Market Competition Measurements and Firms' R&D Responses to Market Competition Pressure

Jun Du and Minjia Chen
Aston Business School
Aston Univeristy

Abstract

The effects of market competition on firms are inconclusive theoretically and lacks of empirical evidence. Particularly, the empirical difference of the effects is possibly because of the imperfections of various measures of competition. This paper tries to undertake a broad comparison between popular competition measures and to test the relationship between measures of competition and measures of innovation using a large panel of Chinese manufacturing firm data. Our results suggest that the consistency between competition measures within the same type is high, but low across different types. We confirm an inverted-U shape relationship between competition and innovation, including innovation input, output and efficiency.

Key words: competition measurement, industry, innovation, China

JEL: L10

1. Introduction

Competition is often thought to be able to improve corporate performance and behaviour. It is generally believed that competition pressurises firms to lower costs, reduces managerial and operational slack, provides incentives to optimise efficiency, and drives innovation. This belief has been influencing policy making across the global, such as the deregulation towards free market, trade liberalisation and, even more broadly, globalisation. However, economists dispute about this hypothesis since the theoretical foundation for it is not solid and empirical evidences are vague. Nevertheless, market competition remains an interesting topic to academia and an important issue for policy makers. Though economists disagree on the mechanism and channels that competition may affect market participants, it is less doubtful that competition has effects on corporate behaviour at micro level and on a country's economic health at macro level.

Further, competition in a particular market forms an important element of industrial as well as firm characteristics, which have been commonly recognised and critical to be taken into account by economists in empirical studies. The competition characteristics in the industry are associated with firms' sunk cost, resource dependence, managerial practices, entry barriers, technology applications, and so forth.

There have been many studies on the relationship between market concentration and industry research and development (R&D) (e.g. Lee, 2005; Aghion *et al.*, 2005). Many studies (e.g. Tang, 2006; Boone, 2000) have also pointed out that the diverse and often conflicting relationships that are found in the literature are mostly due to differences in theoretical models, *ad hoc* empirical specifications and data limitations. Among many others, Nickell (1996) supports a positive effect of competition on innovation from a panel of UK data. Scherer and Huh (1992) find that firms' R&D reactions of US firms to high-tech import penetration vary considerably across industries with no clear behavioural pattern. We argue that even though theoretically models do not suggest a conclusive relationship between market competition and innovation, the variations in the empirical findings in the literature could be influenced by the choice of the measures of competition pressure to a large extent. Market competition is unobservable and economists can, therefore, only use measureable proxies to approximate competition based on theoretical assumptions. Consequently, empirical measures of market competition are subject to measurement errors and interpretations, which can potentially influence the effects of measured competition on firms.

The purpose of this paper is threefold. The first is to use a large Chinese firm dataset to produce popular measures of market competition in the literature, and study their usefulness and consistency with one another. We discuss and construct measures for three major groups of proxies for competition, i.e. market concentration, firms' size inequality and firms' profit margin. To our knowledge, there has not been any previous research comparing many competition measures using a single dataset. Early study by Stigler (1964) compares two closely defined measures of market concentration, namely concentration ratio and Herfindahl-Hirschman index. Especially comparisons between measures across the different groups have been little. We will contribute to the literature by filling this gap. Our paper will provide researchers with useful tools for studying competition related topics and industrial characteristics in relation to competition.

Our methodology to study the consistency between measures is to inspect the correlation between the measures and their statistical significance. Since all the empirical measures of competition have monotonic relations with assumed level of market competition, by adjusting the signs of some measure to let all the measures have negative relations with levels of competition, the correlations can tell how relatively those measures rank the marker competition between industries. Therefore, positive and statistical significant correlations mean consistence between respective measures. We find that the consistencies are relatively high when the measures are closely defined, and are low when the measures do not share the same principle of measurement. The consistencies among measures across different groups are low and even negative. Therefore, our results suggest that it is important for researchers to carefully select the competition measures for their research and to check robustness of their economic implications.

The second purpose is to apply the popular competition measures to study the relationship between competition and innovation, i.e. R&D input, output and efficiency. This relationship has been drawing large attention from both researchers and policy-makers. Most studies in the literature only focus on the effects of competition on R&D inputs, such as the ratio of R&D spending to sales. We take a step forward by also investigating the output of innovation, i.e. the ratio of new product sales to total sales, and the efficiency of innovation, i.e. the ratio of new product sales to R&D expenditure. Using these three aspects of firm innovation and a selection of market competition measures we try to contribute to the existing literature by providing a more complete picture of the impacts of competition on firm innovation. We believe that these three aspects of innovation are of high importance to policy makers.

The rest of the paper is organised as follows. Section 2 reviews and discusses the literature on the effects of market competition on firms. Section 3 analyses some issues with the empirical measurements of competition. Section 4 presents the popular measures of market competition in the literature and discusses their usefulness and limitations. Our data is introduced in section 5 with the correlations of the empirical measures using our data. Section 6 presents the results of the competition and innovation models. Finally, section 7 concludes.

2. The Effects of Competition on Firms

Vickers (1995) argues that the concept of competition should encompass all forms of rivalry (including market trading, auctions etc.), instruments of rivalry (including price, advertising, research and development, merger and acquisition, and efforts), and objects of rivalry (such as profits, market share, corporate control, promotions, and survival). Competition is perceived to have positive effects on the efforts to improve economic efficiency through its ability to encourage firms and individuals to improve productivity and work efforts. For example, Nickell *et al.* (1992) and Nickell (1996) find in UK firms that high level of market competition is associated with high productivity growth. Vickers (1995) and Disney, Haskel, and Heden (2003) also show that competition shifts resources and output from inefficient firms towards efficient ones and forces the former to exit, and therefore increase overall efficiency. However, Green and Mayes (1991), Caves *et al.* (1992) and Schmidt (1997) document inverted U-shaped relationship between competition and efficiency, which shows that increases in competition are most beneficial to efficiency in markets where competition is low to start with. Oligopolists, on the other hand, may tend to spend too much resource on deterring rivals, which can lead directly to production inefficiency (Nickell, 1996). A number of studies demonstrate a positive link between market competition and managerial incentives and efforts (e.g. Nickell, Nicolitsas and Dryden, 1997; Schimdt, 1997; Meyer and Vickers, 1997; Bloom and Van Reenen, 2007; Karuna, 2007).

Competition may mitigate risks and volatility to some degree by improving information availability, particularly concerning technology (Nalebuff and Stiglitz, 1983). However, excessive competition may increase idiosyncratic risks and the volatility of firms' profits (Nalebuff and Stiglitz, 1983; Gaspar and Massa, 2006). Mankiw and Whinston (1986) also pointed out that excessive competition can lead to wasteful duplications and cause incumbent firms to reduce output.

A possible channel through which competition can encourage efficiency in general is that under certain reasonable conditions firms with competition pressure are more likely to invest in R&D and undertake innovative activities than monopoly firms, but Aghion *et al.* (2005) argues that competition and innovation are likely to have inverted-U relationship.

One of the famous debates in economics is the relationship between competition and innovation as inspired by Schumpeter (1934 and 1943) and Arrow (1962). The former argues that monopolies favour innovation because they face less market uncertainty and have larger and stable cash flow to fund innovation activities. On the contrary, the latter suggests that competition pressure is often the main driver of research and development and innovation. It is worth noting that Schumpeterian and Arrow's arguments need not conflict. Schumpeter refers to competition in the post innovation market while Arrow refers to competition in the pre-innovation market. Both sides of the arguments may hold depending on what exactly the relationship is of interest.

Many empirical findings of the existing studies on the effects of various forms of market competition on firm R&D and innovation behaviour are diverse and often conflicting. On the one hand Schumpeterian theories of industrial organisation typically predict that innovation increases as competition increases. For example, Dixit and Stiglitz (1977) models that more intense product market competition, as increases in the substitutability between differentiated products, reduces post-entry rents, and therefore, reduces the equilibrium number of entrants. Consequently, increased product market competition discourages innovation (by reducing post-entry rents). This view is shared by existing endogenous growth model (e.g. Romer, 1990), where an increase in product market competition has a negative effect on productivity growth by reducing the monopoly rents that reward new innovation. Using a long panel of UK data and controlling technological opportunities Geroski (1990) finds evidence that concentration and other measures of monopoly power reduces innovation, and, hence, productivity growth. Both Nickell (1996) and Blundell, Griffith and Van Reenen (1999) find a positive linear effect of competition on innovation on UK firms. Nickell (1996) also argues that in the empirical studies which analysing the relationship between market structure and R&D intensity it is important to control for technological opportunities, which differ substantially across industries and tend to be correlated with market concentration. Such technological opportunities are likely to be fixed effects.

