

Cluster-Based Industrialization in China: Financing and Performance

Cheryl Long
Colgate University

Xiaobo Zhang
International Food Policy Research Institute

September 2009

Abstract

Using firm-level data from China's two recent censuses (Industry Census 1995 and Economic Census 2004) and a new measure of clustering (industry proximity), we show that China's rapid industrialization is marked by increased clustering. A higher degree of clustering is further shown to be associated with greater export growth and higher total factor productivity. We also find supporting evidence that clustering helps ease the credit constraints facing many small and medium enterprises through two mechanisms: (1) within a cluster, finer division of labor lowers the capital barriers to entry and (2) closer proximity makes the provision of trade credit among firms easier. Since both mechanisms reduce the need for external financing, a larger number of firms—and thus greater competition—emerge within clusters, which helps explain the higher levels of exports and total factor productivity.

This cluster-based industrialization model fit particularly well with China's comparative advantage during its initial stage of takeoff, which was marked by scarcity of capital and an inefficient financial system. Hence our findings may be helpful to other developing countries with similar factor endowment patterns that are considering cluster-based development strategies.

Keywords: clustering; industrialization; finance; export; productivity; China

1. Introduction

Many have argued that a well-developed financial system is a key prerequisite for industrial development, as it can help pool disparate savings to finance large lump-sum investments in machineries and factory buildings (Goldsmith 1969; McKinnon 1973; King and Levine 1993; Rajan and Zingales 1998; Ayyagari, Demirgüç-Kunt, and Maksimovic 2006). However, China's rapid industrialization in the past three decades seems to defy conventional wisdom. At the incipient stage of reform in the late 1970s, China's financial system was far from developed, by any existing standards (Allen, Qian, and Qian 2005). In particular, the vast number of small and medium enterprises (SMEs) had little access to credit from state-owned banks (Lin and Li 2001; Wang and Zhang 2003; Lin and Sun 2005). Despite the initial lack of financial development, China has achieved the same degree of industrialization in three decades that took two centuries to occur in Europe (Summers 2007). Paradoxically, the SMEs in rural China have grown much faster than the large firms. How was China able to quickly industrialize in such a credit-constrained environment?

Previous research has suggested reliance on informal financing as the main solution (Allen, Qian, and Qian 2006). However, considering that at the onset of China's reform, a large proportion of rural people were poor (Ravallion and Chen 2007), the amount of local savings available for informal financing would have been rather limited. Without denying the importance of formal and informal financing in overcoming credit constraints, we argue herein that the cost of investment in production technologies may not be as prohibitive as suggested in the literature. The presence of credit constraints has the unintended consequence of inducing entrepreneurs to divide seemingly integrated production technologies into incremental steps by adopting the clustering mode of production. Clustering deepens division of labor, hence lowering capital entry barriers and enabling more entrepreneurs to participate in nonfarm production. An additional benefit of clustering is the consequent closer proximity of firms, which allows more inter-firm trade credit and reduces the need for working capital. Both these channels help lower the barriers of entry to industries, which in turn promotes competition and growth.

To establish the link between clustering, financing, and growth, we start by introducing a new measure of clustering to better assess the pattern of industrialization in China during the last decade. Although there are a large number of measures of regional specialization and industry concentration, they do not capture the interconnectedness among firms. For example, in the planned economic era, China concentrated its heavy industries in only a few locations. The existing measures would undoubtedly indicate a high degree of concentration in these industries at the time. However, this artificial industry concentration with little spillover into the local economy is not the same as the emerging patterns of clustering observed in post-reform China.

As has been reported in the media, China's rapid industrialization has been accompanied by the emergence of numerous "specialty cities" of a particular kind.¹ Thousands of firms, large and small, each specialized in a finely defined production step, are lumped together in a densely populated region, where some particular manufactured consumer good is churned out by the millions (if not billions) annually. Many formerly rural towns in the coastal areas have become so specialized that they boast of themselves as the world's Socks City, Sweater City, Kid's Clothing City, Footwear Capital, and so on. Each of the specialty cities described above fits Porter's concept of an industrial cluster, which is "a geographically proximate group of inter-connected companies (and associated institutions) in a particular field" (Porter 2000, page 16).

Despite the numerous popular media reports of this phenomenon, few studies have been performed to rigorously establish patterns using data covering a large sample and a long time period. Toward this end, we use complete firm-level data from the China Industrial Census 1995 and the China Economic Census 2004 to compute measures of clustering. The measure we focus on, industry proximity, allows us for the first time to explore how firms interact with one another, a key feature of clustering as highlighted by Porter (1998, 2000). Our results suggest that China's rapid industrialization during this time period was marked by closer interactions among firms within the same region.

We further examine the role of clustering on firm financing. At the county level, we calculate both clustering measures and the minimum asset level by industry. We find that clustering is associated with lower minimum capital requirements for industrial investment at the county level. With the finer division of labor implied by clustering, a production process is decomposed into small steps and thereby lowers the minimum capital requirement. Next, based on a panel dataset at the firm level from the two censuses, we document that more trade credit is extended among firms within an industrial cluster, thus reducing the reliance on external financing for working capital. In a word, clustering eases both starting and working capital constraints.

The availability of detailed firm-level data also allows us to correlate the observed patterns of clustering with firm performance. We find that firms in more clustered regions experience higher export and total factor productivity (TFP) growth. This provides supporting evidence that an increase in clustering in the past decade has contributed to improved export performance and productivity for firms in China. We view this as evidence that clusters help better align China's growth with its comparative advantage in labor endowment.

Our paper is unique on several fronts. First, we have access to firm-level data from two time periods for China as a whole, which are more disaggregated and updated than data used in previous studies. Second, we have adopted a new industry proximity measure to capture the evolving patterns of clustering, a key feature of China's industrialization. Third, we quantitatively show the positive impact of clustering on firm finance and correlate clustering with firm performance.

¹ For example, see <http://www.nytimes.com/2004/12/24/business/worldbusiness/24china.html> for a *New York Times* report.

The study of China's industrialization may help shed some light on research on industrialization in general. China's miraculously rapid industrialization provides a unique laboratory enabling us to observe and understand the process of industrialization. While industrialization in Western Europe and North America at the early stages of the Industrial Revolution can now be studied only through the relatively dim mirror of history, industrialization can be viewed directly in the ongoing economic revolution in China. China's experience may be relevant to other developing countries characterized by a high population density and a low capital-to-labor ratio. A clearer understanding of the industrialization processes in China will be of great value in helping propagate these processes to the world's less fortunate regions.

The structure of this paper is as follows: Section 2 reviews the literature on clustering, finance, and industrial development. Section 3 describes the data and the clustering measures, as well as the clustering patterns of China's industrialization. Section 4 examines the role of clustering in firm financing. Section 5 relates the evolving patterns of clustering to firm performance, while Section 6 offers some conclusions.

2. Literature Review on Clustering, Finance, and Industrial Development

Industrialization is often accompanied by clustering (or spatial agglomeration) of industrial activities.² The literature has highlighted the positive externalities of industrial clusters. Marshall (1920) lists three key externalities: better access to the market and suppliers, labor pooling, and easy flow of technology know-how. Porter (1998) argues that clustering is an important way for firms to fulfill their competitive advantage. Fujita, Krugman, and Venables (2001) view spatial clustering as a key feature of industrialization and highlight many of the positive externalities of spatial agglomeration. Two types of clustering have been observed during the industrialization process of developed countries. In the U.K., the decentralized production system scattered in different family workshops was replaced by a large integrated factory system during the Industrial Revolution (Landes 1998). The trend was similar and more evident in the United States during its industrialization (Chandler 1977). For example, the auto industry is highly concentrated in the Detroit metropolitan area, with several dominant large firms. This type of industrial cluster is generally anchored by a few very large firms while other smaller firms act as suppliers.

Italy, Japan, and other East Asian countries and regions experienced a different path of spatial clustering during the course of industrialization, which was led by small and medium enterprises (SMEs). In this business model, a large number of SMEs often cluster together, with comprehensive vertical division of labor. One noted example is the putting-out system, in which a merchant obtained market orders and subcontracted the production to nearby farmers or skilled workers, who usually finished the work in their

² In the literature, various terms for the phenomenon of clustering abound, including *spatial agglomeration*, *industrial district*, *cluster*, *industrial concentration*, and so on. In this paper, we prefer to use *cluster*, as it better captures the interconnectedness among firms in a narrowly concentrated location.

homes or family workshops (Hounshell 1984). The putting-out system was popular in the U.K. prior to its Industrial Revolution and was widely observed in nineteenth-century Japan (Nakabayashi 2006). Outsourcing (or subcontracting), the modern variant of the traditional putting-out system, remains a major feature of industrial production organization in contemporary Japan and Taiwan (Sonobe and Otsuka 2006). Industrial districts in which different workshops and factories clustered together were ubiquitous in France and Italy until the mid-twentieth century and are still viable in some regions of Italy (Piore and Sabel 1984; Porter 1998).

