projects ($I^{*}_{NEW,it}$) and overinvestment in negative NPV projects ($I^{\xi}_{NEW,it}$) as follows:

$$I_{NEW,it} = I^{*}_{NEW,it} + I^{\xi}_{NEW,it} \quad , \tag{7}$$

where the expected component ($I^{*}_{NEW,it}$) varies with firm's growth opportunities, financing constraints, industry affiliation and other factors; and the unexpected or abnormal component ($I^{\xi}_{NEW,it}$) can be either positive or negative, corresponding to overinvestment or underinvestment respectively. The difficulty is one of implementation, i.e. how to measure the expected and unexpected components of new investment.

Following Richardson (2006), we define $I_{MAINTENANCE,it}$ as amortization and depreciation, which proxies the portion of total investment expenditure that is necessary to maintain plant, equipment, and other operating assets. $I^{*}_{NEW,it}$ is the predicted value from the following model:

$$I/K_{i,t} = \beta_0 + \beta_1 I/K_{i,t-1} + \beta_2 cash\,flow_{i,t-1} + \beta_3 sale\ growth_{i,t-1} + \beta_4 leverage_{i,t-1} + \beta_5 firm\ size_{i,t-1} + \beta_6\ firm\ age_{i,t} + \beta_7 ownership\ dummies_{i,t} + v_i + v_t + v_j + v_r + \varepsilon_{i,t}\ , \quad (8)$$

where $I/K_{i,t}$, the net fixed investment divided by the real tangible fixed assets, corresponds to the new investment expenditure ($I_{NEW,it}$). A lagged dependent variable is included to reflect the sluggish adjustment that may characterize corporate investment in response to changes in the explanatory variables. The other variables are the same as those included in equation (4). Because we cannot compute Tobin's Q, which is typically included in investment models to control for investment opportunities, as most firms in our sample are not listed, the growth of real sales is used to capture investment opportunities. This is an imperfect measure of demand and it may generate measurement error. However, this method has been widely used in the investment literature (Konings *et al*., 2003; Guariglia, 2008). The positive values of the residuals ($I^{\xi}_{NEW,it}$) from the model established in equation (8) are then used as our estimate of overinvestment ($OVERINV_{i,t}$).

According to Richardson (2006), free cash flow is the cash flow beyond what is required to maintain assets and finance expected new investments. It can be defined as follows:

$$FCF_{it} = cash\,flow_{it} - I_{MAINTENANCE,it} - I^{*}_{NEW,it}\ , \quad (9)$$

where FCF_{it} is free cash flow, and is equal to the cash flow/asset ratio ($cash\,flow_{it}$) net of amortization and depreciation ($I_{MAINTENANCE,it}$) and expected new investment ($I^{*}_{NEW,it}$).

The residual approach to calculating overinvestment and the use of a reduced form investment model are open to measurement error problems. We therefore utilize an alternative approach, developed by Bates (2005) to calculate an alternative set of overinvestment ($OVERINV2_{i,t}$) and free cash flow ($FCF2_{it}$) measures. Bates (2005) determines whether firms overinvest or not, by comparing the cash and capital expenditure ratios of sample firms in a given year to those of median industry firms in that year. If the difference between net capital investment of a sample firm and that of the median firm in the same industry is positive, then the firm overinvests. If the difference is negative, then the firm underinvests. The positive values of this difference are labelled *OVERINV2*. Free cash flow (*FCF2*) is then defined as the cash flow of a sample firm that exceeds the median cash flow of all firms operating in the same industry. We follow this definition to construct our second set of measures.

To test for the free cash flow hypothesis of overinvestment, we first estimate the following basic regression:

$$OVERINV(2)_{i,t} = \beta_0 + \beta_1 FCF(2)_{i,t-1} + \beta_2 ownership\ dummies_{i,t} + v_i + v_t + v_j + v_r + \varepsilon_{i,t}, \quad (10)$$

where $OVERINV(2)_{i,t}$ and $FCF(2)_{i,t-1}$ are our overinvestment and free cash flow measures constructed using Richardson's (2006) and Bates' (2005) approaches in turn. We further distinguish the effects of free cash flow on the overinvestment of firms owned by different agents, by estimating the following equation:

$$\begin{aligned} OVERINV(2)_{i,t} = \beta_0 + \beta_1 FCF(2)_{i,t-1} * SOE_i + \beta_2 FCF(2)_{i,t-1} * COL_i + \beta_3 FCF(2)_{i,t-1} * PRIV_i + \\ \beta_4 FCF(2)_{i,t-1} * FOR_i + v_i + v_t + v_j + v_r + \varepsilon_{i,t}, \end{aligned} \quad (11)$$

where $FCF(2)_{i,t-1} * SOE_i$, $FCF(2)_{i,t-1} * COL_i$, $FCF(2)_{i,t-1} * PRIV_i$, and $FCF(2)_{i,t-1} * FOR_i$ are the interactions between our free cash flow measures and various ownership dummies.

It is also interesting to examine the extent to which the free cash flow hypothesis holds among firms with different political connections. On the one hand, government intervention may distort firms' investment behaviour, reduce investment efficiency, and lead to overinvestment. In particular, firms with high political affiliation are more likely to engage in investment that does not aim to maximize firm value, but aims to achieve objectives preferred by the government. Using a sample of state-owned listed firms, Chen *et al*. (2010) find evidence for this argument. On the other hand, political connections could be economically beneficial for firms by providing access to key resources such as bank loans and other preferential government policies like tax benefits (Li *et al*., 2008). Given these two offsetting effects, we keep an open view on the coefficients of those variables involving different levels of political affiliation in the following overinvestment equation:

$$\begin{aligned} OVERINV(2)_{i,t} = \beta_0 + \beta_1 FCF(2)_{i,t-1} * HIGHPA_i + \beta_2 FCF(2)_{i,t-1} * MEDIUMPA_i + \\ \beta_3 FCF(2)_{i,t-1} * NOPA_i + v_i + v_t + v_j + v_r + \varepsilon_{i,t}, \end{aligned} \quad (12)$$

where $HIGHPA_i$, $MEDIUMPA_i$ and $NOPA_i$ are dummy variables equal to 1 if firms *i* has high, medium, and no political affiliation respectively, and 0 otherwise[6].