Blundell, Griffith and Van Reenen (1995) find that in their UK listed firm sample the market dominant firms tend to innovate more. However, they also find counteracting results that industry concentration dampens innovative activity. Further, to the extent that growing dominance increases industrial concentration the level of aggregate innovation tends to fall. Their results could suggest that different perspectives of competition can indicate different relationships between competition and innovation. This also gives us some motivation to investigate the role of different empirical measures of competition in the relationship.

Aghion *et al.* (2005) allow for a non-monotonicity and develop an inverted-U relationship between competition and innovation. It can be based on the combination of agency models (Hart, 1983) and Schumpeterian models. They develop a model where competition discourages laggard firms from innovating but encourages incumbent firms operating at similar technological levels to innovate. Their predicted inverted-U shape relationship is supported by their UK panel data.

Some other authors also find that the relationship between competition and innovation is not simply linear, but differentiated across firms and subject to certain conditions. Aghion *et al.* (2009) find in a rich panel of US data that the threat of technologically advanced entry spurs innovation incentives in sectors close to the technology frontier, where successful innovation allows incumbents to survive the threat, but discourages innovation in technologically laggard sectors where the threat of entry reduces incumbents' expected rents from innovation. Lee (2009) develops a model that predicts that a firm' incentives to invest in R&D depends primarily on its level of technological competence, i.e. firms with high levels of technological competence are more likely to increase R&D when facing intensifying market competition pressure, whereas firms with low technological competence tent to reduce R&D. Their model is confirmed by a cross section firm level survey data from a number of countries. They also find that such effects are more evident in high-income countries than in low-income countries. Hahn (2010) develops a model to show that the effects of competition on firms' innovation depend on the switching cost. When the switching cost is low, a competing firm is very likely to innovate while a monopoly firm is not. However, if the switching cost is considerably big, competing firms will be very reluctant to innovate while a monopoly firm will still have incentive to innovate.

Therefore, it seems that there are no conclusive theoretical predictions and empirical evidences on the relationship between market competition and innovation, or broadly the effects of competition on firms in the literature. In this paper, we attempt to explore the empirical relationship between competition and innovation from a new viewpoint by investigating how various empirical measures of competition may affect the estimated relationship. To our knowledge, there has no existing study considered this new angle.

3. Issues with Empirical Measures of Competition

Empirical modelling of market competition structure encounters the difficulty of data unavailability on the key variables in the competitive market theory. Firms' sales price and quantity for products and information on product differentiation may have high variation and often are not likely to be available to economists. Because it is often not clear how best to measure many variables suggested by theory or practically not possible to obtain the data, the most interesting empirical relations are those that are robust to plausible variations in measurement methods as well as to variations in specification.

Empirical economists have developed methods to approximate market competition and its impacts on firms. These proxies are based on the theories industrial organisation. Instead of directly modelling the demand and supply situations in the market conditional on firms' ability of influencing market price, these proxies try to measure the observable market structure as consequences of competitions. The rationale behind this method might be that firms competing in the product market are seeking to maximise profits possibly by maximising sales. Therefore, economists attempt to measure the product market share of sales of firms and firms' profit margins, as they are closely related to the consequences of competition.

Researchers have developed two main streams of proxies to measure the market competition, related to the two aspects of observable firm characteristics, namely, market concentration and profit margin. A highly concentrated market is regarded as having low level of competition. By using market concentration to measure competition researchers implicitly assume that market structure is exogenous, i.e. prices decline as concentration falls. However, Raith (2003) argues that when market structure is considered as endogenous, the relationship between concentration and competition is ambiguous. Similarly, a market in which firms earn a large profit margin is assumed not to have high competition. The economic intuition behind it is that the closer to perfect competition an industry is, the more precise price approximates marginal cost, and the closer to zero profit is. Another less popular group of measures is firms' size inequality. Besides having similar measurement problems mentioned above they actually measure firm size distribution in an industry.

Further, besides the fact that these two types of measures using observable information for competition are only imperfect and indirect proxies, they suffer from a further weakness. Taking into account of product differentiation in reality it is difficult for economists to precisely define a product market, again, mainly due to the limitation of available information and practical complexities. Thus economists usually use official industry classifications to roughly define the market. Such process certainly aggregates over a possibly wide range of product markets, and may misunderstand the competition harshness among firms. Official data construction and definitions vary across countries and over time, which adds more intricacies for researches involving comparison. Karuna (2007) shows some evidence of the ambiguous relation between concentration measures and managerial incentives and cautions researchers against using concentration measures as the sole proxy for market competition.

Also industry-specific measures calculate one figure for each industry. By using such measures one implicitly assumes that all the firms in a particular industry have the same competition pressure. However, given the evident heterogeneity across firms, it is not likely to be the case. Such measure should potentially aggregate over heterogeneous firms and overlook the real competition pressure for individual firms.

Moreover, aggregation at even narrowly defined industry level may neglect the characteristics associated with some geographically-fragmented markets in a single industry. This might be especially the case in some large countries in terms of geographical area. Levels of competition can be greatly different within the same product market across different geographical areas. For example, it is often the case that utility supply companies are the local monopolies though the utility industry in the country may have quite a number of similarly-sized firms. To address this problem, some scholars design questionnaire surveys to ask the firm managers directly how many competitors that they think that they have. This will help to identify the local monopolies as long as the survey answers are credible.

Besides these obvious intrinsic measurement errors in proxies for market competition, Schmalensee (1989) calls attention to another point that might mislead the interpretation of competition, and stresses the importance of using robustness checks. Market competition theory has essential implications on the stability of equilibrium. Schmalensee (1989) points out that if departures from long-term equilibrium are correlated with the independent variables in empirical models, a biased relation among long-run equilibria might show. Such correlations are often plausible. For instance, capital intensive industries tend to be concentrated and to have cyclically sensitive profitability. If researchers are interested in the relationship between competition and profitability, firms' capital intensity may cause biases to the observed relationship. Major technological innovations are also likely to disturb equilibrium in an industry. Schmalensee (1989) suggests that this makes it desirable to control for departure from equilibrium particularly in cross-section studies, to check for robustness and to employ long time period panel data.

Therefore, it should be prudent to always apply a number of alternative measures in the empirical studies to verify the robustness. It may also suggest that different measures of competition could indicate different picture of market competition. We use a large dataset from Chinese firms to apply these measures. The reason to choose Chinese data for this study is that the majority of the existing studies related to market structure are on industrialised countries, especially the US and UK. The phenomenal globalisation process has drawn the attention of academia, the business circle and governments to many developing countries such as China. It will be of substantial interests to study the market structure and related features of these economies. In addition, data from developed countries is commonly known to have better quality than data from developing countries. It might be also important to see how well the competition measures in the literature can be applicable and useful using data with more limitations. Also China is a large country in terms of domestic market size and geographical area, and has high degree of openness in terms of international trade. It has a relatively comprehensive industrial structure in the sense that its industrial sector produces almost all the common types of products. Hence, using Chinese data will allow us to go into detailed industrial categorisations that may not be possible in data from some other developing countries, and to make sharper tests on related issues.

4. Empirical Measures of Competitive Industries

Despite the imperfections of the empirical measures of market competition, researchers have to rely on these measures to look into important economic issues connected to market structure. Some popular indicators of market competition are summarised below, followed by some discussions of the pros and cons of them.

1) Number of firms

A straightforward but naïve way to look at competition level is to count the number of competitors in an industry. The obvious advantage is that the firm number information is often available. This method, however, simply ignores the firm heterogeneity in size. It may only be useful when the number of firms in an industry is small. Nonetheless, more or less firms in an industry, regardless of the size distribution of these firms, may still to some extent make some difference to the participants in the industry.