One key difference between the two types of clustering is firm size and number of firms. In the second type of clustering, an integrated production process is often disaggregated into many small steps that are performed by a large number of SMEs with dispersed ownership. By dividing a production process into incremental stages, a large lump-sum investment can be transformed into many small steps, thereby lowering the capital entry barriers (Schmitz 1995). Therefore, this mode of industrial organization may fit better in countries or regions with scarce capital and less developed financial sectors. Several in-depth case studies and popular media reports (as mentioned at the beginning of the paper) seem to suggest that China followed the second type of cluster-based industrialization path (Sonobe, Hu, and Otsuka 2002, 2004; Ruan and Zhu 2009; Huang, Zhang, and Zhu 2008). For example, Sonobe, Hu, and Otsuka (2002) studied how a garment cluster formed in a rural town in Zhejiang Province starting with small-scale production in family workshops.

The limitation of such case studies, however, is the difficulty of generalizing their findings. To test whether the patterns described in these studies are typical of the whole economy, one must use census-like firm-level data. Furthermore, conventional measures of concentration need to be replaced with other measures that can capture the differences between the two types of clusters discussed above, that is, how firms within a cluster interact with one another. We introduce such a measure, industry proximity, based on the Hausmann-Klinger proximity matrix (Hausmann and Klinger 2006).

The second line of research that our study closely relates to is the literature on the impact of financing constraints on industrial growth and investment. One implicit assumption of the finance and growth literature is that production technologies are indivisible. Because of the high cost to build up a factory and purchase machinery, in the absence of a capital market it is hard for many poor potential entrepreneurs to start their businesses (Banerjee and Newman 1993). Therefore, financial development is regarded as a prerequisite to pool savings to finance the investment projects, thereby placing the first-order importance on economic growth (King and Levine, 1993). However, it is a daunting task to develop a well-functioning capital market. In the case of China, rural industrialization took place largely preceding the major financial reforms. Allen, Qian, and Qian (2005) propose that informal financing must be a key explanatory factor for China's rapid industrialization. Recent empirical evidence throws some doubt on this hypothesis. Based on firm-level data from the recent Investment and Climate Survey conducted by the World Bank, Cull, Xu, and Zhu (2009) find that firms accessing formal financing grow faster than those relying on alternative channels. However, formal financing cannot explain the

puzzle either, as most SMEs, the engine of China's industrial growth, do not have access to it. Thus it is hard to resolve the puzzle by looking at only the literature on finance and growth.

An unintended consequence of credit constraints is that they may induce innovations in production organizations, with the cluster-based production structure being one of them. One key feature of clusters in China and other East Asian economies is that production technologies in a cluster are decomposed into many incremental steps that are undertaken by different entrepreneurs. Huang, Zhang, and Zhu (2008) detailed how the footwear cluster in Wenzhou helped overcome financial, institutional, and technological barriers. Ruan and Zhu (2009) in particular demonstrated that clustering lowers capital entry barriers and enables more entrepreneurs to participate in the production process through vertical division of labor.

These case studies provide insight into how clusters work. To generalize these findings, we use firm-level data from China's two censuses to test the hypothesis that the two following mechanisms help firms in clusters overcome financial constraints. First, the inherent finer division of labor in this kind of cluster helps lower the capital barriers to entry and enables a large number of low-wealth entrepreneurs from rural areas to finance profitable projects in a cluster. And second, the greater proximity and repeated transactions among firms in a cluster facilitate interfirm trade credit, thereby reducing working capital constraints.

3. Data, Proximity Measure, and Patterns of China's Industrialization

We utilize firm-level data from the China Industrial Census 1995 and China Economic Census 2004 for analysis in this paper. Tables 1 and 2 present summary statistics of the gross industrial output based on the census data by industry and by region, respectively. Table 3 compares the sample of our datasets with the published national aggregate statistics for China in 1995 and 2004. As shown in the table, our datasets capture the whole universe of Chinese industrial firms in these two years. Compared to the datasets used in previous studies on China's industrialization patterns (Young 2000; Bai et al. 2004; Wen, 2004; Zhang and Tan, 2004), our datasets have more comprehensive coverage and include industrial firms of all sizes (not only those above a certain scale).

Since the data are at the firm level, we can calculate the degree of clustering at any level of our choice, such as township, county, prefecture, or province, for regional aggregation, and two-, three-, or four-digit industry level for sectoral aggregation. For the main part of the analysis, we chose county and four-digit CIC (China Industry Code) as the levels of aggregation. But for robustness tests, we also used prefecture and provincial levels for geographic aggregation, and three-digit and two-digit CICs for industrial aggregation. When constructing the clustering measures, we first determined the level of aggregation to convert firm-level data to cell-level totals, where each cell is a combination of a certain level of region and a certain level of industry. For example, the most detailed cell is the

combination of four-digit CIC and county. We then create the clustering measure using the cell-level data.

China modified its industry coding system in 2002 (switching from GB1994 to GB2002). Therefore, when studying changes between 1995 and 2004, we match industry codes that changed from 1994 to 2002 as follows: for industry codes that became more disaggregated in the 2002 coding system, we use the 1994 codes as the standard; for those that became more aggregated, we use the 2002 codes as the standard. In other words, we use the more aggregated codes to group and compare industries between 1995 and 2004. During the period between the two censuses (1995–2004), the territories of some counties were also redrawn and the names of others changed. We have carefully tracked these changes to match the counties throughout the time period.

Conventional measures of industrial agglomeration are based on regional specialization or industrial concentration. The market share of a certain number of the largest, say, three firms, in an industry or region is often used as a concentration measure. The advantage of this measure is that it is easy to calculate and interpret, but when the distribution of firms is relatively spread out, it may miss those firms below the cut-off lines. To overcome this problem, the Gini coefficient is often used to calculate the regional variation of output or employment shares for all the firms in an industry. Krugman (1991) modifies the Gini coefficient by accounting for the discrepancy between a region's share of output/employment in a certain industry and its share in all manufacturing industries in calculating the Gini coefficient.

However, these concentration measures do not distinguish between the following two kinds of “agglomeration”: one in which a small number of large firms with minimum interfirm connections are located, versus the other in which a large number of variously sized firms congregate and interact closely with one another. While the first type of agglomeration characterizes cities such as Detroit, the second type of agglomeration seems to better fit the patterns observed in coastal China, where thousands of firms of all sizes are densely populated in a small region, closely intertwined with one another throughout the production processes, all the while churning out thousands of products with breathtaking efficiency.

The second type of agglomeration fits very well into the definition of clusters given by Porter, whose concept of an industrial cluster is summarized as “a geographically proximate group of inter-connected companies (and associated institutions) in a particular field” (Porter 2000, page 16). Although the concept is intuitive and extremely easy to understand, the measurement of interconnectedness seems more elusive. To our knowledge, no previous studies have directly measured it except in case studies in which firms can provide detailed information on how they interact with other firms.

Such detailed information is necessarily absent for large-scale studies like ours. In the absence of the first-best information, we analyze Porter's concept of clustering more carefully to explore alternative ways of measuring the interconnectedness among firms. When delineating the main actors within a cluster, Porter states, “They include, for

example, *suppliers of specialized inputs such as components, machinery, and services as well as providers of specialized infrastructure*. Clusters also often extend downstream to channels or customers and laterally to *manufacturers of complementary products or companies related by skills, technologies, or common inputs*” (Porter 2000, 16–17, italics added by authors). In addition, Porter emphasizes that one main benefit derived from geographically concentrated clusters is that industries in the same cluster share common technologies, skills, knowledge, inputs, and institutions. Previous work has also shown that technology linkages among related industries are an important engine for innovation (Scherer 1982; Feldman and Audresch 1999).

The works cited above suggest one way to measure the interconnectedness in the cluster concept as defined by Porter. If industries and firms use similar combinations of inputs in their production processes, then they are more likely to rely on the same set of suppliers and clients, and thus are more likely to be related by skills, technologies, and other common inputs. The similarity among products of industries can thus be used as a measure for clustering, as defined by Porter. Given Porter’s emphasis on the role of other related institutions in the cluster (such as government agencies, trade associations, universities, and research institutions), we should consider inputs in the broad sense, including not only raw materials, physical assets, and labor (skilled and unskilled), but also knowledge, intermediate inputs, infrastructure, and so on. As a result, information from the input-output table may not suffice. Furthermore, the similarity among goods should be based on the production technology, and thus does not depend on the location of the production process.