Lastly, the following two models are estimated to examine the role of debt in alleviating the overinvestment bias for firms owned by different agents (equation 13) and firms with different degrees of political affiliation (equation 14):

$$OVERINV(2)_{i,t} = \beta_0 + \beta_1 FCF(2)_{i,t-1} + \beta_2 leverage_{i,t-1} * SOE_i + \beta_3 leverage_{i,t-1} * COL_i + \beta 4 leverage i,t-1 * PRIV i + \beta 5 leverage i,t-1 * FOR i + vi + vt + vj + vr + \varepsilon i,t, \quad (13)$$

$$OVERINV(2)_{i,t} = \beta_0 + \beta_1 FCF(2)_{i,t-1} +* \beta_2 leverage_{i,t-1} * HIGHPA_i + \beta_3 leverage_{i,t-1} * MEDIUMPA_i + \beta_4 leverage_{i,t-1} * NOPA_i + v_i + v_t + v_j + v_r + \varepsilon_{i,t}. \quad (14)$$

These equations are estimated using the system GMM estimator discussed in Section 4.1.

5.2 Descriptive statistics

To compute Richardson's (2006) measures, we estimate equation (8) and use it to predict net fixed investment for each firm. The results are reported in Table 4: the model performs well, and the results are consistent with common findings in the investment literature. For example, the lagged dependent variable is significant, implying that that the dynamic model fits the data and correctly captures the smooth adjustment of investment. Both the cash flow term and the growth of real sales attract positive and significant coefficients, suggesting the importance of internal funds and investment opportunities in determining firm investment. The coefficient on leverage is negative and highly significant. This finding is consistent with that of Firth *et al*. (2008), and can be seen as evidence in favour of a disciplinary role of debt on firms' investment decisions. Younger firms and larger firms tend to invest more. Both SOEs and foreign firms invest less than private firms, which is in line with our initial descriptive statistics reported in Table 1.

Table 5 presents descriptive statistics for $I_{NEW,it}$, $I^*_{NEW,it}$ and $I^{\xi}_{NEW,it}$ for the four ownership groups. The pattern of expected new investment ($I^*_{NEW,it}$) is very similar to that of actual new investment ($I_{NEW,it}$): it is highest for private firms, followed by foreign firm and collective firms, and lowest for SOEs. $I^*_{NEW,it}$ is always slightly lower than $I_{NEW,it}$ (by

[6] Our dataset contains a measure of firms' political affiliation, which takes values from 10 to 90. High political affiliation (i.e. affiliation with central or provincial governments) corresponds to a value below 20; medium political affiliation (i.e. affiliation with local governments) corresponds to a value between 20 and 90; and no political affiliation corresponds to a value of 90. 6.4% of the firms in our sample have high political affiliation, 39.82% have medium affiliation, and 53.78% have no affiliation. This distribution is fully documented in Table A3 in the Appendix.

1.5%, 1.0%, 0.9%, and 1.0% for state, collective, private and foreign firms respectively). This gives a relatively small residual term ($I^{\xi}_{NEW,it}$). Hence, by comparing the mean values of these variables, we do not find much evidence of overinvestment, as the discrepancy between actual and optimal levels of investment is small. However, the aggregation of a large number of firms might conceal the fact that some firms overinvest and others underinvest, so that the small average values simply result from the two opposing effects.

We next define our first measure of overinvestment ($OVERINV$) as the positive values of $I^{\xi}_{NEW,it}$. We also calculate the percentage of firms that overinvest ($\%OVERINV$), which is a binary variable taking a value of one if the firm overinvests (i.e. if $I^{\xi}_{NEW,it} > 0$), and zero otherwise. We then construct free cash flow (FCF) using equation (9). Following Bates' (2005) definition, we construct a comparable set of measures of overinvestment ($OVERINV2$), percentage of firms that overinvest ($\%OVERINV2$), and free cash flow ($FCF2$). The descriptive statistics of the two sets of variables together with the leverage ratio are presented for each ownership group in Table 6.

Overinvestment ($\%OVERINV$) characterises 22.1% of all firm-year observations in our sample. Although the percentage is slightly lower for private (20.2%) and collective firms (20.9%), and slightly higher for SOEs (25.5%) and foreign firms (28.2%), overinvestment is quite common in all types of firms. Once underinvesting firms are removed, overinvestment calculated following Richardson's (2006) definition and expressed as a proportion of tangible fixed assets ($OVERINV$) is given by 24.8% for the full sample. It is highest for private firms (26.5%) and lowest for SOEs (19.1%). The ratio of free cash flow to tangible fixed assets (FCF) is 14.6% for the full sample; it is lowest for SOEs (3.2%) and highest for collective firms (21.9%).

Bates' (2005) definitions yield very similar results to Richardson's (2006) for both overinvestment ($OVERINV2$) and free cash flow ($FCF2$). Yet, there is a discrepancy in the proportion of firms that overinvest ($\%OVERINV2$): Bate's (2005) ratios are much higher than those based on Richardson's (2006) definition: 44.6% of firm-year observations show overinvestment, with private firms exhibiting the highest figure (46.3%) and SOEs, the lowest (35.0%).

Descriptive statistics for the leverage ratio are also presented in Table 6. The ratio is 57.8% for the full sample. It is highest for SOEs (63.1%) and lowest for foreign firms (48.5%). This evidence of lax lending to SOEs echoes the findings in Section 4.

5.3 Regression analysis

i. The free cash flow hypothesis of overinvestment

According to Jensen (1986), managers of firms with funds in excess of those required to finance positive NPV projects are likely to overinvest. We test this hypothesis using equation (10). The results are reported in Table 7. Based on Richardson's (2006) definitions (column 1), the free cash flow term is found to be positive and significant in the overinvestment equation: a one percentage point rise in the free cash flow ratio raises a firm's overinvestment ratio by 0.08 percentage points. This positive relationship constitutes evidence in favour of the free cash flow hypothesis. Compared with the benchmark group of private firms, all other ownership groups (SOEs, collective firms, and foreign firms) tend to exhibit lower overinvestment: the coefficients on the ownership dummies are in fact all negative and precisely determined. The results using Bates' (2005) definitions confirm our finding of support for the free cash flow hypothesis of overinvestment (column 2).