4.1 Measures of Market Concentration

1) Market share

Market share is usually calculated as the share of a firm's sale in the total sales of an industry. Like number of firms market share is an easy measure to obtain but its role is actually unclear. First of all, market share may depend on other factors that are difficult to control, such as potential competition. If there is threat of entry or expansion by existing competitors, firms may spend much effort to deter potential entrants and maintain their market share. In such case firms may face high competition pressure, but their market share may appear unchanged. Secondly, if the denominator of the market share measure is an aggregate level industry sales, such as a 3-digit industry by the standard industrial classification (SIC) classification, then it might be far from representing a correct market. Therefore, such a market share measure may have little value as a cross-section measure of market power. However, Nickell (1996) suggests that if market share is used as a time-series measure of market power, the inaccuracy problems mentioned above would be less serious. This is because the unobservable and uncontrollable factors that may also influence market share are likely to be relatively constant over time, which implies that one might expect to observe some correlation between changes in the measure of market share and changes in a true measure of market power. Thus, Nickell (1996) proposes to use lagged changes of market share as an inverse measure of competition, i.e. $\Delta s_{i,t-k}$, where s is the share of firm i in the industry's total sales, assets or employment, t is time period, k = 1,2,3....

2) Concentration ratio (CR)

CR = $\sum_{i=1}^{n} s_i$, where s_i is the share of the i^{th} largest firm in the total industry sales, assets or employment and n is the n largest firms in the industry. n is often 3, 4, 5 or 8 in the literature (see surveys by Lipczynski, Wilson, and Goddard, 2005; and Schmalensee, 1989)[1]. The higher CR the more concentrated a market possibly is. However, this measure does not take into account the size distribution within these top n firms and the size distribution of the firms outside the n firms. Therefore, CR may only provide a very rough idea of the competition level.

3) Herfindahl-Hirschman index (HHI)

HHI = $\sum_{i=1}^{N} s_i^2$, where s_i is the market share of firm i and N is the total number of firms in the industry. max(HHI) = 1 when there is a single monopoly producer. min(HHI) = 1/N when the industry consists of N equal-sized firms. HHI depicts the market concentration better than the concentration ratio but also requires much more data than CR. Stigler (1964) finds that the HHI outperforms the 4-firm concentration ratio using US data in 1953-1957.

4) Hannah and Kay index:

HK(α) = $\sum_{i=1}^{N} s_i^{\alpha}$, where α is a parameter to be selected. $\alpha > 0$ and $\alpha \neq 1$. s_i is the market share of firm i and N is the total number of firms in the industry. HK(α) index is a weighted market share using market share itself as the weight and is a generalisation of the HHI index. The weight can be defined as $w_i = s_i^{\alpha-1}$. Therefore HK(α) = $\sum_{i=1}^{N} w_i s_i = \sum_{i=1}^{N} s_i^{\alpha-1} s_i = \sum_{i=1}^{N} s_i^{\alpha}$. If $\alpha = 2$, HK(2) = HHI. If $\alpha < 2$, HK(α) attaches relatively more weight to smaller firms and relatively less weight to larger firms, and conversely if $\alpha > 2$. The high HK the lower the competition level is[2].

An advantage of HHI and HK(α) indices is that even if in practice individual data of firm size is not available for the smaller firms, a reasonable approximation of HHI or HK(α) can be obtained using data on the larger firms only (Lipczynski, Wilson, and Goddard, 2005).

However, the various measures of market concentration as proxies for competition level in an industry may have overemphasised the role of concentration, since other dimensions of market competition structure such as entry barrier and product differentiation are overlooked probably due to the difficulty to quantify in empirical research. The usefulness of these measures is also directly dependent on a proper definition of a particular market, which hinges primarily on the notion of substitutability.

Earlier studies as surveyed by Schmalensee (1989) find mixed results of the relationship between market concentration and firm profitability using US and west European countries data. Schmalensee (1989) also notices that most of the studies have not detected sharp differences among the frequently discussed and highly correlated market concentration ratios.

[1] In our empirical measures presented in section 6, we choose n = 4.
[2] In our empirical measures presented in section 6, we choose α = 1.5.

4.2 Measures of size inequality

Most market concentration measures in the previous section do not take the size distribution in an industry into account. The general assumption is that the higher firm size inequality in an industry, the more likely that the competition level is low in the industry. However, these measures are most precisely interpreted as measures of inequality in firm size distribution, as their best known application in household income. They are relatively less popular in the literature with only a handful of studies using them (such as Hannan, 1997; Barla, 2000; de Guevara, *et al.*, 2007; Fedderke, and Szalontai, 2009).

1) Coefficient of variation

COV = $\sqrt{[\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2]/\bar{x}}$, where x_i is measures of firm i's size, $\bar{x}$ is the mean of x, and n is the number of firms in the industry. The expression in [·] is the variance of firm size in an industry with n firms. COV scales the variance by the size mean which avoids the problem of mean-dependence.

2) Gini coefficient

Gini = $\{[\sum_{n=1}^{N}\sum_{i=1}^{n}x_i]/[0.5(N+1)\sum_{i=1}^{N}x_i]\}-1$, where x_i is measures of the size of firm i, n is the largest n firms and N is the total number of firms in the industry [3].

3) Relative mean deviation

RMD = $\frac{1}{2\bar{x}n}\sum_{i=1}^{n}|x_i-\bar{x}|$, where x_i is measures of firm i's size, $\bar{x}$ is the mean of x, and n is number of firms in the industry. RMD shows the extent to which individual firm size in an industry of size n differ from the mean. It suggests the inequality between those above the mean and those below the mean. However, it is insensitive to reflect the inequality between firm size on the same side of the mean.

4) Generalised entropy measures

$GE^{\alpha}=\frac{1}{\alpha^2-\alpha}\sum_{i=1}^{n}[(x_i/\bar{x})^{\alpha}-1]$, where $\alpha\in(-\infty,+\infty)$ and is a parameter that capture the distributional sensitivity. For a large and positive α the GE is sensitive to changes in the distribution that affect the upper tail; for a negative α GE is sensitive to changes in the distribution that affect the lower tail. Depending on the choice of parameter α, popular GE^{α} is also referred to as Entropy Index (α = -1), Mean Log Deviation (MLD) (α = 0), Theil Index (α = 1) and Half (α = 2).

5) Entropy coefficient

[3] In our empirical measures presented in section 6, we choose n = 4.

$E=\sum_{i=1}^{N} s_i \ln(1/s_i)$ is another weighted sum concentration measure. The weights $\ln(1/s_i)$ are inversely related to firms' market shares. E is small for a highly concentrated industry and large for an industry with low concentration. min(E) = 0 when the only monopoly producer has 100 per cent of market share, and max(E) = ln(N) for an industry comprising N equal-sized firms. However, to make it straightforward to compare concentration of industry with different sizes a relative entropy coefficient, which does not depend on the number of firms, is more convenient. RE = $E/\ln(N)=[1/\ln(N)]\sum_{i=1}^{N} s_i \ln(1/s_i)$. Therefore, min(RE) = 0 for a monopoly and max(RE) = 1 for N firms with equal share in an industry.

6) Variance of the logarithms of firm size

VL = $\frac{1}{N}\sum_{i=1}^{N}[\ln(x_i)-\bar{x}]^2$, where $\bar{x}=\frac{1}{N}\sum_{i=1}^{N}\ln(x_i)$ and x_i is measures of the size of firm *i*.

4.3 Measures of Profit Margins

This group of measures try to capture firms' or industries' profit margin, which is expected to have a negative relationship with competition. Researchers have proposed a number of ways to calculate it. These measures are relatively the most sophisticated. They can measure not only industry level competition, but also the competition individual firms may face. Therefore, they can be useful when firm heterogeneity is considered.

1) Monopoly rent $=\frac{1}{N}\sum_{n=1}^{N}$[(profit$_i$ – capital cost$_i$)/value added$_i$] (Nickell, 1996)

where profit = profit before tax$_i$ + depreciation$_i$ + interest payment$_i$; capital cost = capital cost$_i$ × capital stock$_i$. The capital stock$_i$ is the replacement cost of capital stock at current price. capital cost$_i$ = $r+\delta+\lambda\rho$, where r is real interest rate, δ is the rate of depreciation assumed to be constant at 0.04 by Nickell (1996), ρ is the risk premium and λ is a weight ($0\le\lambda\le 1$). r is the annual real gross redemption yield on 2 percent Treasury index linked securities. ρ is firm's average stock market return over the sample period less the average short-term interest rate over the same period. Note that this measure implies a negative relationship between rents and stock market returns. If firms with market power have higher returns, it will generate a negative correlation between rents and market power. Therefore, to apply this measure, the underlying stock market must be an efficient market to ensure that firms with market power do not have higher returns, on average at least.