New results obtained by Hausmann and Klinger (2006) allow us to implement the above measure of interconnectedness among industries (and participating firms) in a cluster. In their 2006 paper (Hausmann and Klinger 2006), the authors constructed a proximity matrix for all four-digit SITC products, in which the proximity between any two goods captures their similarity in the following sense: If the two goods need the same combination of inputs (or endowments and capabilities) to produce, then there is a higher probability that a country has a comparative advantage in both. In other words, if we have trade data from all countries on all goods, we can then compute the probability of a country simultaneously having a comparative advantage in any pair of goods to measure how close the two goods are to each other in terms of input or capability requirements.

Specifically, Hausmann and Klinger propose the proximity between two products i and j to be computed as follows: $p_{i,j} = \min\{P(x_i | x_j), P(x_j | x_i)\}$, where $x_i=1$ if a country has the revealed comparative advantage in product i (or if $RCA_i > 1$), and 0 otherwise, while the conditional probabilities $P(x_i | x_j), P(x_j | x_i)$ are computed using trade information on all countries.

To get to the intuition of the formula, consider the pair of goods of ostrich meat (good i) and metal ores (good j). Some countries such as Australia export both goods. The formula implies that the probability of exporting metal ores given that a country exports ostrich meat is large, but the probability that a country exports ostrich meat given that it exports metal ores is very low, since although Australia exports both, Chile, Peru, and Zambia do

not export ostrich meat but do export metals. The proximity between ostrich meat and metal ores will thus be low, because the formula requires the minimum of the two conditional probabilities, with $P(x_i | x_j)$ being low, despite a high $P(x_j | x_i)$. Thus, the formula is superior to a simple conditional probability $P(x_i | x_j)$ or $P(x_j | x_i)$.

The proximity measure also isolates the degree of similarity between the two goods from how prevalent they are in different countries. An alternative measure for proximity is the joint probability $P(x_i \cap x_j)$, which Hausmann and Klinger rejected for the following reason. Consider ostrich meat and ostrich eggs, two goods with extremely high similarity, because every single country that exports ostrich eggs also exports ostrich meat. But if only three countries in the world export these two goods, “then the joint probability for any single country exporting the two would be small, instead of large.” (Hausmann and Klinger 2006, page 10) The problem with using the joint probability to measure proximity is that it combines the degree of similarity between the two goods and their prevalence in different parts of the world.

The proximity measure has some additional nice features: (1) as a characteristic of the production technology based on export/import information from all countries, it applies to all countries, be they open or closed; (2) computed as the minimum between two conditional probabilities, it is a symmetric measure; and (3) by focusing on countries with a revealed comparative advantage in product i (i.e., $x_i=1$ if a country has the revealed comparative advantage in product i [or if $RCA_i > 1$] and 0 otherwise), the measure captures all the significant exports but leaves aside the noise.

Since products more likely to be exported together depend on the same comparative advantages, the proximity between two products captures the degree of similarity between these products in terms of how much their production processes need the same endowments and production capabilities. Firms and industries that produce these products are then more likely to interact with one another in various ways, including dependence on similar inputs (be they raw materials, labor, or machinery), reliance on similar technologies and research and development, and even dependence on the same supply or marketing facilities. Thus, those industries producing products that are more proximate in the Hausmann-Klinger space are likely to be more interconnected in the Porter sense. As a result, this proximity measure can be used to provide a gauge for how closely interconnected industries and their participating firms are within a specific region.

We begin with the product proximity matrix constructed by Hausmann and Klinger (2006). Because the proximity matrix is computed for products at the four-digit SITC level, we have made a concerted effort to convert the CICs first to ISICs and then to SITCs based on the manuals obtained from China’s National Bureau of Statistics as well as correspondence tables from Eurostat and the United Nations. Given that both CICs and ISICs are industry-level codes, while SITCs are based on products, there are many cases in which we have one industry correspond to more than one product. In such cases, we give equal weights (that sum up to 1) to the products produced by the industry. Specifically, our procedures are as follows: (1) Aggregate output, asset, and employment

to the cell level, where the cell is defined as a combination of county and a four-digit CIC industry. (2) Convert the CIC first to ISIC and then to SITC based on the manuals obtained from China's National Bureau of Statistics as well as correspondence tables from Eurostat and the United Nations. (3) For each industry in a cell, calculate its average proximity to all industries located in the same region, using the Hausmann-Klinger product proximity matrix, which gives the proximity (or the inverse distance) between each pair of products (and between each pair of industries through the conversion procedures above). The average proximity for each industry (for a certain region) is computed as a weighted average using the size of the other industry in each pair as the weight. (4) Finally, the average industry proximity for each region is computed as the average of the proximities of all the industries in that region, weighted by the size of each industry.

The proximity measure can be based on assets, employment, or output. In other words, the weights discussed above that are used to adjust for the size of each industry can be assets, employment, or output. We use all these measures, as they may provide different angles of clustering. An illustration follows. Consider a region with three industries: steel, automobiles, and rubber. Intuitively, the automobile industry has a high proximity to both the steel and rubber industries, while the proximity between the other two is low. Now suppose that the region has experienced faster growth in the auto industry than in the other industries. Following the procedures describe above, we see that the average proximity of the auto industry has not changed, since the relative weights of the other two industries have not changed. But the average industry proximity of the whole region has increased, because the industry that is closer to the others, the auto industry in this case, has grown faster. Now consider the role of the weight. If the growth of the auto industry is in its output relative to those of the other industries, then the greater interconnectedness among industries in the region will be reflected in a greater proximity using output as the weight. Proximity measures weighted by asset or employment can be understood accordingly.

These three proximities may therefore measure different kinds of interconnectedness, which in turn imply different kinds of cluster effects. Marshall (1920) outlined three types of advantages from agglomeration or clusters: labor market pooling, specialized supplies, and technological spillovers. Large populations of skilled laborers enter the area and are able to exchange knowledge, ideas, and information. In addition, there is increased access to the specialized goods and services provided for the clustering firms, which provides increasing returns to scale for each of the firms located within that area because of the proximity to the available sources needed for production. Finally, clustering in specific fields leads to quicker diffusion of ideas and adoption of ideas.

Although likely to contribute to all three of these advantages, output-weighted proximity is probably more conducive to technological spillovers, since the output can be used as input in the production of other industries in the same region, while employment-weighted proximity implies more labor-market pooling, and asset-weighted proximity implies more specialized supplies, especially in capital goods. All these effects of agglomeration will lead to higher productivity at the firm level.

In addition, we emphasize in this paper another effect of agglomeration that has not drawn enough attention previously, namely, its impact on firm finances. We argue that industrial clusters help alleviate firms' financial constraints through two channels: (1) the finer division of labor among firms within an industrial cluster lowers capital requirements for these firms and (2) trade credit extended among firms within an industrial cluster helps diminish the need for external financing. As financial transactions permeate the whole production process, including labor hiring, asset purchasing, and product sales, we expect all three measures of proximity to play a role in helping overcome firms' financial constraints.

Using the proximity measures described above, we found that within each region, the proximity among industries increased significantly between 1995 and 2004.³ Table 4 presents the industry proximity measure for each of the Chinese provinces in 1995 and 2004, based on output. The measures constructed at the prefecture and the county levels give the same pattern of higher industry proximity in each region in the latter year, as shown in Figures 1 and 2.

In addition, we find that the degree of clustering is correlated with firm size distribution in some interesting ways. At the county level, a higher degree of clustering is correlated with larger average firm size. But counties with a greater degree of clustering also tend to have more evenly distributed firm sizes among different industrial sectors as well as more evenly distributed numbers of firms among sectors.⁴ Appendix Tables 1 and 2 show how the degree of clustering affects the distribution of firm size and number of firms, using firm-level panel data. In other words, the type of clustering measured by the proximity index is different from the Detroit type, in which a small number of very large firms emerge as the dominant players. Rather, it portrays a pattern similar to the East Asian cluster-based industrialization model, in which a large number of firms are present, often of small and medium size.