In contrast with conventional thinking, it is the private sector rather than the state sector that has overinvested most in recent years. One possible explanation is that the rising profitability in the non-state sectors generates abundant free cash flow, which leads to excessive investment. In contrast, SOEs do not have much free cash flow at hand because they are less profitable, and this restricts their proclivity for overinvestment. Another possibility is that SOEs have divested to get rid of obsolete capital in the face of increasing competition, and that this restructuring has curbed their tendency to overinvest (Ding *et al*., 2010).

Table 8 reports the estimates of equations (11) and (12), which are aimed at testing the effects of free cash flow on overinvestment for firms owned by different agents and firms with different degrees of political affiliation. In panel A, we can see that based on the definitions of both Richardson (2006) and Bates (2005), the coefficient of free cash flow is positive and significant for all types of ownership groups. Yet, there is no clear pattern in the magnitude of the free cash flow coefficient across ownership groups. A similar story holds for political connections: free cash flow has a significantly positive effect on the overinvestment of firms with high, medium and no political affiliations, and the magnitude of the interacted coefficients is quite similar (Panel B). Hence, the sensitivity of overinvestment to free cash flow, although always positive, appears to be very similar across firms owned by different agents and firms characterized by different levels of political affiliation.

ii. The disciplinary role of debt in limiting overinvestment

Both Jensen (1986) and Stulz (1990) emphasize the role of debt as an efficient governance instrument to mitigate the agency costs associated with overinvestment. They argue that high leverage can discourage management from undertaking non-profitable investments: debt pre-commits firms to pay cash as interest and principal and such commitments can reduce managerial discretion over free cash flow that may otherwise be allocated to negative NPV projects.

Table 9 shows estimates of equations (13) and (14), which aim at testing this hypothesis. In column 1 of panel A, which is based on Richardson's (2006) definitions, after introducing the interaction terms of leverage and ownership dummies, the coefficient of the free cash flow term remains significantly positive, which suggests that our support for the free cash flow hypothesis is robust. Leverage is found to reduce overinvestment for private firms only: a one percentage point rise in the debt to assets ratio reduces private firms' overinvestment by 0.09 percentage points. No significant debt effects are found for SOEs, collective and foreign firms. Hence, the disciplinary pressures from banks help to curb any tendency to overinvestment in the private sector. In contrast, SOEs, which have enjoyed relatively easy access to formal finance (reflected by their high leverage ratio), are unlikely to face strict screening and monitoring pressures from banks.

The results based on Bates' (2005) measures support the free cash flow hypothesis of overinvestment, but not the disciplinary role of debt in limiting overinvestment: none of the interaction terms is significant (column 2).

Lastly, we compare the effects of debt on overinvestment among firms with different degrees of political affiliation. The free cash flow hypothesis is supported again using the definitions of both Richardson (2006) and Bates (2005). Interesting results are found in terms of the effects of debt. Focusing on Richardson's definitions (2006), leverage has a significant and positive impact on overinvestment in firms with high political affiliation, but a significantly negative effect in firms with medium or no political affiliation (column 1). When firms have high political affiliation, banks' incentives to exert disciplinary pressures on them may be compromised. Without sound monitoring, external funds may lead to more overinvestment in firms with high political affiliation. In contrast, the disciplinary role of debt is found for firms with medium or no political affiliation. Interestingly, the biggest effect of debt in reducing overinvestment is found for firms with medium political affiliation rather than for those without political affiliation. This confirms our hypothesis that, in China, a certain degree of political connections may be beneficial for firms in order to gain access to external finance and other opportunities, but too much government intervention may distort incentives and harm investment efficiency.

Using Bates' (2005) definitions (column 2), a similar story is found for firms with high political affiliation: a one percentage point rise in the leverage ratio raises the overinvestment ratio by 0.10 percentage points. Yet, the coefficients of debt for firms with medium and no political affiliations are negative but not statistically significant.

6. Conclusion

We have adopted various methodological approaches to examine investment efficiency and the overinvestment behaviour of a large sample of Chinese firms over the period 2000-2007. We first looked at overinvestment indirectly, by calculating the average and marginal revenue product of capital as measures of investment efficiency. The initial descriptive statistics suggested that firms that invest most (such as private and foreign firms, operating in the electronic and transport equipment sectors, and located in the coastal region) also have the highest investment efficiency. Furthermore, investment efficiency was found to rise consistently over time. Regressions testing the determinants of investment efficiency indicated that internal funds and investment opportunities contribute positively to investment efficiency. Leverage was found to improve the investment efficiency of private firms but to reduce that of SOEs. This might be attributed to soft budget constraints in the state sector, which limit the disciplinary role of debt. The differences in investment efficiency among state-owned, private, and collective enterprises were found to be important.

In order to provide some direct evidence of overinvestment, we adopted two methods (one attributable to Richardson, 2006, and the other to Bates, 2005) to construct measures of overinvestment and free cash flow. Looking at the mean values of the sample as a whole, the phenomenon of overinvestment did not appear to be important, as we found little discrepancy between actual and optimal investment. However, focusing on different groups of firms, we found significant disparities among firms, with some overinvesting and others underinvesting. When we examined only the overinvesting firms, we found a ratio of overinvestment to tangible fixed asset of about 25% in the full sample, which was highest for private firms and lowest for SOEs. Depending on the methodology used, we found that between 22% and 45% of the firms in our sample overinvest.

Our regression analyses strongly supported the free cash flow hypothesis of Jensen (1986) according to which firms are likely to overinvest when cash flow exceeds its optimal level. This offers an explanation for the overinvestment of private firms. Their rising profitability in recent years has generated significant free cash flow that has induced overinvestment. The disciplinary role of debt on overinvestment (Jensen, 1986; Stulz, 1990) was found to hold for private firms, but not for SOEs. This helps to explain overinvestment in the state sector: despite the gradual financial sector reforms, banks still impose fewer restrictions on SOEs' borrowing and investment decisions, which creates an overinvestment bias. As for political affiliation, we found that debt can limit the overinvestment bias only for firms with medium or no political affiliations.