This measure has to rely on stock market information, which limits the research on listed firms only.

2) Lerner index

Lerner index is originally proposed by Lerner (1934) to measure market power. It has advantages over indicators of market concentration, since it does not rely on a strict definition of product market. However, in practice it is difficult to gather the necessary information on prices and marginal costs to calculate this original Lerner index. In empirical studies, researchers have modified and developed the Lerner index to be more applicable. Bloom and Van Reenen (2007) apply the Lerner index as

$$\text{L} = \frac{1}{N_j - 1} \sum_{\substack{k=1..i-1,i+1..\\ k \in j}}^{N} \left(1 - \frac{profit_k}{sales_k}\right)$$, where N is the number of firms in industry j.[4]

This Lerner index measures the mean of firms' (·) within each industry except firm i's own. Similarly, Aghion *et al.* (2005) also measure this price-cost margin by

$c_{jt} = 1 - \frac{1}{N_{jt}} \sum_{i \in j} L_{it}$, where L = (operating profit – depreciation – provisions – estimated financial cost of capital) / sales, and N_{jt} is the number of firms in industry j in year t.[5]

Aghion *et al.* (2005) assume that the financial cost of capital is a fixed proportion (0.085) of the capital stock, which is measured using the perpetual inventory method. They also alternatively exclude the financial cost from the Lerner measure since it is relatively small and stable over time, and find robust results. A value of 1 of c_{jt} indicates perfect competition while values below 1 suggest some degree of market power.

Gaspar and Massa (2006) refers to the Lerner index as price-cost margin (PCM) and applies the following three measures

PCM_i = operating $profits_i$ (before depreciation, interest, special items and taxes) / ($sales_i$ – cost of goods $sold_i$ – selling, general and administrative $expenses_i$),

$$excessPCM_i = PCM_i - \frac{1}{N} \sum_{i=1}^{N} PCM_i \ ,$$

and industry value-weighted $excessPCM_i = PCM_i - \sum_{i=1}^{N} (PCM_i \times value_i / \sum_{i=1}^{N} value_i)$. N is the total number of firms in the industry. The higher firms' profit to sales ratio relative the industry average is the more market power the firms may have.

3) Product $substitutability_j$ = industry $sales_j$ / industry operating $cost_j$ (Karuna, 2007) [6]

Depending on the available data structure, operating cost, for example, can be computed as the sum of cost of goods sold, selling, general and administrative expense, and depreciation, depletion and amortisation. Or it may be the sum of the cost of materials and production workers' wage. This is also a measure of price-cost margin. Price elasticity of demand has a positive relationship with the extent of product substitutability. Low (high) levels of price-cost margin signify high (low) levels of substitutability, and thus high (low) levels of competition. This measure is much simpler than those above.

---

[4] In our empirical measures presented in section 6, we define $profit_k$ = firms' total profit.

[5] In our empirical measures presented in section 6, we define L = (operational profit – depreciation – 0.06*capital stock using perpetual inventory method) / sales.

[6] In our empirical measures presented in section 6, we define Product $substitutability_j$ = industry $sales_j$ / industry sales costs + sales fees + sales tax and other fees + management fees + wages + employee $benefits_j$

4.4 Other Measures of Competitive Industries

1) Import penetration = $imports_j$ / home $sales_j$ (Nickell, 1996)
   or = ln($import_j$ / domestic $production_j$) (Bloom and Van Reenen, 2007)
   or = $imports_j$ /(domestic $sales_j$ + $imports_j$) (Konings, *et al.*, 2005)

where *j* indexes industries. While most of those measures of competition above only consider the domestic market, import penetration takes into account competition pressure related to international market. This measure, strictly speaking, is a simple proxy of the possible impact of imports on the domestic market and do not really reflect the competition severity in this product market.

2) Survey questions on managers' perception of firms' number of competitors

Nickell (1996) and Bloom and Van Reenen (2007) both use a benchmark of 5 competitors to distinguish high and low levels of competition environment. This method may not be useful if the researcher intends to generate implications at an aggregate level, but it might be more useful if the researcher is interested in perceived competition pressure on firms. It will also isolate the problem of a geographical separated market, in which case firm managers may not consider non-local producers as their competitors.

There have been inconclusive arguments on which measure best captures the market competition. It is often the case that researchers apply a number of alternative measures to verify the robustness. Particularly, measures of domestic and international market competition are put together and/or interacted (e.g. Konings, *et al.*, 2005) to have a more complete measure.

5. Data and Summary Statistics

We use our data to calculate most of the measures discussed above and try to compare the consistency between them using correlation matrix. Where an industry is concerned, we always use the SIC 3-digit industry classification. We and many previous studies stress the importance of using alternative measures of competition to verify the robustness, since the empirical measures are intrinsically imperfect. To select the appropriate measures to check the robustness one must first know the comparison and consistency between these measures. To achieve this goal our rational and approach is to use the correlations and their significance, which effectively we compare how consistently each measure ranks the competition level of industries.

If an industry is considered as with high (low) level of competition by a number of alternative measures, this industry probably has high (low) level of competition indeed. Similarly if a group of alternative competition measures have positive and significant correlations with each other, it may suggest that these measures are relatively consistent in ranking competition level. If they have insignificant correlations, it may mean low consistency between the measures. Negative and significant correlations would mean opposite indications of competition from the measures.

Table 1 lists the measures that we calculate from our data. Due to our data availability and the high similarity between a few measures we choose the measures as listed in table 1. All the market concentration measures and firm size inequality measures are calculated using three variables, i.e. number of employees, sales and total assets. Some basic statistics of each measure and their relationship with the level of competition are presented. A direct comparison of their figures would not tell meaningful stories. Except the number of firms and Lerner indices all other measures have negative relationships with the level of competition. Therefore, when we estimate the correlations and regressions, we invert the firm number and Lerner index measures by adding negative signs to all the values.

The correlation matrix is presented in table 2. The correlations are calculated using sales of each market concentration and size equality measures, since sales is more important in market competition for firms than number of employees or assets. Indeed, firms' sales, total assets and employment measure very different aspects of firm size, and their relative comparison with others vary considerably across industries. Intuitively and generally, firms' sales may matter more to their competitors than their assets or employee numbers, as their main competition with each other would eventually be how much profit they can generate out of their sales. For particular industries at certain time period possessing important assets and skilled work force would mean critical future increase in firm value and profits. Also firms in capital intensive industries may value assets more than labour force, whereas labour intensive firms may consider work force more important than fixed assets. Therefore, the choice among the three variables should depend on the specific industries that researchers are interested in. Otherwise, different choice of the variable will not tell the competition level quite in the same way when applying these popular market concentration measures. The correlations that are calculated using employment and total assets are reported in appendix.

Firm number does not seem to be similar with any other measures due to its simplicity. It is still an often used measure due to data requirement. The correlations in column 1 of table 2 show that it has positive and significant correlations with all the market concentration measures and Lerner_A index, negative and significant correlations with pcm measures and insignificant correlation with the rest. Therefore, firm numbers in an industry does seem to be consistent with market concentration, but not with firm size inequality, and has no clear relationship with profit margin measure in general.

5.1 Market concentration measures

These market concentration related measures are often calculated using firms' sales, total assets or the number of employees in the literature depending on the information availability and research purposes. Table 2 columns 2 to 5 and rows (b) to (g) show the correlations between market concentration measures as well as (negative) firm numbers. We use the lagged changes of market sales share following Nickell (1996), and four-firm CR due to its popularity. We use 1.5 as the α parameter for HK, which gives more weights to small firms. All the correlations are positive and highly significant at less than 1% level, which suggest that these measures are greatly consistent with each other. Particularly, the correlation coefficients between the concentration measures are all above 0.8 and the largest coefficient is 0.979 between CR(4) and HK(1.5).

5.2 Size inequality measures

The inequality measures are mainly initiated in the income inequality studies in the literature. Market competition literature borrows these measures and ideas to use as alternatives. Strictly speaking, they measure the inequality of firm size distribution in an industry rather than competition itself. Economists assume that firms' size inequality has a negative relationship with the level of competition in an industry when using inequality measures to indicate competition. This assumption is generally sensible, but not perfect.