4. Clustering and Firm Financing

We now turn to explore the effects of such increased industry proximity in a geographical location, with our first focus on firm finances. Table 5 provides summary statistics of variables used in the analysis in this and the next sections. As discussed previously, there are two potential mechanisms through which clusters may help alleviate financial constraints for firms located in the clusters. First, because firms are more interconnected, finer division of labor becomes feasible, which reduces the capital requirement for firms on average. Second, the interconnectedness among firms in a narrow location may facilitate inter-firm financing through trade credit.

³ Interestingly, we find similar results using other conventional concentration measures, including the Hirsch index and Gini coefficient. See Long and Zhang (2009) for details.

⁴ In contrast, a higher degree of clustering has no significant impact on the average number of firms.

To explore the effect of clustering on firms' capital requirements, we look at the minimum level of assets among firms within a certain region. It is thus crucial that our sample does not exclude firms due to their small size. The 1995 and 2004 censuses that include all industrial firms provide the ideal data for computing for each county the minimum level of assets and testing the hypothesis. Table 6 shows results from the following regression:

$$\min(asset_{c,2004}) = \hat{a} + \hat{a}_1 * \min(asset_{c,1995}) + \hat{a}_2 * P_{c,1995} + \hat{a}_c, \quad (1)$$

where c indicates county, $\min(asset_{c,2004})$ is the minimum level of assets in 2004, $\min(asset_{c,1995})$ is the minimum level of assets in 1995, $P_{c,1995}$ is the industry proximity in 1995, and ε is the random error term. Therefore, the coefficient β_2 shows the effect of industry proximity in a region on the minimum requirement of capital for firms located in that region.

As shown in Table 6, higher proximity, measured in both assets and employment, is correlated to a lower level of minimum assets, consistent with the argument that clustering facilitates a finer division of labor and thus reduces the capital requirement for firms. The proximity in output is also negatively related to minimum assets at the county level, but the effect is insignificant. The effects of proximity in assets and employment are economically important as well. In particular, a standard-error increase in industry proximity (0.03) will lead to a reduction in the minimum capital requirement of RMB 50,000 for a typical county, which accounts for about 39% of the average minimum capital requirement. As expected, the minimum capital requirement in 1995 is positively correlated with that in 2004. Column 4 shows that when industry proximity is dropped from the regression, the 1995 minimum capital requirement remains significant but the R-square drops by 2%, which is the additional explanatory power of industry proximity.

To study the effects of industry proximity on trade credit, we use firm-level data from 1995 and 2004. Since detailed accounting information is provided for only a subsample of firms even in the census years of 1995 and 2004, we cannot aggregate the data into county level as for the minimum-asset data. Instead, we construct a balanced panel of firms for which information is available.⁵ The estimation regression is as follows:

$$trade\ credit_{ict} = \hat{a}_i + \hat{a}_t + \hat{a} * P_{ct} + \tilde{a} Z_{ict} + \hat{a}_{ict}, \quad (2)$$

where i , c , and t indicate firm, county, and year, respectively; P is the proximity measure at the county level by year; Z is a vector of firm characteristics; and ε is the random error term. Therefore, the coefficient β shows the effect of industry proximity in a region on the provision of trade credit among firms located in that region. We use two measures for trade credit: accounts payable / total debt, and accounts receivable / revenue. While the former measures the proportion of the firm's debt that is financed by its trading partners, the latter indicates the degree to which the firm provides credit to its business partners.

⁵ Given that the panel covers only two years, the singletons in the unbalanced panel are dropped out in the fixed effect estimation due to the demeaning process. Thus the results based on the unbalanced panel give the same results as the balanced panel.

Table 7 shows that all three proximity measures are positively correlated with both measures of trade credit.⁶ Specifically, for accounts payable as a percentage of total debt, proximity measures weighted by assets, employment, and output all have positive and significant effects. For a standard-error increase in industry proximity (about 0.025), the ratio of accounts payable to total debt increases by about 0.7 percentage point, which amounts to about 3% of the average ratio of accounts payable to total debt. If proximity increases from the lowest to the highest level, then the ratio of accounts payable to total debt increases by about 8 percentage point, which amounts to almost 40% of the average ratio of accounts payable to total debt. Therefore, these effects are of non-negligible magnitude.

The coefficient for the asset variable is significantly positive for firms extending trade credit and negative for firms receiving trade credit. This is consistent with the lending policies of state banks and the findings in the literature (Cull, Xu, and Zhu 2009). In China, the state banks are more likely to extend credit to larger firms, in particular state-owned enterprises, for two reasons. First, because of their multifunctional role in providing many social services, the large state-owned enterprises received preferential treatment in accessing state bank credit. Second, large firms can use their fixed assets as collateral to secure bank loans, while SMEs often lack collateral. In large part because of their difficulty in accessing formal credit, SMEs rely more on trade credit from larger firms. Through trade credit, SMEs can indirectly gain access to state credit to help alleviate the constraints of insufficient working capital.

5. Clustering and Firm Performance

Our next analysis addresses firm exports and TFP growth. As discussed above, increased proximity may lead to productivity improvements due to increased pooling of labor, easier access to specialized inputs, and technological spillovers. In addition, as shown in the previous section, by lowering the capital barriers, clustering enables more potential entrepreneurs scattered in rural areas to engage in more productive industrial production. The wide availability of trade credit inherent in clusters also eases firms' working capital constraints and boosts output. Thus, we will look at the effects of the proximity measures on firm performance.

We first examine the impact of proximity and geographical concentration on export growth, using the following estimation specification:

⁶ Based on the investment climate survey conducted by the World Bank, Cull, Xu, and Zhu (2009) find that trade credit does not play a significant role in firm performance among Chinese firms. There are two possible reasons for the difference between their findings and ours. First, the firm size in their sample is larger than that in the industrial and economic censuses used in this paper. Because large firms are more likely to access formal bank credit, their demand for trade credit is lower than that of smaller firms. Second, they do not relate trade credit with cluster development. Our point is that clustering facilitates the extension of trade credit. Therefore, trade credit is more likely to be observed in areas with industrial clusters than in those without clusters.

$$export_{ict} = \hat{a}_i + \hat{a}_t + \hat{a} * P_{ct} + \tilde{a}Z + \hat{a}_{ict}, \quad (3)$$

where i , c , and t indicate firm, county, and year, respectively; $export$ is share of export value in total sales or a dummy variable indicating whether a firm exported in a certain year; P is the clustering measure at the county level by year; Z is a vector of firm characteristics including firm age, firm scale (measured by the ratio between firm revenue and the industry average revenue), and firm ownership types; and \hat{a} is the random error term. Therefore, the coefficient \hat{a} shows the effect of industry proximity in a region on the labor productivity of firms located in that region.

We next study the relationship between clustering and TFP based on the following estimation⁷:

$$\log(Y_{ict}) = \hat{a}_i + \hat{a}_t + \hat{a}_1 * \log(K_{ict}) + \hat{a}_2 * \log(L_{ict}) + \hat{a}_3 * P_{ct} + \tilde{a}Z + \hat{a}_{ict}, \quad (4)$$

where i , c , and t denote firm, county, and year, respectively; Y is value added; K and L refer to assets and labor; P is a clustering measure at the county level by year; Z is a vector of firm characteristics (including firm age and firm ownership type); and \hat{a} is the random error term. The coefficient \hat{a}_3 measures the effect of industry proximity in a region on the TFP of firms located in that region. To allow the possibility that the production function may have changed between 1995 and 2004, we also include the year 2004 dummy as well as its interaction terms with the logs of K and L .⁸

Tables 8 and 9 present results from the above estimation. All three proximity measures are found to have positive effects on both the export and total factor productivity of the firms, and the effects are also economically important. Specifically, an increase in industry proximity of a standard deviation (0.03) will lead to a 0.65 percentage point increase in the export-to-sales ratio, amounting to about 12% of the average export-to-sales ratio. The corresponding effect on the likelihood of the firm being an exporter is 0.3 percentage point, which is about 4% of the firms that are exporters. In terms of the effect on TFP, a standard-deviation increase in industry proximity leads to a 1.8 percentage point increase in TFP. In summary, even after controlling for capital intensity and other firm characteristics, we still find evidence showing the positive and non-negligible effects of greater industry proximity on firm performance. Firms in clusters are more productive and more competitive in the international market. As the results are obtained after controlling for firm fixed effects in a balanced panel, they reflect effects of clustering on firms that existed in both years, instead of those on firms that are new entrants.⁹

6. Conclusions

⁷ Truncation of the negative values of value added may be a concern. But by comparing the sample sizes in column 1 and column 4, the reduction in sample size is only 3,500 out of 69,000 (about 5%), which does not seem a major concern to us.