This research informs the current debate about the policy challenges stemming from China's remarkably high investment. China's urban-oriented, highly capital-intensive industrial development strategy, starting from the 1990s, and the corresponding high investment rate generated massive macroeconomic imbalance. The share of profits in national income rose markedly, so that the share of wage income kept falling. This contributed to rising income inequality and constrained household expenditure. Our research helps to provide some concrete policy suggestions towards the goal highlighted in the recent Chinese Communist Party (CPC)'s Central Committee Meeting (the Fifth Plenum of the 17th CPC), i.e. to shift China's growth pattern from investment-driven towards consumption-driven during the period 2011-15. For instance, tax policies should be introduced to reduce the excess cash flow of firms, especially some monopolistic firms with extraordinary profits. Further financial reforms are needed to strengthen the role of banks or other creditors in corporate governance, and to enhance the role of stock markets in improving market discipline. Institutional reforms are also necessary to reduce the degree of political intervention in firms' investment decisions and in banks' lending decisions.

References

Aghion, P., Dewatripont, M., & Rey, P. (1999). Competition, Financial Discipline and Growth. *Review of Economic Studies , 66* (4), 825-852.

Ahn, S., Denis, D. J., & Denis, D. K. (2006). Leverage and Investment in Diversified Firms. *Journal of Financial Economics , 79*, 317-337.

Aivazian, V., Ge, Y., & Qiu, J. (2005). The Impact of Leverage on Firm Investment: Canadian Evidence. *Journal of Corporate Finance , 11*, 277-291.

Allen, F., Qian, J. and M. Qian. (2005). Law, Finance, and Economic Growth in China. *Journal of Financial Economics , 77*, 57-116.

Arellano, M. and O. Bover. (1995). Another Look at the Instrumental Variable Estimation of Error-Components Models. *Journal of Econometrics , 68*, 29-52.

Aziz, J., & Dunaway, S. (2007). China's Rebalancing Act. *Finance and Development , 44* (3), 27-31.

Bai, C. E., Hsieh, C. T., & Qian, Y. Y. (2006). The Return to Capital in China. *Brookings Papers on Economic Activity , 2*, 61-88.

Barnett, S., & Brooks, R. (2006). What's Driving Investment in China? *IMF Working Paper No. WP/06/265* .

Bates, T. W. (2005). Asset Sales, Investment Opportunities, and the Use of Proceeds. *Journal of Finance , 60* (1), 105-135.

Benito, A. (2005). Financial Pressure, Monetary Policy Effects and Inventories: Firm-Level Evidence from a Market-Based and a Bank-Based Financial System, Economica, Vol. 72, No. 286, 201-24.

Benito, A. and Hernando, I. (2007). Firm Behaviour and Financial Pressure: Evidence from Spanish Panel Data, Bulletin of Economic Research, Vol. 59, No. 4, 283-311.

Blundell, R. and S. Bond. (1998). Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. *Journal of Econometrics , 87* (1), 115-143.

Blundell, R., Bond, S. and Windmeijer, F. (2000). Estimation in Dynamic Panel Data Models: Improving on the Performance of the Standard GMM Estimators, In: Fomby, T. B. and R. C. Hill (eds.), Nonstationary Panels, Panel Cointegration, and Dynamic Panels (Advances in Econometrics, Vol. 15), Amsterdam: JAI Press, Elsevier Science, pp.53-91.

Bond, S., & Van Reenen, J. (2007). Microeconometric Models of Investment and Employment. In *Handbook of Econometrics* (Vol. 6A, pp. 4418-4498). London: Elsevier.

Bond, S., Elston, J. A., Mairesse, J., & Mulkay, B. (2003). Financial Factors and Investment in Belgium, France, Germany, and the United Kingdom: A Comparison Using Company Panel Data. *Review of Economics and Statistics , 85* (1), 153-165.

Brandt, L., Biesebroeck, J. V., & Zhang, Y. F. (2009). Creative Accounting or Creative Destruction? Firm-Level Productivity Growth in Chinese Manufacturing. *NBER Working Paper 15152* .

Brown, J., & Petersen, B. (2009). Why has the Investment-Cash Flow Sensitivity Declined to Sharply? Rising R&D and Equity Market Developments. *Journal of Banking and Finance , 33*, 971-984.

Carpenter, R. E., & Guariglia, A. (2008). Cash Flow, Investment, and Investment Opportunities: New Tests Using UK Panel Data. *Journal of Banking and Finance , 32,* 1894-1906.

Chen, S., Sun, Z., Yang, S., & Wu, D. (2010). Government Intervention and Investment Efficiency: Evidence from China. *Journal of Corporate Finance,* forthcoming.

Ding, S., & Knight, J. (2010). Why has China Grown so fast? The Role of Physical and Human Capital Formation. *Oxford Bulletin of Economics and Statistics,* forthcoming.

Ding, S., & Knight, J. (2008). Why has China Grown so Fast? The Role of Structural Change. *Department of Economics, University of Oxford, Discussion Paper No. 415 .*

D'Mello, R., & Miranda, M. (2010). Long-Term Debt and Overinvestment Agency Problem. *Journal of Banking and Finance , 34*, 324-335.

Dollar, D., & Wei, S.-J. (2007). Das (Wasted) Capital: Firm Ownership and Investment Efficiency in China. *IMF Working Paper WP/07/9 .*

European Chamber. (2009). *Overcapacity in China: Causes, Impacts and Recommendations.* <www.europeanchamber.com.cn>.

Gong, G., & Lin, J. (2008). Deflationary Expension: An Overshooting Perspective to the Recent Business Cycle in China. *China Economic Review , 19*, 1-17.

Guariglia, A. (2008). Internal Financial Constraints, External Financial Constraints, and Investment Choice: Evidence from A Panel of UK Firms. *Journal of Banking and Finance , 32*, 1795-1809.