Table 2 columns 6 to 12 and rows (f) to (l) show the correlations among the sales inequality measures. Again the consistency within inequality measures is high. Most of the correlations are positive and highly significant and three correlations (i.e. between entropy index and COV, MLD and Theil index respectively) are insignificant. Most of the significant coefficients are large, with the highest between Half and COV at 0.985.

5.3 Measures of profit margin

The above measures are all industry specific which do not vary across firms within the same industry. It potentially overlooks firm heterogeneity in terms of the actual effects of competition pressure on individual firms. The measures of profit margin can be firm specific or industry specific depending on the purpose of research. Firms' profit margin can be aggregated to obtain an industry level figure to signify the possible general level of competition that all firms in the industry potentially face. To make these measures comparable with other industry level measures in this study we use industry level profit margins, which is the mean of firms' profit margin figures.

We present our calculations and comparison of Bloom and Van Reenen (2007)'s Lerner index, Lerner_B, Aghion *et al.* (2005)'s Lerner index, Lerner_A, Gaspear and Massa (2006)'s price-cost margins and Karuna (2007)'s product substitutability in table 2 columns 13 to 17 and rows (m) to (q). When we calculate Lerner_B we use firms' total profit. Due to our data limitations we do not have the provisions variable over the whole sample period to calculate Lerner_A, and therefore we omit this term. We have also tried to use other replacements such as firms' income tax, provision tax or total profit before tax, but the one omitting the term seems to be most reasonable comparing to the calculations in the literature. We assume the financial cost of capital is a fixed proportion of 0.06 of the capital stock using the perpetual inventory method. We follow Gaspar and Massa (2006)'s method to calculate the industry average PCM, pcm_mean, and weighted excessive PCM, pcmw_ex.

The correlations among this group of measures are mixed. Particularly, Lerner_B and Lerner_A have negative and significant correlations with each other. Only the pcm_mean and pcmw_ex have a large correlation of 0.995 with less than 1% of significance. pcm_mean has positive and significant correlations with both Lerner indices. pcmw_ex only has positive and significant correlation with Lerner_A and insignificant correlation with Lerner_B. Product substitutability has no significant correlation with any other profit margin measures at all. Such results may indicate that these measures are inconsistent in evaluating competition level. Though the ideas of these measures to capture firms' profit margin is similar, our correlation results suggest that they do not imply the level of competition quite similarly.

5.4 Comparison between groups of competition measures

Table 2 also presents correlations between measures across groups. Columns 2 to 5 and rows (f) to (l) compare the correlations between market concentration and inequality measures.

Entropy index has a negative and 10% significant correlation with CR(4), and negative but insignificant correlations with all other concentration measures. This is also possible because entropy index with a negative α gives more weight to the changes to the lower tail of the distribution, but those concentration measures mainly measure the upper tail of the sales distribution. All the rest of correlations between each pair of concentration and inequality measures are all positive and significant at less than 1% level. These correlations may suggest that the measurement consistency between market concentration and inequality measures is mostly high.

We next compare the consistency between market concentration measures and profit margin measures, as shown in columns 2 to 5 and rows (m) to (q). Generally, for these two groups the consistency is low, since there are many negative and significant, or insignificant correlation coefficients. Lerner_B has positive and significant correlation with HHI and insignificant correlations with the rest. Lerner_A has positive and significant correlations with all the concentration measures. Pcm_mean and pcmw_ex have negative correlations with all the concentration measures, though insignificant with HHI. Product substitutability has no significant correlation with any of the concentration measures. This means that market concentration and profit margin could suggest competition level in the opposite way.

Finally, we look at the correlations between measures of inequality and profit margin as shown in columns 6 to 12 and rows (m) to (q). The consistency between these two groups also appears to be low. Many of the correlations are p and significant. Lerner_B has insignificant correlations with COV and Half, and positive and significant correlations with all others. Lerner_A has negative and significant correlation with entropy index, and positive and significant correlations with the rest. Again pcm_mean and pcmw_ex have insignificant correlations with MLD, and negative and significant correlations with the rest. Product substitutability has positive and significant correlation with Gini, and insignificant correlation with all others. Thus, it seems that inequality measures and profit margin measures do not picture the relative competition level of an industry in quite a similar way.

Overall, measures within the concentration group and inequality group and measures between these two groups are consistent with one another, but measures in the profit margin group and between it and the other two groups are not. Nonetheless, these observations are based on the relative ranks of competition level that each measure indicates.

6. Competition and Firm Innovation

We next test the relationship between market competition and innovation. Aghion *et al.* (2005) show both theoretically and empirically that competition measured by Lerner index has an inverted-U shape relationship with innovation measured by number of patents and R&D expenditure. We take a step forward by looking into innovation in greater detail, i.e. specifying innovation input, output and efficiency, and looking into the most popular alternative measures of competition, i.e. both Lerner index and concentration measures. We do not use inequality measures in these tests due to their low popularity in the literature and relative less usefulness.

We start with basic specification as follows

$$innovation_{jt} = \alpha + \beta_1 comp_{jt} + \beta_2 comp_{jt}^{\ 2} + \nu_t + \nu_j + \varepsilon_{jt} \qquad (1)$$

where $innovation_{jt}$ includes innovation input, which is industry *j*'s R&D expenditure over sales ratio, innovation output, which is industry *j*'s new product sales over total sales ratio, and innovation efficiency, which is industry *j*'s new product sales to R&D expenditure ratio at time *t*. comp_{jt} is industry level measures of competition, which include Lerner_A, Lerner_B, HHI_isv, HK(1.5)_isv an CR_isv that we summarised in section 5. This simple model is to capture the quadratic effects of competition on innovation.

Estimation results using Tobit regression are presented in table 3 to 5 for three aspects of innovation respectively. A universal feature of the results is that all the estimations suggest an inverted-U shape relationship between measures of competition and measures of innovation, which is in line with the prediction of Aghion *et al.* (2005). In each model, competition measured by five indicators has positive and significant effects on innovation and the squared term of competition has negative and significant effects on innovation. They suggest that when competition is intermediate innovation tend to be high, and when competition is either low or high innovation tend to be relatively low.

We then further verify the inverted-U shape relationship by estimating models with various additional controls and using alternative regression methods. The results are reported in table 6. In these models we choose Lerner_A as our main measure of competition, since it is relatively the most sophisticated among those summarised above. In columns 1 to 3 we use Tobit regressions, columns 4 and 5 bivariate Tobit regressions and columns 6 to 8 trivariate Tobit regressions. For the innovation input models in columns 1, 4 and 6 besides Lerner_A and its squared term we also include the lagged excess pcm as an additional measure of competition. We further add some firm characteristics control variables, i.e. firms lagged size, age and lagged labour quality, which are closely related to firms' likelihood to innovate. The results in columns 1, 4 and 6 show that all the explanatory variables are highly significant and an inverted-U shape relationship between Lerner_A and innovation input is valid. These additional explanatory variables suggest that generally larger and older firms with high labour quality are more likely to spend more on R&D.

The innovation output models in columns 2 and 5 are similar with the innovation input models with an additional lagged R&D intensity variable. Again the results show highly significant coefficients for all the independent variables and valid inverted-U shapes. However, the specification is not valid for trivariate Tobit regression. We only reduce the model to add lagged R&D intensity alone in column 7, which shows an inverted-U shape as well. Larger and older firms with high labour quality and high R&D intensity are more likely to generate new products out of their R&D spending.

Finally, for the innovation efficiency models in columns 3 and 8, we add industries' lagged private share and lagged foreign share as control variables. Still the relationship between Lerner_A and innovation efficiency is inverted-U shaped. Interestingly, private share has positive effects on innovation efficiency, but foreign share has negative effects. Given our Chinese data such results are likely. Private firms are relative disadvantaged in the market, as they do not enjoy state support as state owned firms do and neither receive preferential policy treatment as foreign firms do. Thus, it is understandable that they may have strong incentives to put extra efforts to make the most out of their R&D spending. On the other hand, foreign firms do not seem to be efficient in term of R&D.

7. Conclusions

In this paper we summarise and discuss various popular empirical measures of market competition. Since the true market competition is unobservable, empirical measures can only try to approximate competition from some observable features of market and firms. Therefore, given the obvious imperfections of the market competition measures the preferred and convincing method is to apply alternative measures and verify the robustness of any economic conclusions. The aim of this paper is to compare and analyse the consistency of the popular empirical measures of market competition and provide researchers with an overview of these measures, and to test the relationship between competition and innovation with various measures.