⁸ Estimation without the interaction terms obtains similar effects of proximity on TFP.

⁹ We thank our referee for pointing this out to us.

Using census data at the firm level from 1995 and 2004, we have shown in this paper that China's industrialization has been accompanied by increasing interactions among industries within regions. The pattern of industrial clustering to a large degree resembles the industrialization path in some European countries and other East Asian economies.

In addition, our results indicate that the number of firms is growing faster and firm size is not significantly larger in clustered areas than in nonclustered regions, while at the same time there is a finer division of labor and closer technological affinity among firms. This pattern is similar to the East Asian cluster-based industrialization model led by numerous SMEs but differs from the observed patterns in the United States, where regional agglomeration and industrial districts were mainly driven by the presence of large firms.

This cluster-based industrialization dominated by SMEs may have fit well with China's comparative advantage. This business model makes more use of entrepreneurs and labor, and less of capital, compared to nonclustered large factories, and thus may have emerged as the choice of Chinese firms over time, leading to more clustered industries in China.

One key benefit of cluster-based industrialization in China is that it helps lessen the credit constraints facing the vast number of SMEs. With lower minimum capital requirements, many low-wealth entrepreneurs can start businesses despite the constrained credit environment. Close proximity and intense competition among firms within a cluster may also reduce the temptation to act dishonestly, making frequent trade credit among firms within a cluster possible. All these factors help ease the reliance on external financing.

It is worth emphasizing, however, that the results obtained do not necessarily indicate that financial-sector development is less important. Rather, clustering may be a second-best solution to the financing problem when the local conditions do not permit easy access to regular financing. One potentially fruitful line of research would be to study whether financial constraints indeed have induced the emergence of industrial clusters in China. A related issue is whether the substitution of clustering as an alternative mechanism for formal financing has affected the formation and growth of firms. For example, do firms tend to be larger in the absence of financial constraints as compared to those in industrial clusters? And do these differences in firm size distribution cause any efficiency loss? These topics are beyond the scope of the current paper, but are subjects we plan to research in the near future.

Nonetheless, given that the ideal conditions for economic development are rarely in existence, the organization innovations embodied in clustering are essential, especially for developing countries, for which economic growth is particularly important.

Despite the positive role of clustering in industrial development, governments should be cautious in promoting cluster-based development in industrial sectors that do not make use of comparative advantage (Rodríguez-Clare 2007). In the case of China, clusters have been proven to be more productive and more export-oriented. However, this may be due to the fact that most clusters in China are based on labor-intensive production technologies, which are in line with China's comparative advantage. A region with overly

concentrated capital-intensive industries, in contrast, may not experience the same rate of growth and exports. We leave a more in-depth study of this topic for future investigation.

References

- Allen, F., J. Qian, and M. Qian. 2005. Law, finance, and economic growth in China. *Journal of Financial Economics* 77: 57–116.
- Ayyagari, M., A. Demirgüç-Kunt, and V. Maksimovic. 2006. How important are financing constraints? The role of finance in the business environment. Policy Research Working Paper No. 3820. Washington, D.C.: World Bank.
- Bai, C.-E., Y. Duan, Z. Tao, and S. T. Tong. 2004. Local protectionism and regional specialization: Evidence from China's industries. *Journal of International Economics* 63 (2): 397–417.
- Banerjee, A., and A. F. Newman. 1993. Occupational choice and the process of development. *Journal of Political Economy* 101 (2): 274–298.
- Chandler, A., Jr. 1977. *The visible hand: The managerial revolution in American business*. Cambridge, Mass., U.S.A.: Harvard University Press.
- China, National Bureau of Statistics. Various years. *China statistical yearbooks*. Beijing: China Statistical Publishing House.
- China, National Bureau of Statistics. 1995. *China industrial census*. Beijing: China Statistical Publishing House.
- China, National Bureau of Statistics. 2004. *China economic census*. Beijing: China Statistical Publishing House.
- Cull, Robert & Xu, Lixin Colin & Zhu, Tian, 2009. "Formal finance and trade credit during China's transition," *Journal of Financial Intermediation*, 18(2): 173-192.
- Feldman, M, and Audresch DB.1999. Innovation in cities: science-base diversity, specialization and localized competition. *European Economic Review* 43: 409–429.
- Fujita, M., P. Krugman, and A. J. Venables. 2001. *The spatial economy: Cities, regions, and international trade*. Cambridge, Mass., U.S.A.: MIT Press.
- Goldsmith, Raymond W., 1969. *Financial Structure and Development*. New Haven, CT: Yale University Press.
- Hausmann, R., and B. Klinger. 2006. Structural transformation and patterns of comparative advantage. Center for International Development Working Paper No. 128. Cambridge, Mass., U.S.A.: Harvard University.
- Hounshell, David A. 1984. *From the American System to Mass Production, 1800-1932*. Baltimore: The Johns Hopkins University Press.
- Huang, Z., X. Zhang, and Y. Zhu. 2008. The role of clustering in rural industrialization: A case study of Wenzhou's footwear industry. *China Economic Review* 19: 409–420.
- King, Robert G. and Ross Levine, 1993. "Finance and Growth: Schumpeter Might Be Right." *The Quarterly Journal of Economics*, 108 (August), 717–737.
- Krugman, P. 1991. *Geography and trade*. Cambridge, Mass., U.S.A.: MIT Press.
- Landes, D. S. 1998. *The wealth and poverty of nations*. New York: W.W. Norton and Co.
- Lin, Yifu and Yongjun Li. 2001, "Zhongxiao Jinrong Jigou Fazhan yu Zhongxiao Qiye Rongzi" (Small and Medium Financial Institutions and Small and Medium Enterprises Financing). *Jingji Yanjiu* (Economic Study) 1 (January): 10–18.

- Lin, Yifu, and Xifang Sun. 2005, "Xinxi, Feizhenggui Jinrong yu Zhongxiao Qiye Rongzi" (Information, Non-formal Financing and Small and Medium Enterprises Financing). *Jingji Yanjiu (Economic Study)* 7 (July): 35–44.
- Long, C., and X. Zhang. 2009. The evolving patterns of industrial clustering in China. Memo. Colgate University and International Food Policy Research Institute.
- Mckinnon, Ronald I. 1973. *Money and Capital in Economic Development*. Washington DC: Brooking Institution.
- Marshall, A. 1920. *Principles of Economics* (8th ed.) London: Macmillan. (Original work published 1890)
- Nakabayashi, M. 2006. Flexibility and diversity: The putting-out system in the silk fabric industry of Kiryu, Japan. Discussion Paper in Economic and Business No. 06-10. Osaka, Japan: Osaka University.
- Piore, M. J., and C. F. Sabel. 1984. *The second industrial divide: Possibilities for prosperity*. New York: Basic Books.
- Porter, E. M. 1998. Clusters and the new economics of competition. *Harvard Business Review* 76 (6): 77–90.
- Porter, E. M. 2000. Location, competition, and economic development: Local clusters in a global economy. *Economic Development Quarterly* 14 (1): 15–34.
- Rajan, Raghuram G. and Luigi Zingales, 1998. "Financial Dependence and Growth." *American Economic Review*, 88 (June), 559–87.
- Ravallion, Martin, and Shaohua Chen. 2007. "China's (Uneven) Progress against Poverty," *Journal of Development Economics* 82 (1): 1-42.
- Rodríguez-Clare, A. 2007. Clusters and comparative advantage: Implications for industrial policy. *Journal of Development Economics* 82: 43–57.
- Ruan, J., and X. Zhang. 2009. Finance and cluster-based industrial development in China. *Economic Development and Cultural Change*: 58:143–164, October 2009.
- Scherer, Frederic M., 1982. "Demand-Pull and Technological Innovation: Schmoekler Revisited," *Journal of Industrial Economics* 30(3): 225-238.
- Schmitz, H. 1995. Collective efficiency: Growth path for small-scale industry. *Journal of Development Studies* 31 (4): 529–566.
- Sonobe, T., D. Hu, and K. Otsuka. 2002. Process of cluster formation in China: A case study of a garment town. *Journal of Development Studies* 39 (1): 118–139.
- Sonobe, T., D. Hu, and K. Otsuka. 2004. From inferior to superior products: An inquiry into the Wenzhou model of industrial development in China. *Journal of Comparative Economics* 32: 542–563.
- Sonobe, T., and K. Otsuka. 2006. *Cluster-based industrial development: An East Asia model*. New York: Palgrave MacMillan.
- Summers, L. 2007. The rise of Asia and the global economy. *Research Monitor* (Special Issue): 4–5.
- Wang, Xiao, and Jie Zhang. 2003. "Yinhang Xindai Peiji yu Zhongxiao Qiye Daikui: Yige Neishenghua Diyapin he Qiye Guimo de Lilun Moxing" (Credit Bank Loan Provision and Small and Medium Enterprises Borrowing: A Theoretical Model on Endogenous Mortgage and Scale of Enterprises). *Jingji Yanjiu (Economic Study)* 7 (July): 68–75.
- Wen, M. 2004. Relocation and agglomeration of Chinese industry. *Journal of Development Economics* 73 (1): 329–347.