Guariglia, A., & Poncet, S. (2008). Could Financial Distortions be No Impediment to Economic Growth After All? Evidence from China. *Journal of Comparative Economics , 36* (4), 633-657.

Guariglia, A., Liu, X., and L. Song (2011). "Is the growth of Chinese firms constrained by internal finance?" *Journal of Development Economy* (forthcoming).

Haggard, S., & Huang, Y. S. (2008). The Political Economy of Private-Sector Development in China. In L. Brandt, & T. G. Rawski, *China's Great Economic Transformation* (pp. 337-374). New York: Cambridge University Press.

Hsieh, C. T., & Klenow, P. J. (2009). Misallocation and Manufacturing TFP in China and India. *Quarterly Journal of Economics , CXXIV* (4), 1403-1448.

Huang, Y. (2003). *Selling China: Foreign Direct Investment during the Reform Era.* New York: Cambridge University Press.

Jensen, M. C. (1988). Agency Costs of Free Cash Flow, Corporate Finance, adn the Market for Takeovers. *American Economic Review , 76*, 323-329.

Knight, J., & Ding, S. (2010). Why Does China Invest So Much? *Asian Economic Papers*, 9(3), 87–117.

Konings, J., Rizov, M., & Vandenbussche, H. (2003). Investment and Financial Constraints in Transition Economies: Micro Evidence from Poland, the Czech Republic, Bulgaria, and Romania. *Economics Letters , 78*, 253-258.

Lang, L., Ofek, E., & Stulz, R. M. (1996). Leverage, Investment and Firm Growth. *Journal of Financial Economics , 40*, 3-29.

Li, H., Meng, L., Wang, Q., & Zhou, L. A. (2008). Political Connections, Financing, and Firm Performance: Evidence from Chinese Private Firms. *Journal of Development Economics* , *87*, 283-299.

Liang, H. (2006). *China's Investment Strength is Sustainable.* Global Economics Paper No. 146.

Liu, Q., & Siu, A. (2006). Institutions, Financial Development, and Corporate Investment: Evidence from an Implied Return on Capital in China. *SSRN: http://ssrn.com/abstract=965631* .

Liu, Q., & Xiao, G. (2004). Look who are Disguising Profits: An Application to Chinese Industrial Firms. *Mimeograph, University of Hong Kong* .

Lu, F., Song, G., Tang, J., Zhao, H., & Liu, L. (2008). Profitability of China's Industrial Firms (1978-2006). *China Economic Journal* , *1* (1), 1-31.

Myers, S. C. (1984). The Capital Structure Puzzle. *Journal of Finance* , *34* (3), 572-592.

Nickell, S. and Nicolitsas, D. (1999). How does Financial Pressure Affect Firms?, European Economic Review, Vol. 43, 1435-56.

Qin, D., & Song, H. Y. (2009). Source of Investment Inefficiency: The Case of Fixed -Asset Investment in China. *Journal of Development Economics* , *90*, 94-105.

Rawski, T. (2002). Will Investment Behaviour Constrain China's Growth? *China Economic Review* , *13*, 361-372.

Richardson, S. (2006). Overinvestment of Free Cash Flow. *Review of Accounting Studies* , *11*, 159-189.

Riedel, J., Jin, J. and J. Gao. (2007). *How China Grows: Investment, Finance and Reform.* Princeton: Princeton University Press.

Stulz, R. M. (1990). Managerial Discretion and Optimal Financing Policies. *Journal of Financial Economics* , *26*, 3-28.

Wei, Z., Xie, F., & Zhang, S. (2005). Ownership Structure and Firm Value in China's Privatized Firms. *Journal of Financial and Quantitative Analysis* , *40*, 87-108.

Zhang, J. (2003). Investment, Investment Efficiency, and Economic Growth in China. *Journal of Asian Economics* , *14*, 713-734.

Table 1. Descriptive statistics of investment efficiency measures

Panel A. By ownership

	SOEs	*Collective firms*	*Private firms*	*Foreign firms*
Fixed investment/tangible fixed assets (*I/K*)	0.041 (0.021)	0.072 (0.056)	0.106 (0.087)	0.099 (0.069)
Average revenue product of capital (*ARPK*)	0.527 (0.296)	1.092 (0.666)	0.981 (0.606)	1.108 (0.636)
Marginal revenue product of capital (*MRPK*)	0.192 (0.104)	0.541 (0.294)	0.468 (0.279)	0.451 (0.257)
Observations	48,689	52,427	399,079	113,469

Note: mean and median (in parentheses) values of each variable are reported. The ownership classification is based on the majority rule.

Panel B. By industry

	Metal & metal product	*Non-metal product & petroleum processing*	*Chemical & plastic*	*Machinery & equipment*	*Electronic equipment*
Fixed investment/tangible fixed assets (*I/K*)	0.105 (0.085)	0.071 (0.047)	0.097 (0.076)	0.103 (0.084)	0.115 (0.093)
Average revenue product of capital (*ARPK*)	0.957 (0.611)	0.781 (0.447)	0.828 (0.530)	0.994 (0.658)	1.205 (0.727)
Marginal revenue product of capital (*MRPK*)	0.466 (0.278)	0.384 (0.215)	0.452 (0.269)	0.493 (0.289)	0.526 (0.305)
Observations	56,013	58,283	107,052	75,434	82,945
	Transport equipment	*Food & tobacco*	*Textile*	*Leather, timber & furniture*	*Mining & logging*
Fixed investment/tangible fixed assets (*I/K*)	0.115 (0.095)	0.075 (0.046)	0.092 (0.068)	0.089 (0.059)	0.081 (0.058)
Average revenue product of capital (*ARPK*)	0.923 (0.585)	0.630 (0.380)	1.275 (0.727)	0.923 (0.529)	0.834 (0.475)
Marginal revenue product of capital (*MRPK*)	0.434 (0.256)	0.351 (0.200)	0.431 (0.254)	0.397 (0.235)	0.467 (0.236)
Observations	31,428	24,758	95,480	59,913	47,920

Note: mean and median (in parentheses) values of each variable are reported.