Using a large panel of Chinese firm data we calculate measures of market competition. Generally, the consistency indicated by correlations within groups of measures, i.e. market concentration, firm size inequality and profit margin, is mainly high. The consistency between market concentration group and size inequality group is also high, but the consistencies between profit margin group and the other two groups are low. Though these competition measures all try to capture some market and firm features that are related to competition, they give very different pictures of competition.

We then select the most popular and sophisticated competition measures to test the relationships between competition and innovation, and find inverted-U shape relationship as predicted by Aghion et al. (2005). This relationship seems to exist for innovation input, output and efficiency, and is robust to some firm characteristics, regression methods and model specifications.

Our results highlight that it is important to check the robustness of the impact of competition. It is also critical that researchers select the appropriate measures of competition for their research, as the choice of the measure may have significant impact on the research. This research may help researchers to make their selection decisions. It also has important and robust implications on innovation policies. While competition may encourage innovation to some extent, excessive competition could discourage innovation too.

References

Aghion, P., Bloom, N., Blundell, R., Griffith, R. and Howitt, P. (2005), “Competition and innovation: An inverted-U relationship”, *The Quarterly Journal of Economics*, Vol. 120, No. 2, 701-28.

Aghion, P., Blundell, R., Griffith, R., Howitt, P. and Prantl, S. (2009), “The Effects of Entry on Incumbent Innovation and Productivity”, *Review of Economics and Statistics*, Vol. 91, No. 1, 20-32.

Arrow, K. J. (1962), “Economic Welfare and the Allocation of Resources for Innovation”, Ed. Nelson, R., *The Rate and Direction of Inventive Activity: Economic and Social Factors*, Princeton University Press, Princeton.

Barla, P. (2000), "Firm Size Inequality and Market Power", *International Journal of Industrial Organization,* Vol. 18, No.5, 693-722.

Bloom, N. and Van Reenen, J. (2007), “Measuring and Explaining Management Practices across Firms and Countries”, *The Quarterly Journal of Economics*, Vol. 122, No. 4, 1351-408.

Blundell, R., Griffith, R. and Van Reenen J. (1995), “Dynamic Count Data Models of Technological Innovation”, *The Economic Journal*, Vol. 105, No. 429, 333-44.

Blundell, R., Griffith, R. and Van Reenen J. (1999), “Market Share, Market Value and Innovation in a Panel of British Manufacturing Firms”, *Review of Economic Studies*, Vol. 66, No. 3, 529-54.

Boone, J. (2000), “Competitive Pressure: the Effects on Investment in Product and Process Innovation”, *Rand Journal of Economics*, Vol. 31, No. 3, 549-69.

Caves, R. E. (1992), (Ed) *Industrial Efficiency in Six Nations*, MIT Press, Cambridge, Massachusetts.

De Guevara, J. F., Maudos, J. and Perez, F. (2007), "Integration and Competition in the European Financial Markets", *Journal of International Money and Finance* Vol. 26, No.1, 26-45.

Disney, R., Haskel, J. and Heden, Y., (2003), “Restructuring and Productivity Growth in UK Manufacturing”, *Economic Journal*, Vol. 113, No. 489, 666-94.

Dixit, A. K. and Stiglitz, J. E. (1977), “Monopolistic Competition and Optimum Product Diversity”, *American Economic Review*, Vol. 67, No. 3, 297-308.

Fedderke, J. and Szalontai, G. (2009), "Industry Concentration in South African Manufacturing Industry: Trends and Consequences, 1972-96", *Economic Modelling,* Vol. 26, No. 1, 241-50.

Gaspar, J. M. and Massa, M. (2006), “Idiosyncratic Volatility and Product Market Competition”, *Journal of Business*, Vol. 79, No. 6, 3125-52.

Geroski, P. A. (1990), “Innovation, Technological Opportunities, and Market Structure”, *Oxford Economic Papers*, Vol. 42, No.3, 586-602.

Green, A. and Mayes, D. (1991), “Technical Inefficiency in Manufacturing-Industries”, *The Economic Journal*, Vol. 101, No. 406, 523-38.

Hahn, S. (2010), “Competition, Comparison and Innovation”, *Economics Letters*, Vol. 106, No. 2, 122-4.

Hannan, T. H. (1997), "Market Share Inequality, the Number of Competitors, and the HHI: An Examination of Bank Pricing", *Review of Industrial Organization*, Vol. 12, No. 1, 23-35.

Hart, O. (1983), “The Market Mechanism as an Inventive Scheme”, Bell *Journal of Economics*, Vol. 14. No. 2, 366-82.

Karuna, C. (2007), “Industry Product Market Competition and Managerial Incentives”, *Journal of Accounting and Economics*, Vol. 43, No. 2-3, 275-97.

Koning, J., Van Cayseele, P. and Warzynski, F. (2005), “The Effects of Privatisation and Competitve Pressure on Firms’ Price-Cost Margins: Micro Evidence from Emerging Economies”, *The Review of Economics and Statistics*, Vol. 87, No. 1 124-34.

Lee, C.-Y. (2005), “A New Perspective on Industry R&D and Market Structure”, *Journal of Industrial Economics*, Vol. 53, No. 1, 101-22.

Lee, C.-Y. (2009), “Competition Favours the Prepared Firm: Firms’ R&D Responses to Competitive Market Pressure”, *Research Policy*, Vol. 38, No.5, 861-70.

Lerner, A. P. (1934), "The Concept of Monopoly and the Measurement of Monopoly Power", *The Review of Economic Studies*, Vol. 1, No. 3, 157-75.

Lipczynski, J. Wilson, J.J., and Goddard, J. (2005), *Industrial Organization: Competition, Strategy, Policy*, Pearson Education UK.

Mankiw, N. G. and Whinston, M. D. (1986), “Free Entry and Social Inefficiency”, *The RAND Journal of Economics*, Vol. 17, No. 1, 48-58.

Meyer, M. A. and Vickers, J. (1997) “Performance Comparisons and Dynamic Incentives”, *Journal of Political Economy*, Vol. 105, No.3, 547-81.
Nalebuff, B. J. and Stiglitz, J. E. (1983), “Information, Competition and Markets”, *The American Economic Review*, Vol. 73, No. 2, 278-83.
Nickell, S. J. (1996), “Competition and Corporate Performance”, *The Journal of Political Economy*, Vol. 104, No.4, 724-46.
Nickell, S. J., Nicolitsas, D. and Dryden, N. (1997), “What Makes Firms Perform Well?”, *European Economic Review*, Vol.41, No. 3-5, 783-96.
Nickell, S. J., WadHwani, S. B. and Wall, M. (1992), “Productivity Growth in United Kindom Companies, 1975-1986”, *European Economic Review*, Vol. 36, No. 5, 1055-85.
Raith, M. (2003), “Competition, Risk and Managerial Incentives”, *The American Economics Review*, Vol. 93, No. 4, 1425-36.
Romer, P. M. (1990), “Endogenous Technological Change”, *Journal of Political Economy*, Vol. 98, No. 5, 71-102.
Schmalensee, R. (1989), “Inter-Industry Studies of Structure and Performance”, *Handbook of Industrial Organisation*, Vol. 2, Ed. Schmalensee, R. and Willing, R. D., Elsevier Science, North Holland.
Schmidt, K. M. (1997), “Managerial Incentives and Product Market Competition”, *Review of Economic Studies*, Vol. 64, No. 2, 191-213.
Schumpeter, J. A. (1934), *The Theory of Economic Development*, Harvard University Press, Cambridge.
Schumpeter, J. A. (1943), *Capitalism, Socialism, and Democracy*, Allen and Unwin, London.
Stigler, G. J. (1964), “A Theory of Oligopoly”, *Journal of Political Economy*, Vol. 77, No. 1, 44-61.
Tang, J. (2006), “Competition and Innovation Behaviour”, *Research Policy*, Vol. 35, No. 1, 68-82.