- Young, A. 2000. The razor's edge: Distortions and incremental reform in the People's Republic of China. *Quarterly Journal of Economics* 115 (4): 1091–1135.
- Zhang, X., and K.-Y. Tan. 2007. Incremental reform and distortions in China's product and factor markets. *World Bank Economic Review* 21 (2): 279–299.

Table 1. Summary statistics of gross industrial output by two-digit industry

Industry	1995			2004		
	Mean	SD	No. of firms	Mean	SD	No. of firms
Coal Mining & Dressing	9,664	96,605	11,953	17,643	224,515	26,822
Petroleum & Natural Gas Extraction	1,066,011	4,254,463	134	962,613	5,758,893	481
Ferrous Metals Mining & Dressing	5,228	25,019	2,141	9,554	46,737	10,256
Nonferrous Metals Mining & Dressing	8,554	36,978	3,766	14,919	104,298	6,075
Non-Metallic Minerals Mining & Dressing	3,087	11,114	11,820	3,293	19,150	34,945
Other Minerals Mining & Dressing	2,515	4,604	149	3,948	23,920	263
Foodstuff Processing Industry	10,042	40,142	30,962	13,719	105,609	69,521
Foodstuff Manufacturing Industry	5,856	34,955	16,313	11,026	87,351	29,811
Beverage Manufacturing Industry	7,852	43,376	14,719	10,740	96,640	25,485
Tobacco Processing	237,406	902,942	423	885,794	2,275,005	281
Spinning Industry	18,002	53,656	24,459	14,029	100,348	83,011
Manufacturers of Clothes & Other Fiber Products	7,453	25,219	18,937	9,671	58,140	48,250
Leather, Fur, Feather & Other Products	9,308	30,236	10,468	13,816	59,022	22,677
Timber Processing & Bamboo, Cane, Palm, Straw Products	2,620	12,020	15,480	5,072	32,941	37,028
Furniture	2,580	9,564	8,760	6,255	39,320	23,892
Paper Makers & Paper Products	7,303	27,232	13,890	10,005	83,103	39,669
Printing & Record Medium Reproduction	2,553	9,898	16,763	4,234	21,719	44,070
Teaching & Sport Products for Daily Use	8,575	31,219	5,356	9,702	42,732	14,711
Oil Processing & Refining	73,925	716,204	2,744	126,789	1,373,127	7,146
Chemical Material & Products	13,750	116,761	26,872	19,175	239,355	69,120
Pharmaceutical & Medicine Manufacturing	16,527	61,057	6,051	29,861	148,194	11,271
Chemical Fibers	76,277	474,156	1,034	59,128	329,805	3,372
Rubber Products	13,294	74,614	4,663	13,490	128,384	15,178
Plastic Products	5,856	20,254	19,255	7,573	44,341	69,729
Non-Metallic Mineral Products	4,926	18,678	61,278	6,306	32,082	157,734
Smelting & Pressing of Ferrous Metals	44,108	479,425	8,429	84,284	961,069	20,494
Smelting & Pressing of Non-Ferrous Metals	29,697	143,612	4,621	41,174	259,277	15,162
Metal Products	5,534	24,407	26,744	7,849	52,709	80,976
Common Machines	7,719	41,075	31,474	9,032	80,815	113,691
Special Equipment	10,805	69,347	18,391	10,556	84,887	55,095
Traffic Equipment	17,009	216,986	19,522	27,664	475,230	51,844
Manufacture of Electrical Machinery and Apparatus	13,206	74,917	18,928	20,680	249,967	54,979
Electrical Machines & Equipment	34,343	250,416	5,489	74,793	1,048,297	15,211
Electronic & Communication Equipment	12,552	75,699	9,735	40,516	532,205	35,203
Instruments, Culture & Office Devices	4,576	17,211	12,127	6,966	38,783	26,627
Recycling of Material Waste and Scrap	2,799	10,227	4,440	4,491	28,720	6,156
Electricity, Steam, Thermal Power Production & Supply	19,369	107,370	12,600	60,653	1,052,128	24,568
Coal Gas Production & Supply	20,474	82,663	372	30,310	130,000	1,445
Tap Water Production & Supply	3,545	29,942	5,147	5,058	34,173	11,035
Total	10,763	134,410	506,409	16,198	310,686	1,363,284

Note: The gross industrial output is reported in thousands of RMB at current prices.

Table 2. Summary statistics of gross industrial output by province

Province	Province code	1995			2004		
		Mean	SD	No. of firms	Mean	SD	No. of firms
Beijing	BJ	15,067	256,247	9,623	18,926	418,528	31,364
Tianjin	TJ	13,638	155,565	10,735	23,949	500,129	25,432
Hebei	HEB	9,216	80,343	23,592	15,789	201,284	64,062
Shanxi	SX	8,538	91,723	11,416	14,490	216,240	28,641
Neimeng	NM	5,559	83,740	9,432	19,689	223,747	11,759
Liaoning	LN	10,184	171,331	29,435	16,844	365,899	54,115
Jilin	JL	7,751	167,292	13,100	22,085	542,689	16,037
Heilongjiang	HLJ	8,524	308,216	18,745	19,613	738,370	20,101
Shanghai	SH	23,260	273,997	16,690	26,263	499,886	55,315
Jiangsu	JS	15,815	106,861	41,582	15,618	262,800	187,212
Zhejiang	ZJ	10,363	58,130	32,725	11,236	173,580	187,588
Anhui	AH	6,912	64,940	23,474	10,808	189,114	38,827
Fujian	FJ	8,080	47,063	19,038	15,126	225,394	49,532
Jiangxi	JX	4,528	45,443	18,253	9,331	146,981	29,144
Shandong	SD	17,466	179,590	26,980	20,477	303,314	119,699
Henan	HEN	9,703	75,318	23,119	12,065	164,347	76,292
Hubei	HUB	10,432	139,562	20,881	18,191	359,619	28,937
Hunan	HUN	5,738	66,417	23,720	9,668	145,047	43,529
Gongdong	GD	17,715	114,052	34,536	22,969	473,908	136,606
Gongxi	GX	7,719	42,786	12,312	11,870	155,918	18,753
Hainan	HAIN	9,932	52,781	1,278	21,086	198,547	2,025
Chongqing	CQ	6,676	82,141	11,456	12,677	149,313	20,359
Sichuan	SC	6,675	74,866	26,380	12,137	168,971	43,325
Guizhou	GZ	5,500	48,962	7,450	13,831	178,497	10,996
Yunnan	YN	13,970	223,904	6,267	16,157	239,845	14,271
Tiebet	TB	2,343	6,554	295	7,004	22,096	354
Shaanxi	SAX	6,182	60,000	12,950	12,251	209,434	25,573
Gansu	GS	8,260	109,848	7,140	14,648	305,597	11,549
Qinghai	QH	8,193	77,821	1,446	17,524	218,482	2,168
Ningxia	NX	9,011	56,606	1,706	15,132	151,483	3,984
Xinjiang	XJ	9,990	187,435	5,077	28,813	441,176	5,735
Total		10,725	134,909	500,833	16,198	310,686	1,363,284

Note: The gross industrial output is reported in thousands of RMB at current prices.

Table 3. Comparing sample with aggregate data

	Gross Industrial Output (trillions of RMB, at current prices)		
	Sample (1)	Statistical Yearbook (2)	(1)/(2)*100%
1995	5.495	5.526	99.438
2004	20.174	18.722	107.754

Note: The official figures for gross industrial output and industrial value added are from the *China Statistical Yearbook* for 1996 and 2005. However, the official figures in the *China Statistical Yearbook* 2005 do not include non-state-owned small enterprises below a certain scale. Therefore, the ratio of the tabulated to official figures exceeds one in 2004.