Panel C. By year

	2000	*2001*	*2002*	*2003*	*2004*	*2005*	*2006*	*2007*
Fixed investment/tangible fixed assets (*I/K*)	0.151 (0.122)	0.093 (0.066)	0.107 (0.079)	0.101 (0.069)	0.052 (0.039)	0.113 (0.086)	0.101 (0.081)	0.083 (0.058)
Average revenue product of capital (*ARPK*)	0.772 (0.463)	0.853 (0.492)	0.893 (0.530)	0.930 (0.559)	0.967 (0.591)	1.012 (0.621)	1.078 (0.661)	1.200 (0.732)
Marginal revenue product of capital (*MRPK*)	0.360 (0.209)	0.378 (0.216)	0.406 (0.236)	0.431 (0.249)	0.438 (0.255)	0.471 (0.278)	0.502 (0.296)	0.559 (0.328)
Observations	49,639	66,241	78,640	93,330	92,291	91,147	87,147	80,947

Note: mean and median (in parentheses) values of each variable are reported.

Panel D. By region

	Coastal region	*Inner region*	*Western region*
Fixed investment/tangible fixed assets (*I/K*)	0.101 (0.079)	0.083 (0.061)	0.073 (0.045)
Average revenue product of capital (*ARPK*)	1.062 (0.647)	0.762 (0.441)	0.657 (0.390)
Marginal revenue product of capital (*MRPK*)	0.483 (0.286)	0.384 (0.195)	0.303 (0.173)
Observations	481,756	87,736	69,890

Note: mean and median (in parentheses) values of each variable are reported. The coastal region includes Liaoning, Hebei, Tianjin, Shandong, Jiangsu, Shanghai, Zhejiang, Fujian, Guangdong, and Hainan, plus Beijing (11 provinces). The inner region includes Shanxi, Inner Mongolia, Heilongjiang, Jilin, Anhui, Jiangxi, Henan, Hubei, and Hunan (9 provinces). The Western region includes Chongqing, Gansu, Guangxi, Guizhou, Ningxia, Qinghai, Shaanxi, Sichuan, Tibet, Xinjiang and Yunnan (11 provinces).

Table 2. Basic model of investment efficiency

	ARPK (1)	*MRPK* (2)
$cash\ flow_{i,t-1}$	1.153** (0.044)	0.896** (0.023)
$sales\ growth_{i,t-1}$	0.153** (0.046)	0.043* (0.023)
$leverage_{i,t-1}$	0.199** (0.072)	-0.036 (0.036)
$firm\ age_{i,t}$	0.121** (0.011)	0.004 (0.004)
$firm\ size_{i,t-1}$	-0.169** (0.014)	0.006 (0.006)
$SOE\ dummy_{i,t}$	-0.118** (0.028)	-0.090** (0.005)
$collective\ dummy_{i,t}$	0.012 (0.024)	0.016** (0.005)
$foreign\ dummy_{i,t}$	0.199** (0.018)	-0.059** (0.006)
m3 test (p value)	0.986	0.331
Hansen J test (p value)	0.000	0.000
Observations	286,548	286,548

Note: The dependent variables are the average revenue product of capital (*ARPK*) and the marginal revenue product of capital (*MRPK*) respectively in columns 1 and 2. All specifications were estimated using a sustem GMM estimator. Heteroskedasticity-consistent standard errors are reported in parentheses. We treat $cash\ flow_{i,t-1}$,$sales\ growth_{i,t-1}$,$leverage_{i,t-1}$ and $firm\ size_{i,t-1}$ as potentially endogenous variables: levels of these variables lagged 3 times or more are used as instruments in the first-differenced equations, and first-differences of these same variables lagged twice are used as additional instrument in the level equations. m3 is a test for third-order serial correlation in the first-differenced residuals, asymptotically distributed as N(0,1) under the null of no serial correlation. The Hansen J test of over-identifying restrictions is distributed as Chi-square under the null of instrument validity. Time dummies, industry dummies, and regional dummies are included in both the regressions and the instrument set. The ownership classification is based on the majority rule. ** and * indicate significance at the 5 and 10 percent level respectively.

Table 3. The role of debt in investment efficiency

	ARPK (1)	*MRPK* (2)
$cash\ flow_{i,t-1}$	1.284** (0.093)	0.985** (0.021)
$sales\ growth_{i,t-1}$	-0.021 (0.138)	0.130** (0.051)
$firm\ age_{i,t}$	0.133** (0.032)	0.016* (0.009)
$firm\ size_{i,t-1}$	-0.167** (0.021)	0.014* (0.007)
$leverage_{i,t-1} * SOE_i$	-0.513** (0.227)	-0.184** (0.072)
$leverage_{i,t-1} * COL_i$	0.602 (0.453)	-0.016 (0.167)
$leverage_{i,t-1} * PRIV_i$	0.302** (0.115)	0.091* (0.049)
$leverage_{i,t-1} * FOR_i$	1.078** (0.286)	0.053 (0.058)
H0:impact of $leverage_{i,t-1}$ on *ARPK* or *MRPK* is the same across SOEs and collective firms (*p* value)	0.016	0.395
H0:impact of $leverage_{i,t-1}$ on *ARPK* or *MRPK* is the same across SOEs and private firms (*p* value)	0.000	0.000
H0:impact of $leverage_{i,t-1}$ on *ARPK* or *MRPK* is the same across SOEs and foreign firms (*p* value)	0.000	0.000
m3 test (*p* value)	0.930	0.360
Hansen *J* test (*p* value)	0.961	0.005
Observations	286,548	286,548

Note: SOE$_i$, *PRIV*$_i$, *COL*$_i$, and *FOR*$_i$ are dummy variables equal to 1 if firm *i* is owned respectively by the state, private, collective, and foreign agents, and 0 otherwise. We treat $cash\ flow_{i,t-1}$, $sales\ growth_{i,t-1}$, $firm\ size_{i,t-1}$, $leverage_{i,t-1} * SOE_i$, $leverage_{i,t-1} * COL_i$, $leverage_{i,t-1} * PRIV_i$, and $leverage_{i,t-1} * FOR_i$ as potentially endogenous variables: levels of these variables lagged three times or more are used as instruments in the first-differenced equations, and first-differences of these same variables lagged twice are used as additional instrument in the level equations. The *p* values associated with Chi-square tests for general restrictions are reported. Also see Notes to Table 2.