Table 1: Summary Statistics of Competition Measures

			Parameter	Definition	Measures	min	max	mean	median	Relationship between measures and competition
Firm number				Firm number		12	8819	1170	879	+
mkshr_emp	measure of market concentration	Nickell (1996)		Lagged changes of market share	employment	0.000	0.083	0.004	0.002	-
mkshr_isv					sales	0.000	0.083	0.004	0.002	
mkshr_ta					total assets	0.000	0.083	0.004	0.002	
CR_emp		CR		Sum of the market shares of the 4 largest firms	employment	0.005	0.899	0.091	0.055	-
CR_isv					sales	0.003	0.972	0.072	0.034	
CR_ta					total assets	0.010	0.936	0.131	0.087	
HHI_emp		HHI		Sum of squared market share	employment	0.000	0.315	0.012	0.005	-
HHI_isv					sales	0.000	0.449	0.013	0.004	
HHI_ta					total assets	0.001	0.485	0.020	0.009	
HK_emp		HK	α = 1.5	Sum of market share to the power of 1.5	employment	0.016	0.525	0.077	0.059	-
HK_isv					sales	0.016	0.630	0.078	0.060	
HK_ta					total assets	0.025	0.647	0.098	0.079	
cov_emp	Measure of size inequality	COV		Coefficient of variation	employment	0.418	2.653	1.531	1.540	-
cov_isv					sales	0.596	2.392	1.389	1.367	
cov_ta					total assets	0.725	3.578	2.123	2.155	
gini_emp		Gini		Gini coefficient	employment	0.217	0.771	0.582	0.586	-
gini_isv					sales	0.314	0.809	0.624	0.620	
gini_ta					total assets	0.402	0.838	0.689	0.693	
rmd_emp		RMD		Relative mean deviation	employment	0.183	0.651	0.441	0.440	-
rmd_isv					sales	0.277	0.656	0.482	0.476	
rmd_ta					total assets	0.322	0.707	0.539	0.539	
entropy_emp			α = -1	Entropy index	employment	0.420	5.592	1.716	1.514	-
entropy_isv					sales	2.79	1465.99	281.08	175.36	
entropy_ta					total assets	0.736	29.739	4.766	3.356	
mld_emp			$\alpha = 0$	Mean log deviation	employment	0.180	1.357	0.668	0.644	-
mld_isv					sales	0.505	2.463	1.180	1.136	
mld_ta					total assets	0.412	1.904	1.028	0.994	
theil_emp			$\alpha = 1$	Theil index	employment	0.110	1.278	0.659	0.660	-
theil_isv					sales	0.218	1.370	0.696	0.677	
theil_ta					total assets	0.299	1.739	1.001	1.012	
half_emp			α=2	Half	employment	0.087	3.463	1.208	1.186	-
half_isv					sales	0.178	2.859	0.986	0.934	
half_ta					total assets	0.263	6.396	2.332	2.320	
Lerner_B	Measure of profit margin	Bloom & VanReenen (2007)		Lerner index		0.692	0.991	0.870	0.869	+
Lerner_A		Aghion et al. (2005)				0.831	0.992	0.968	0.970	
pcm_mean		Gaspar & Massa (2006)		Price cost margin = operating profit / sales		-0.181	0.100	0.006	0.014	-
pcmw_ex				Weighted excessive pcm		-0.182	0.098	0.008	0.017	
Product substitutability (ps2)		Karuna (2007)		= Sales / (sales costs+sales fees+sales tax and other fees+management fees+wages+employee benefits)		0.658	20.069	1.403	1.052	-

**Notes*:

Table 2: Correlations between Competition Measures – Sales

		Firm number	Market share	CR	HHI	HK(1.5)	COV	Gini	RMD	Entropy (α = -1)	MLD (α = 0)	Theil (α = 1)	Half (α = 2)	Lerner_B	Lerner_A	pcm_mean	pcmw_ex	Product substitutability
			sales															
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
Firm number	(a)	1.000																
Market share	(b)	0.320***	1.000															
CR	(c)	0.381***	0.950***	1.000														
HHI	(d)	0.247***	0.955***	0.905***	1.000													
HK(1.5)	(e)	0.480***	0.954***	0.979***	0.912***	1.000												
COV	(f)	0.062	0.272***	0.405***	0.370***	0.344***	1.000											
Gini	(g)	0.040	0.229***	0.322***	0.323***	0.284***	0.866***	1.000										
RMD	(h)	0.082	0.289***	0.353***	0.370***	0.336***	0.743***	0.952***	1.000									
Entropy (α=-1)	(i)	-0.022	-0.105	-0.125*	-0.096	-0.104	-0.071	0.161**	0.204***	1.000								
MLD (α = 0)	(j)	0.072	0.244***	0.299***	0.294***	0.286***	0.599***	0.828***	0.899***	0.337***	1.000							
Theil (α = 1)	(k)	0.069	0.250***	0.367***	0.330***	0.321***	0.917***	0.954***	0.920***	0.080	0.850***	1.000						
Half (α = 2)	(l)	0.079	0.276***	0.407***	0.370***	0.356***	0.985***	0.823***	0.719***	-0.087	0.601***	0.910***	1.000					
Lerner_B	(m)	0.033	0.096	0.014	0.153**	0.082	-0.064	0.180**	0.266***	0.436***	0.485***	0.135*	-0.059	1.000				
Lerner_A	(n)	0.157**	0.231***	0.268***	0.229***	0.239***	0.312***	0.168**	0.143*	-0.170**	0.154**	0.292***	0.362***	-0.179**	1.000			
Pcm_mean	(o)	-0.129*	-0.124*	-0.218***	-0.090	-0.182**	-0.249***	-0.161**	-0.118*	0.174**	0.048	-0.156**	-0.248***	0.597***	0.182**	1.000		
Pcmw_ex	(p)	-0.128*	-0.127*	-0.221***	-0.097	-0.186***	-0.241***	-0.153**	-0.114	0.179**	0.057	-0.147**	-0.243***	0.605	0.274***	0.995***	1.000	
product substitutability	(q)	0.040	-0.022	-0.003	-0.018	0.008	0.094	0.123*	0.113	-0.050	0.051	0.099	0.079	-0.014	0.015	-0.045	-0.054	1.000

Notes*: In column (2) to (12) we use firms' sales to calculate measures of market concentration and inequality. *, **, and * indicate significance at the 10%, 5%, and 1% level respectively.

Table 3: Innovation Input – R&D/sales

Tobit	Lerner_A	Lerner_B	HHI_isv	HK_isv1.5	CR_isv
	(1)	(2)	(3)	(4)	(5)
Comp	0.535**	0.066***	0.123***	0.065***	0.025***
	(0.236)	(0.008)	(0.007)	(0.003)	(0.002)
Comp_sqr	-0.321***	-0.051***	-0.156***	-0.120***	-0.035***
	(0.122)	(0.005)	(0.02)	(0.011)	(0.005)
Observations	1189655	1086461	1198483	1198483	1198474
Innovation-competition relationship	∩	∩	∩	∩	∩

Notes: The dependent is R&D expenditure over sales ratio in all the regressions. Independent variables are the 5 selected competition measures and their squares listed in the top row. Standard errors are in parentheses. A constant term is included in all the regressions. Lerner_A is the Lerner index in Aghion *et al*. (2005) and Lerner_B is the Lerner index in Bloom & Van Reenen (2007). HHI_isv, HK_isv1.5 and CR_isv are calculated using industry sales. *** indicates p<0.01, ** p<0.05, and * p<0.1.

Table 4: Innovation Output – new product sales/total sales

Tobit	Lerner_A	Lerner_B	HHI_isv	HK_isv1.5	CR_isv
	(1)	(2)	(3)	(4)	(5)
Comp	17.410**	3.606***	3.410***	2.256***	0.725***
	(8.629)	(0.357)	(0.271)	(0.102)	(0.070)
Comp_sqr	-10.620**	-2.580***	-4.727***	-4.528***	-1.323***
	(4.451)	(0.221)	(0.839)	(0.435)	(0.187)
Observations	1246276	1371250	1383765	1383765	1383757
Innovation-competition relationship	∩	∩	∩	∩	∩

Notes: The dependent is new product sales over total sales ratio in all the regressions. Please also see notes to table 3.