Table 4. Regional specialization and proximity (output)

Province	1995	2004
Beijing	0.206	0.220
Tianjin	0.194	0.208
Hebei	0.212	0.219
Shanxi	0.207	0.208
Neimeng	0.198	0.214
Liaoning	0.204	0.205
Jilin	0.206	0.220
Heilongjiang	0.186	0.197
Shanghai	0.222	0.219
Jiangsu	0.210	0.210
Zhejiang	0.211	0.220
Anhui	0.204	0.211
Fujian	0.208	0.202
Jiangxi	0.200	0.206
Shandong	0.200	0.205
Henan	0.201	0.209
Hubei	0.207	0.216
Hunan	0.201	0.210
Guangdong	0.209	0.215
Guangxi	0.208	0.214
Hainan	0.201	0.207
Chongqing	0.206	0.197
Sichuan	0.198	0.202
Guizhou	0.188	0.196
Yunnan	0.187	0.197
Tibet	0.223	0.238
Shaanxi	0.191	0.192
Gansu	0.199	0.205
Qinghai	0.197	0.217
Ningxia	0.215	0.22
Xinjiang	0.190	0.199
Weighted sample average	0.206	0.211
Difference		0.005 (0.001)***

Table 5. Summary statistics of variables used in regressions

Variable	Mean	SD	Min	Max	N
<i>Panel A: county-level variables</i>					
Proximity 2004 (w=output)	0.226	0.038	0.000	0.631	2,833
Proximity 2004 (w=employment)	0.220	0.032	0.000	0.403	2,834
Proximity 2004 (w=asset)	0.226	0.035	0.091	0.397	2,833
Proximity 1995 (w=output)	0.217	0.030	0.000	0.495	2,764
Proximity 1995 (w=employment)	0.222	0.037	0.000	0.495	2,756
Proximity 1995 (w=asset)	0.218	0.031	0.000	0.495	2,765
Minimum asset 2004 (in millions)	0.178	1.536	0.000	33.001	2,860
Minimum asset 1995 (in millions)	0.101	1.454	0.000	57.603	2,791
<i>Panel B: firm-level variables</i>					
Firm age	17.601	14.358	0.000	99.000	104,324
Private%	0.146	0.340	0.000	1.000	104,324
HMT%	0.062	0.216	0.000	1.000	104,324
Other foreign%	0.025	0.139	0.000	1.000	104,324
Log(value added)	7.357	1.973	-2.591	17.253	104,324
Log(value added1)	7.396	1.992	-2.461	17.309	103,016
Log(asset)	8.933	1.941	0.693	18.235	104,324
Log(labor)	4.339	1.791	0.000	13.317	104,324
Export/sales	0.060	0.203	0.000	1.000	152,122
Exporter	0.118	0.322	0.000	1.000	152,260
Accounts receivable/revenue	0.257	0.287	0.000	1.999	93,792
Accounts payable/total debt	0.204	0.247	0.000	1.187	112,321
Debt/asset	0.639	0.316	0.000	2.997	112,321
Fixed asset/asset	0.383	0.222	0.000	1.000	112,321

Note: *HMT* stands for firms owned by Hong Kong, Marco, and Taiwan. *Year04* is a dummy variable for 2004.

Table 6. Minimum level of assets and proximity at county level

	(1)	(2)	(3)	(4)
			Dependent variable=minimum asset	
Minimum asset in 1995	0.450*** (0.040)	0.452*** (0.039)	0.448*** (0.040)	0.260*** (0.028)
Proximity in 1995 (w=asset)	-1.569** (0.785)			
Proximity in 1995 (w=labor)		-3.061*** (0.622)		
Proximity in 1995 (w=output)			-0.787 (0.801)	
Observations	2,761	2,752	2,760	2,785
R-squared	0.05	0.05	0.04	0.03

Note: *Minimum asset* is the lowest amount of assets among firms at the county level in 2004 (in millions of RMB). The symbols *, **, and *** stand for significance level at 10%, 5%, and 1%, respectively. Robust standard errors are in parentheses.

Table 7. Trade credit and proximity at firm level

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable=accounts receivable/revenue			Dependent variable=accounts payable/total debt		
Proximity_asset	0.165** (0.066)			0.222*** (0.083)		
Proximity_labor		0.297*** (0.060)			0.244*** (0.073)	
Proximity_output			0.192*** (0.064)			0.165** (0.081)
Firm age	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)
Log(sales)	-0.007*** (0.001)	-0.007*** (0.001)	-0.007*** (0.001)	-0.096*** (0.002)	-0.096*** (0.002)	-0.095*** (0.002)
Debt/asset	-0.098*** (0.005)	-0.097*** (0.005)	-0.098*** (0.005)			
Fixed asset/total asset	-0.148*** (0.007)	-0.148*** (0.007)	-0.148*** (0.007)			
Private share%	0.067*** (0.002)	0.069*** (0.002)	0.067*** (0.002)	0.065*** (0.003)	0.067*** (0.003)	0.065*** (0.003)
HMT share%	-0.035*** (0.004)	-0.035*** (0.004)	-0.035*** (0.004)	0.015*** (0.005)	0.015*** (0.005)	0.015*** (0.005)
Other foreign share%	0.230*** (0.008)	0.229*** (0.008)	0.230*** (0.008)	0.018 (0.046)	0.017 (0.046)	0.018 (0.046)
Year04	0.422*** (0.008)	0.420*** (0.008)	0.422*** (0.008)	0.012 (0.036)	0.012 (0.036)	0.012 (0.036)
Constant	0.292*** (0.019)	0.262*** (0.018)	0.286*** (0.019)	1.004*** (0.023)	0.998*** (0.022)	1.016*** (0.023)
Observations	112,324	112,321	112,324	93,793	93,792	93,793
R-squared	0.80	0.80	0.80	0.87	0.87	0.87

Note: Sample includes only firms that are surveyed in both censuses. *Accounts receivable/revenue* and *accounts payable/total debt* are used as two different measures of trade credit among firms. *HMT* stands for firms owned by Hong Kong, Marco, and Taiwan. The symbols *, **, and *** stand for significance level at 10%, 5%, and 1%, respectively. Robust standard errors are in parentheses.

Table 8. Proximity and export

	(1)	(2)	(3)	(4)	(5)	(6)
		Dependent variable=export/sales			Dependent variable=exporter	
Proximity_asset	0.215*** (0.032)			0.140*** (0.049)		
Proximity_labor		0.292*** (0.029)			0.099** (0.044)	
Proximity_output			0.207*** (0.032)			0.095** (0.048)
Firm age	0.001*** (0.000)	0.001*** (0.000)	0.001*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)	-0.000*** (0.000)
Log(sales)	0.006*** (0.001)	0.006*** (0.001)	0.006*** (0.001)	0.025*** (0.001)	0.025*** (0.001)	0.025*** (0.001)
Private share%	0.005*** (0.002)	0.005*** (0.002)	0.005*** (0.002)	0.031*** (0.003)	0.031*** (0.003)	0.032*** (0.003)
HMT share%	0.208*** (0.005)	0.207*** (0.005)	0.208*** (0.005)	0.041*** (0.007)	0.041*** (0.007)	0.041*** (0.007)
Other foreign share%	0.188*** (0.005)	0.187*** (0.005)	0.188*** (0.005)	-0.112*** (0.007)	-0.112*** (0.007)	-0.112*** (0.007)
Year04	0.048*** (0.001)	0.050*** (0.001)	0.048*** (0.001)	0.064*** (0.002)	0.065*** (0.002)	0.064*** (0.002)
Constant	-0.087*** (0.008)	-0.106*** (0.008)	-0.085*** (0.008)	-0.146*** (0.013)	-0.137*** (0.012)	-0.136*** (0.012)
Observations	152,126	152,122	152,126	152,126	152,122	152,126
R-squared	0.75	0.75	0.75	0.78	0.78	0.78