Table 4. Investment equation

Dependent variable: $I/K_{i,t}$	
$I/K_{i,t-1}$	0.361** (0.042)
$cash\ flow_{i,t-1}$	0.029* (0.016)
$sales\ growth_{i,t-1}$	0.203** (0.021)
$leverage_{i,t-1}$	-0.074** (0.032)
$firm\ age_{i,t}$	-0.010** (0.003)
$firm\ size_{i,t-1}$	0.021** (0.006)
SOE_i	-0.023** (0.004)
COL_i	-0.001 (0.003)
FOR_i	-0.027** (0.005)
m3 test (*p* value)	0.169
Hansen *J* test (*p* value)	0.000
Observations	286,379

Note: We treat $cash\ flow_{i,t-1}$,$sales\ growth_{i,t-1}$,$leverage_{i,t-1}$ and $firm\ size_{i,t-1}$ as potentially endogenous variables: levels of these variables lagged three times or more are used as instruments in the first-differenced equations, and first-differenced of these same variables lagged twice are used as additional instrument in the level equations. Also see Notes to Tables 2 and 3.

Table 5. Descriptive statistics for $I_{NEW,it}$, $I^*_{NEW,it}$ and $I^{\xi}_{NEW,it}$

	SOEs	*Collective firms*	*Private firms*	*Foreign firms*
$I_{NEW,it}$	0.041 (0.021)	0.071 (0.087)	0.107 (0.087)	0.099 (0.069)
$I^*_{NEW,it}$	0.026 (0.027)	0.061 (0.062)	0.098 (0.098)	0.089 (0.088)
$I^{\xi}_{NEW,it}$	0.015 (0.007)	0.010 (0.002)	0.009 (-0.002)	0.010 (-0.005)
Observations	45,695	47,852	349,994	103,072

Note: mean and median (in parentheses) values of each variable are reported. The ownership classification is based on the majority rule.

Table 6. Descriptive statistics of overinvestment and free cash flow measures

	Full sample	*SOEs*	*Collective firms*	*Private firms*	*Foreign firms*
Richardson (2006)'s definitions					
$OVERINV$	0.248 (0.160)	0.191 (0.111)	0.244 (0.147)	0.265 (0.178)	0.230 (0.146)
$\%OVERINV$	0.221 (0.000)	0.255 (0.000)	0.209 (0.000)	0.202 (0.000)	0.282 (0.000)
FCF	0.146 (0.024)	0.032 (-0.005)	0.218 (0.059)	0.142 (0.022)	0.169 (0.035)
Bates (2005)'s definitions					
$OVERINV2$	0.257 (0.191)	0.214 (0.144)	0.249 (0.181)	0.269 (0.206)	0.237 (0.169)
$\%OVERINV2$	0.446 (0.000)	0.350 (0.000)	0.426 (0.000)	0.463 (0.000)	0.443 (0.000)
$FCF2$	0.153 (0.000)	-0.043 (-0.102)	0.224 (0.031)	0.155 (0.005)	0.192 (0.028)
leverage ratio					
$leverage$	0.578 (0.592)	0.631 (0.643)	0.591 (0.605)	0.597 (0.614)	0.485 (0.481)
Observations	639,382	48,689	52,427	399,079	113,469

Note: mean and median (in parentheses) values of each variable are reported. $OVERINV$, $\%OVERINV$, and FCF are Richardson's (2006) measures of the amount of overinvestment, the percentage of firms that overinvest, and free cash flow, respectively; $OVERINV2$, $\%OVERINV2$, and $FCF2$ are Bates' (2005) definitions of the same variables. *OVERINV* and *OVERINV2* are calculated on the samples of firms that actually overinvest. The ownership classification is based on the majority rule.

Table 7. The free cash flow hypothesis of overinvestment (basic equation)

	Richardson's (2006) definitions (1)	*Bates' (2005) definitions* (2)
$FCF(2)_{i,t-1}$	0.075** (0.038)	0.059** (0.005)
SOE_i	-0.058** (0.004)	-0.049** (0.002)
COL_i	-0.019** (0.004)	-0.028** (0.001)
FOR_i	-0.030** (0.002)	-0.035** (0.001)
m3 test (*p* value)	0.170	0.763
Hansen *J* test (*p* value)	0.271	0.000
Observations	107,128	254,142

Note: The dependent variables are $OVERINV_{i,t}$ or $OVERINV2_{i,t}$ respectively in columns 1 and 2. *FCF* is used as independent variable in column 1, and *FCF2,* in column 2. We treat $FCF_{i,t-1}$and $FCF2_{i,t-1}$ as potentially endogenous variables: levels of these variables lagged three times or more used as instruments in the first-differenced equations, and first-differences of these same variables lagged twice are used as additional instrument in the level equations. Also see Notes to Tables 2, 3, and 6.

Table 8. The free cash flow hypothesis of overinvestment (robustness tests)

Panel A. Effects on firms owned by different agents

	Richardson's (2006) definitions (1)	*Bates' (2005) definitions* (2)
$FCF(2)_{i,t-1} * SOE_{\mathrm{i}}$	0.048** (0.017)	0.069** (0.016)
$FCF(2)_{i,t-1} * COL_i$	0.098** (0.018)	0.051** (0.007)
$FCF(2)_{i,t-1} * PRIV_i$	0.068** (0.007)	0.068** (0.003)
$FCF(2)_{i,t-1} * FOR_i$	0.021** (0.007)	0.032** (0.004)
m3 test (*p* value)	0.139	0.752
Hansen *J* test (*p* value)	0.001	0.163
Observations	107,128	254,142