Table 5: Innovation Efficiency – new product sales/R&D

Tobit	Lerner_A	Lerner_B	HHI_isv	HK_isv1.5	CR_isv
	(1)	(2)	(3)	(4)	(5)
Comp	681.3***	47.84***	26.19***	12.76***	3.379***
	(152.8)	(6.178)	(5.566)	(1.567)	(1.243)
Comp_sqr	-359.4***	-31.89***	-46.81***	-21.27***	-5.307*
	(79.160)	(3.817)	(15.25)	(5.816)	(3.001)
Observations	131597	131342	131597	131597	131593
Innovation-competition relationship	∩	∩	∩	∩	∩

Notes: The dependent is new product sales over R&D expenditure ratio in all the regressions. Please also see notes to table 3.

Table 6: Effects of Competition on Innovation – Alternative Models and Methods

Model	Tobit			Bivariate Tobit		Trivariate Tobit		
Dependent variable	Innovation Input	Innovation Output	Innovation Efficiency	Innovation Input	Innovation Output	Innovation Input	Innovation Output	Innovation Efficiency
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lerner_A	1.584***	144.0***	878.7***	1.597***	141.0***	1.436***	95.64***	694.1***
	(0.304)	(15.54)	(176.9)	(0.304)	(15.29)	(0.301)	(14.95)	(115.5)
Lerner_Asqr	-0.900***	-77.74***	-464.7***	-0.906***	-76.45***	-0.823***	-52.53***	-366.1***
	(0.158)	(8.070)	(91.69)	(0.158)	(7.942)	(0.156)	(7.764)	(59.92)
Excess pcm $_{t-1}$	0.019***	0.331***		0.019***	0.424***	0.018***		
	(0.000)	(0.015)		(0.000)	(0.015)	(0.000)		
Size $_{t-1}$	0.004***	0.116***		0.004***	0.130***	0.004***		
	(0.000)	(0.001)		(0.000)	(0.001)	(0.000)		
Age	5.28e-05***	0.003***		5.28e-05***	0.003***	3.96e-05***		
	(2.76e-06)	(0.000)		(2.76e-06)	(0.000)	(2.38e-06)		
Labour quality $_{t-1}$	0.016***	0.210***		0.016***	0.304***	0.014***		
	(0.000)	(0.008)		(0.000)	(0.008)	(0.000)		
R&D intensity $_{t-1}$		23.91***			6.352***		5.547***	
		(0.239)			(0.286)		(0.144)	
Private share $_{t-1}$			0.216***					0.017
			(0.075)					(0.030)
Foreign share $_{t-1}$			-2.317***					-0.395***
			(0.156)					(0.063)
Observations	670948	652345	101777	670948	670948	709441	709441	709441

Notes:

Table A1: Correlations between Competition Measures – Employment

		firm number	Market share	CR	HHI	HK(1.5)	COV	Gini	RMD	Entropy index	Mean log deviation ($\alpha = 0$)	Theil index	Half	Bloom & VanReenen 2007	Aghion *et al.* (2005)	Gaspar & Massa (2006)		Karuna (2007), product substitutability
			employment													pcm_mean	pcmw_ex	p2
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
firm number	(a)	1.000																
Market share	(b)	0.321***	1.000															
CR	(c)	0.451***	0.916***	1.000														
HHI	(d)	0.292***	0.947***	0.922***	1.000													
HK(1.5)	(e)	0.518***	0.923***	0.980***	0.922***	1.000												
COV	(f)	-0.066	-0.092	0.087	0.101	0.034	1.000											
Gini	(g)	-0.002	0.014	0.095	0.160**	0.107	0.777***	1.000										
RMD	(h)	0.025	0.056	0.110	0.187***	0.147**	0.707***	0.982***	1.000									
Entropy index (α=-1)	(i)	0.129*	0.247***	0.204***	0.314***	0.291***	0.217***	0.640***	0.721***	1.000								
Mean log deviation ($\alpha = 0$)	(j)	0.073	0.144**	0.179**	0.270***	0.236***	0.604***	0.917***	0.963***	0.849***	1.000							
Theil index (α = 1)	(k)	-0.012	-0.019	0.096	0.153**	0.105	0.889***	0.946***	0.939***	0.564***	0.889***	1.000						
Half (α = 2)	(l)	-0.039	-0.077	0.111	0.122*	0.073	0.979***	0.716***	0.662***	0.207***	0.588***	0.871***	1.000					
Bloom & VanReenen 2007	(m)	0.033	0.098	0.059	0.131*	0.077	0.145**	0.170**	0.187***	0.325***	0.251***	0.207***	0.129*	1.000				
Aghion *et al.* (2005)	(n)	0.157**	0.229***	0.251***	0.228***	0.241***	-0.047	-0.078	-0.069	-0.172**	-0.103	-0.077	-0.041	-0.179**	1.000			
Gaspar & Massa	(o)	-0.129*	-0.123*	-0.083	-0.049	-0.125*	0.238***	0.137*	0.111	0.081	0.096	0.191***	0.189***	0.597***	0.182**	1.000		
(2006)	(p)	-0.128*	-0.126*	-0.088	-0.056	-0.131*	0.237***	0.140**	0.113	0.077	0.093	0.190***	0.186***	0.605***	0.274***	0.995***	1.000	
Karuna (2007), product substitutability	(q)	0.040	-0.022	0.023	0.028	0.049	0.288***	0.158**	0.158**	0.068	0.157**	0.242***	0.330***	-0.014	0.015	-0.045	-0.054	1.000

Notes*: In column (2) to (12) we use firms' number of employees to calculate measures of market concentration and inequality. *, **, and * indicate significance at the 10%, 5%, and 1% level respectively.

Table A2: Correlations between Competition Measures – Total Assets

		firm number	Market share	CR	HHI	HK(1.5)	COV	Gini	RMD	Entropy index	Mean log deviation ($\alpha = 0$)	Theil index	Half	Bloom & VanReenen 2007	Aghion *et al.* (2005)	Gaspar & Massa (2006)		Karuna(2007), product substitutability
			Total assets													pcm_mean	pcmw_ex	p2
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
firm number	(a)	1.000																
Market share	(b)	0.330***	1.000															
CR	(c)	0.499***	0.866***	1.000														
HHI	(d)	0.294***	0.916***	0.855***	1.000													
HK(1.5)	(e)	0.512***	0.909***	0.952***	0.925***	1.000												
COV	(f)	-0.239***	-0.217***	0.006	0.026	-0.070	1.000											
Gini	(g)	-0.138*	-0.102	-0.026	0.088	0.017	0.675***	1.000										
RMD	(h)	-0.101	-0.075	-0.034	0.081	0.038	0.564***	0.975***	1.000									
Entropy index (α=-1)	(i)	0.068	0.247***	0.150**	0.252***	0.232***	0.017	0.413***	0.466***	1.000								
Mean log deviation ($\alpha = 0$)	(j)	-0.001	0.093	0.080	0.193***	0.181**	0.352***	0.854***	0.921***	0.663***	1.000							
Theil index (α = 1)	(k)	-0.182**	-0.207***	-0.058	0.007	-0.048	0.845***	0.933***	0.901***	0.293***	0.753***	1.000						
Half (α = 2)	(l)	-0.212***	-0.197***	0.025	0.038	-0.040	0.982***	0.611***	0.514***	0.014	0.333***	0.818***	1.000					
Bloom & VanReenen 2007	(m)	0.033	0.109	-0.019	0.120*	0.065	-0.128*	0.046	0.077	0.124*	0.139*	-0.003	-0.132*	1.000				
Aghion *et al.* (2005)	(n)	0.157**	0.205***	0.287***	0.148*	0.123	-0.010	-0.162**	-0.175**	-0.177**	-0.198***	-0.131*	0.003	-0.179**	1.000			
Gaspar&	(o)	-0.129*	-0.126*	-0.119*	-0.029	-0.124*	0.156**	0.070	0.065	-0.116	0.020	0.135*	0.119*	0.597***	0.182**	1.000		
Massa (2006)	(p)	-0.128*	-0.128*	-0.124*	-0.040	-0.132*	0.149**	0.065	0.059	-0.131*	0.008	0.128*	0.110	0.605***	0.274***	0.995***	1.000	
Karuna (2007), product substitutability	(q)	0.040	-0.023	0.023	0.018	0.053	0.187***	0.184**	0.191***	0.090	0.188***	0.225***	0.202***	-0.014	0.015	-0.045	-0.054	1.000

Notes*: In column (2) to (12) we use firms' total assets to calculate measures of market concentration and inequality. *, **, and * indicate significance at the 10%, 5%, and 1% level respectively.