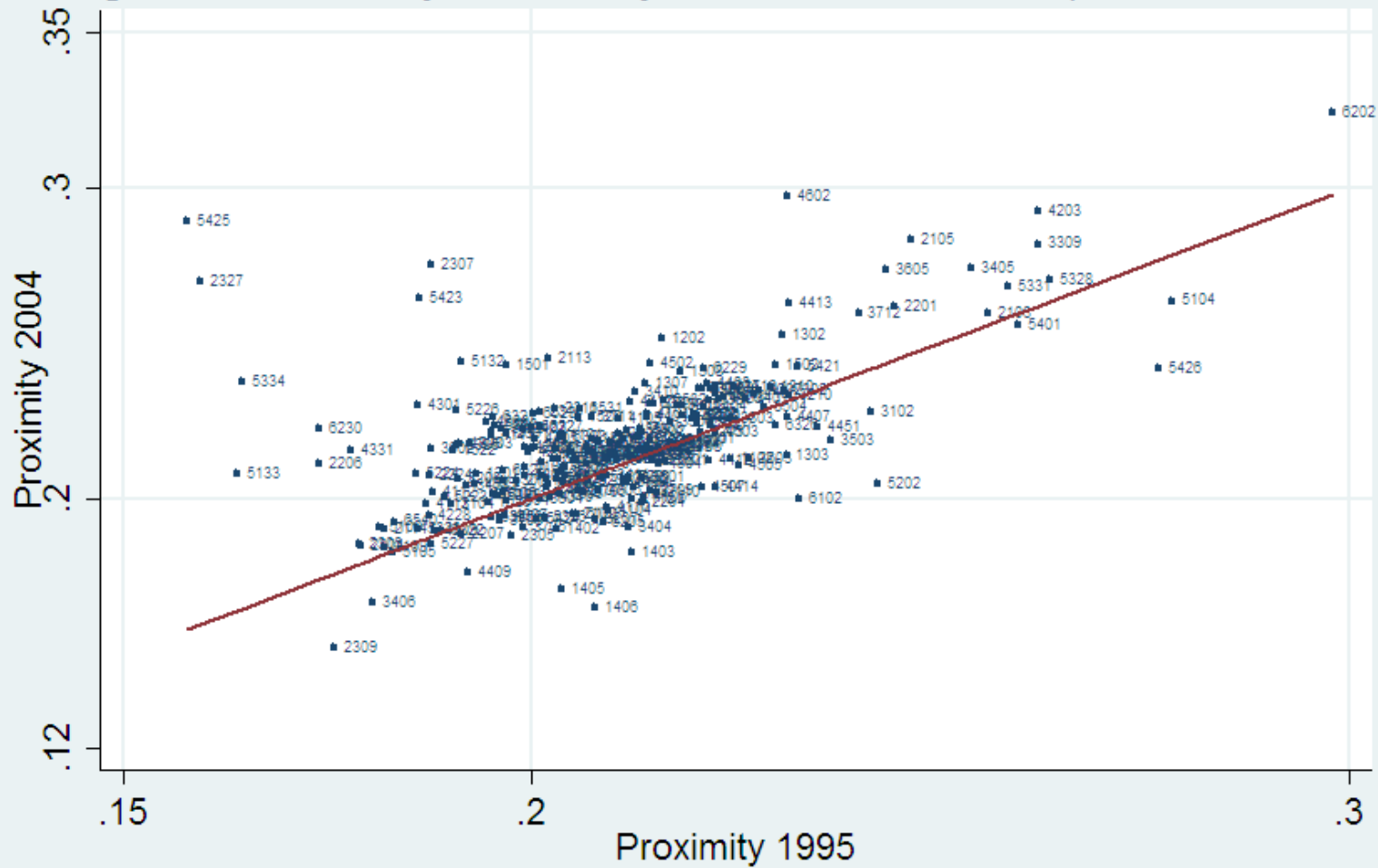
Note: Sample includes only firms that are surveyed in both censuses. *Exporter* is a dummy variable indicating whether a firm exported in a certain year. *HMT* stands for firms owned by Hong Kong, Marco, and Taiwan. *Year04* is a dummy variable for 2004. The symbols *, **, and *** stand for significance level at 10%, 5%, and 1%, respectively. Robust standard errors are in parentheses.

Table 9. Proximity and total factor productivity

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable=log(value added1)			Dependent variable=log(value added2)		
Proximity_asset	0.618** (0.250)			0.553** (0.268)		
Proximity_labor		0.536** (0.229)			0.384 (0.245)	
Proximity_output			0.564** (0.250)			0.598** (0.268)
Log(labor)	0.068*** (0.004)	0.068*** (0.004)	0.068*** (0.004)	0.066*** (0.004)	0.065*** (0.004)	0.065*** (0.004)
Log(asset)	0.800*** (0.007)	0.801*** (0.007)	0.801*** (0.007)	0.804*** (0.008)	0.804*** (0.008)	0.804*** (0.008)
Log(labor)*year04	0.262*** (0.008)	0.262*** (0.008)	0.262*** (0.008)	0.239*** (0.009)	0.240*** (0.009)	0.239*** (0.009)
Log(asset)*year04	-0.179*** (0.006)	-0.179*** (0.006)	-0.179*** (0.006)	-0.155*** (0.007)	-0.155*** (0.007)	-0.155*** (0.007)
Firm age	0.001** (0.000)	0.001** (0.000)	0.001** (0.000)	0.001** (0.001)	0.001** (0.001)	0.001** (0.001)
Private share%	0.083*** (0.015)	0.083*** (0.015)	0.084*** (0.015)	0.074*** (0.016)	0.074*** (0.016)	0.075*** (0.016)
HMT share%	0.303*** (0.030)	0.302*** (0.030)	0.303*** (0.030)	0.277*** (0.032)	0.276*** (0.032)	0.277*** (0.032)
Other foreign share%	0.683*** (0.029)	0.682*** (0.029)	0.683*** (0.029)	0.660*** (0.031)	0.660*** (0.031)	0.660*** (0.031)
Year04	0.545*** (0.038)	0.551*** (0.038)	0.546*** (0.038)	0.415*** (0.041)	0.420*** (0.041)	0.415*** (0.041)
	-0.059 (0.084)	-0.043 (0.081)	-0.047 (0.084)	-0.011 (0.090)	0.025 (0.087)	-0.021 (0.090)
Observations	104,437	104,437	104,437	103,128	103,128	103,128
R-squared	0.95	0.95	0.95	0.94	0.94	0.94

Note: Sample includes only firms that are surveyed in both censuses. *Value added1* is computed as a weighted average of value added constructed from the output approach and value added computed from the income approach; *value added2* is computed based only on the output approach. *HMT* stands for firms owned by Hong Kong, Marco, and Taiwan. *Year04* is a dummy variable for 2004. The symbols *, **, and *** stand for significance level at 10%, 5%, and 1%, respectively. Robust standard errors are in parentheses.

Figure 1: Industry Proximity within Prefecture (1995 vs. 2004)



Data source: Authors' calculation based on CIC 1995 and CEC 2004

Appendix Table 1. Effects of proximity on firm size distribution

	(1)	(2)	(3)	(4)	(5)	(6)
	Avg firm size_output	Avg firm size_labor	Avg firm size_asset	Firm size Gini_output	Firm size Gini_labor	Firm size Gini_asset
Proximity_output	13,838.610** (6,492.289)			-0.209*** (0.051)		
Proximity_labor		322.774*** (61.811)			-0.316*** (0.051)	
Proximity_asset			12,097.348 (11,206.631)			-0.368*** (0.050)
Yr2004	2,729.672*** (244.740)	-49.390*** (2.418)	3,075.987*** (398.439)	0.043*** (0.002)	-0.072*** (0.002)	0.047*** (0.002)
Constant	1,773.701 (1,416.099)	17.436 (13.822)	4,464.986* (2,460.894)	0.825*** (0.011)	0.893*** (0.011)	0.855*** (0.011)
Observations	5,499	5,495	5,500	5,464	5,495	5,500
R-squared	0.71	0.58	0.68	0.76	0.72	0.81

Note: The symbols *, **, and *** stand for significance level at 10%, 5%, and 1%, respectively. Robust standard errors are in parentheses. **[correct?]**

Appendix Table 2. Effects of proximity on firm number distribution

	(1)	(2)	(3)	(4)	(5)	(6)
	Avg firm number_output	Avg firm number_labor	Avg firm number_asset	Firm number Gini_output	Firm number Gini_labor	Firm number Gini_asset
Proximity_output	-278.058 (428.162)			-0.345*** (0.062)		
Proximity_labor		281.563 (400.453)			-0.268*** (0.058)	
Proximity_asset			-66.682 (450.255)			-0.368*** (0.065)
Yr2004	304.859*** (16.140)	303.502*** (15.668)	302.955*** (16.008)	0.144*** (0.002)	0.140*** (0.002)	0.143*** (0.002)
Constant	243.717*** (93.391)	120.951 (89.545)	198.066** (98.873)	0.465*** (0.014)	0.450*** (0.013)	0.471*** (0.014)
Observations	5,499	5,495	5,500	5,499	5,495	5,500
R-squared	0.65	0.65	0.65	0.86	0.86	0.86

Note: The symbols *, **, and *** stand for significance level at 10%, 5%, and 1%, respectively. Robust standard errors are in parentheses. **[correct?]**