Panel B. Effects on firms with degrees of political affiliation

	Richardson(2006)'s definitions (1)	*Bates(2005)'s definitions* (2)
$FCF(2)_{i,t-1} * HIGHPA_{\mathrm{i}}$	0.034** (0.015)	0.047** (0.008)
$FCF(2)_{i,t-1} * MEDIUMPA_{\mathrm{i}}$	0.059** (0.007)	0.054** (0.003)
$FCF(2)_{i,t-1} * NOPA_i$	0.055** (0.006)	0.064** (0.002)
m3 test (*p* value)	0.150	0.728
Hansen *J* test (*p* value)	0.001	0.018
Observations	107,128	254,149

Note: The dependent variables are $OVERINV_{i,t}$ and $OVERINV2_{i,t}$ respectively in columns 1 and 2. *FCF* (interacted with relevant dummy variables) is used as independent variable in column 1, and *FCF2* (interacted with relevant dummies), in column 2. We treat all regressors as potentially endogenous variables: levels of these variables lagged three times or more are used as instruments in the first-differenced equations, and first-differenced of these same variables lagged twice are used as additional instrument in the level equations. Also see notes to Tables 2, 3, and 6.

Table 9. The role of debt in alleviating the overinvestment bias

Panel A. Effects on firms owned by different agents

	Richardson's (2006) definitions (1)	*Bates' (2005) definitions* (2)
$FCF(2)_{i,t-1}$	0.071** (0.026)	0.066** (0.017)
$leverage_{i,t-1} * SOE_{\mathrm{i}}$	-0.032 (0.049)	-2.080 (1.491)
$leverage_{i,t-1} * COL_i$	0.022 (0.111)	1.305 (1.657)
$leverage_{i,t-1} * PRIV_i$	-0.089** (0.042)	0.551 (0.492)
$leverage_{i,t-1} * FOR_{\mathrm{i}}$	-0.090 (0.068)	1.299 (1.134)
m3 test (*p* value)	0.177	0.160
Hansen *J* test (*p* value)	0.015	0.379
Observations	107,128	254,142

Panel B. Effects on firms with different degrees of political affiliation

	Richardson's (2006) definitions (1)	*Bates' (2005) definitions* (2)
$FCF(2)_{i,t-1}$	0.094** (0.026)	0.062** (0.010)
$leverage_{i,t-1} * HIGHPA_{\mathrm{i}}$	0.176** (0.094)	0.097** (0.058)
$leverage_{i,t-1} * MEDIUMPA_{\mathrm{i}}$	-0.084** (0.050)	-0.002 (0.029)
$leverage_{i,t-1} * NOPA_{\mathrm{i}}$	-0.046** (0.039)	-0.009 (0.024)
m3 test (*p* value)	0.175	0.809
Hansen *J* test (*p* value)	0.199	0.010
Observations	107,128	254,149

Note: The dependent variables are $OVERINV_{i,t}$ and $OVERINV2_{i,t}$ respectively in columns 1 and 2. *FCF* (together with variables involving leverage) is used as independent variable in column 1, and *FCF2*, in column 2. We treat all regressors as potentially endogenous variables: levels of these variables lagged three times or more are used as instruments in the first-differenced equations, and first-differences of these same variables lagged twice are used as additional instrument in the level equations. Also see Notes to Tables 2, 3, and 6.

Appendix

Table A1. Structure of our unbalanced panel

Panel I.

Year	*Number of observations*	*Percent*	*Cumulative*
2000	49,639	7.76	7.76
2001	66,241	10.36	18.12
2002	78,640	12.30	30.42
2003	93,330	14.60	45.02
2004	92,291	14.43	59.45
2005	91,147	14.26	73.71
2006	87,147	13.63	87.34
2007	80,947	12.66	100.00
Total	639,382	100.00	

Panel II.

Number of obs. per firm	*Number of observations*	*Percent*	*Cumulative*
5	154,645	24.19	24.19
6	140,316	21.95	46.13
7	153,685	24.04	70.17
8	190,736	29.83	100.00
Total	639,382	100.00	

Table A2. Distribution of observations by ownership

Panel I. By the majority rule

	SOEs	*Collective firms*	*Private firms*	*Foreign firms*	*Mixed ownership*	*Total*
2000	11.80	11.06	52.04	19.49	5.61	100.00
2001	9.49	9.62	58.00	18.20	4.69	100.00
2002	8.65	8.90	60.89	17.23	4.33	100.00
2003	7.57	8.04	63.36	17.25	3.77	100.00
2004	7.36	7.83	63.56	17.53	3.71	100.00
2005	6.75	7.62	64.42	17.47	3.73	100.00
2006	6.27	7.21	65.18	17.69	3.65	100.00
2007	5.28	6.93	66.25	17.99	3.55	100.00
Average	7.62	8.20	62.42	17.75	4.02	100.00

Note: all numbers are in percentage terms.

Panel II. By the 100% rule

	SOEs	*Collective firms*	*Private firms*	*Foreign firms*	*Mixed ownership*	*Total*
2000	5.89	3.58	23.53	10.54	56.45	100.00
2001	4.75	3.13	31.18	10.04	50.90	100.00
2002	4.27	2.96	35.43	9.62	47.73	100.00
2003	3.71	2.71	39.57	9.96	44.05	100.00
2004	3.68	2.69	40.00	10.21	43.41	100.00
2005	3.25	2.57	40.52	10.21	43.45	100.00
2006	2.95	2.40	41.14	10.39	43.13	100.00
2007	2.23	2.27	42.04	10.62	42.85	100.00
Average	3.69	2.73	37.67	10.18	45.72	100.00

Note: all numbers are in percentage terms.

Table A3. Distribution of firms by degree of political affiliation

Political affiliation level	*Original values*	*Number of observations*	*Percent*	*Cumulative*
High political affiliation	10	12,270	1.92	1.92
	20	28,674	4.48	6.40
Medium political affiliation	40	53,838	8.42	14.82
	50	73,605	11.51	26.34
	61	11,535	1.80	28.14
	62	48,363	7.56	35.70
	63	29,964	4.69	40.39
	71	2,087	0.33	40.72
	72	35,183	5.50	46.22
No political affiliation	90	343,863	53.78	100.00
Total		639,382	100.00	100.00